Some Properties of Bit Decoding Algorithms Over A Generalized Channel Model

Ali Abedi and A. K. Khandani
Coding & Signal Transmission Lab.
Dept. of Elec. and Comp. Eng.,
University of Waterloo,
Waterloo, ON, N2L 3G1
{ali,khandani}@cst.uwaterloo.ca

Abstract — **In this paper[1], we study certain properties of the bit decoding algorithms for the case of linear binary block codes. Our focus is on the probability distribution function (*pdf*) of the bit Log-Likelihood-Ratio (*LLR*). We consider a general channel model 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 origin). For the case of channels with single bit input a sufficient condition for two bit positions to have the same *pdf* is presented.**

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 are known to minimize the probability of the *Frame Error Rate (FER)* under the mild condition that the code-words occur with equal probability. 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. This results in minimizing the value of the *Bit Error Rate (BER)*. Note that unlike ML algorithms, in the case of the bit decoding algorithms the collection of decoded bits do not necessarily form a valid code-word.

The straightforward approach to bit (or more generally symbol in the case of non binary codes) decoding is based on summing up the probabilities of different code-words according to the value of their component in a given position under consideration. A number of research works have addressed the problem of finding bit decoding algorithms of a reduced complexity (as compared to an exhaustive method) assuming a soft decision at the channel output. An optimum symbol decoding rule is proposed in [1] which is still exhaustive, but uses the set of code-words of the dual code in the decoding process. This method results in a lower complexity as compared to an exhaustive search if the dual code has a smaller number of code-words. Two modifications of the basic exhaustive method are presented in [2]. It gives a set of necessary and sufficient conditions for achieving minimum symbol error probability decoding and uses these conditions to derive a non-exhaustive optimum decoding algorithm of a reduced complexity. Bit-by-bit soft-decision decoding of binary cyclic codes is considered in [3] where the authors have modified the optimum decoding rule so as to reduce the complexity while maintaining good performance.

The problem of decoding linear block code used over a binary symmetrical channel with a given cross over probability is considered in [4–7] where the objective is to minimize the probability of an information bit error. Reference [4] gives an optimal rule for selecting coset leaders. Seguin [5] notes that the probability of an information symbol being in error is a function of the generator matrix chosen when the decoder is fixed. Seguin shows how to choose the optimal generator matrix when a fixed standard array code has been selected. A more difficult problem of simultaneously choosing the generator matrix and decoder that minimizes the probability of an information bit error is considered by Dunning [6]. Tolhuizen and van Gils [7] show that the large number of computations required for Dunning's procedure can be reduced somewhat by using the automorphism group of the code. Some authors [8,9] have considered specific coset leader rules for use when cross over probability of BSC is small and the encoding is systematic. Elia and Prati [8] give a decoding strategy, and note that for some codes it outperforms minimum weight coset leaders for small cross over probability. Montgomery and Vijaya Kumar [9] give another improved (though still sub-optimal) decoding strategy.

In an early paper Posner examines the information bit error probability obtained by using a linear block code over an Additive White Gaussian Noise (AWGN) channel with low signal to noise ratio and hard decision decoding [10]. More recently, in [11] the performance of linear block codes is examined when used on AWGN channel with soft decision decoding. Some asymptotic expressions are derived in [12] for bit error probability under optimum decoding for the AWGN channels. There have been also some works on bounds and approximation on the bit error probabilities of decoding convolutional codes [13] and trellis codes [14].

Maximum Likelihood decoding algorithms have been the subject of numerous research activities, while bit decoding algorithms have received much less attention in the past. The reason being that the bit decoding algorithms are known to offer a BER performance very close to that of ML algorithms, while they have a substantially higher level of decoding complexity. More recently bit decoding have received increasing

[1] This work is financially supported by Natural Sciences and Engineering Research Council of Canada (NSERC) and by Communications and Information Technology Ontario (CITO).

attention, mainly due to the fact that they deliver soft output decision (reliability information) which can be advantageously exploited in both uncoded and coded systems. In 1993, a new class of channel codes, called Turbo-codes [15], were announced which have an astonishing performance and at the same time allow for a simple bit decoding algorithm. Due to the importance of Turbo-codes, there has been a growing interest among communication researchers to work on the bit decoding algorithms. Reference [16] provides a method (known as BCJR) to compute the bit probabilities of a given code using its trellis diagram. An efficient exact APP decoding algorithm based on coset decoding principle proposed in [17]. There are also some special optimum methods for bit decoding of linear block codes based on sectionalized trellis diagrams [18] and based on using the code-words of the dual code [19]. The main simplification of BCJR has been the SOVA (soft output Viterbi Algorithm) of Hoeher and Hagenauer [20] which is a sub-optimum solution. A reduced-search BCJR algorithm is also proposed in [21]. Other researches have been done on reducing complexity of bit decoding like early detection and trellis splicing in [22].

