ISIT'2013 1569719315

Equidistant Codes Meeting the Plotkin Bound are Not Optimal on the Binary Symmetric Channel

Po-Ning Chen, Hsuan-Yin Lin, and Stefan M. Moser
Department of Electrical & Computer Engineering
National Chiao Tung University (NCTU)
Hsinchu, Taiwan
Email: qponing@mail.nctu.edu.tw, {lin.hsuanyin, stefan.moser}@ieee.org

***Abstract*—In this paper, we re-introduce from our previous work [1] a new family of *nonlinear* codes, called *weak flip codes*, and show that its subfamily *fair weak flip codes* belongs to the class of equidistant codes, satisfying that any two distinct codewords have identical Hamming distance. It is then noted that the fair weak flip codes are related to the *binary nonlinear Hadamard codes* as both code families maximize the minimum Hamming distance and meet the Plotkin upper bound under certain blocklengths. Although the fair weak flip codes have the largest minimum Hamming distance and achieve the Plotkin bound, we find that these codes are by no means optimal in the sense of average error probability over binary symmetric channels (BSC). In parallel, this result implies that the equidistant Hadamard codes are also nonoptimal over BSCs. Such finding is in contrast to the conventional code design that aims at the maximization of the minimum Hamming distance.**

The results in this paper are proved by examining the exact error probabilities of these codes on BSCs, using the *column-wise* analysis on the codebook matrix.

I. Introduction

In 1948, Shannon [2] ingeniously established the ultimate limit of a reliable transmission rate over noisy channels and baptized it as *channel capacity*. From then on, a new research trend for finding good codes that operate close to the channel capacity began. Implicitly from Shannon's random coding proof, such good codes call for large blocklength. Since linearity does not inhibit the achievability of channel capacity, but simplifies the analysis and implementation of codes, coding theory and practice that follow often restrict themselves to *linear codes*. Motivated by the agreement between Hamming weights of codewords and Hamming distances between codewords for linear codes, and also by the union bound that converts the global error probability into pairwise error probabilities, it then becomes common to use the *minimum Hamming distance* as a quality criterion [3] for code design.

On the other hand, due to the analytical obstacle on the determination of exact error probability, information theorists usually resort to bounds such as the random coding bound. These bounds were often derived based on certain simplifications and are by no means accurate unless the blocklength of codes is sufficiently large. In our previous work, we attempted to break away these simplifications and instead focused on *the exact error probability* of codes for practically finite blocklength [1]. This new attempt could give a practical code design and remark on whether the implication from minimum-Hamming-distance code design is consistent with the true behavior of the error performance of codes.

In this paper, we re-introduce a new class of codes from our previous work, called *fair weak flip codes* [1] [4] [5], and confirm that they are equidistant. We further show that they can achieve the Plotkin upper bound and hence have the largest minimum Hamming distance among all (possibly nonlinear) codes of equal length. We then investigate whether this optimality in minimum Hamming distance can be extended to the error performance.

Note that there exists another class of binary nonlinear codes that also meet the Plotkin bound, called *binary nonlinear Hadamard codes*. This class of binary nonlinear codes are constructed with the help of Hadamard matrices and Levenshtein's theorem [6, Ch. 2], from which the codes are named. It is thus essential to clarify the relation between the fair weak flip codes and the Hadamard codes. A simple comparison gives that if for the parameters (M, n) of a given fair weak flip code there exists a Hadamard code, then these two codes are equivalent.[1] In this sense we can consider the fair weak flip codes to be a subclass of Hadamard codes. However, there is no guarantee that for every choice of parameters (M, n) for which fair weak flip codes exist, there also exists a corresponding Hadamard code. By this, the fair weak flip codes can also be regarded as an extension of the Hadamard codes.

The foundations of our insights lie in a new powerful way of creating and analyzing both linear and nonlinear block-codes. As is quite common, we use the *codebook matrix* containing the codewords in its rows to describe our codes. However, for our code construction and performance analysis, we look at this codebook matrix not row-wise, but *column-wise*. All our proofs and also our analysis for an equidistant code are fully based on this new approach to a code. This is another fundamental difference between our results and the binary nonlinear Hadamard codes that are constructed based on Hadamard matrices and Levenshtein's theorem [6].

