Nonuniform Codes for Correcting Asymmetric Errors

Hongchao Zhou
Electrical Engineering Department
California Institute of Technology
Pasadena, CA 91125
Email: hzhou@caltech.edu

Anxiao (Andrew) Jiang
Computer Science and Engineering Department
Texas A&M University
College Station, TX 77843
Email: ajiang@cse.tamu.edu

Jehoshua Bruck
Electrical Engineering Department
California Institute of Technology
Pasadena, CA 91125
Email: bruck@caltech.edu

***Abstract*—Codes that correct asymmetric errors have important applications in storage systems, including optical disks and Read Only Memories. The construction of asymmetric error correcting codes is a topic that was studied extensively, however, the existing approach for code construction assumes that every codeword could sustain t asymmetric errors. Our main observation is that in contrast to symmetric errors, where the error probability of a codeword is context independent (since the error probability for 1s and 0s is identical), asymmetric errors are context dependent. For example, the all-1 codeword has a higher error probability than the all-0 codeword (since the only errors are $1 \to 0$). We call the existing codes uniform codes while we focus on the notion of nonuniform codes, namely, codes whose codewords can tolerate different numbers of asymmetric errors depending on their Hamming weights. The goal of nonuniform codes is to guarantee the reliability of every codeword, which is important in data storage to retrieve whatever one wrote in. We prove an almost explicit upper bound on the size of nonuniform asymmetric error correcting codes and present two general constructions. We also study the rate of nonuniform codes compared to uniform codes and show that there is a potential performance gain.**

I. Introduction

Asymmetric error-correcting codes have important applications in storage and communication systems, such as optical fibers, optical disks, VLSI circuits and Read Only Memories. In such systems, the error probability from 1 to 0 is significantly higher than the error probability from 0 to 1, which is modeled by binary asymmetric channel (the $Z-$channel) where the transmitted sequences only suffer one type of errors, say $1 \to 0$. Asymmetric error-correcting codes have been widely studied: In [1], Kløve summarized and presented several such codes. In addition, a large amount of effort is contributed to the design of systematic codes [2], [3], constructing single or multiple error-correcting codes [4]–[6], increasing the lower bounds [7]–[9] and applying LDPC codes in the context of asymmetric channels [10].

However, the existing approach for code construction is similar to the approach taken in the construction of symmetric error correcting codes, namely, it assumes that every codeword could sustain t asymmetric errors. As a result, different codewords might have different reliability. To see this, let's consider errors to be i.i.d., where every bit that is a 1 can change to a 0 by an asymmetric error with crossover probability $p > 0$. For a codeword $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \{0,1\}^n$, let $w(\mathbf{x}) = |\{i\ : 1 \leq i \leq n, x_i = 1\}|$ denote the Hamming weight of $\mathbf{x}$. Then the probability for $\mathbf{x}$ to have at most t asymmetric errors is $P_t(x) = P(t, w(\mathbf{x}), p)$, where

$$P(t, m, p) \triangleq \sum_{i=0}^{t} \binom{m}{i} p^i (1-p)^{m-i}.$$

Since $\mathbf{x}$ can correct t errors, $P_t(x)$ is the probability of correctly decoding $\mathbf{x}$ (assuming codewords with more than t errors are uncorrectable). It can be readily observed that the reliability of codewords decreases when their Hamming weights increase.

Different from telecommunication applications, in data storage we care about the worst-case performance, namely, we need guarantee that every codeword can be correctly decoded with very high probability. In this case, it is not desired to let all the codewords tolerate the same number of asymmetric errors, since the codeword with the highest Hamming weight will become a 'bottleneck' and limit the code rate. This motivated us to propose the concept of *nonuniform* codes, whose codewords can tolerate different numbers of asymmetric errors based on their Hamming weights. The objective is to guarantee the reliability of every codeword. That is, we consider the worst-case instead of the average-case reliability of the codewords. Given this constraint, we would like to maximize the size of the code. Specifically, let $q_e < 1$ to be maximal tolerated error probability for each codeword and let $t(x)$ denote the number of asymmetric errors that x can correct. Then given a code C, for every codeword $x \in C$, we have $P(t(x), w(x), p) \geq 1 - q_e$, so that every erroneous codeword can be corrected with probability at least $1 - q_e$.

