Heavy Weight Codes

Gérard Cohen Patrick Solé Aslan Tchamkerten

***Abstract*—Motivated by certain recent problems in asynchronous communication, we introduce and study $B(n,d,w)$, defined as the maximum number of length n binary codewords with minimum distance d, and such that each codeword has weight *at least* w. Specifically, we investigate the asymptotic exponential growth rate of $B(n,d,w)$ with respect to n and with fixed ratios $\delta = d/n$ and $\omega = w/n$. For $\omega \in [0,1/2]$, this growth rate function $b(\delta,\omega)$ is shown to be equal to $a(\delta)$, the asymptotic exponential growth rate of $A(n,d)$ — the maximum number of length n binary codewords with minimum distance d. For $\omega \in (1/2,1]$, we show that $b(\delta,\omega) \le a(\delta,\omega) + f(\omega)$, where $a(\delta,\omega)$ denotes the asymptotic exponential growth rate of $A(n,d,w)$, the maximum number of length n binary codewords with minimum distance d and constant weight w, and where $f(w)$ is a certain function that satisfies $0 < f(\omega) < 0.088$ and $\lim_{\omega\to 1} f(\omega) = \lim_{\omega\to 1/2} f(\omega) = 0$. Based on numerical evidence, we conjecture that $b(\delta,\omega)$ is actually equal to $a(\delta,\omega)$ for $\omega \in (1/2,1]$. Finally, lower bounds on $B(n,d,w)$ are obtained via explicit code constructions.**

***Index Terms*—constant weight codes, asynchronous communication**

I. Introduction

Synchronization and information transmission have traditionally been considered separately. Codewords usually start with a common preamble with good detection properties that helps the decoder locate the subsequent information bits.

Recent results [11], [3] that consider bursty sources of information show that this separation architecture need not be optimal in the high rate regime. In this regime, a combination of synchronization and information transmission can lead to a significant reduction in error probability.

To understand how codewords can carry information while also acting as 'information flag,' let us consider a simple asynchronous channel model recently introduced in [11] and depicted in Fig. 1.

The authors are with Telecom ParisTech (Networks and Computer Science dept. for G. Cohen and Communications and Electronics dept. for P. Solé and A. Tchamkerten), 46 rue Barrault, 75634 Paris Cedex 13, France. A. Tchamkerten is partly supported by an Excellence Chair grant from the French National Research Agency (ANR, ACE project).

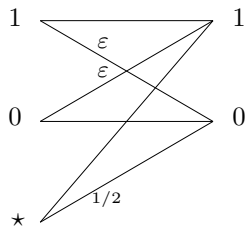

Fig. 1: Binary symmetric channel with 'external' noise.

The transmitter starts sending a randomly selected length n codeword across a binary symmetric channel with crossover probability $\varepsilon \in [0,1/2)$, at a random time ν uniformly distributed over $\{1,2,\ldots,A\}$. The parameter $A \ge 1$ defines the level of asynchronism between the transmitter and the receiver — if $A = 1$, the channel is synchronous. Information transmission ends at time $\nu + N - 1$. Outside the information transmission period, i.e., before time ν and after time $\nu + N - 1$, the receiver observes 'noise,' which we model as zeros and ones occurring independently with equal probability. Note that, outside the information transmission period, one may view the idle transmitter as if it were effectively sending a third symbol (in Fig. 1 given by '$\star$') that produces 0 or 1 with equal probability.

Without knowing ν but knowing A, the goal of the receiver is to reliably *detect and isolate* the sent message (for a precise formulation, we refer to [11, Section IV]).

To achieve good message detection and good message isolation requires different code properties. Message isolation refers to the decoder M-hypothesis test used for discriminating the sent message from the other $M-1$, given that the codeword is correctly located. And it is well known that increasing the minimum distance of a code, in general, lowers the probability of mis-isolation.

In contrast, message detection refers to the 2-hypothesis test for discriminating *a sent message* from noise. Here the main parameter is the (Kullback-Leibler) distance between the output distribution (over n bits) generated by a randomly se-

lected message,[1] and the output distribution induced by noise, in our example given by the binomial $(n, 1/2)$ distribution.

Hence, to reliably detect the sent codeword requires to bias the codewords' weight distribution so that to produce noise atypical symbols at the channel output. This bias, in turn, may reduce the minimum distance of the code.