The remainder of this paper is structured as follows. After some comments about our notations, we will introduce the

[1] Two codes are said to be *equivalent* if permuting the columns of the codebook matrix of one code leads to the codebook matrix of the other. For the definitions of the parameters (M, n) and of the codebook matrix, see Def. 1 and (5), respectively.

channel model and review some common definitions in Section II. We will illustrate a nontrivial example for the BSC in Section III. In Section IV we introduce the new family of *weak flip codes* and also its subfamily of *fair weak flip codes*. In Section V, we review some previous results of [5] [1] to give a comparison and original motivation for this paper. The main results are then summarized and discussed in Section VI.

As is common in coding theory, vectors (denoted by bold face Roman letters, e.g., $\mathbf{x}$) are row-vectors. However, for simplicity of notations and to avoid a large number of transpose-signs, we slightly misuse this notational convention for one special case: any (codebook) vector $\mathbf{c}$ is a column-vector. It should be always clear from the context because these vectors are used to build codebook matrices and are therefore also conceptually quite different from the transmitted codeword $\mathbf{x}$ or the received sequence $\mathbf{y}$. Moreover, we use a bar $\bar{\mathbf{x}}$ to denote the flipped version of $\mathbf{x}$, i.e., $\bar{\mathbf{x}} \triangleq \mathbf{x} \oplus \mathbf{1}$, where $\oplus$ denotes the componentwise XOR operation.

II. Channel Model and Coding Schemes

We quickly review a few common definitions.

Definition 1: An (M, n) *coding scheme* for a discrete memoryless channel (DMC) consists of a codebook $\mathscr{C}^{(\mathsf{M},n)}$ with M codewords of length n, an encoder that maps every message m into its corresponding codeword $\mathbf{x}_m$, and a decoder that makes a decoding decision $g(\mathbf{y}) \in \{1, \ldots, \mathsf{M}\}$ for every received binary n-vector $\mathbf{y}$.

We will always assume that the M possible messages are equally likely and that the decoder is a *maximum likelihood (ML) decoder:*[2]

$$g(\mathbf{y}) \triangleq \operatorname*{argmax}_{1 \le m \le \mathsf{M}} P_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}_m). \tag{1}$$

Hence, we are going to be lazy and directly concentrate on the set of codewords $\mathscr{C}^{(\mathsf{M},n)}$, called (M, n) *codebook* or usually simply (M, n) *code*. Sometimes we follow the custom of traditional coding theory and use three parameters: (M, n, d) *code*, where the third parameter d denotes the *minimum Hamming distance*, i.e., the minimum number of components in which any two codewords differ.

Definition 2: The *average error probability* $P_{\mathrm{e}}^{(n)}$ of an (M, n) code is defined as

$$P_{\mathrm{e}}^{(n)}\big(\mathscr{C}^{(\mathsf{M},n)}\big) \triangleq \frac{1}{\mathsf{M}} \sum_{m=1}^{\mathsf{M}} \Pr[g(\mathbf{Y}) \neq m | \mathbf{X} = \mathbf{x}_m]\,. \tag{2}$$

Sometimes it will be more convenient to focus on the probability of not making any error: the *average success probability*[3] $P_{\mathrm{c}}^{(n)}$ is defined as

$$P_{\mathrm{c}}^{(n)}\big(\mathscr{C}^{(\mathsf{M},n)}\big) \triangleq \frac{1}{\mathsf{M}} \sum_{m=1}^{\mathsf{M}} \Pr[g(\mathbf{Y}) = m | \mathbf{X} = \mathbf{x}_m]\,. \tag{3}$$

[2]Note that the ML decoder is optimal in the sense that for a given codebook and DMC and under the assumption of equally likely messages, it minimizes the average error probability as defined in (2).

[3]The subscript "c" stands for "correct."

