TABLE I
THE $(n_0, k_0) = (2, 1)$ FIXED $(T = 1)$ NONCATASTROPHIC CONVOLUTIONAL ENCODER WITH MEMORY $M = 4$ AND MAXIMUM FREE DISTANCE TOGETHER WITH SOME BETTER TIME - VARYING ENCODERS[1]

T	G_0^*	G_1^*	G_2^*	G_3^*	G_4^*	d_∞	N_∞	I_∞
1	3	2	2	1	3	7	2	4
2	E	A	C			7	3/2	3
	3	B	7					
3	3A	1C	00			7	4/3	8/3
	0E	3B	00					
	03	2D	30					
4	E9	C0				7	3/2	13/4
	3B	70						
	0D	EC						
	03	B7						
5	3A7	000				7	6/5	13/5
	0ED	300						
	037	1C0						
	00D	3B0						
	003	2DC						

[1] d_∞ The fixed encoder has maximum free distance and is optimum in the sense of smallest number of weight d_∞ paths per time instant N_∞ and in the sense of smallest average number of information bit errors along these paths I_∞. The encoders are specified by the (n_0T, k_0T) matrices G_j^* of (7); the rows of G_j^* are written in hexadecimal form; e.g., the upper row of G_0^* for $T = 2$ is specified as "E" and hence is the row $[1\ 1\ 1\ 0]$.

where G_j^* for $0 \leqslant j \leqslant M^*$ is the $(Tk_0) \times (Tn_0)$ matrix

$$G_j^* = \begin{bmatrix} G_{jT}(0) & G_{jT+1}(1) & \cdots & G_{jT+T-1}(T-1) \\ G_{jT-1}(0) & G_{jT}(1) & \cdots & G_{jT+T-2}(T-1) \\ \vdots & \vdots & & \vdots \\ G_{jT-T+1}(0) & G_{jT-T+2}(1) & \cdots & G_{jT}(T-1) \end{bmatrix} \tag{7}$$

where by way of convention $G_j(u) = 0$ for $j > M$ and for $j < 0$ and where M^* is the smallest integer equal to or greater than M/T, i.e.,

$$M^* = \lceil M/T \rceil. \tag{8}$$

Thus, every (n_0, k_0) periodic convolutional encoder with period T and memory M can be considered to be an (n_0^*, k_0^*) fixed convolutional encoder with memory M^* given by (8) and with

$$k_0^* = Tk_0 \tag{9a}$$

$$n_0^* = Tn_0. \tag{9b}$$

This equivalence permits upper bounds on code distance, as well as software, developed for FCE's to be applied to PCE(T)'s.

III. NEW CODES

We now report some positive results from a computer-aided search for noncatastrophic PCE(T)'s with $(n_0, k_0) = (2, 1)$ that are superior to the best noncatastrophic FCE with the same parameters n_0, k_0, and M. The search was concentrated on the case $M = 4$ as this is the smallest M such that Heller's upper bound [4] on d_∞, namely $d_\infty = 8$, is not achieved by any FCE. The results of the search are given in the Table I. The $T = 1$ entry is the fixed convolutional encoder found by Oldenwalder [5] that has maximum d_∞, namely $d_\infty = 7$, and is optimum both in the sense of minimum N_∞ and in the sense of minimum I_∞. For $T > 1$, the codes are time-varying but are specified by the corresponding fixed encoding matrices G_j^*, $0 \leqslant j \leqslant M^*$, defined by (7).

In Table I, the encoders with period $T = 2$ and 3 were found by an exhaustive search to be optimal in the sense of minimizing N_∞; the codes for $T = 4$ and 5 were the best found in a heuristic nonexhaustive search. The codes given in Table I for $T = 2$, 3, 4, and 5 are all superior to the best fixed code $(T = 1)$ both in the sense of smaller N_∞ and also in the sense of smaller I_∞.

It is somewhat disappointing that no time-varying codes with larger d_∞ than the best fixed code were found. It seems likely that no such superiority is possible for $M = 4$ when $(n_0, k_0) = (2, 1)$; the next M for which such superiority is possible is $M = 7$ where Heller's bound gives $d_\infty \leqslant 11$ but the best fixed code has $d_\infty = 10$. It is encouraging, however, that periodic codes superior to the best fixed codes could be found at all, as no such instances could be found in the prior literature.