The above considerations about minimum distance and weight distribution motivates the following basic problem which is the focus of this paper: find the largest number of length n binary codewords, with minimum pairwise Hamming distance d, and such that each codeword has weight at least w.[2]

Our main result is a relation between $B(n, d, w)$, and two fundamental functions in coding theory, namely $A(n, d)$, the maximum number of length n binary codewords with minimum distance d, and $A(n, d, w)$, the maximum number of length n binary codewords, minimum distance d, and weight exactly equal to w [1], [2].

Denoting by $b(\delta, \omega)$, $a(\delta)$, and $a(\delta, \omega)$ the asymptotic exponential growth rates with respect to n with fixed $\delta = d/n$ and $\omega = w/n$, of $B(n, d, w)$, $A(n, d)$, and $A(n, d, w)$, respectively, we show that $b(\delta, \omega) = a(\delta)$ for $\omega \in [0, 1/2]$. For $\omega \in (1/2, 1]$, we show that $b(\delta, \omega) \leq a(\delta, \omega) + f(\omega)$ for some function $f(\omega)$ such that $0 < f(\omega) < 0.088$ and $\lim_{\omega \to 1/2} f(\omega) = \lim_{\omega \to 1} f(\omega) = 0$.

We also prove that $b(\delta, \omega) = a(\delta, \omega)$ for $\omega > 1/2$ under the conjecture that $A(n, d, w)$ is unimodal around $w = n/2$, for any fixed $n \geq 1$ and $d \geq 0$, a property consistent with tables in [1], [2]. Finally, lower bounds on $B(n, d, w)$ are obtained via three code construction techniques: expurgation, translation, and concatenation.

The paper is organized as follows. Section II provides basic combinatorial bounds on $B(n, d, w)$. Section III presents the main results and provides some challenging conjectures. Section IV provides code constructions and thereby additional lower bounds on $B(n, d, w)$.

[1]Recall that there is a uniform prior on the sent message.

[2]For the channel depicted in Fig. 1, to reliably detect the sent codeword, codewords must have a weight distribution either biased towards the all zero sequence or biased towards the all one sequence. In this context, it is meaningful to consider, $B^s(n, d, w)$, the symmetric version of $B(n, d, w)$, defined as the maximum number of binary codewords of length n, minimum distance d, and such that each codeword has a weight either above w or below $n - w$ (with $w \geq n/2$). We however choose to investigate $B(n, d, w)$ instead of $B^s(n, d, w)$ because of the apparent simplicity of dealing with $B(n, d, w)$.

II. Preliminaries

In this section we establish a few basic relations between $B(n, d, w)$ and $A(n, d, w)$.

Note first that $B(n, d, w)$ is increasing in n, and decreasing in d and w. Further, by definition of $B(n, d, w)$, we have

$$B(n, d, w) \geq A(n, d, j) \quad \text{for } j \geq w\,. \tag{1}$$

By taking weight classes sufficiently far apart so that they do not overlap, we get

$$B(n, d, w) \geq \sum_{h=0}^{\lfloor \frac{n-w}{d} \rfloor} A(n, d, w + hd) \tag{2}$$

where $\lfloor x \rfloor$ denotes the largest integer not exceeding x.

Since any code is a disjoint union of constant weight codes, we have

$$B(n, d, w) \leq \sum_{j=w}^{n} A(n, d, j)\,. \tag{3}$$

Removing the weight constraint can only improve the size, hence

$$B(n, d, w) \leq A(n, d) = B(n, d, 0)\,. \tag{4}$$

Finally, the following Gilbert type lower bound is immediate:

Proposition 1: For all $n \geq 1$, $d \leq n$, and $w \leq n$

$$B(n, d, w) \geq \frac{\sum_{i=w}^{n} \binom{n}{i}}{\sum_{i=0}^{d-1} \binom{n}{i}}\,.$$

III. Asymptotics

For fixed $\delta, \omega \in [0, 1]$, we denote by $b(\delta, \omega)$ the asymptotic exponent of $B(n, d, w)$ with respect to n with $d = d(n) = \lfloor \delta n \rfloor$ and $w = w(n) = \lfloor \omega n \rfloor$, i.e.

$$b(\delta, \omega) = \limsup_{n \to \infty} \left(\frac{1}{n} \log B(n, d(n), w(n)) \right)$$

where the logarithm is to the base 2. The asymptotic exponents of $A(n, d, w)$ and $A(n, d)$ are defined similarly and are denoted by $a(\delta, \omega)$ and $a(\delta)$, respectively.