Definition 3: For a given code $\mathscr{C}^{(\mathsf{M},n)}$, we define the *decoding region* $\mathcal{D}_m$ corresponding to the mth codeword $\mathbf{x}_m$ as follows:

$$\mathcal{D}_m \triangleq \{\mathbf{y} \colon g(\mathbf{y}) = m\}. \tag{4}$$

Usually, the codebook $\mathscr{C}^{(\mathsf{M},n)}$ is written as an $\mathsf{M} \times n$ *codebook matrix* with the M rows corresponding to the M codewords:

$$\mathscr{C}^{(\mathsf{M},n)} = \begin{pmatrix} — \ \mathbf{x}_1 \ — \\ \vdots \\ — \ \mathbf{x}_{\mathsf{M}} \ — \end{pmatrix} = \begin{pmatrix} | & | & & | \\ \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_n \\ | & | & & | \end{pmatrix}. \tag{5}$$

However, it turns out to be much more convenient to consider the codebook *column-wise* rather than row-wise! We denote the column-vectors of the codebook by $\mathbf{c}$.

The minimum Hamming distance is a well-known and often used quality criterion of a code. Unfortunately (as can be seen from the results presented in this paper and also from [5]), a design based on the minimum Hamming distance can be strictly suboptimal. We therefore define a slightly more general and more concise description of a code: the *pairwise Hamming distance vector*.

Definition 4: Given a code $\mathscr{C}^{(\mathsf{M},n)}$ with codewords $\mathbf{x}_m$ we define the *pairwise Hamming distance vector* $\mathbf{d}^{(\mathsf{M},n)}$ of length $\frac{(\mathsf{M}-1)\mathsf{M}}{2}$ as

$$\begin{aligned}\mathbf{d}^{(\mathsf{M},n)} \triangleq \Big(&d_{12}^{(n)},\ d_{13}^{(n)}, d_{23}^{(n)},\ d_{14}^{(n)}, d_{24}^{(n)}, d_{34}^{(n)}, \ldots, \\ &d_{1\mathsf{M}}^{(n)}, d_{2\mathsf{M}}^{(n)}, \ldots, d_{(\mathsf{M}-1)\mathsf{M}}^{(n)}\Big)\end{aligned} \tag{6}$$

with $d_{mm'}^{(n)} \triangleq d_{\mathrm{H}}(\mathbf{x}_m, \mathbf{x}_{m'})$, $1 \le m < m' \le \mathsf{M}$, where $d_{\mathrm{H}}(\cdot,\cdot)$ is the well known Hamming distance function. The *minimum Hamming distance* $d_{\min}$ is defined as the minimum component of the pairwise Hamming distance vector $\mathbf{d}^{(\mathsf{M},n)}$.

Definition 5: An (M, n) code is called *equidistant* if all components of the pairwise Hamming distance vector are equal, i.e., $d_{mm'}^{(n)} = \text{const} = d_{\min}$ for all $m \neq m'$.

Finally, we quickly recall an important bound that holds for any (M, n, d) code.

Lemma 6 (Plotkin Bound [6]): The minimum distance of an (M, n) binary code $\mathscr{C}^{(\mathsf{M},n)}$ always satisfies

$$d_{\min}\big(\mathscr{C}^{(\mathsf{M},n)}\big) \le \begin{cases} \frac{n \cdot \frac{\mathsf{M}}{2}}{\mathsf{M}-1} & \mathsf{M} \text{ even}, \\ \frac{n \cdot \frac{\mathsf{M}+1}{2}}{\mathsf{M}} & \mathsf{M} \text{ odd}. \end{cases} \tag{7}$$

Note that if an equidistant code meets the Plotkin bound (7), then this code maximizes the minimum Hamming distance.

III. An Example

To show that the minimal Hamming distance is not necessarily an optimal quality criterion for code design on BSC, we would like to give a simple example before we summarize our main results.