ACKNOWLEDGMENT

The author is very grateful to Prof. James L. Massey who not only suggested this work but also devoted much time and patience in supervising it.

REFERENCES

[1] J. L. Massey, "Error bounds for tree codes, trellis codes and convolutional codes with encoding and decoding procedures," in *Coding and Complexity*, G. Longo, Ed., C.I.S.M. Courses and Lectures No. 216. New York: Springer-Verlag, 1974, pp. 1–57.

[2] A. J. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. Inform. Theory*, vol. IT-13, pp. 260–269, Apr. 1967.

[3] ——, "Convolutional codes and their performance in communication systems," *IEEE Trans. Commun. Technol.*, vol. COM-19, pp. 751–771, Oct. 1971.

[4] J. A. Heller, "Sequential decoding: Short constraint length convolutional codes," Jet Propulsion Lab., California Inst. of Technology, Pasadena, Space Program Summary 37-54, vol. 3, pp. 171–174, Dec. 1968.

[5] J. P. Oldenwalder, "Optimal decoding of convolutional codes," Ph.D. dissertation, School of Engineering and Applied Science, University of California, Los Angeles, CA, 1970.

A Generalized Method for Encoding and Decoding Run-Length-Limited Binary Sequences

G. F. M. BEENKER AND K. A. SCHOUHAMER IMMINK

Abstract—Many modulation systems used in magnetic and optical recording are based on binary run-length-limited codes. We generalize the concept of *dk*-limited sequences of length *n* introduced by Tang and Bahl by imposing constraints on the maximum number of consecutive zeros at the beginning and the end of the sequences. It is shown that the encoding and decoding procedures are similar to those of Tang and Bahl. The additional constraints allow a more efficient merging of the sequences. We demonstrate two constructions of run-length-limited codes with merging rules of increasing complexity and efficiency and compare them to Tang and Bahl's method.

I. INTRODUCTION

Many baseband modulation systems applied in magnetic and optical recording are based on binary run-length-limited codes [1], [2], [3]. A string of bits is said to be run-length-limited if the

---

Manuscript received April 5, 1982; revised November 30, 1982. This work was partially presented at the IEEE International Symposium on Information Theory, Les Arcs, France, June 21–25, 1982.

The authors are with the Philips Research Laboratories, 5600 MD Eindhoven, The Netherlands.

0018-9448/83/0900-0751$01.00 ©1983 IEEE

number of consecutive zeros between adjacent ones is bounded between a certain minimum and a certain maximum value. The upper run-length constraint guarantees a transition within a specified time interval needed for clock regeneration at the receiver. The lower run-length constraint is imposed to reduce the intersymbol interference. The latter constraint appears to have a bearing on the spectral properties of the sequences [4].

In this correspondence we consider an encoding and decoding procedure for a special class of run-length-limited codes. The constraints we consider can be defined as followed. Let n, d, k, l, and r be given, where $l \leqslant k$ and $r \leqslant k$. A binary sequence of length n is called a *dklr*-sequence if it satisfies the following constraints:

a) d-constraint: any two ones are separated by a run of at least d consecutive zeros;
b) k-constraint: any run of consecutive zeros has a maximum length k;
c) l-constraint: the number of consecutive leading zeros of the sequence is at most l;
d) r-constraint: the number of consecutive zeros at the end of the sequence is at most r.

A sequence satisfying the d- and k-constraints is called a *dk*-sequence.

Given certain desired run-length constraints, it is not trivial how to map uniquely the input data stream onto the encoded output data stream. A systematic design procedure on the basis of fixed length sequences has been given by Tang and Bahl [1]. Their method is based on mapping *dk*-sequences of length n onto consecutive integers and vice versa. In this correspondence we intend to generalize their results to *dklr*-sequences. We are also going to show that the application of *dklr*-sequences enables the sequences of length n to be merged efficiently without violation of the d- and k-constraints at the boundaries.