The asymptotic Plotkin bound [7, Theorem 2.10.2], shows that $a(\delta) = 0$ for $\delta \in [1/2, 1]$. Hence, by (4) $b(\delta, \omega) = 0$ for all $\delta \in [1/2, 1]$ and all $\omega \in [0, 1]$. In fact, the support of $b(\delta, \omega)$ can be completely characterized.

Proposition 2: $b(\delta, \omega) > 0$ if and only if $\delta < 2\omega(1 - \omega)$.

Proof of Proposition 2: If $\delta < 2\omega(1 - \omega)$, then $a(\delta, \omega) > 0$ by the 'Gilbert lower bound' [10, p.160, right column, bottom] FIXME: SHOULD WE RESTRICT $\delta < 1/2$

$$a(\delta, \omega) \geq h(\omega) - \omega h(\delta/2\omega) - (1-\omega)h(\delta/2(1-\omega)) \, .$$

Hence, $b(\delta, \omega) > 0$ by (1).

Now, restating a classical lemma of Elias [7, Lemma 2.5.1] yields

$$B(n, d, w) \leq \frac{nd}{nd - 2w(n - w)}$$

whenever $nd > 2w(n-w)$. Hence, by letting $d \simeq \delta n$ and $w \simeq \omega n$ with $\delta > 2\omega(1 - \omega)$, we get

$$\limsup_{n\to\infty} B(n, d, w) \leq \frac{\delta}{\delta - 2\omega(1 - \omega)} \, ,$$

implying that $b(\delta, \omega) = 0$ whenever $\delta > 2\omega(1 - \omega)$. ∎

Theorem 1: For any $\delta \in [0, 1]$ and $\omega \in [0, 1/2]$ we have $b(\delta, \omega) = a(\delta)$.

Proof of Theorem 1: The Elias-Bassalygo bound [10, equation (2.8)]

$$\frac{A(n, d)}{2^n} \leq \frac{A(n, d, w)}{\binom{n}{w}} \tag{5}$$

together with the trivial inequality $A(n, d, w) \leq A(n, d)$ shows that the asymptotic exponents of $A(n, d)$ and $A(n, d, n/2)$ are the same.

The result then follows by combining the bounds (1) and (4) to obtain

$$A(n, d, n/2) \leq B(n, d, w) \leq A(n, d)$$

for $w \leq n/2$. ∎

Clearly $b(\delta, \omega) \geq a(\delta, \omega)$ by (1). The following theorem provides a bound on $b(\delta, \omega) - a(\delta, \omega)$ that is uniform in δ.

Theorem 2: For any $\delta \in [0, 1]$ and $\omega \in (1/2, 1]$,

$$b(\delta, \omega) - a(\delta, \omega) \leq f(\omega)$$

where

$$f(\omega) \triangleq \min\{g(\omega), 1 - h(\omega)\} \, ,$$

$$g(\omega) \triangleq \omega(\log \omega - \log e) + \log e \, ,$$

and

$$h(\omega) \triangleq -\omega \log \omega - (1 - \omega) \log(1 - \omega)$$

is the binary entropy function.

It can be checked that the maximum of $f(\omega)$ is ≈ 0.088 and occurs at $\omega \approx 0.672$, that $f(\omega)$ is unimodal around its maximum, and that $f(\omega) \to 0$ as $\omega \to 1/2$ and as $\omega \to 1$.

Proof of Theorem 2: We first show that

$$b(\delta, \omega) - a(\delta, \omega) \leq 1 - h(\omega) \, . \tag{6}$$

From the Elias-Bassalygo bound (5), we get the following inequality in terms of growth rates

$$a(\delta, \omega) \geq a(\delta) + h(\omega) - 1 \, .$$

Hence,

$$a(\delta) \geq b(\delta, \omega) \geq a(\delta, \omega) \geq a(\delta) + h(\omega) - 1 \, ,$$

from which (6) follows.

To establish that

$$b(\delta, \omega) - a(\delta, \omega) \leq g(\omega) \, , \tag{7}$$

we need the following two Lemmas.