We consider a BSC with cross probability $\epsilon = 0.4$ and the following two codes with $\mathsf{M} = 4$ codewords and a blocklength

$n = 4$ each:[4]

$$\mathscr{C}_1^{(4,4)} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}, \quad \mathscr{C}_2^{(4,4)} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}. \tag{8}$$

We observe that while both codes are linear (i.e., any sum of two codewords is also a codeword), the first code has a minimum Hamming distance 1, and the second has a minimum Hamming distance 2. It is quite common to believe that $\mathscr{C}_2^{(4,4)}$ shows a better performance. This intuition is based on Gallager's famous performance bound [7, Exercise 5.19]:

$$P_{\mathrm{e}}\big(\mathscr{C}^{(\mathsf{M},n)}\big) \le (\mathsf{M}-1)e^{-d_{\min}(\mathscr{C}^{(\mathsf{M},n)})\log\frac{1}{\sqrt{4\epsilon(1-\epsilon)}}}. \tag{9}$$

The exact average error probability as given in (2), however, is evaluated to $P_{\mathrm{e}}\big(\mathscr{C}_1^{(4,4)}\big) \approx 0.6112$ and $P_{\mathrm{e}}\big(\mathscr{C}_2^{(4,4)}\big) = 0.64$. Hence, even though the minimum Hamming distance of the first codebook is smaller, its overall performance is superior to the second codebook!

IV. Weak Flip Codes and Hadamard Codes

We next introduce some special families of binary codes. We start with a code with two codewords.

Definition 7: The *weak flip code* with $\mathsf{M} = 2$ codewords is defined by the following codebook matrix $\mathscr{C}^{(2,n)}$:

$$\mathscr{C}^{(2,n)} \triangleq \begin{pmatrix} \mathbf{x} \\ \bar{\mathbf{x}} \end{pmatrix} = \begin{pmatrix} 0 & \cdots & 0 \\ 1 & \cdots & 1 \end{pmatrix}. \tag{10}$$

Defining the column vector

$$\left\{ \mathbf{c}_1^{(2)} \triangleq \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \tag{11}$$

we see that the weak flip code with two codewords is given by a codebook matrix that consists of n columns $\mathbf{c}_1^{(2)}$.

Note that the bits of $\mathbf{c}_1^{(2)}$ are flipped versions of each other, therefore also the name of the code.

We have shown in [5] that for any blocklength n the weak flip code with two codewords is optimal among all possible codes with two codewords for the BSC and the Z-channel.

Definition 8: The *weak flip code of type* (t_2, t_3) for $\mathsf{M} = 3$ or $\mathsf{M} = 4$ codewords is defined by a codebook matrix $\mathscr{C}_{t_2,t_3}^{(\mathsf{M},n)}$ that consists of $t_1 \triangleq n - t_2 - t_3$ columns $\mathbf{c}_1^{(\mathsf{M})}$, t_2 columns $\mathbf{c}_2^{(\mathsf{M})}$, and t_3 columns $\mathbf{c}_3^{(\mathsf{M})}$, where

$$\left\{ \mathbf{c}_1^{(3)} \triangleq \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \mathbf{c}_2^{(3)} \triangleq \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \mathbf{c}_3^{(3)} \triangleq \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\} \tag{12}$$

or

$$\left\{ \mathbf{c}_1^{(4)} \triangleq \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \mathbf{c}_2^{(4)} \triangleq \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \mathbf{c}_3^{(4)} \triangleq \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\}, \tag{13}$$

respectively. We often describe a weak flip code of type (t_2, t_3) by the *code parameters* $[t_1, t_2, t_3]$.

[4]We will see in Sec. IV that both codes are *weak flip codes*.

The columns given by the sets (12) and (13) are called *candidate columns*.

Lemma 9: The pairwise Hamming distance vector of a weak flip code of type (t_2, t_3) is

$$\mathbf{d}^{(3,n)} = (t_2 + t_3, t_1 + t_3, t_1 + t_2),$$
$$\mathbf{d}^{(4,n)} = (t_2 + t_3, t_1 + t_3, t_1 + t_2, t_1 + t_2, t_1 + t_3, t_2 + t_3).$$

To be able to generalize the definition of weak flip codes to an arbitrary M, we give the following definition [1].