More recently, in [23], Abedi and Khandani present an analytical method for performance evaluation of binary linear block codes using an AWGN channel model with Binary Phase Shift Keying (BPSK) modulation. The computation is based on using the probability distribution function of the bit *LLR* which is expressed in terms of the Gram-Charlier series expansion. This expansion requires knowledge of the statistical moments of the bit *LLR*. An analytical method is presented for calculating these moments using some straight-forward recursive calculations involving certain weight enumerating functions of the code. It is shown that the estimate of the bit error probability provided by the proposed method will asymptotically converge to the true bit error performance. Numerical results are provided for the (15,11) Cyclic code which demonstrate close agreement with the simulation results.

This paper is organized as follows, in section II the model used to analyze the problem is presented. All notations and assumptions are in this section. Some examples of our general channel model are presented in section III. We prove some theorems on bit decoding algorithms in section IV. This work is a continuation of [24] in which the case of AWGN channel with BPSK modulation is considered.

II. Modeling

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

$$\begin{aligned} \forall \mathbf{c}^i \in \mathcal{C} : &\text{if } \ c_k^i = 0 \Longrightarrow \mathbf{c}^i \in \ C_k^0 \\ &\text{if } \ c_k^i = 1 \Longrightarrow \mathbf{c}^i \in \ C_k^1 \\ C_k^0 \cup C_k^1 = \mathcal{C}, &\quad C_k^0 \cap C_k^1 = \varnothing \end{aligned} \tag{1}$$

We define bit wise binary addition of two code-words on our code book as, $\mathbf{c}^i \oplus \mathbf{c}^j$. Note that the sub-code C_k^0 is closed under binary addition and all code-words are equally probable to be sent. Each code-word will be partitioned into L blocks of m bits, i.e., $N = mL$, 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 $(\mathbf{I}_1^i, \mathbf{I}_2^i, \ldots, \mathbf{I}_L^i)$. Note that 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 crucial condition is that the channels in different transmissions should be independent of each other (memoryless channel). For simplicity of modeling, we use fixed length m-blocks in our following discussions and 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 $\mathcal{I}$ form a group under binary addition.

The channel has 2^m discrete input and discrete or continuous output as shown in Figure 1. For the continuous output channels, $\mathcal{O} \subset \Re^n$, where $\Re$ is the set of real numbers and small $p(.)$ stands for *pdf*. For channels with discrete output, $\mathcal{O}$ is a set of discrete alphabets and small $p(.)$ stands for probability mass function (*pmf*) [2].

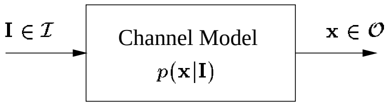

$\mathbf{I} \in \mathcal{I}$
Channel Model
$p(\mathbf{x}|\mathbf{I})$
$\mathbf{x} \in \mathcal{O}$

Figure 1: Channel Model

Consider the situation of sending a code-word $\tilde{\mathbf{c}}$ through the channel. Each m-block of the code-word $\tilde{\mathbf{c}}$ will be transmitted and a symbol $\mathbf{x}_j$ will be received at each channel use.

A common tool to express the bit probabilities in the bit decoding algorithms is based on using the so-called Log-Likelihood-Ratio (*LLR*). The *LLR* of the k^{th} bit is defined by the following equation.

$$LLR(k) = \log \frac{p(\tilde{c}_k = 1|\mathbf{x}_1 \ldots \mathbf{x}_L)}{p(\tilde{c}_k = 0|\mathbf{x}_1 \ldots \mathbf{x}_L)} = \log \frac{\sum\limits_{\mathbf{c}^i \epsilon C_k^1} p(\mathbf{c}^i|\mathbf{x}_1 \ldots \mathbf{x}_L)}{\sum\limits_{\mathbf{c}^i \epsilon C_k^0} p(\mathbf{c}^i|\mathbf{x}_1 \ldots \mathbf{x}_L)} \tag{2}$$

where $\tilde{c}_k$ is the value of k^{th} bit in the transmitted code-word, $\tilde{\mathbf{c}}$. For the case of equally likely code-words, we have,

$$LLR(k) = \log \frac{\sum\limits_{\mathbf{c}^i \epsilon C_k^1} p(\mathbf{x}_1 \ldots \mathbf{x}_L|\mathbf{c}^i)}{\sum\limits_{\mathbf{c}^i \epsilon C_k^0} p(\mathbf{x}_1 \ldots \mathbf{x}_L|\mathbf{c}^i)} = \log \frac{\sum\limits_{\mathbf{c}^i \epsilon C_k^1} \prod\limits_{j=1}^{L} p(\mathbf{x}_j|\mathbf{I}_j^i)}{\sum\limits_{\mathbf{c}^i \epsilon C_k^0} \prod\limits_{j=1}^{L} p(\mathbf{x}_j|\mathbf{I}_j^i)} \tag{3}$$

