VI. Conclusion

The distributions of loop lengths in the graphical models for turbo decoding were analyzed and simulated. Short loops (e.g., of length $c \leq 8$) occur with relatively low probability at any randomly chosen node. As the loop length increases, there is a threshold effect and the probability of finding a loop of length c or less approaches 1 (e.g., for $c > 20$). For turbo codes, as K, the number of information bits, becomes large, the probability that a loop of length c or less exists at any randomly chosen node behaves approximately as $e^{-\frac{2^{C-1}-4}{K}}$, $c \geq 4$. In summary, the results in this correspondence demonstrate that the loop lengths in the graphical models of turbo codes and LDPC codes have a specific distributional character. We hope that this information can be used to further understand the workings of iterative decoding.

Acknowledgment

The authors wish to thank R. J. McEliece and the coding group at Caltech for many useful discussions and feedback.

References

[1] C. Berrou, A. Glavieux, and P. Thitimajshima, “Near Shannon limit error-correcting coding and decoding: Turbo codes,” in *Proc. IEEE Int. Conf. Communications*, 1993, pp. 1064–1070.

[2] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Mateo, CA.: Morgan Kaufmann, 1988.

[3] R. J. McEliece, D. J. C. MacKay, and J.-F. Cheng, “Turbo decoding as an instance of Pearl’s ‘belief propagation’ algorithm,” *IEEE J. Select. Areas Commun.*, vol. 16, pp. 140–152, Feb. 1998.

[4] F. R. Kschischang and B. J. Frey, “Iterative decoding of compound codes by probability propagation in graphical models,” *IEEE J. Select. Areas Commun.*, vol. 16, pp. 219–230, Feb. 1998.

[5] R. J. McEliece, E. R. Rodemich, and J.-F. Cheng, “The turbo-decision algorithm,” in *Proc. 33rd Annu. Allerton Conf. Communication, Control, and Computing*, Oct. 4–6, 1995, pp. 366–379.

[6] S. Dolinar and D. Divsalar, “Weight distributions for turbo codes using random and nonrandom permutations,” Jet Propulsion Lab., Pasadena, CA, Tech. Rep., Aug. 1995.

[7] X. Ge, D. Eppstein, and P. Smyth. (1999, Mar.) The distribution of cycle lengths in graphical models for iterative decoding. Univ. Calif. Irvine, Tech. Rep. UCI-ICS 99-10. [Online]. Available: http://www.datalab.uci.edu/papers/trcycle.pdf.

[8] R. G. Gallager, *Low-Density Parity-Check Codes*. Cambridge, MA: MIT Press, 1963.

[9] D. J. C. MacKay and R. M. Neal, “Near Shannon limit performance of low density parity check codes,” *Electron. Lett.*, vol. 32, no. 18, pp. 1645–1646, Aug. 1996. Reprinted in *Electron. Lett.*, vol. 33, no. 6, March 13, pp. 457–458, 1997.

[10] B. J. Frey, *Graphical Models for Machine Learning and Digital Communication*. Cambridge, MA: MIT Press, 1998.

Constructions for Perfect 2-Burst-Correcting Codes

Tuvi Etzion, *Senior Member, IEEE*

***Abstract*—In this correspondence, we present two constructions. In the first construction, we show how to generate perfect linear codes of length 2^{r-1}, $r \geq 5$, and redundancy r, which correct a single burst of length 2. In the second construction, we show how to generate perfect linear codes organized in bytes of length 2^{r-1}, $r \geq 5$, with redundancy divisible by r, which correct a single burst of length 2 within the bytes.**

***Index Terms*—Byte-organized memory, perfect codes, syndromes, 2-burst-correcting codes.**

I. Introduction

In many memory systems the most common error is a burst of errors. Some memories are byte-oriented and in these memories the burst can occur only within the bytes. In this correspondence, we consider constructions of perfect 2-burst-correcting codes of these types.