Definition 10: Given a number of codewords M, a length-M candidate column $\mathbf{c}$ is called a *weak flip column* if its first component is 0 and its Hamming weight equals to $\left\lfloor \frac{\mathsf{M}}{2} \right\rfloor$ or $\left\lceil \frac{\mathsf{M}}{2} \right\rceil$. The collection of all possible weak flip columns is called *weak flip candidate columns set* and is denoted by $\mathcal{C}^{(\mathsf{M})}$. Note that (12) and (13) correspond to $\mathcal{C}^{(3)}$ and $\mathcal{C}^{(4)}$, respectively. We see that a weak flip column contains an almost equal number of zeros and ones.

For the remainder of this paper, we introduce the shorthands

$$\ell \triangleq \left\lceil \frac{\mathsf{M}}{2} \right\rceil, \tag{14}$$

and

$$\mathsf{L} \triangleq \left|\mathcal{C}^{(\mathsf{M})}\right| = \binom{2\ell - 1}{\ell}, \tag{15}$$

where L represents the cardinality of the corresponding weak flip candidate columns set.

We are now ready to generalize Definition 7.

Definition 11: A *weak flip code* is a codebook that is constructed only by weak flip columns.

Definition 12: A weak flip code is called *fair* if it is constructed by an equal number of all possible weak flip candidate columns in $\mathcal{C}^{(\mathsf{M})}$. Hence, the blocklength of a fair weak flip code is always a multiple of L.

Note that fair weak flip codes have been used by Shannon *et al.* for the derivation of error exponents [8].

Lemma 13 (Weak Flip Codes, Plotkin Bound, and Equidistance): A code that achieves the Plotkin bound (7) must be a weak flip code. Moreover, fair weak flip codes always meet the Plotkin bound, and they are equidistant.

Proof: See [1]. ∎

Related to the weak flip codes and the fair weak flip codes are the families of Hadamard codes [6, Ch. 2].

Definition 14: For an even integer n, a (*normalized*) *Hadamard matrix* H_n of order n is an $n \times n$ matrix with entries $+1$ and -1 and with the first row and column being all $+1$, such that

$$\mathsf{H}_n \mathsf{H}_n^{\mathsf{T}} = n\mathsf{I}_n, \tag{16}$$

if such a matrix exists. Here I_n is the identity matrix of size n. If the entries $+1$ are replaced by 0 and the entries -1 by 1, H_n is changed into the *binary Hadamard matrix* A_n.

Note that a necessary (but not sufficient) condition for the existence of H_n (and the corresponding A_n) is that n is 1, 2, or a multiple of 4 [6, Ch. 2].

Definition 15: The binary Hadamard matrix A_n gives rise to three families of *Hadamard codes*:

1) The $\left(n, n-1, \frac{n}{2}\right)$ *Hadamard code* $\mathscr{H}_{1,n}$ consists of the rows of A_n with the first column deleted. The codewords in $\mathscr{H}_{1,n}$ that begin with 0 form the $\left(\frac{n}{2}, n-2, \frac{n}{2}\right)$ *Hadamard code* $\mathscr{H}'_{1,n}$ if the initial zero is deleted.
2) The $\left(2n, n-1, \frac{n}{2}-1\right)$ *Hadamard code* $\mathscr{H}_{2,n}$ consists of $\mathscr{H}_{1,n}$ together with the complements of all its codewords.
3) The $\left(2n, n, \frac{n}{2}\right)$ *Hadamard code* $\mathscr{H}_{3,n}$ consists of the rows of A_n and their complements.

Further Hadamard codes can be created by an arbitrary combination of codebook matrices of different Hadamard codes. Note that every Hadamard code is a weak flip code. Also note that for a given number of codewords M and a blocklength n, the existence of a Hadamard code is not guaranteed. For a more detailed discussion of Hadamard codes, see [1].