Given the value of bit *LLR*, decision on the value of bit k is made by comparing the *LLR* with a threshold of zero. 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 resulting probability distribution as long as the value of the k^{th} bit remains unchanged.

We consider two different channel models both with discrete input alphabet:

[2] Note that without loss of generality we have combined modulation and channel into a new channel model.

- Channels with Geometrical Representation
- Channels without Geometrical Representation

We define an endomorphism $\phi_{\mathbf{c}}$ on code $\mathcal{C}$ which permutes the code-words by adding code-word $\mathbf{c}$ to them,

$$\begin{aligned} \phi_{\mathbf{c}} &: \mathcal{C} \longrightarrow \mathcal{C} \\ \phi_{\mathbf{c}} &: \mathbf{c}^i \longrightarrow \mathbf{c}^i \oplus \mathbf{c} \end{aligned} \qquad (4)$$

This mapping also change each m-block within a code-word to another m-block. The following sufficient conditions are required to carry out the proofs.

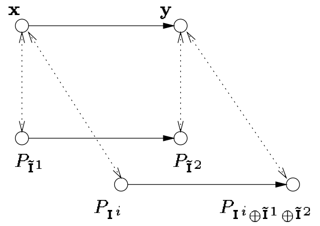
$$\begin{aligned} &\forall \tilde{\mathbf{I}}^1, \tilde{\mathbf{I}}^2 \in \mathcal{I},\ \forall \mathbf{x} \in \mathcal{O},\ \exists \mathbf{y} \in \mathcal{O} : \\ &p(\mathbf{x}|\mathbf{I} \oplus \tilde{\mathbf{I}}^1 \oplus \tilde{\mathbf{I}}^2) = p(\mathbf{y}|\mathbf{I}),\ \ \forall \mathbf{I} \in \mathcal{I} \end{aligned} \qquad (5)$$

In the following subsections we inspect the above mentioned different channel models.

A Channels with Geometrical Representation

In this section, we consider channels for which input and output symbols belong to $\Re^n$. We use the notation $\mathbf{P}_{\mathbf{I}^i} \in \Re^n$ to refer to the channel input symbols representing $\mathbf{I}^i$.

In other words, the m-blocks are just labels of the points in the Euclidean space. We assume that the signal set at the channel input is geometrically uniform [25]. 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 scenario shown in Figure 2 is valid for the corresponding labels.



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

Figure 2: Mapping of points with an isometry

It is easy to see that under the following conditions, the condition given in 5 will be satisfied, if

(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 $\|\mathbf{x} - \mathbf{P}_{\tilde{\mathbf{I}}^1}\|$

A good example for a channel satisfying condition (ii) is the AWGN channel.

B Channels without Geometrical Representation

In this section, we assume that the channel output set is a discrete set composed of elements $\mathbf{x}^j \in \mathcal{O}$. We define matrix $\mathbf{A}$ whose elements a_{ij} are conditional *pdf* of channel output given its input.

$$\mathbf{A}_{u \times v} = [a_{ij}], \quad a_{ij} = p(\mathbf{x}^j|\mathbf{I}^i), \quad u = 2^m = |\mathcal{I}|,\ v = |\mathcal{O}| \quad (6)$$

We can satisfy the condition given in (5) if after shuffling 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 shuffled in the same order as the corresponding m-blocks.

III. Examples of Discrete Channel Model

In this section, we present some examples of the discrete channel models which satisfy the required condition stated in (5).

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

$$\mathbf{A} = \left[\begin{array}{c|cccc} & \mathbf{x}^1 & \mathbf{x}^2 & \mathbf{x}^3 & \mathbf{x}^4 \\ \hline 0 & 1/2-\epsilon_1 & \epsilon_1 & 1/2-\epsilon_2 & \epsilon_2 \\ 1 & \epsilon_1 & 1/2-\epsilon_1 & \epsilon_2 & 1/2-\epsilon_2 \end{array}\right] \quad (7)$$

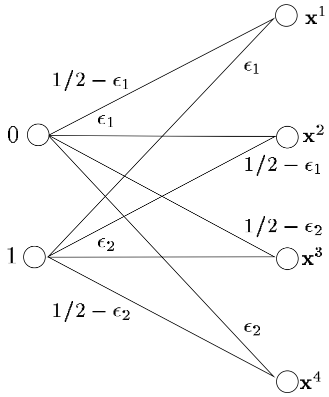


Figure 3: Channel model for example 1.