The rest of the paper is organized as follows. In Section II, we provide some definitions and properties related to nonuniform codes. In Section III, we give an almost explicit upper bound for the size of nonuniform codes. Two general constructions, based on multiple layers or bit flips, are proposed in Section IV and Section V. Finally, Section VI studies the asymptotic rates of nonuniform codes and uniform codes (both upper bounds and lower bounds). An extended version of this paper with detailed proofs and explanations is given in [16].

978-1-4577-0594-6/11/$26.00 ©2011 IEEE

II. Definitions and Properties

A code C is called a *nonuniform* (n,p,q_e) *code* if for each codeword $x \in C$, it can correct $t(w(x))$ asymmetric errors, where

$$t(w) = \min\{s \in N | P(s,w,p) \geq 1 - q_e\}. \quad (1)$$

That implies each codeword in C can be recovered with probability at least $1-q_e$. The maximum size of a nonuniform (n,p,q_e) code is denoted by $B_\beta(n,p,q_e)$.

As comparison, most existing error-correcting codes are *uniform* codes. For a code C of codeword length n, the Hamming weight of its codewords is at most n. (And in many existing asymmetric error-correcting codes, the maximum codeword weight indeed equals n [1].) So we define C to be an *uniform* (n,p,q_e) *code* if every codeword can correct t asymmetric errors, where

$$t = t(n) = \min\{s \in N | P(s,n,p) \geq 1 - q_e\}.$$

The maximum size of an uniform (n,p,q_e) code is denoted by $B_\alpha(n,p,q_e)$.

Lemma 1. *For any* $0 < p, q_e < 1$ *and integer* w *in* $[0,n]$*, we have* $0 \leq t(w+1) - t(w) \leq 1$ *for a nonuniform* (n,p,q_e) *code.*

Given two binary vectors $\mathbf{x} = (x_1,\dots,x_n)$ and $\mathbf{y} = (y_1,\dots,y_n)$, we say $\mathbf{x} \leq \mathbf{y}$ if and only if $x_i \leq y_i$ for all $1 \leq i \leq n$. Let $S_s(\mathbf{x})$ be the set of vectors obtained by changing at most s 1's in $\mathbf{x}$ into 0's, i.e.,

$$S_s(\mathbf{x}) = \{\mathbf{v} \in \{0,1\}^n | \mathbf{v} \leq \mathbf{x} \text{ and } N(\mathbf{x},\mathbf{v}) \leq s\}.$$

where

$$N(\mathbf{x},\mathbf{y}) \triangleq |\{i : x_i = 1, y_i = 0\}|.$$

Let $S_{s',s}(\mathbf{x})$ be the set of vectors obtained by changing at most s' 0's in $\mathbf{x}$ into 1's *or* at most s 1's in $\mathbf{x}$ into 0's, i.e.,

$$\begin{aligned} S_{s',s}(x) &= \{\mathbf{v} \in \{0,1\}^n | \mathbf{v} \leq \mathbf{x} \text{ and } N(\mathbf{x},\mathbf{v}) \leq s\} \\ &\quad \bigcup \{\mathbf{v} \in \{0,1\}^n | \mathbf{x} \leq \mathbf{v} \text{ and } N(\mathbf{v},\mathbf{x}) \leq s'\} \end{aligned}$$

Note that $S_s(\mathbf{x}) = S_{0,s}(\mathbf{x})$.

The following properties of nonuniform codes can be easily proved, as the generalizations of those for uniform codes, including Lemmas 2.2, 2.3, 3.2, 3.3 in [1].

Lemma 2. *Code* C *is a nonuniform* (n,p,q_e) *code if and only if* $S_{t(w(\mathbf{x}))}(\mathbf{x}) \bigcap S_{t(w(\mathbf{y}))}(\mathbf{y}) = \emptyset$ *for all* $\mathbf{x},\mathbf{y} \in C$ *with* $\mathbf{x} \neq \mathbf{y}$.