Lemma 1: For any $\delta \in [0, 1]$ and $\omega \in [0, 1]$

$$b(\delta, \omega) = \sup\{a(\delta, \rho), \ \omega \leq \rho \leq 1\} \, .$$

Proof of Lemma 1: We have

$$\begin{aligned} \max_{j\in\{w,w+1,\ldots,n\}} A(n, d, j) &\leq B(n, d, w) \\ &\leq (n - w + 1) \max_{j\in\{w,w+1,\ldots,n\}} A(n, d, j) \end{aligned}$$

by (1) for the first inequality and by (3) for the second inequality. The lemma then follows, after some algebra. ∎

The following lemma provides a weaker proof of the Johnson type of bound obtained in [8, Ch. 17, Thm 4].

Lemma 2: For $w \leq n$ we have

$$A(n, d, w) \leq \frac{n}{w} A(n, d, w - 1) \, .$$

Proof: Let C be a constant weight code realizing $A(n, d, w)$, and consider the matrix whose rows are the codewords of C. The average weight W of

a column is given by the total number of 1's in the matrix divided by n, i.e.,

$$W = \frac{wA(n,d,w)}{n}.$$

Now, say column l has weight at least W (one such column clearly exists). Pick the subcode of C given by the codewords of C that have a 1 in the l-th position. Modify this subcode by changing 1 into 0 in the l-th component of each codeword. If we denote by C' the code obtained after the above two procedures, we conclude that $W \leq |C''| \leq A(n,d,w-1)$. The lemma follows. ∎

To prove (7), we first iterate Lemma 2 u times to obtain

$$A(n,d,r) \leq \frac{n^u}{r(r-1)\cdots(r-u+1)} A(n,d,w)\,,$$

with $r = w + u$. Therefore,

$$A(n,d,r) \leq A(n,d,w) n^{r-w} \frac{w!}{r!}.$$

By setting $r = \lfloor \rho n \rfloor$ ρn and $w = \lfloor \omega n \rfloor$, with $1/2 < \omega \leq \rho \leq 1$, and using Stirling's approximation, we get

$$a(\delta,\rho) \leq a(\delta,\omega) + \omega(\log\omega - \log e) + \rho(\log e - \log\rho)\,.$$

The maximum of $\rho(\log e - \log\rho)$ for $\rho \in [\omega, 1]$ is achieved for $\rho = 1$. This combined with Lemma 1 gives $b(\delta,\omega) - a(\delta,\omega) \leq g(\omega)$. ∎

The following conjecture — given in [8, Appendix A, A5] as a research problem — is consistent with the tables in [1], [2].

Conjecture 1: For fixed n and d, the function $A(n,d,w)$ is unimodal in w, with a maximum at $w = n/2$ for n even, and with two maxima at $w = (n-1)/2$ and $w = (n+1)/2$ for n odd.

The following conjecture is an asymptotic, hence weaker version, of Conjecture 1.

Conjecture 2: For fixed δ, the function $a(\delta,\omega)$ is unimodal with a maximum at $\omega = 1/2$.

The following conjecture provides a stronger statement than Theorem 2.

Conjecture 3: If $\omega > 1/2$, then $b(\delta,\omega) = a(\delta,\omega)$.

Proposition 3: Conjectures 2 and 3 are equivalent.

Proof: Lemma 1 together with Conjecture 2 gives Conjecture 3. Conjecture 3 implies Conjecture 2 since $b(\delta,\omega)$ is non-increasing in ω. ∎

IV. Constructions

Three well studied code construction techniques are expurgation, translation, and concatenation. The first is perhaps mostly of theoretical interest as a good decoding algorithm needs not, in general, provide a good decoding algorithm for a subcode. In contrast, the other two techniques also provide practical decoding algorithms.

A. Expurgation

The following result shows that, for $w \leq d$, $B(n,d,w)$ and $A(n,d)$ are essentially the same (recall that $B(n,d,w) \leq A(n,d)$).

Proposition 4: For $1 \leq w \leq d \leq n$, we have

$$B(n,d,w) \geq A(n,d) - 1.$$

Proof: Let C be a code achieving $A(n,d)$. By first translating this code so that to include the all-zero codeword, then by removing the all-zero codeword, we get a new code of size $A(n,d)-1$, with minimum distance and weight both at least equal to d. The proposition follows. ∎

Theorem 3: For all large enough and even n, all $w \leq n/2$, and all $d \leq nh^{-1}(1/2)$,[3] we have

$$B(n,d,w) \geq 2^{(n-2)/2}.$$

Proof: Pick a self dual code above the Gilbert bound [9]. This code being binary self-dual, contains the all-one codeword, and is therefore self-complementary. Hence, half of its codewords at least have weight at least $n/2$. ∎

B. Translation

The following result fills the gap FIXME: BE MORE PRECISE left by Proposition 4. We assume that the reader has some familiarity with the covering radius concept [5].