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

$$\mathbf{A} = \left[\begin{array}{c|ccccc} & \mathbf{x}^1 & \mathbf{x}^2 & \mathbf{x}^3 & \mathbf{x}^4 & \mathbf{x}^5 \\ \hline 00 & e_0 & e_1 & \epsilon & 0 & 0 \\ 01 & e_1 & e_0 & \epsilon & 0 & 0 \\ 10 & 0 & 0 & \epsilon & e_0 & e_1 \\ 11 & 0 & 0 & \epsilon & e_1 & e_0 \end{array}\right] \quad (8)$$

where $e_0 + e_1 + \epsilon = 1$. This matrix has the required symmetry properties stated in (5).

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

$$\mathbf{A} = \left[\begin{array}{c|cccccccc} & \mathbf{x}^1 & \mathbf{x}^2 & \mathbf{x}^3 & \mathbf{x}^4 & \mathbf{x}^5 & \mathbf{x}^6 & \mathbf{x}^7 & \mathbf{x}^8 \\ \hline 00 & e_0 & e_1 & e_2 & e_3 & e_4 & e_3 & e_2 & e_1 \\ 01 & e_2 & e_1 & e_0 & e_1 & e_2 & e_3 & e_4 & e_3 \\ 10 & e_4 & e_3 & e_2 & e_1 & e_0 & e_1 & e_2 & e_3 \\ 11 & e_2 & e_3 & e_4 & e_3 & e_2 & e_1 & e_0 & e_1 \end{array}\right] \quad (9)$$

where $e_0 + 2(e_1 + e_2 + e_3) + e_4 = 1$. In this example the signal points are labeled using Ungerboeck set partitioning [26]. It is easy to see that the required condition for the columns of the probability matrix can be satisfied.

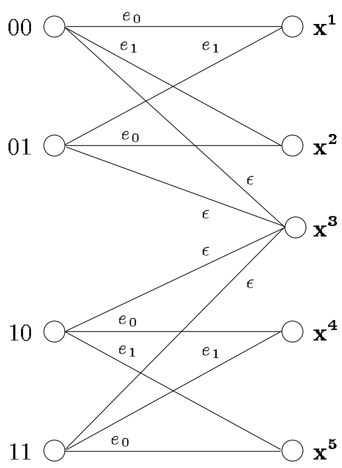


Figure 4: Channel model for example 2.

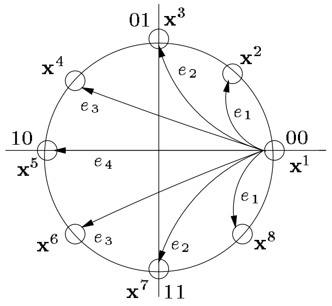


Figure 5: Channel model for example 3: The values of error probabilities which are not shown follow the same pattern as the values specified on the figure.

IV. Main Results

Using the above definitions and notation, we have the following theorems.

Theorem 1 *The probability distribution of $LLR(k)$ 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: If we take two different code-words $\tilde{\mathbf{c}}^1, \tilde{\mathbf{c}}^2$ and transmit them over the channel by partitioning them to $(\tilde{\mathbf{I}}_1^1, \ldots, \tilde{\mathbf{I}}_L^1)$ and $(\tilde{\mathbf{I}}_1^2, \ldots, \tilde{\mathbf{I}}_L^2)$, we will receive $(\mathbf{x}_1 \ldots \mathbf{x}_L)$ and $(\mathbf{y}_1 \ldots \mathbf{y}_L)$, respectively. We form their bit LLR for the k^{th} bit position using (3),

$$LLR^1(k) = \log \frac{\sum_{\mathbf{c}^i \in C_k^1} \prod_{j=1}^{L} p(\mathbf{x}_j|\mathbf{I}_j^i)}{\sum_{\mathbf{c}^i \in C_k^0} \prod_{j=1}^{L} p(\mathbf{x}_j|\mathbf{I}_j^i)} \tag{10}$$

$$LLR^2(k) = \log \frac{\sum_{\mathbf{c}^i \in C_k^1} \prod_{j=1}^{L} p(\mathbf{y}_j|\mathbf{I}_j^i)}{\sum_{\mathbf{c}^i \in C_k^0} \prod_{j=1}^{L} p(\mathbf{y}_j|\mathbf{I}_j^i)} \tag{11}$$