Lemma 3. *There always exists a nonuniform* (n,p,q_e) *code of the maximum size that contains the all-zero codeword.*

Given a nonuniform code C, let C_r denote the number of codewords with Hamming weight r in C, i.e.

$$C_r = |\{\mathbf{x} \in C | w(\mathbf{x}) = r\}|.$$

Lemma 4. *Let* C *be a nonuniform* (n,p,q_e) *code and* $t(w)$ *is defined in (1). For integer* r *in* $[0,n]$*, let* s *be an integer such that* $0 \leq s \leq t(r-s)$ *and let* $k = \max\{z | 0 \leq z \leq n, z - (t(z) - s) \leq r\}$*, then we have*

$$\sum_{j=1}^{s} \binom{n-r+j}{j} C_{r-j} + \sum_{j=0}^{t(k)-s} \binom{r+j}{j} C_{r+j} \leq \binom{n}{r}.$$

Note that in Lemma 4, if we let $s = 0$, then we can get

$$\sum_{j=0}^{t(k)} \binom{r+j}{j} C_{r+j} \leq \binom{n}{r} \quad (2)$$

where $k = \max\{z | 0 \leq z \leq n, z - t(z) \leq r\}$. This inequality will be used to get an almost explicit upper bound for the size of nonuniform codes.

III. An Almost Explicit Upper Bound

We now derive an almost explicit upper bound for the size of nonuniform codes, followed the idea of Kløve [11] for uniform codes. First, we define

$$\overline{h}(r) = \max\{w | 0 \leq w \leq n, w - t(w) = r\},$$

$$\underline{h}(r) = \min\{w | 0 \leq w \leq n, w - t(w) = r\}.$$

And let $M_\beta(n,p,q_e) = \max \sum_{r=0}^{n} z_r$, where the maximum is taken over the following constraints:

1) z_r are non-negative real numbers;
2) $z_0 = 1$;
3) $\sum_{j=0}^{t(\overline{h}(r))} \binom{r+j}{j} z_{r+j} \leq \binom{n}{r}$ for $r \geq 0$.

Then $M_\beta(n,p,q_e)$ is an upper bound for $B_\beta(n,p,q_e)$. Here, condition 2) is given by Lemma 3, and condition 3) is given by Equ. (2) from Lemma 4. Our goal in this section is to find an almost explicit way to express $M_\beta(n,p,q_e)$.

Lemma 5. *Assume* $\sum_{r=0}^{n} z_r$ *is maximized over* $z_0, z_1, ..., z_n$ *in the problem above. Let*

$$Z_r = \sum_{j=0}^{min\{n-r, t(\overline{h}(r))\}} z_{r+j} \binom{r+j}{j}.$$

Then $Z_r = \binom{n}{r}$ *for* $r \leq n - t(n)$.

Proof: Suppose that $Z_r < \binom{n}{r}$ for some $r \leq n - t(n)$. Let $g = \overline{h}(r)$ and $k = \min\{w | z_w > 0, w > g\}$.

Let $m = \max\{w | k - t(k) > w\}$. Then it can be proved that for all $r < w \leq m$, $Z_w < \binom{n}{w}$.

Now, we construct a new group of real numbers $z_0^*, z_1^*, ..., z_n^*$ such that

1) $z_g^* = z_g + \Delta$
2) $z_k^* = z_k - \delta$
3) $z_r^* = z_r$ for $r \neq h, r \neq k$

with

$$\Delta = \min(\{\frac{\binom{n}{w} - Z_w}{\binom{g}{w}} | r \leq w \leq m\} \bigcup \{\frac{\binom{k}{w}}{\binom{g}{w}} z_k | m < w \leq g\}),$$

$$\delta = \frac{1}{\min\{\frac{\binom{k}{w}}{\binom{g}{w}} | m < w \leq g\}} \Delta.$$

For such Δ, δ, it is not hard to prove that $Z_r^* = \binom{n}{r}$ for $0 \le r \le n$. On the other hand,

$$\sum_{r=0}^{n} z_r^* = \sum_{r=0}^{n} z_r + \Delta - \delta > \sum_{r=0}^{n} z_r,$$

which contradicts our assumption that $\sum_{r=0}^{n} z_r$ is maximized over the constrains. So the lemma is true. ■

Similarly, using the same idea as above, we can get the following lemma.

Lemma 6. *Assume $\sum_{r=0}^{n} z_r$ is maximized over $z_0, z_1, ..., z_n$ in the problem above. Let*

$$Y_r = \sum_{j=0}^{min\{n-r, t(\underline{h}(r))\}} z_{r+j} \binom{r+j}{j}.$$

Then $Y_r = \binom{n}{r}$ for $r \le n - t(n)$.