V. Previous Results

In [5] it is shown that for $\mathsf{M} = 3$ or $\mathsf{M} = 4$, weak flip codes with code parameters

$$[t_1^*, t_2^*, t_3^*] = \begin{cases} [k+1, k-1, k] & \text{if } n \bmod 3 = 0, \\ [k+1, k, k] & \text{if } n \bmod 3 = 1, \\ [k+1, k, k+1] & \text{if } n \bmod 3 = 2, \end{cases} \tag{17}$$

where we use

$$k \triangleq \left\lfloor \frac{n}{3} \right\rfloor, \tag{18}$$

are optimal codes for a BSC, i.e., they achieve the globally best possible average error probability. The corresponding pairwise Hamming distance vectors (see Lemma 9) are

$$\begin{cases} (2k-1, 2k, 2k+1) & \text{if } n \bmod 3 = 0, \\ (2k, 2k+1, 2k+1) & \text{if } n \bmod 3 = 1, \\ (2k+1, 2k+2, 2k+1) & \text{if } n \bmod 3 = 2. \end{cases} \tag{19}$$

From [1] we know that for $\mathsf{M} = 3$ or $\mathsf{M} = 4$, weak flip codes with code parameters

$$[t_1^*, t_2^*, t_3^*] = \begin{cases} [k, k, k] & \text{if } n \bmod 3 = 0, \\ [k+1, k, k] & \text{if } n \bmod 3 = 1, \\ [k+1, k, k+1] & \text{if } n \bmod 3 = 2 \end{cases} \tag{20}$$

and with corresponding pairwise Hamming distance vectors

$$\begin{cases} (2k, 2k, 2k) & \text{if } n \bmod 3 = 0, \\ (2k, 2k+1, 2k+1) & \text{if } n \bmod 3 = 1, \\ (2k+1, 2k+2, 2k+1) & \text{if } n \bmod 3 = 2 \end{cases} \tag{21}$$

are optimal for a BEC. Hence, we see that apart from $n \bmod 3 = 0$, the optimal codes for a BSC are identical to the optimal codes for a BEC for $\mathsf{M} = 3$ or $\mathsf{M} = 4$ codewords.

It is interesting to note that for $n \bmod 3 = 0$ the optimal codes for the BEC are fair and therefore maximize the minimum Hamming distance, while this is not the case for the (very symmetric!) BSC. Hence, we learn that the fair weak flip codes with $\mathsf{M} = 3$ or 4 are not optimal on a BSC. The question therefore arises whether a code with maximal minimum Hamming distance is always strictly suboptimal on a BSC. In the next section, we are going to investigate this question.

VI. Main Results

We start with the following lemma.

Lemma 16: Fix some arbitrary integers $\mathsf{M} \geq 2$, $n \geq 1$, and $\gamma \geq 1$. Consider a DMC and an (M, n) code $\mathscr{C}^{(\mathsf{M},n)}$ for this DMC with M codewords and blocklength n, and create a new code $\mathscr{C}^{(\mathsf{M},n+\gamma)}$ by appending γ arbitrary column vectors to the codebook matrix of $\mathscr{C}^{(\mathsf{M},n)}$. Then the success probability of this new code cannot be smaller than the success probability of the original code:

$$P_{\mathrm{c}}\big(\mathscr{C}^{(\mathsf{M},n+\gamma)}\big) \geq P_{\mathrm{c}}\big(\mathscr{C}^{(\mathsf{M},n)}\big). \tag{22}$$

We omit the proof and refer to the journal version that is under preparation.

The main result of this paper is that for many blocklengths n, equidistant codes that achieve the Plotkin bound (7) with equality, i.e., they maximize the minimum Hamming distance, are strictly suboptimal. We will first state this result for the special case of fair weak flip codes, and afterwards generalize it to arbitrary equidistant codes.