Proposition 5: Fix two integers $n \geq 1$ and $d \geq 1$, and let $e = \lfloor (d-1)/2 \rfloor$. If $w \leq e$, or, if $w \leq e+1$ when no perfect code realizes $A(n,d)$, then

$$B(n,d,w) = A(n,d).$$

[3] $h^{-1}(\cdot)$ denotes the inverse function of the binary entropy over the range $[0, 1/2]$.

Proof: Pick a code C realizing $A(n,d)$. There exists a translate of C of weight w as long as w is below the covering radius of C, in particular as long as w is below the packing radius of C, or the packing radius plus one if C is not perfect. This gives $B(n,d,w) \geq A(n,d)$. The reverse inequality is (4). ■

C. Concatenation

Consider a binary code of length n, size 2^m, minimum weight w, and distance d. If we concatenate this code with a code of length N, minimum weight W, and minimum distance D over $GF(2^m)$, we get a binary code of length $N2^m$, weight at least wW, and minimum distance dD. Hence

$$B(Nn, dD, wW) \geq B_{2^m}(N, D, W)\,,$$

where $B_q(\cdot,\cdot,\cdot)$ is the natural generalization of $B(\cdot,\cdot,\cdot)$ to an alphabet of size q.

Efficient decoding algorithms for concatenated codes can be found in [6].

V. Concluding Remarks

Motivated by certain problems arising in the context of asynchronous communication, we introduced $B(n,d,w)$, defined as the maximum number of length n binary codewords with minimum distance d and weight at least w. $B(n,d,w)$ has close ties with $A(n,d,w)$ and the main problems at this stage are whether they are actually equal or at least asymptotically equal. To handle the latter problem, and in view of the difficulty of Conjecture 1, open for more than thirty years, one may want to try proving Conjecture 2 independently.

VI. Acknowledgments

We are grateful to Iiro Honkala for helpful discussions, and to the anonymous referees for helpful comments.

References

[1] Best, M. R.; Brouwer, A. E.; MacWilliams, F. Jessie; Odlyzko, Andrew M.; Sloane, Neil J. A. Bounds for binary codes of length less than 25. IEEE Trans. Information Theory IT-24 (1978), no. 1, 81–93.

[2] Brouwer, A. E.; Shearer, James B.; Sloane, N. J. A.; Smith, Warren D. A new table of constant weight codes. IEEE Trans. Inform. Theory IT-36 (1990), no. 6, 1334–1380.

[3] Chandar, Venkat; Tchamkerten, Aslan; Wornell, Gregory, Training-based schemes are suboptimal for high rate asynchronous communication, Information Theory Workshop (ITW),Taormina, Italy, October 2009.

[4] Brouwer, A. E., Cohen, A. M.; Neumaier, A. *Distance-regular graphs.* Ergebnisse der Mathematik und ihrer Grenzgebiete (3) , 18. Springer-Verlag, Berlin, 1989.

[5] G.Cohen, I.Honkala, S.Litsyn, A.Lobstein, *Covering Codes*, Elsevier, 1997.

[6] Dumer, Ilya I., Concatenated codes and their multilevel generalizations in *Handbook of coding theory*, Vol. II, V. Pless and W.C Huffman, eds, 1911–1988, North-Holland, Amsterdam, 1998.

[7] W. Cary Huffman, Vera Pless *Fundamentals of error correcting codes*, Cambridge (2003).

[8] MacWilliams, F. J.; Sloane, N. J. A, *The theory of Error Correcting Codes*, North Holland (1977).

[9] MacWilliams, F. J.; Sloane, N. J. A.; Thompson, J. G. Good self dual codes exist. Discrete Math. 3 (1972), 153–162.

[10] McEliece, Robert J.; Rodemich, Eugene R.; Rumsey, Howard, Jr.; Welch, Lloyd R. New upper bounds on the rate of a code via the Delsarte-MacWilliams inequalities. IEEE Trans. Information Theory IT-23 (1977), no. 2, 157–166.

[11] Tchamkerten, Aslan; Chandar, Venkat; Wornell, Gregory, Communication under strong asynchronism, IEEE Trans. Information Theory, IT-55 (2010), no. 10, 4508-4528.