We assume $r \geqslant d$. Theorems 1 and 2 of this correspondence remain valid for $r < d$ and Theorem 3 can be generalized accordingly. The proofs, however, are less elegant, the reason being that for $r < d$ there are no *dklr*-sequences $(x_{n-1}, \cdots, x_0)$ having a one at positions j where $r < j \leqslant d$.

II. ENCODING AND DECODING

In this section we consider a way of mapping the set of all *dklr*-sequences onto a set of consecutive integers and vice versa. Our results are similar to those obtained by Tang and Bahl for *dk*-sequences.

Let A_n be the set of all *dklr*-sequences of length n. A_n can be embedded in a larger set $\mathcal{A}_n$ consisting of the all-zero sequence of length n and of all binary sequences of length n satisfying the d-, k-, and r-constraints where the number of consecutive leading zeros is allowed to be greater than k.

The set $\mathcal{A}_n$ can be ordered lexicographically as follows: if $x = (x_{n-1}, \cdots, x_0)$ and $y = (y_{n-1}, \cdots, y_0)$ are elements of $\mathcal{A}_n$ then y is called less than x, $y <\!\circ\, x$, if there exists an $i, 0 \leqslant i < n$, such that $y_i < x_i$ and $x_j = y_j$ for $i < j < n$. The position of x in the lexicographical ordering of $\mathcal{A}_n$ is denoted by $r(x)$; i.e., $r(x)$ is the number of all y in $\mathcal{A}_n$ with $y <\!\circ\, x$. Consequently $r(\mathbf{0}) = 0$.

For the sake of convenience we introduce the residual of a vector. Let $y = (y_{n-1}, \cdots, y_0) \in \mathcal{A}_n$, $y \neq \mathbf{0}$, and let t be such that $y_t = 1$ and $y_j = 0$ if $t < j < n$. Then the residual of y, $\operatorname{res}(y)$, is defined as follows:

$$\operatorname{res}(y) := y - \Delta_t, \qquad \text{where } \Delta_{ti} = \begin{cases} 1, & \text{if } i = t; \\ 0, & \text{elsewhere,} \end{cases}$$

$$\operatorname{res}(\mathbf{0}) := \mathbf{0}.$$

It can easily be seen that $y \in \mathcal{A}_n$ implies $\operatorname{res}(y) \in \mathcal{A}_n$. The following observation is basic to the proof of Theorem 1. Let $x, u \in \mathcal{A}_n$ and assume that $x_j = u_j = 0$ $(t < j < n)$ and $x_t = u_t = 1$ $(0 \leqslant t < n)$. Then it is not difficult to show that $r(u) - r(x) = r(\operatorname{res}(u)) - r(\operatorname{res}(x))$.

Let $N_r(i)$, $i > 0$, be the number of *dklr*-sequences with $l = k$ of length i and let $N_r(0) = 1$.

Theorem 1: Let $x = (x_{n-1}, \cdots, x_0) \in \mathcal{A}_n$. Then

$$r(x) = \sum_{j=0}^{n-1} x_j N_r(j).$$

Proof: Let the nonzero coordinates of x be indexed by $i_1 < i_2 < \cdots < i_q$, i.e., $x_i = 1$ if and only if $i \in \{i_1, \cdots, i_q\}$. Let u be the smallest element of $\mathcal{A}_n$ with the property that $u_{i_q} = 1$. Then it is not difficult to see that the second 1 of u occurs at position $i_q - k - 1$, if $i_q - k - 1 \geqslant 0$ (otherwise u has only one 1, if $r \geqslant i_q$, or also $u_0 = 1$ if $r < i_q$). Here we recall that the all-zero sequence of length n is a *dklr*-sequence if and only if $n \leqslant \min\{l, r\}$. Let $\epsilon(i_q, r) = 1$ if $r \geqslant i_q$, and $\epsilon(i_q, r) = 0$ if $r < i_q$. Then we obtain

$r(u) - r(\operatorname{res}(u))$ = the number of *dklr*-sequences with $l = k$ of length i_q with their leftmost 1 at position j where $\max\{0, i_q - k - 1\} \leqslant j < i_q + \epsilon(i_q, r) = N_r(i_q)$.