where $\mathbf{I}_j^i$ are m-blocks of $\mathbf{c}^i$. As long as the value of the k^{th} bit remains unchanged both $\tilde{\mathbf{c}}^1$, $\tilde{\mathbf{c}}^2$ are in the same subset namely C_k^0 or C_k^1. No matter they are in which subset, $\tilde{\mathbf{c}}^1 \oplus \tilde{\mathbf{c}}^2$ will be in sub-code C_k^0. Noting that $\tilde{\mathbf{c}}^1 \oplus \tilde{\mathbf{c}}^2 \in C_k^0$, we use $\phi_{\tilde{\mathbf{c}}^1 \oplus \tilde{\mathbf{c}}^2}$ as defined in (4), to map the sub-code C_k^0 onto itself as seen in Figure 6.

$$\begin{aligned} \phi_{\tilde{\mathbf{c}}^1 \oplus \tilde{\mathbf{c}}^2} &: C_k^0 \longrightarrow C_k^0 \\ \phi_{\tilde{\mathbf{c}}^1 \oplus \tilde{\mathbf{c}}^2} &: \mathbf{c}^i \longrightarrow \mathbf{c}^i \oplus \tilde{\mathbf{c}}^1 \oplus \tilde{\mathbf{c}}^2 \end{aligned} \tag{12}$$

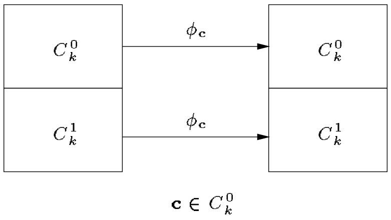


Figure 6: Mapping of C_k^0 onto itself

Consider this mapping applied to the denominator of (10).

$$\sum_{\mathbf{c}^i \in C_k^0} \prod_{j=1}^{L} p(\mathbf{x}_j|\mathbf{I}_j^i) \longrightarrow \sum_{\mathbf{c}^i \in C_k^0} \prod_{j=1}^{L} p(\mathbf{x}_j|\mathbf{I}_j^i \oplus \tilde{\mathbf{I}}_j^1 \oplus \tilde{\mathbf{I}}_j^2) \tag{13}$$

where $\tilde{\mathbf{I}}_j^1 \oplus \tilde{\mathbf{I}}_j^2$ is the m-block which is added to each m-block of $\mathbf{c}^i$ as a direct result of the mapping $\phi_{\tilde{\mathbf{c}}^1 \oplus \tilde{\mathbf{c}}^2}$. Assuming that the properties of channel defined in section II hold, there exists a $\mathbf{y}_j$ such that,

$$p(\mathbf{y}_j|\mathbf{I}_j^i) = p(\mathbf{x}_j|\mathbf{I}_j^i \oplus \tilde{\mathbf{I}}_j^1 \oplus \tilde{\mathbf{I}}_j^2) \tag{14}$$

This means,

$$\sum_{\mathbf{c}^i \in C_k^0} \prod_{j=1}^{L} p(\mathbf{x}_j|\mathbf{I}_j^i) \longrightarrow \sum_{\mathbf{c}^i \in C_k^0} \prod_{j=1}^{L} p(\mathbf{y}_j|\mathbf{I}_j^i) \tag{15}$$

The right hand side expression is the denominator of (11). Applying the mapping $\phi_{\tilde{\mathbf{c}}^1 \oplus \tilde{\mathbf{c}}^2}$ to the numerator of (10), the elements of coset shuffle with the same permutation and the numerator of (10) will be mapped to the numerator of (11). From (5) we obtain,

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

Noting that each transmission is independent of the others we conclude that,

$$\prod_{j=1}^{L} p(\mathbf{x}_j|\tilde{\mathbf{I}}_j^1) = \prod_{j=1}^{L} p(\mathbf{y}_j|\tilde{\mathbf{I}}_j^2) \tag{17}$$

$$p(\mathbf{x}_1 \ldots \mathbf{x}_L|\tilde{\mathbf{I}}_1^1 \ldots \tilde{\mathbf{I}}_L^1) = p(\mathbf{y}_1 \ldots \mathbf{y}_L|\tilde{\mathbf{I}}_1^2 \ldots \tilde{\mathbf{I}}_L^2) \tag{18}$$

$$p(\mathbf{x}_1 \ldots \mathbf{x}_L|\tilde{\mathbf{c}}^1) = p(\mathbf{y}_1 \ldots \mathbf{y}_L|\tilde{\mathbf{c}}^2) \tag{19}$$

Note that the value of $LLR^1(k)$ is uniquely determined by vector $(\mathbf{x}_1 \ldots \mathbf{x}_L)$ and the value of $LLR^2(k)$ is uniquely determined by vector $(\mathbf{y}_1 \ldots \mathbf{y}_L)$. It is shown that for each $(\mathbf{x}_1 \ldots \mathbf{x}_L)$ there exists a $(\mathbf{y}_1 \ldots \mathbf{y}_L)$ with the same conditional probability as given in (19) which preserves the value of LLR. This means (10), (11) posses the same *pdf* independent of the transmitted code-word. ■

The following theorem explains the effect of a change in the specific value taken by bit k on the probability distribution of $LLR(k)$.