Proposition 17: Consider a BSC with the conditional channel probability

$$P_{Y|X}(y|x) = \begin{cases} 1-\epsilon & \text{if } y = x, \\ \epsilon & \text{if } y \neq x, \end{cases} \qquad x, y \in \{0, 1\}, \tag{23}$$

with crossover probability $0 < \epsilon < \frac{1}{2}$. For a fair weak flip code $\mathscr{C}_{\text{fair}}^{(\mathsf{M},n)}$ with a blocklength satisfying

$$n \bmod \mathsf{L} = 0, \tag{24}$$

let $\mathscr{C}_{\text{reduced}}^{(\mathsf{M},n-1)}$ be a code that is created from $\mathscr{C}_{\text{fair}}^{(\mathsf{M},n)}$ by deleting an arbitrary column in the codebook matrix. Then

$$P_{\mathrm{c}}\Big(\mathscr{C}_{\text{fair}}^{(\mathsf{M},n)}\Big) = P_{\mathrm{c}}\Big(\mathscr{C}_{\text{reduced}}^{(\mathsf{M},n-1)}\Big). \tag{25}$$

Moreover, let $\mathscr{C}_{\text{unfair}}^{(\mathsf{M},n)}$ be a code that is created by appending a weak flip column to $\mathscr{C}_{\text{reduced}}^{(\mathsf{M},n-1)}$ such that it is not a fair weak flip code. Then

$$P_{\mathrm{c}}\Big(\mathscr{C}_{\text{unfair}}^{(\mathsf{M},n)}\Big) > P_{\mathrm{c}}\Big(\mathscr{C}_{\text{reduced}}^{(\mathsf{M},n-1)}\Big). \tag{26}$$

We again omit the proof.

Theorem 18: Fair weak flip codes with an arbitrary number of codewords M and with a blocklength n such that $n \bmod \mathsf{L} = 0$, are strictly suboptimal on a BSC.

Proof: From Proposition 17, we see that for any fair weak flip code we can construct a codebook that does not achieve the Plotkin bound (7), but whose average success probability is strictly larger than the the average success probability of the fair weak flip code. ■

Proposition 17 and Theorem 18 can be extended to general equidistant codes that meet the Plotkin bound (7). The clue of the proof is the observation from Lemma 13 saying that all codes that achieve the Plotkin bound with equality are

weak flip codes. The main difficulty lies in the choice of column to delete when creating $\mathscr{C}_{\text{reduced}}^{(M,n-1)}$ and the choice of column to add when creating a new code that is strictly better (compare with (25) and (26)). Furthermore, the condition on the blocklength n (24) also will change. We omit the details and only summarize the main statement.

Theorem 19: For many blocklengths n, all equidistant codes that meet the Plotkin bound (7) with equality (in particular, all equidistant Hadamard codes) are strictly suboptimal on a BSC.

ACKNOWLEDGMENT

This work was supported by the National Science Council under NSC 100–2221–E–009–068–MY3.

REFERENCES

[1] P.-N. Chen, H.-Y. Lin, and S. M. Moser, "Weak flip codes and applications to optimal code design on the binary erasure channel," in *Proc. 50th Allerton Conf. Commun., Contr. and Comput.*, Monticello, IL, USA, Oct. 1–5, 2012.

[2] C. E. Shannon, "A mathematical theory of communication," *Bell System Techn. J.*, vol. 27, pp. 379–423 and 623–656, Jul. and Oct. 1948.

[3] S. Lin and D. J. Costello, Jr., *Error Control Coding*, 2nd ed. Upper Saddle River, NJ: Prentice Hall, 2004.

[4] P.-N. Chen, H.-Y. Lin, and S. M. Moser, "Ultra-small block-codes for binary discrete memoryless channels," in *Proc. IEEE Inf. Theory Workshop*, Paraty, Brazil, Oct. 16–20, 2011, pp. 175–179.

[5] ——, "Optimal ultra-small block-codes for binary discrete memoryless channels," 2013, to app. in *IEEE Trans. Inf. Theory*. [Online]. Available: http://moser.cm.nctu.edu.tw/publications.html

[6] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. Amsterdam: North-Holland, 1977.

[7] R. G. Gallager, *Information Theory and Reliable Communication*. New York: John Wiley & Sons, 1968.

[8] C. E. Shannon, R. G. Gallager, and E. R. Berlekamp, "Lower bounds to error probability for coding on discrete memoryless channels," *Inform. Contr.*, pp. 522–552, May 1967, part II.