Furthermore, on the basis of the above mentioned observation it holds that

$$\begin{aligned} r(x) &= r(u) + r(x) - r(u) \\ &= r(u) + r(\operatorname{res}(x)) - r(\operatorname{res}(u)) \\ &= r(\operatorname{res}(x)) + N_r(i_q). \end{aligned}$$

The theorem then follows by induction. Q.E.D.

We have found a simple method for mapping the elements of $\mathcal{A}_n$ onto consecutive integers. For practical applications we need a mapping of the elements of A_n onto the set $\{0, 1, \cdots, |A_n| - 1\}$, where $|A_n|$ is the cardinality of A_n. Obviously the set A_n consists of the $|A_n|$ largest elements of $\mathcal{A}_n$. In addition it is clear that the number α of elements of $\mathcal{A}_n$ that are smaller than all the elements of A_n is equal to $r(a)$; i.e., $\alpha = r(a)$, where a is the smallest element of A_n. In this way we have proved the following theorem.

Theorem 2: The transformation t: $A_n \to \mathbb{N} \cup \{0\}$ defined by $t(x) = r(x) - \alpha$ for all $x \in A_n$ is a one-to-one mapping from A_n onto the set $\{0, 1, \cdots, |A_n| - 1\}$ which preserves the ordering of A_n, i.e., $x <\!\circ\, y$ if and only if $t(x) < t(y)$.

The number α can also be expressed in another way. In order to do so we define $N_r^0(j)$, $j > 0$, to be the number of *dk*-sequences of length j with their leftmost element equal to 1 and satisfying the r-constraint, $N_{r_0}(0) := 1$. It is not difficult to show that

$$\alpha = \sum_{j=1}^{n-1-l} N_r^0(j) + 1 = \sum_{j=0}^{n-1-l} N_r^0(j).$$

The numbers $N_r(j)$ and $N_r^0(j)$ are easily computed. They can be found by a straightforward computer search. A more sophisticated approach to finding these numbers can be based on the finite state transition matrix corresponding to the *dklr*-sequences [5].

The conversion from integers to *dklr*-sequences of length n is also analogous to Tang and Bahl's method and can be carried out as follows. Let $T_0, \cdots, T_{n-1}$ be integers defined by

T_i = the number of elements y in $\mathcal{A}_n$ smaller than the smallest element u in $\mathcal{A}_n$ with $u_i = 1$ and $u_j = 0$ for $j > i$,

i.e.,

$$T_i = \sum_{j=1}^{i} N_r^0(j) + 1 = \sum_{j=0}^{i} N_r^0(j).$$

TABLE I
MERGING OF *dklr*-SEQUENCES WHERE 0^j STANDS FOR j CONSECUTIVE ZEROS

s, t	Merging Bits
$s + t + d < k + 1$	0^d
$s + t + d \geqslant k + 1$	
if $s \leqslant d$	$0^{d-s}10^{s-1}$
if $s > d$	10^{d-1}

From the definition and our assumption $r \geqslant d$ it immediately follows that $T_0 < T_1 < \cdots < T_{n-1}$. These integers are used for mapping consecutive integers onto *dklr*-sequences of length n as shown in the following theorem.

Theorem 3: Let $\boldsymbol{x} = (x_{n-1}, \cdots, x_0)$ be a *dklr*-sequence of length n. Then a) $x_t = 1$ and $x_j = 0$ for $t < j < n \Leftrightarrow T_t \leqslant r(\boldsymbol{x}) < T_{t+1}$. Furthermore, if $x_t = 1$ and $x_j = 0$ for $t < j < n$, then b) $T_{t-k-1} \leqslant r(\boldsymbol{x}) - N_r(t) < T_{t-d}$.

Proof: a) This statement follows from the definition of T_i. b) Let $\boldsymbol{x}$ be a *dklr*-sequence of length n with $x_t = 1$ and $x_j = 0$ for $t < j < n$. Then $T_{t-k-1} \leqslant r(\text{res}(\boldsymbol{x})) < T_{t-d}$. Hence b) follows, since $r(\boldsymbol{x}) - r(\text{res}(\boldsymbol{x})) = N_r(t)$. Q.E.D.