Now let $y_0, y_1, ..., y_n$ be a group of optimal solutions to $z_0, z_1, ..., z_n$ that maximize $\sum_{r=0}^{n} z_r$. Then $y_0, y_1, ..., y_n$ satisfy the condition in Lemma 6. We see that $y_0 = 1$. Then based on Lemma 6, we can get $y_1, ..., y_n$ uniquely by iteration. Hence, we have the following theorem for the upper bound $M_\beta(n, p, q_e)$.

Theorem 7. *Let $y_0, y_1, ..., y_n$ be defined by*

1) $y_0 = 1$;
2) $y_r = 0, \quad \forall 1 \le r \le \max\{s | 1 \le s \le n, s \le t(s)\}$;
3) $y_r = \frac{1}{\binom{r}{t(r)}}[\binom{n}{r-t(r)} - \sum_{j=1}^{t(r)} y_{r-j} \binom{r-j}{t(r)-j}]$,
$\forall \max\{s | 1 \le s \le n, s \le t(s)\} < r \le n$.

Then $B_\beta(n, p, q_e) \le M_\beta(n, p, q_e) = \sum_{r=0}^{n} y_r$.

This theorem provides an almost explicit expression for the upper bound $M_\beta(n, p, q_e)$, which is much easier to calculate than the equivalent expression defined at the beginning of this section.

IV. CONSTRUCTIONS BASED ON MULTIPLE LAYERS

In [1], Kløve summarized some constructions of uniform codes for correcting asymmetric errors. The code of Kim and Freiman was the first code constructed for correcting multiple asymmetric errors. Varshamov [12] and Constrain and Rao [13] presented some constructions based group theory. Later, Delsarte and Piret [14] proposed a construction based on 'expurgating/puncturing' with some improvements given by Weber et. al. [15]. In this section, we propose a general construction of nonuniform codes based on multiple layers.

From the definition of nonuniform codes, we know that $t(w)$ can be easily and uniquely determined by p, q_e. So a question arises: if $n, t(w)$ (for $0 \le w \le n$) are given, how to construct a nonuniform code efficiently? Intuitively, we can divide all the codewords of a nonuniform code into at most $t(n)+1$ layers such that all the codewords in the i^{th} layer (with $0 \le i \le t(n)$) can tolerate at least i asymmetric errors. In other words, the code is the combination of up to $t(n)+1$ uniform codes, each of which corrects a different number of asymmetric errors. However, we cannot design such a code by constructing codewords independently for different layers, because a simple combination of several independent codes may violate the error correction requirements of the nonuniform codes, due to the interference between two neighbor layers. Our idea is simple: let's first construct a code which can tolerate $t(n)$ asymmetric errors. Then we add some codewords to the lowest $t(n)$ layers such that the codewords in the top layer keep unchanged and they still can tolerate $t(n)$ asymmetric errors, and the codewords in the other layers can tolerate up to $t(n)-1$ asymmetric errors. Iteratively, we can continue to add many codeword into the lowest $t(n)-1$ layers ... Based on this idea, given $n, t(w)$, we construct layered codes as follows.

Theorem 8 (Layered Codes)**.** *Let $k = t(n)$ and let $C_0, C_1, ..., C_k$ be $k+1$ binary codes of codeword length n, where $C_0 \supset C_1 \supset ... \supset C_k$ and for $0 \le t \le k$, the code C_t can correct t asymmetric errors. Let*

$$C = \{\mathbf{x} \in \{0,1\}^n | \mathbf{x} \in C_{t'(w(\mathbf{x}))}\},$$

where

$$t'(w(\mathbf{x})) = t(\max\{w' | w' - t(w') \le w(\mathbf{x})\}).$$

Then for all $\mathbf{x} \in C$, $\mathbf{x}$ can tolerate $t(w(\mathbf{x}))$ asymmetric errors.

Proof: We prove that for all $\mathbf{x}, \mathbf{y} \in C$ with $\mathbf{x} \ne \mathbf{y}$, $S_{t(w(\mathbf{x}))}(\mathbf{x}) \bigcap S_{t(w(\mathbf{y}))}(\mathbf{y}) = \emptyset$. W.l.o.g., we assume $w(\mathbf{x}) \ge w(\mathbf{y})$.

If $w(\mathbf{x}) - t(w(\mathbf{x})) > w(\mathbf{y})$, the conclusion is true.