All codes considered in this correspondence are binary codes. Let (n, k) code denote a linear code of length n and dimension k. The redundancy r of such a code is $r = n - k$. A code is called a *b-burst-correcting code* if it can correct any single burst of length b or less. A code is called *cyclic b-burst-correcting code* if it can correct any single cyclic burst of length b or less. Note that in other papers, e.g., [1], a b-burst-correcting code can correct cyclic bursts. Reiger [8] has proved that an $(n, n-r)$ b-burst-correcting code (or cyclic b-burst-correcting code) must satisfy

$$r \geq 2b. \tag{1}$$

For an $(n, n-r)$ cyclic b-burst-correcting code, there are $n2^{b-1}$ different cyclic bursts of length up to b, each one corresponding to a different nonzero syndrome. Since the number of nonzero vectors of length r is $2^r - 1$, it follows that $2^r - 1 \geq n2^{b-1}$. Abramson [2] noted that this inequality, along with the fact that n is an integer, implies

$$n \leq 2^{r-b+1} - 1. \tag{2}$$

A cyclic b-burst-correcting code which satisfies (2) with equality is said to be *optimum*. Abramson [2] has constructed optimum cyclic 2-burst-correcting codes for any $r \geq 3$. Elspas and Short [5] constructed such codes for $b = 3$ and $b = 4$, Abdel-Ghaffar, McEliece, Odlyzko, and van Tilborg [1] have shown that infinitely many optimum codes exist for each $b \geq 1$. It is clear that if a code satisfies (2) with equality then $2^r - 1 > n2^{b-1}$ (when $b > 1$), i.e., not all nonzero vectors of length r are syndromes obtained as sums of at most b adjacent columns. Therefore, these optimum codes are not perfect. For an $(n, n-r)$ b-burst-correcting code, there are $(n - b + 2)2^{b-1} - 1$ different bursts of length up to b, each one corresponding to a different nonzero syndrome. Again, since the number of nonzero vectors of length r is $2^r - 1$, it follows that $2^r - 1 \geq (n - b + 2)2^{b-1} - 1$, i.e.,

$$n \leq 2^{r-b+1} + b - 2. \tag{3}$$

A b-burst-correcting code which satisfies (3) with equality is said to be *perfect* (or “complete”as called by Reiger [8]). No perfect b-burst-

Manuscript received October 15, 2000; revised April 20, 2001. This work was supported in part by the Israeli Science Foundation under Grant 88/99-1.

The author is with the Computer Science Department, Technion–Israel Institute of Technology, Haifa 32000, Israel (e-mail: etzion@cs.technion.ac.il).

Communicated by S. Litsyn, Associate Editor for Coding Theory.
Publisher Item Identifier S 0018-9448(01)07019-5.

correcting code is known to date. In this correspondence, we will give constructions for perfect 2-burst-correcting codes. Some of these codes will be byte-oriented codes. For more information on byte-oriented error-correcting codes, burst-correcting codes, and their applications, the reader is referred to [3], [4], [7].

The rest of this correspondence is organized as follows. In Section II, we will construct a $(2^{r-1}, 2^{r-1} - r)$ perfect 2-burst-correcting code for each $r \geq 5$. In Section III, we construct perfect codes which correct a single burst of length 2 within bytes of sizes 2^{r-1}, $r \geq 5$, for each redundancy which is divisible by r. We also construct perfect codes which correct a single burst of length 2 within bytes of length 3 for each redundancy divisible by 4. Conclusion and a list of open problems are given in Section IV.

II. Perfect Burst-Correcting Codes

In this section, we show that for each $r \geq 5$ there exists a $(2^{r-1}, 2^{r-1} - r)$ perfect 2-burst-correcting code. For a parity-check matrix $H = [h_1, h_2, \ldots, h_n]$ of an (n, k) code C, let $S(H)$ denote the set of syndromes obtained from single columns and two adjacent columns, i.e.,

$$S(H) = \{h_i : 1 \leq i \leq n\} \cup \{h_i + h_{i+1} : 1 \leq i \leq n-1\}.$$

C is a perfect 2-burst-correcting code if every nonzero column vector of length r belongs to $S(H)$ and

$$\{h_i : 1 \leq i \leq n\} \cap \{h_i + h_{i+1} : 1 \leq i \leq n-1\} = \emptyset.$$

By (1) there is no 2-burst-correcting code for redundancy less than 4. It can be easily verified by simple backtracking that an $(8, 4)$ 2-burst-correcting code does not exist. Perfect 2-burst-correcting codes for redundancies 5 and 6, i.e., perfect $(16, 11)$ and perfect $(32, 26)$ 2-burst-correcting codes, respectively, were obtained by computer search. Their parity-check matrices are given below.

$$\begin{bmatrix} 1000101010000101 \\ 0100001001101011 \\ 0010101000110010 \\ 0001000110101101 \\ 0000010101010110 \end{bmatrix}$$

$$\begin{bmatrix} 10001001110101110011110010011100 \\ 01000101000110110000100101101111 \\ 00101001000000001011011010101010 \\ 00010100100100000101101011011111 \\ 00000010010010101010111011111110 \\ 00000000001001010101010101010101 \end{bmatrix}.$$