The conversion from integers to *dk*-sequences can therefore also be generalized to the conversion from integers to *dklr*-sequences. The following simple encoding algorithm for *dklr*-sequences can be derived from this theorem. Given an integer I in the set $R = \{r(\boldsymbol{x}) | \boldsymbol{x} \in A_n\}$ ($I = 0$ if $\boldsymbol{x} = \boldsymbol{0}$), we first locate the largest possible t, $0 \leqslant t < n$, such that $T_t \leqslant I < T_{t+1}$ and we make $x_t = 1$. Subtracting the contribution of x_t in I, we get a new integer $I - N_r(t)$. Theorem 3 can be used again to find the next nonzero component of $\boldsymbol{x}$. The second part of Theorem 3 assures us that x_t will be followed by at least d, but no more than k zeros.

III. THE EFFICIENCY OF *dklr*-SEQUENCES

In modulation systems the *dklr*-sequences of length n cannot in general be cascaded without violating the *dk*-constraint at the boundaries. Inserting a number β of merging bits between adjacent n-sequences makes it possible to preserve the d- and k-constraints for the cascaded output sequence. The *dk*-sequences need $\beta = d + 2$ merging bits [1], whereas only $\beta = d$ merging bits are required for *dklr*-sequences, provided that the parameters l and r are suitably chosen. Hence this method is more efficient, especially for small values of n. We shall now demonstrate two constructions of codes with merging rules of increasing complexity and efficiency.

Construction 1: Choose d, k, r, and n such that $r + d \leqslant k$. Let $l = k - d - r$ and $\beta = d$. Then the *dklr*-sequences of length n can be cascaded without violating the d- and k-constraints if the merging bits are all set to zero.

Construction 2: Choose d, k, and n such that $2d - 1 < k$. Let $r = l = k - d$ and $\beta = d$. Then the *dklr*-sequences of length n can be cascaded without violating the d- and k-constraints if the merging bits are determined by the following rules. Let an n-sequence end with a run of s zeros ($s \leqslant r$) while the next n-sequence start with t ($t \leqslant l$) leading zeros. Table I shows the merging rule for the $\beta = d$ merging bits.

The number m of data bits that can be represented uniquely by a *dklr*-sequence of length n is given simply by

$$m = \lfloor \log_2 |A_n| \rfloor,$$

where $\lfloor x \rfloor$ is the greatest integer not greater than x. The ratio R of the number of data bits and the number of needed channel bits is called the information rate of the code. For example, the information rate of the codes based on the two above-mentioned constructions equals $R = m/(n + d)$. The asymptotic information rate is the capacity C of Shannon's discrete noiseless run-length-limited channel [6], [7],

$$C = \lim_{n \to \infty} \frac{\log_2 |A_n|}{n}.$$

The efficiency η can be defined as the ratio of the information rate R and the capacity C of the noiseless run-length-limited channel,

$$\eta = R/C.$$

TABLE II
BLOCK CODES BASED ON CONSTRUCTION 1

d	k	n	R	C	$\eta = R/C$
1	7	12	8/13	0.68	0.91
2	17	14	8/16	0.55	0.91
3	14	17	8/20	0.46	0.87
4	18	19	8/23	0.40	0.87

TABLE III
BLOCK CODES BASED ON CONSTRUCTION 2

d	k	n	R	C	$\eta = R/C$
1	5	12	8/13	0.65	0.95
2	10	14	8/16	0.54	0.92
3	10	17	8/20	0.45	0.90
4	12	19	8/23	0.39	0.90

TABLE IV
BLOCK CODES BASED ON TANG AND BAHL'S CONSTRUCTION

d	k	n	R	C	$\eta = R/C$
1	5	12	8/15	0.65	0.82
2	9	14	8/18	0.54	0.83
3	8	17	8/22	0.43	0.86
4	10	19	8/25	0.38	0.85

In order to get some insight into the efficiency of the codes based on Constructions 1 and 2 we have considered some examples. For $m = 8$ and for $d = 1, 2, 3, 4$ and $k = 2d, \cdots, 20$ we have determined n in such a way that the information rate R was maximized. In order to compare our two constructions to Tang and Bahl's method we have calculated the corresponding capacities C and efficiencies η. The capacity of the noiseless run-length-limited channels was calculated by a method given in [1].