If $w(\mathbf{x}) - t(w(\mathbf{x})) \le w(\mathbf{y})$ and $w(\mathbf{x}) \ge w(\mathbf{y})$, we have $S_{t(w(\mathbf{x}))}(\mathbf{x}) \bigcap S_{t(w(\mathbf{y}))}(\mathbf{y}) \subseteq S_{t'(w(\mathbf{y}))}(\mathbf{x}) \bigcap S_{t'(w(\mathbf{y}))}(\mathbf{y})$. However, we know that $\mathbf{x} \in C_{t'(w(\mathbf{x}))} \subseteq C_{t'(w(\mathbf{y}))}$ and $\mathbf{y} \in C_{t'(w(\mathbf{y}))}$, therefore $S_{t'(w(\mathbf{y}))}(\mathbf{x}) \bigcap S_{t'(w(\mathbf{y}))}(\mathbf{y}) = \emptyset$. Furthermore, we have $S_{t(w(\mathbf{x}))}(\mathbf{x}) \bigcap S_{t(w(\mathbf{y}))}(\mathbf{y}) = \emptyset$. ■

We see that the constructions of layered codes are based on the provided group of codes $C_0, C_1, ..., C_k$ such that $C_0 \supset C_1 \supset ... \supset C_k$ and for $0 \le t \le k$, the code C_t can correct t asymmetric errors. Examples of such codes include Varshamov codes [12], BCH codes, etc. One constructions of BCH codes can be described as follows: Let $(\alpha_0, \alpha_1, ..., \alpha_{n-1})$ be n distinct nonzero elements of G_{2^m} with $n = 2^m - 1$. For $0 \le t \le k$, let

$$C_t := \{\mathbf{x} \in \{0,1\}^n | \sum_{i=1}^{n} x_i \alpha_i^{(2l-1)} = 0 \text{ for } 1 \le l \le t\}.$$

In the above example, assume $\mathbf{x}$ is a codeword in C_t and $\mathbf{y} = \mathbf{x} + \mathbf{e}$ is a received word with error $\mathbf{e}$, then there is an efficient algorithm to decode $\mathbf{y}$ into a codeword, which is denoted by $D_t(\mathbf{y})$. If $\mathbf{y}$ has at most t asymmetric errors, then $D_t(\mathbf{y}) = \mathbf{x}$. In the following theorem, we show that the layered codes proposed above also have an efficient decoding algorithm if $D_t(\cdot)$ (for $0 \le t \le k$) are provided and efficient.

Theorem 9 (Decoding of Layered Codes)**.** *Let C be a layered code, let $\mathbf{x} \in C$ be a codeword, and let $\mathbf{y} = \mathbf{x} + \mathbf{e}$ be a received word such that $|\mathbf{e}| = N(\mathbf{x}, \mathbf{y}) \le t(w(\mathbf{x}))$. (Here $\mathbf{e}$ is the asymmetric-error vector.) Then there exists at least one integer t such that*

1) $t'(w(\mathbf{y})) \leq t \leq t'(w(\mathbf{y}) + t'(w(\mathbf{y})))$;
2) $D_t(\mathbf{y}) \in C$;
3) $\mathbf{y} \leq D_t(\mathbf{y})$ *and* $N(D_t(\mathbf{y}), \mathbf{y}) \leq t(w(D_t(\mathbf{y})))$.

For such t, we have $D_t(\mathbf{y}) = \mathbf{x}$.

Proof: If we let $t = t'(w(\mathbf{x}))$, then we can get that t satisfies the conditions and $D_t(\mathbf{y}) = \mathbf{x}$. So such t exists.

Now we only need to prove that once there exists t satisfying the conditions in the theorem, we have $D_t(\mathbf{y}) = \mathbf{x}$. We prove this by contradiction. Assume there exists t satisfying the conditions but $\mathbf{z} = D_t(\mathbf{y}) \neq \mathbf{x}$. Then $N(\mathbf{z}, \mathbf{y}) \leq t(w(\mathbf{z}))$ and $N(\mathbf{x}, \mathbf{y}) \leq t(w(\mathbf{x}))$, which contradicts the property of the layered codes. ∎

According to the above theorem, to decode a noisy word $\mathbf{y}$, we can check all the integers between $t'(w(\mathbf{y}))$ and $t'(w(\mathbf{y}) + t'(w(\mathbf{y})))$ to find the value of t. Once we find the integer t satisfying the conditions in the theorem, we can decode $\mathbf{y}$ into $D_t(\mathbf{y})$ directly. (Note that $t'(w(\mathbf{y}) + t'(w(\mathbf{y}))) - t'(w(\mathbf{y}))$ is normally much smaller than $w(y)$. It is approximately $\frac{p^2}{(1-p)^2} w(y)$ when $w(y)$ is large.) We see that this decoding process is efficient if $D_t(.)$ is efficient for $0 \leq t \leq k$.