Given an $r \times 2^{r-1}$ parity-check matrix H for a perfect 2-burst-correcting code we construct the following four matrices:

$$\tilde{H}_1 = \begin{bmatrix} H \\ T_1 \end{bmatrix} \text{ where } T_1 = \begin{bmatrix} 1 & 1 & 1 & 1 & \cdots & 1 & 1 \\ 0 & 1 & 0 & 1 & \cdots & 0 & 1 \end{bmatrix}$$

$$\tilde{H}_2 = \begin{bmatrix} H \\ T_2 \end{bmatrix} \text{ where } T_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

$$\tilde{H}_3 = \begin{bmatrix} H \\ T_3 \end{bmatrix} \text{ where } T_3 = \begin{bmatrix} 1 & 0 & 1 & 0 & \cdots & 1 & 0 \\ 1 & 1 & 1 & 1 & \cdots & 1 & 1 \end{bmatrix}$$

$$\tilde{H}_4 = \begin{bmatrix} H \\ T_4 \end{bmatrix} \text{ where } T_4 = \begin{bmatrix} 0 & 1 & 0 & 1 & \cdots & 0 & 1 \\ 1 & 0 & 1 & 0 & \cdots & 1 & 0 \end{bmatrix}$$

Lemma 1: Each column vector of length $r+2$ which does not begin with r zeros appears exactly once either as a column of one of the $\tilde{H}_i$'s or as a sum of two adjacent columns from one of the $\tilde{H}_i$'s.

Proof: The proof follows immediately from the following observations.

- In each position the four T_i's have exactly the four possible 2-tuples.
- In each pair of adjacent positions, the adjacent columns in the four T_i's sum exactly to the four possible 2-tuples.
- Each nonzero column vector of length r appears exactly once either as a column or as sum of two adjacent columns from H. $\square$

For an $r \times n$ matrix

$$F = [\, f_1 \quad f_2 \quad f_3 \quad \cdots \quad f_{n-2} \quad f_{n-1} \quad f_n \,]$$

let

$$\begin{aligned} p(F) &= [\, f_1 + f_2 \quad f_2 \quad f_3 \quad \cdots \quad f_{n-2} \quad f_{n-1} \quad f_n \,] \\ s(F) &= [\, f_1 \quad f_2 \quad f_3 \quad \cdots \quad f_{n-2} \quad f_{n-1} \quad f_{n-1} + f_n \,] \\ e(F) &= [\, f_1 + f_2 \quad f_2 \quad f_3 \quad \cdots \quad f_{n-2} \quad f_{n-1} \quad f_{n-1} + f_n \,] \\ F^R &= [\, f_n \quad f_{n-1} \quad f_{n-2} \quad \cdots \quad f_3 \quad f_2 \quad f_1 \,]. \end{aligned}$$

The following lemma is a simple observation.

Lemma 2: For any $r \times n$ parity-check matrix H

$$S(H) = S(p(H)) = S(s(H)) = S(e(H)) = S(H^R).$$

We are now in a position to state our main result.

Theorem 1: If $H = [h_1\, h_2 \cdots h_n]$ is an $r \times n$, $n = 2^{r-1}$, parity-check matrix of a perfect 2-burst-correcting code then the matrix shown at the bottom of the page is an $(r+2) \times 2^{r+1}$ parity-check matrix of a perfect 2-burst-correcting code.

Proof: Clearly, we need only to prove that every nonzero vector of length $r+2$ is a column of $\tilde{H}$ or the sum of two adjacent columns of $\tilde{H}$.

Let $\tilde{X} = (x_1, \ldots, x_r, x_{r+1}, x_{r+2})$ be a nonzero vector of length $r+2$ and $X = (x_1, \ldots, x_r)$. Let Y^t denote the transpose of Y. We distinguish between two cases.

Case 1: If X is a nonzero vector then by Lemmas 1 and 2 $\tilde{X}^t$ is either a column of $\tilde{H}$ or the sum of two adjacent columns of $\tilde{H}$.

Case 2: X is the allzero vector.

- If $(x_{r+1}, x_{r+2}) = (1, 1)$ then $\tilde{X}^t$ is the sum of the last column of $\tilde{H}_1$ and the first column of $s(\tilde{H}_2^R)$.
- If $(x_{r+1}, x_{r+2}) = (1, 0)$ then $\tilde{X}^t$ is the sum of the last column of $s(\tilde{H}_2^R)$ and the first column of $e(\tilde{H}_3)$.
- If $(x_{r+1}, x_{r+2}) = (0, 1)$ then $\tilde{X}^t$ is the sum of the last column of $e(\tilde{H}_3)$ and the first column of $p(\tilde{H}_4^R)$. $\square$

Combining Theorem 1, the perfect $(16, 11)$ 2-burst-correcting code, and the perfect $(32, 26)$ 2-burst-correcting code given above we have the following.