Theorem 2 *The probability distribution of $LLR(k)$ for value of bit $k = 0$ or 1 are the reflections of one another through the origin (threshold point).*

Proof: Having chosen two different code-words $\tilde{\mathbf{c}}^1$, $\tilde{\mathbf{c}}^2$ (which have different values in their kth bit position), we form their bit LLR for the k^{th} bit position as given in (10),(11). Assume that the elements of C_k^0 are mapped by adding a code-word $\tilde{\mathbf{c}}^1 \oplus \tilde{\mathbf{c}}^2$ to them which contains a 1 in position k. This will change the value of bit k from zero to one. This operation results in each component in the set of the code-words $\mathbf{c}^i \in C_k^0$, to be exchanged with a counterpart element within the set of the code-words $\mathbf{c}^i \in C_k^1$. In this case, we exchange the values of numerator and denominator. Changing the value of the k^{th} bit means $\tilde{\mathbf{c}}^1$, $\tilde{\mathbf{c}}^2$ are in two different subsets. No matter which one is in which subset, $\tilde{\mathbf{c}}^1 \oplus \tilde{\mathbf{c}}^2$ will be in coset C_k^1. Noting that $\tilde{\mathbf{c}}^1 \oplus \tilde{\mathbf{c}}^2 \in C_k^1$, we use $\phi_{\tilde{\mathbf{c}}^1 \oplus \tilde{\mathbf{c}}^2}$ as defined in (4), to map the sub-code C_k^0 onto coset C_k^1 as seen in Figure 7.

$$\begin{aligned} \phi_{\tilde{\mathbf{c}}^1 \oplus \tilde{\mathbf{c}}^2} &: C_k^0 \longrightarrow C_k^1 \qquad (20) \\ \phi_{\tilde{\mathbf{c}}^1 \oplus \tilde{\mathbf{c}}^2} &: \mathbf{c}^i \longrightarrow \mathbf{c}^i \oplus \tilde{\mathbf{c}}^1 \oplus \tilde{\mathbf{c}}^2 \end{aligned}$$

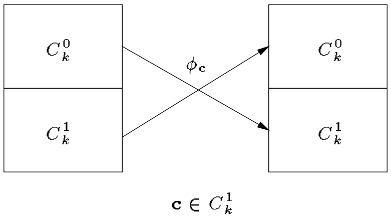


Figure 7: Mapping of C_k^0 onto C_k^1

Consider this mapping applied to the denominator of (10).

$$\sum_{\mathbf{c}^i \epsilon C_k^0} \prod_{j=1}^{L} p(\mathbf{x}_j | \mathbf{I}_j^i) \longrightarrow \sum_{\mathbf{c}^i \epsilon C_k^1} \prod_{j=1}^{L} p(\mathbf{x}_j | \mathbf{I}_j^i \oplus \tilde{\mathbf{I}}_j^1 \oplus \tilde{\mathbf{I}}_j^2) \qquad (21)$$

Assuming that the properties of channel defined in section II hold, there exists $\mathbf{y}_j$ such that,

$$p(\mathbf{y}_j | \mathbf{I}_j^i) = p(\mathbf{x}_j | \mathbf{I}_j^i \oplus \tilde{\mathbf{I}}_j^1 \oplus \tilde{\mathbf{I}}_j^2) \qquad (22)$$

This means,

$$\sum_{\mathbf{c}^i \epsilon C_k^0} \prod_{j=1}^{L} p(\mathbf{x}_j | \mathbf{I}_j^i) \longrightarrow \sum_{\mathbf{c}^i \epsilon C_k^1} \prod_{j=1}^{L} p(\mathbf{y}_j | \mathbf{I}_j^i) \qquad (23)$$

The right hand side expression is the numerator of (11). Applying the mapping $\phi_{\tilde{\mathbf{c}}^1 \oplus \tilde{\mathbf{c}}^2}$ to the numerator of (10), the elements of coset shuffle with the same permutation and the numerator of (10) will be mapped to the denominator of (11).