V. Constructions Based on Bit Flips

Many non-linear codes designed to correct asymmetric errors do not yet have efficient encoding algorithms. Namely, it is not easy to find an efficient encoding function $f : \{0,1\}^k \to C$ with $k \simeq \lfloor \log |C| \rfloor$. On the other hand, in [12], Varshamov showed that linear codes have nearly the same ability to correct asymmetric errors and symmetric errors (for the uniform code case). In this subsection, we focus on the approach of designing nonuniform codes for asymmetric errors with efficient encoding schemes, by utilizing the well studied linear codes for symmetric errors.

We can use a linear code to correct $t(n)$ asymmetric errors directly, but this method is inefficient not only because the decoding sphere for symmetric errors is greater than the sphere for asymmetric errors (and therefore an overkill), but also because for low-weight codewords, the number of asymmetric errors they need to correct can be much smaller than $t(n)$.

Our idea is to build a "flipping code" that uses only low-weight codewords (specifically, codewords of Hamming weight no more than $\sim \frac{n}{2}$), because they need to correct fewer asymmetric errors and therefore can increase the code's rate. In the rest of this section, we present two different constructions.

A. First Construction

First, construct a linear code C (like BCH codes) of length n with generator matrix G that corrects $t(\lfloor \frac{n}{2} \rfloor)$ symmetric errors. Assume the dimension of the code is k. For any binary message $\mathbf{u} \in \{0,1\}^k$, we can map it to a codeword $\mathbf{x}$ in C such that $\mathbf{x} = \mathbf{u}G$. Next, let $\overline{\mathbf{x}}$ denote a word obtained by flipping all the bits in $\mathbf{x}$ such that if $x_i = 0$ then $\overline{x}_i = 1$ and if $x_i = 1$ then $\overline{x}_i = 0$; and let y denote the final codeword corresponding to u. We check whether $w(\mathbf{x}) > \lfloor \frac{n}{2} \rfloor$ and construct y in the following way:

$$y = \begin{cases} x00...0 & \text{if } w(\mathbf{x}) > \lfloor \frac{n}{2} \rfloor \\ \overline{x}11...1 & \text{otherwise} \end{cases}$$

Here, the auxiliary bits (0's or 1's) are added to distinguish that whether x has been flipped or not, and they form a repetition code to tolerate errors.

The corresponding decoding process is straightforward: Assume we received a word y'. If there is at least one 1 in the auxiliary bits, then we "flip" the word by changing all 0's to 1's and all 1's to 0's; otherwise, we keep the word unchanged. Then we apply the decoding scheme of the code C to the first n bits of the word. Finally, the message u can be successfully decoded if y' has at most $t(\lfloor \frac{n}{2} \rfloor)$ errors in the first n bits.

B. Second Construction

In the previous construction, several auxiliary bits are needed to protect one bit of information, which is not very efficient. In this section, we try to move this bit into the message part of the codewords in C. This motivates us to give the following construction.

Let C be a linear code with length n that corrects t' symmetric errors (we will specify t' later). Assume the dimension of the code is k. Now, for any binary message $\mathbf{u} \in \{0,1\}^{k-1}$ of length $k-1$, we get $u' = 0u$ by adding one bit 0 in front of u. Then we can map u' to a codeword $\mathbf{x}$ in C such that

$$x = (0u)G = 0uv$$

where G is the generator matrix of C in systematic form and the length of v is $n-k$. Let α be a codeword in C such that the first bit $\alpha_1 = 1$ and its weight is the maximal one among all the codeword in C, i.e.,

$$\alpha = \arg \max_{x \in C, x_1 = 1} w(x)$$

Generally, $w(\alpha)$ is very close to n. In order to reduce the weights of the codewords, we use the following operations: Calculate the relative weight

$$w(x|\alpha) = |\{1 \leq i \leq n | x_i = 1, \alpha_i = 1\}|$$

Then we get the final codeword

$$y = \begin{cases} x + \alpha & \text{if } w(x|\alpha) > \frac{w(\alpha)}{2} \\ x & \text{otherwise} \end{cases}$$