Corollary 1: For each $r \geq 5$, there exists a perfect $(2^{r-1}, 2^{r-1} - r)$ 2-burst-correcting code.

$$\tilde{H} = [\, \tilde{H}_1 \quad s(\tilde{H}_2^R) \quad e(\tilde{H}_3) \quad p(\tilde{H}_4^R) \,]$$

$$= \left[\begin{array}{ccccc|ccccc|ccccc|ccccc} h_1 & h_2 & \cdots & h_{n-1} & h_n & h_n & h_{n-1} & \cdots & h_2 & h_2 + h_1 & h_1 + h_2 & h_2 & \cdots & h_{n-1} & h_{n-1} + h_n & h_n + h_{n-1} & h_{n-1} & \cdots & h_2 & h_1 \\ 1 & 1 & \cdots & 1 & 1 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 1 & 1 & 1 & 0 & \cdots & 1 & 0 \\ 0 & 1 & \cdots & 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 & \cdots & 1 & 0 & 1 & 1 & \cdots & 0 & 1 \end{array} \right]$$

III. Codes for Byte-Oriented Memories

As said in the Introduction, we are also interested in perfect burst-correcting codes within bytes of size β. If the codes are constructed for a byte-oriented memory with bytes of size β, then when we say that the code is b-burst-correcting meaning that it corrects a single burst of length b within the bytes of size β. One example is a perfect $(9, 5)$ 2-burst-correcting code which corrects a single burst of length 2 or less within bytes of size 3. Its parity-check matrix is

$$\begin{bmatrix} 110 & 110 & 100 \\ 101 & 101 & 000 \\ 111 & 001 & 010 \\ 011 & 110 & 001 \end{bmatrix}.$$

We first want to give a necessary condition for the existence of perfect 2-burst-correcting codes with bytes of length β. The number of different bursts of length 2 or less within a byte is $2\beta - 1$. The total number of nonzero vectors of length r is $2^r - 1$ and hence the total number of bytes in such a code with redundancy r is $\frac{2^r-1}{2\beta-1}$. Therefore,

$$2^r - 1 \equiv 0 \,(\mathrm{mod}\, 2\beta - 1). \qquad (4)$$

In the most interesting case, when β is a power of 2, say $\beta = 2^\rho$, the necessary condition (4) implies that $2^r - 1$ is divisible by $2^{\rho+1} - 1$, i.e., $\rho + 1$ divides r. We will show that, in this case, the necessary condition is also sufficient.

Assume now that $2^r - 1 = m(2\beta - 1)$ and there exists a perfect 2-burst-correcting code of length $m\beta$ with m bytes of size β, and redundancy r. We will describe now how to construct the parity-check matrix of a perfect 2-burst-correcting code of length $mt\beta$ and redundancy rs, where $t = (2^{sr} - 1)/(2^r - 1)$, with mt bytes of size β, for any $s > 1$.

Let $H = [h_1\, h_2 \cdots h_n]$, $n = m\beta$, be the parity-check matrix of a perfect $(m\beta, m\beta - r)$ 2-burst-correcting code with bytes of size β. Viewing the columns of H as elements of GF$\,(2^r)$ we find a vector $(a_1\, a_2 \cdots a_n)$, where $a_i \in$ GF$\,(2^r)$, such that the sums of at most two adjacent elements within the bytes of size β give each of the nonzero elements of GF$\,(2^r)$ exactly once.

Let α be a primitive element of GF$\,(2^{rs})$. Then it has the subfield

$$\mathrm{GF}\,(2^r) = \{0\} \cup \{\alpha^{jt} :\, 0 \le j \le 2^r - 2\}.$$

If we denote $F^* = \mathrm{GF}\,(2^r) \setminus \{0\}$, then

$$\mathrm{GF}\,(2^{rs}) \setminus \{0\} = F^* \cup \alpha F^* \cup \alpha^2 F^* \cup \cdots \cup \alpha^{t-1} F^*.$$

But, then the vector

$$v = (v_0\, v_1\, \cdots\, v_{t-1}) \in \mathrm{GF}\ (2^{rs})^{mt\beta}$$

where $v_i = (\alpha^i a_1\, \alpha^i a_2\, \cdots\, \alpha^i a_n)$, for each i, $0 \le i \le t-1$, has the property that the sums of at most two adjacent elements within the byte of size β give each of the nonzero elements of GF$\,(2^{rs})$ exactly once. Viewing the columns as binary vectors of dimension rs we obtain the following result.