It is shown that for each $(\mathbf{x}_1 \dots \mathbf{x}_L)$ there exists a $(\mathbf{y}_1 \dots \mathbf{y}_L)$ with the same conditional probability as given in (19) which exchanges the value of numerator and denominator of $LLR^1(k)$. This means (10), (11) posses the same *pdf* independent of the transmitted code-word, while their values are different only in their signs. Under these conditions, it is easy to see that, the given value of $LLR(k)$ if k^{th}bit $= 0$ occurs with the same probability as $-LLR(k)$ if k^{th}bit $= 1$, and vice versa. Therefore, changing the value of bit k in the transmitted code-word is equivalent to inverting the sign of the random variable corresponding to $LLR(k)$.

■

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 (no other conditions of the form given in 5 are required). Note that unlike in 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. We define a permutation $\mathcal{P}$ which simply permutes the elements of each code-word. The set of permutations which maps the code-book $\mathcal{C}$ onto itself form a group called Automorphism group of code $\mathcal{C}$.

Theorem 3 *Consider two bit positions of a code-word, a, b such that $1 \leq a, b \leq N$, $a \neq b$. If there exists a permutation $\mathcal{P}$ within Automorphism group of code $\mathcal{C}$ which transfers bit position a to b, then $LLR(a)$ and $LLR(b)$ possess the same probability distribution.*

Proof: Consider the situation of sending $\tilde{\mathbf{c}}^1$ bit by bit and receiving $\mathbf{x}_j$ for bit $\tilde{\mathbf{I}}_j^1$. The permutation mentioned in the theorem will be used as follows,

$$\begin{aligned} \mathcal{P} &: \tilde{\mathbf{c}}^1 \longrightarrow \tilde{\mathbf{c}}^2 & (24) \\ \mathcal{P} &: (\tilde{\mathbf{I}}_1^1, \dots, \tilde{\mathbf{I}}_L^1) \longrightarrow (\tilde{\mathbf{I}}_1^2, \dots, \tilde{\mathbf{I}}_L^2) & (25) \\ \mathcal{P} &: (\mathbf{x}_1, \dots, \mathbf{x}_L) \longrightarrow (\mathbf{y}_1, \dots, \mathbf{y}_L) & (26) \end{aligned}$$

Using the permutation $\mathcal{P}$, the arguments of the summation in the numerator of (10) will be mapped to the arguments of the summation in the numerator of (11), and similarly the arguments of the summation in the denominators of (10) and (11) will be mapped to each other. Noting that the channel is memoryless and is not changing with time, we have,

$$\prod_{j=1}^{L} p(\mathbf{x}_j | \tilde{\mathbf{I}}_j^1) = \prod_{j=1}^{L} p(\mathbf{y}_j | \tilde{\mathbf{I}}_j^2) \qquad (27)$$

The rest of the proof follows similar to theorem 1. ■

Note that set of permutations form a group, it is clear that inverse of $\mathcal{P}$ exists and transfers bit position b to a. The existence of the permutation to yield two bit positions with the same probability distribution for their LLR is our next concern. We apply this result to the class of cyclic codes as a good example for checking the existence of the desired permutation.

Theorem 4 *The permutation mentioned in theorem 3 exists for the class of cyclic codes.*

Proof: Transferring bit position a to b ($a \leq b$) is achievable by shifting elements of the code-words $b - a$ times to the right. It is the property of cyclic codes that any such shift results in another code-word. Hence, this permutation in automorphism group of the code $\mathcal{C}$ exists for the case of cyclic codes. ■

V. Conclusion

In this paper the probabilistic behavior of 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 of 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 value of a bit position will not change the shape of the *pdf* of that bit LLR, while make it reflect through the origin. Some examples have been provided for channels with discrete input and output which have the required symmetry properties. For the case of channels with binary input, it is shown that a sufficient condition for two bit positions, a, b, to have the same *pdf* is the existence of a permutation in the automorphism group of the code which can transfer bit position a to b. We showed that this permutation exists for the class of cyclic codes.

References

[1] C.R.P. Hartmann, L.D. Rudolph ,"An Optimum Symbol-by-Symbol Decoding Rule for Linear Codes," *IEEE Transactions on Information Theory*, Vol.22, No.5, pp. 514-517, September 1976.

[2] T.Y. Hwang ,"Decoding Linear Block Codes for Minimizing Word Error Rate," *IEEE Transactions on Information Theory*, Vol.25, No.6, pp. 733-737, November 1979.

[3] L.D. Rudolph, C.R.P. Hartmann, T.Y. Hwang, N.Q. Duc ,"Algebraic Analog Decoding of Linear Binary Codes," *IEEE Transactions on Information Theory*, Vol.25, No.4, pp. 430-440, July 1979.