where $+$ is the binary sum, so $x + \alpha$ is to flip the bits in x corresponding the ones in α. So far, we see that the maximal weight for y is $\lfloor n - \frac{w(\alpha)}{2} \rfloor$. That means we need to select t' such that

$$t' = t(\lfloor n - \frac{w(\alpha)}{2} \rfloor).$$

In the above encoding process, for different binary messages, they have different codewords. And for any codeword y, we have $y \in C$. That is because either $y = x$ or $y = x + \alpha$, where both x and α are codewords in C and C is a linear code. The decoding process is very simple: Given the received word $y' = y + e$, we can always get y by applying the decoding scheme if $|e| \leq t'$. If $y_1 = 1$, that means x has been flipped based on α, so we have $x = y + \alpha$; otherwise, $x = y$. Then the initial message $u = x_2 x_3 ... x_k$.

	Lower Bound	Upper Bound
$\eta_\alpha(n,p,q_e)_{n\to\infty}$	$[1-H(2p)]I_{0\leq p\leq\frac{1}{4}}$	$(1+p)[1-H(\frac{p}{1+p})]$
$\eta_\beta(n,p,q_e)_{n\to\infty}$	$\max_{0\leq\theta\leq 1-p} H(\theta)-\theta H(p)-(1-\theta)H(\frac{p\theta}{1-\theta})$	$\max_{0\leq\theta\leq 1} H((1-p)\theta)-\theta H(p)$

TABLE I

C. Comments

When n is sufficiently large, the codes based on flips above become nearly as efficient as a linear codes correcting $t(\lfloor\frac{n}{2}\rfloor)$ symmetric errors. (We define the codes' efficiency in Section VI.) It is much more efficient than designing a linear code correcting $t(n)$ symmetric errors. Note that when n is large and p is small, these codes can have very good performance on efficiency. That is because when n is sufficiently large, the efficiency of an optimal nonuniform code is dominated by the codewords with the same Hamming weight w_d ($\leq\frac{n}{2}$), and w_d approaches $\frac{n}{2}$ as p gets close to 0. We can intuitively understand it based on two facts when n is sufficiently large: (1) There are at most $n2^{n(H(\frac{w_d}{n})+\delta)}$ codewords in this optimal nonuniform code. (2) When p becomes small, we can get a nonuniform code with at least $2^{n(1-\delta)}$ codewords. So when n is sufficiently large and p is small, we have $w_d\to\frac{n}{2}$. Hence, the optimal nonuniform code has almost the same asymptotic efficiency with an optimal weight-bounded code (Hamming weight is at most n/2), which corrects $t(n/2)$ errors.

Beside simplicity and efficiency, another advantage of these codes is that they do not require the Z-channel to be perfect, i.e., it is allowed to have $0\to 1$ errors with very small probability (as long as this probability is smaller than the probability of $1\to 0$ errors). All these properties make these codes very useful in practice. However, when p is not small, how to design efficient nonuniform codes with simple encoding/decoding schemes is still an open problem.

VI. Bounds on the Rate

Given (n,p,q_e), we can define the efficiency of uniform codes as $\eta_\alpha(n,p,q_e)\triangleq\frac{\log_2 B_\alpha(n,p,q_e)}{n}$ and define the efficiency of nonuniform codes as $\eta_\beta(n,p,q_e)\triangleq\frac{\log_2 B_\beta(n,p,q_e)}{n}$. In this section, given $0<p,q_e<1$, we study the asymptotic behavior of $\eta_\alpha(n,p,q_e)$ and $\eta_\beta(n,p,q_e)$ as $n\to\infty$. Table I summarizes the upper bounds and lower bounds of $\eta_\alpha(n,p,q_e)_{n\to\infty}$ and $\eta_\beta(n,p,q_e)_{n\to\infty}$ obtained in our full paper [16]. We plot them in Fig. 1. The gap between the bounds for the two codes indicates the potential improvement in efficiency by using the nonuniform codes (compared to using uniform codes) when the codeword length is large.

Acknowledgment

This work was supported in part by the NSF CAREER Award CCF-0747415, the NSF grant ECCS-0802107, and by an NSF-NRI award.