Theorem 2: If there exists a perfect $(m\beta, m\beta - r)$ 2-burst-correcting code with bytes of size β, then there exists a perfect $(mt\beta, mt\beta - rs)$ 2-burst-correcting code with bytes of size β, where $t = (2^{sr} - 1)/(2^r - 1)$, for any $s > 1$.

Combining Theorem 2 with the perfect $(2^{r-1}, 2^{r-1} - r)$ 2-burst-correcting codes, $r \ge 5$, obtained in Section II and the perfect $(9, 5)$ 2-burst-correcting code with bytes of size 3 obtained in this section we have the following.

Corollary 2: There exists a perfect $(n, n - \rho)$ 2-burst-correcting code with bytes of size 2^{r-1}, $r \ge 5$, for each redundancy ρ, $\rho \equiv 0 \,(\mathrm{mod}\, r)$.

Corollary 3: There exists a perfect $(n, n - \rho)$ 2-burst-correcting code with bytes of size 3, for each redundancy ρ, $\rho \equiv 0 \,(\mathrm{mod}\, 4)$.

Finally, we want to remark that the construction given and Theorem 2 apply also for perfect b-burst-correcting code with $b > 2$. For example, it can be applied when the length of the burst is the same as the length of the byte (see [6]).

IV. Conclusion

We gave a construction for perfect $(2^{r-1}, 2^{r-1} - r)$ 2-burst-correcting code for each $r \ge 5$. We also gave a construction for a perfect 2-burst-correcting codes for byte-oriented memories. The construction yields codes which are organized in bytes of length 2^{r-1}, $r \ge 5$, for each redundancy ρ, $\rho \equiv 0 \,(\mathrm{mod}\, r)$. The construction also yields perfect 2-burst-correcting codes with bytes of size 3 for each redundancy ρ divisible by 4.

The bound given in (3) does not exclude the existence of perfect b-burst-correcting codes for each $b \ge 2$, and each redundancy r, $r \ge 2b$, as implied by (1). We conjecture that for $r = 2b$ no such code exists, but it is an intriguing question whether such codes for $r > 2b$ and $b > 2$ exist.

Another interesting question is whether there exist perfect 2-burst-correcting codes for nonbinary alphabets. Two problems should be solved in this context. The first is to find some seed codes and the second is to give a recursive construction for such codes. We note that the recursive construction given in Section II does not work for nonbinary alphabets.

Acknowledgment

The author wishes to thank M. Biberstein who found the first two $(16, 11)$ and $(32, 26)$ 2-burst-correcting codes. Their existence and a discussion with R. Roth and G. Seroussi urged the author to find the construction given in this correspondence. The author would also like to thank two anonymous referees for their constructive suggestions.

References

[1] K. A. S. Abdel-Ghaffar, R. J. McEliece, A. M. Odlyzko, and H. C. A. van Tilborg, "On the existence of optimum cyclic burst correcting codes," *IEEE Trans. Inform. Theory*, vol. IT-32, pp. 768–775, Nov. 1986.

[2] N. M. Abramson, "A class of systematic codes for nonindependent errors," *IRE Trans. Inform. Theory*, vol. IT-5, pp. 150–157, Dec. 1959.

[3] C. L. Chen, "Error-correcting codes with byte error detection capabilities," *IEEE Trans. Comput.*, vol. C-32, pp. 615–621, 1983.

[4] ——, "Byte oriented error-correcting code for semiconductor memory systems," *IEEE Trans. Comput.*, vol. C-35, pp. 646–648, 1986.

[5] B. Elspas and R. A. Short, "A note on optimum burst-error-correcting codes," *IRE Trans. Inform. Theory*, vol. IT-8, pp. 39–42, Jan. 1962.

[6] T. Etzion, "Perfect byte-correcting codes," *IEEE Trans. Inform. Theory*, vol. 44, pp. 3140–3146, Nov. 1998.

[7] T. R. N. Rao and E. Fujiwara, *Error-Control Coding for Computer System*. Englewood Cliffs, NJ: Prentice-Hall, 1989.

[8] S. H. Reiger, "Codes for the correction of 'clustered' errors," *IRE Trans. Inform. Theory*, vol. IT-6, pp. 16–21, Mar. 1960.