[4] A.B. Kiely, J.T. Coffey, M.R. Bell ,"Optimal Information Bit Decoding of Linear Block Codes," *IEEE Transactions on Information Theory*, Vol.41, No.1, pp. 130-140, January 1995.

[5] G. Seguin ,"Optimal Symbol Error Rate Encoding," *IEEE Transactions on Information Theory*, Vol.32, No.2, pp. 319-322, March 1986.

[6] L.A. Dunning ,"Encoding and Decoding for the Minimization of Message Symbol Error Rates in Linear Block Codes," *IEEE Transactions on Information Theory*, Vol.33, No.1, pp. 91-104, January 1987.

[7] L.M.G.M. Tolhuizen, W.J. van Gils ,"A Large Automorphism Group Decreases the Number of Computations in the Construction of an Optimal Encoder/Decoder Pair for a Linear Block Code," *IEEE Transactions on Information Theory*, Vol.34, No.2, pp. 333-338, March 1988.

[8] M. Elia , G. Prati ,"On the Complete Decoding of Binary Linear Codes," *IEEE Transactions on Information Theory*, Vol.31, No.4, pp. 518-520, July 1985.

[9] B.L. Montgomery, B.V.K. Vijaya Kumar ,"On Decoding Rules to Minimize the Probability of Information Bit Errors," *IEEE Transactions on Information Theory*, Vol.34, No.4, pp. 880-881, July 1988.

[10] E.C. Posner,"Properties of Error-Correcting Codes at Low Signal-to-Noise Ratios," *SIAM J. of App. Math.*, Vol.15, No.4, pp. 775-798, July 1967.

[11] Chi-Chao Chao, R.J. McEliece, Laif Swanson ,E.R. Rodemich ,"Performance of Binary Block Codes at Low Signal-to-Noise Ratios," *IEEE Transactions on Information Theory*, Vol. 38, No.6, pp. 1677-1687, November 1992.

[12] C.R.P. Hartmann, L.D. Rudolph, K.G. Mehrotra ,"Asymptotic Performance of Optimum Bit-by-Bit Decoding for the White Gaussian Channel," *IEEE Transactions on Information Theory*, Vol.23, No.4, pp. 520-522, July 1977.

[13] Steven S. Pietrobon ,"On the Probability of Error of Convolutional Codes," *IEEE Transactions on Information Theory*, Vol.42, No.5, pp. 1562-1568, September 1996.

[14] E. Baccarelli, R. Cusani, G.D. Blasio ,"Performance Bound and Trellis-Code Design Criterion for Discrete Memoryless Channels and Finite-Delay Symbol-by-Symbol Decoding," *IEEE Transactions on Communications*, Vol.45, No.10, pp. 1192-1199, October 1997.

[15] C. Berrou, A. Glavieux, and P. Thitimajshima. ,"Near Shannon Limit Error-Correcting Coding and Decoding: Turbo-Codes (1)," *Proceedings IEEE International Conference on Communications*, Geneva, Switzerland, pp. 1064-1070, May 1993.

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

[17] 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, No.7, pp. 2558-2562, November 1999.

[18] Ye Liu, Shu Lin, M.P.C. Fossorier ,"MAP Algorithms for Decoding Linear Block Codes Based on Sectionalized Trellis Diagrams," *IEEE Transactions on Communications*, Vol.48, No.4, pp. 577-586, April 2000.

[19] Sven 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, No.2, pp. 175-185, February 1998.

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

[21] V. Franz, J.B. Anderson ,"Concatenated Decoding With a Reduced-Search BCJR Algorithm," *IEEE Journal on Selected Areas in Communications*, Vol.16, No.2, pp. 186-195, February 1998.

[22] B.J. Frey, F.R.Kschischang ,"Early Detection and Trellis Splicing: Reduced-Complexity Iterative Decoding," *IEEE Journal on Selected Areas in Communications*, Vol.16, No.2, pp. 153-159, February 1998.

[23] Ali Abedi and A. K. Khandani, "An Analytical Method for Performance Evaluation of Binary Linear Block Codes," Technical Report, UW-E&CE 2002-01 (available from www.cst.uwaterloo.ca)

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

[25] G. David Forney, Jr., "Geometrically Uniform Codes," *IEEE Transactions on Information Theory*, Vol.37, No.5, pp. 1241-1260, September 1991.

[26] E. Biglieri, D. Divsalar, P.J. McLane, M.K. Simon, *Introduction to Trellis-Coded Modulation with Applications*, Macmillan, USA, 1991.