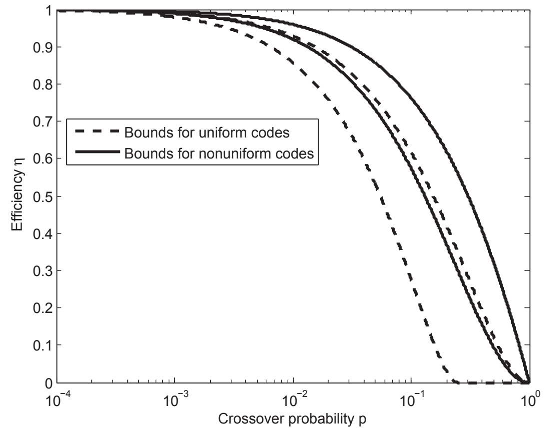

Fig. 1. Bounds to $\eta_\alpha(n,p,q_e)_{n\to\infty}$ and $\eta_\beta(n,p,q_e)_{n\to\infty}$. The dashed curves represent the lower and upper bounds to $\eta_\alpha(n,p,q_e)_{n\to\infty}$, and the solid curves represent the lower and upper bounds to $\eta_\beta(n,p,q_e)_{n\to\infty}$.

References

[1] T. Kløve, "Error correcting codes for the asymmetric channel," Technical Report, Dept. of Informatics, University of Bergen, 1981. (Updated in 1995.)

[2] K. A. S. Abdel-Ghaffar and H. C. Ferreira, "Systematic encoding of the Varshamov-Tenengol'ts codes and the Constantin-Rao codes," *IEEE Trans. Inform. Theory*, vol. 44, pp. 340-345, Jan. 1998.

[3] B. Bose and S. Al-Bassam, "On systematic single asymmetric errorcorrecting codes," *IEEE Trans. Inform. Theory*, vol. 46, pp. 669-672, Mar. 2000.

[4] S. Al-Bassam, R. Venkatesan, and S. Al-Muhammadi, "New single asymmetric error-correcting codes," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1619-1623, Sept. 1997.

[5] Y. Saitoh, K. Yamaguchi, and H. Imai, "Some new binary codes correcting asymmetric/unidirectional errors," *IEEE Trans. Inform. Theory*, vol. 36, pp. 645-647, May 1990.

[6] L. G. Tallini, B. Bose, "On a new class of error control codes and symmetric functions", *IEEE International Symposium on Information Theory (ISIT)*, pp. 980-984, 2008.

[7] T. Etzion, "Lower bounds for asymmetric and unidirectional codes," *IEEE Trans. Inform. Theory*, vol. 37, pp. 1696-1704, Nov. 1991.

[8] Z. Zhang and X. Xia, "New lower bounds for binary codes of asymmetric distance two," *IEEE Trans. Inform. Theory*, vol. 38, pp. 1592-1597, Sept. 1992.

[9] F. Fu, S. Ling, and C. Xing, "New lower bounds and constructions for binary codes correcting asymmetric errors," *IEEE Trans. Inform. Theory*, vol. 49, pp. 3294-3299, Dec. 2003.

[10] C. Wang, S. R. Kulkarni, and H. V. Poor, "Density evolution for asymmetric memoryless channels," *IEEE Trans. Inform. Theory*, vol. 51, pp. 4216-4236, Dec. 2005.

[11] T. Kløve, "Upper bounds on codes correcting asymmetric errors," *IEEE Trans. Inform. Theory*, vol. 27, no. 1, pp. 128-131, 1981.

[12] R. R. Varshamov, "A class of codes for asymmetric channels and a problem from the additive theory of numbers," *IEEE Trans. Inform. Theory*, vol. 19, no. 1, pp. 92-95, 1973.

[13] S. D. Constantin and T. R. N. Rao, "On the theory of binary asymmetric error-correcting codes," *Inform. Contr.*, vol. 40, pp. 20-36, 1979.

[14] P. Delsarte and P. Piret, "Bounds and constructions for binary asymmetric error-correcting codes," *IEEE Trans. Inform. Theory*, vol. 27, pp. 125-128, 1981.

[15] J. H. Weber, C. de Vroedt, and D. E. Boekee, "Bounds and constructions for binary codes of length less than 24 and asymmetric distance less than 6," *IEEE Trans. Inform. Theory*, vol. 34, pp. 1321-1331, Sept. 1988.

[16] H. Zhou, A. Jiang, and J. Bruck, "Nonuniform codes for correcting asymmetric errors in data storage", Technical Report, California Institute of Technology, *http://paradise.caltech.edu/etr.html*, 2011.