Our results can be summarized as follows. For small values of k, i.e., $2d \leqslant k < 3d$, Construction 2 is only slightly better than Tang and Bahl's method (approximately 5 percent), while the efficiency of Construction 1 was worse (5 to 10 percent). For larger values of k, however, Constructions 1 and 2 are clearly better. For those values of k the gain of Construction 2 compared to Tang and Bahl's method is most significant for $d = 1, 2$ (12 to 15 percent), while for $d = 3, 4$ the gain is equal to 9 percent. For large values of k, Constructions 1 and 2 have the same efficiency; for the other values of k, Construction 2 has a better efficiency than Construction 1. Tables II, III, and IV give the results for $m = 8$ and $d = 1$, 2, 3, and 4; in order to limit the length of the tables, we have restricted k and n to those values which maximize the information rate R. We note that rates up to 95 percent of the channel capacity can be achieved. On average we observe a slight difference in the rates obtained by Constructions 1 and 2, approximately 5 percent in favor of Construction 2.

IV. CONCLUSION

Methods are described for the construction of run-length-limited codes on the basis of sequences of fixed length. Additional constraints on the maximum number of zeros at the beginning and end of a sequence, a generalization of Tang and

Bahl's work, allow a more efficient merging of the sequences. For short lengths in particular, our method yields better efficiencies than those of Tang and Bahl.

REFERENCES

[1] D. T. Tang and L. R. Bahl, "Block codes for a class of constrained noiseless channels," *Inform. Contr.*, vol. 17, pp. 436–461, 1970.

[2] H. Kobayashi, "A survey of coding schemes for transmission of recording of digital data," *IEEE Trans. Comm. Tech.*, vol. COM-19, pp. 1087–1100, 1971.

[3] K. A. Immink, "Modulation systems for digital audio disks with optical read out," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, pp. 587–589, 1981.

[4] M. G. Pelchat and J. M. Geist, "Surprising properties of two-level bandwidth compaction codes," *IEEE Trans. Comm.*, vol. COM-23, pp. 878–883, 1973.

[5] P. A. Franaszek, "Sequence-state methods for run-length-limited coding," *IBM J. Res. Dev.*, vol. 14, pp. 376–383, 1970.

[6] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, 1948.

[7] C. V. Freiman and A. D. Wyner, "Optimum block codes for noiseless input restricted channels," *Inform. Contr.*, vol. 7, pp. 398–415, 1964.

On The Source Matching Approach for Markov Sources

DONG H. LEE

Abstract—**The source matching approach is a universal noiseless source coding scheme used to approximate the solution of minimax coding. A numeric solution of the source matching approach is presented for the class of binary first-order Markov sources.**

I. INTRODUCTION

Noiseless source coding, as applied to sources whose statistics are either partially or completely unknown, is called universal noiseless source coding. Based on previous results, Davisson [1] formulated game-theoretic definitions for the problems of universal coding. According to him, the problem of minimax coding is to find a code minimizing the maximal redundancy over a given class of sources. In [2], Davisson and Leon-Garcia presented a coding scheme called "source matching" to approximate the solutions of minimax codes over an arbitrary class of stationary sources. A numerical solution for this source matching approach was obtained for the extended class of Bernoulli-like sources. This correspondence presents a numerical solution for the source matching approach for the class of binary first-order Markov sources.

II. SOURCE MATCHING APPROACH

In this section, the results of [2] are briefly reviewed for reference. Let $A = \{0, 1, \cdots, t\}$ be a source alphabet set. We consider fixed-length to variable-length binary block encoding on source message blocks $\boldsymbol{x} = (x_1, x_2, \cdots, x_N)$ on N-tuple alphabet space A^N. Let $P_N^\theta(\boldsymbol{x})$ be the probability density function (pdf) of $\boldsymbol{x}$ conditioned for each θ taking values in some index set Λ. Let W be the set of all the possible pdf's w on Λ.

Let Δ be the set of all the real-valued $(t+1)^N$-dimensional pdf's Q_N. The source matching approach finds the minimax redundancy R_s over Δ,

$$R_s = \min_{Q_N \in \Delta} \sup_{\theta \in \Delta} H(P_N^\theta : Q_N),$$

where $H(P_N^\theta : Q_N)$ is the relative entropy between P_N^θ and Q_N, represented by

$$H(P_N^\theta : Q_N) = N^{-1} \sum_{\boldsymbol{x} \in A^N} P_N^\theta(\boldsymbol{x}) \log \left(P_N^\theta(\boldsymbol{x}) / Q_N(\boldsymbol{x}) \right).$$

Henceforth, all the logarithms are of base 2. The interpretation of the source matching approach, as an approximation to the minimax solution over the set of uniquely decodable codes, is well demonstrated in [2, th. 2].

A numerical approach to the source matching solution is based on the next relationship [2, th. 3],

$$\min_{Q_N \in \Delta} \sup_{\theta \in \Lambda} H(P_N^\theta : Q_N) = \sup_{w \in \Lambda} \int_\Lambda H(P_N^\theta : Q_N^*)\, dw(\theta), \tag{1}$$

where $Q_N^*(\boldsymbol{x})$ is given by

$$Q_N^*(\boldsymbol{x}) = \int_\Lambda P_N^\theta(\boldsymbol{x})\, dw(\theta), \qquad \text{for all } \boldsymbol{x} \in A^N. \tag{2}$$

Note that the right side of (1) is merely the channel capacity between the source parameter space Λ and the source output space A^N and hence the source matching problem can be converted to the channel capacity problem. It is also known [3, p. 96] that the number of values of the source parameter θ with nonzero probabilities achieving the channel capacity is no larger than the cardinality of the source output space, which is $(t+1)^N$ for alphabet A^N. The capacity of a finite-input to finite-output channel can be obtained numerically by applying the Blahut–Arimoto algorithm [4], [5]. Furthermore, if any sufficient statistic of A^N is found, its application will reduce the computational complexity of the numerical solution since its cardinality is much smaller than $(t+1)^N$ in most cases.

III. MARKOV SOURCES

In this section we consider the source matching approach for the class of first-order Markov sources with binary alphabet $A = \{0, 1\}$. It is assumed that Markov source is stationary with stochastic matrix

$$\begin{bmatrix} \theta_{00} & \theta_{01} \\ \theta_{10} & \theta_{11} \end{bmatrix} = \begin{bmatrix} 1-\theta_0 & \theta_0 \\ \theta_1 & 1-\theta_1 \end{bmatrix},$$

where θ_{ij} for $i, j = 0, 1$ represents the transition probability from the previous state i to the present state j. The stationary pdf $\pi = (\pi_0, \pi_1)$ is uniquely expressed as

$$\pi = \left(\theta_1 (\theta_0 + \theta_1)^{-1}, \theta_0 (\theta_0 + \theta_1)^{-1} \right).$$

The domain Λ of the source parameter $\theta = (\theta_0, \theta_1)$ becomes the Cartesian product $[0, 1] \times [0, 1]$ over the closed interval $[0, 1]$.

We now present a sufficient statistic defined on A^N. Let n be the Hamming weight of 1's of the source message block $\boldsymbol{x} = (x_1, \cdots, x_N)$. With n fixed, let a_i for $i = 1, 2, \cdots, n$ be the number of runs of 1's with run length i. As an example, $N = 6$, $n = 4$, $a_1 = 2$, $a_2 = 1$, and $a_3 = a_4 = 0$ for $\boldsymbol{x} = (101011)$. The total number of all the runs of 1's is upper-bounded as

$$D_n = \sum_{i=1}^{n} a_i \leqslant N - n + 1.$$

With x_1 and x_N, the first and last digits of $\boldsymbol{x} = (x_1, \cdots, x_N)$, respectively, (n, D_n, x_1, x_N) uniquely specifies $P_N^\theta(\boldsymbol{x})$ and is an eligible sufficient statistic. Realizing that the probability of $\boldsymbol{x}$ with $(n, D_n, 0, 1)$ is identical to that with $(n, D_n, 1, 0)$, the sufficient

---

Manuscript received July 6, 1982; revised January 11, 1983. This paper is a part of a dissertation submitted to the University of Maryland, College Park, in partial fulfillment of the requirements for Ph.D. degree.

The author is with Hewlett-Packard Laboratories, Palo Alto, CA 94304.

0018-9448/83/0900-0754$01.00 ©1983 IEEE