GPO PRICE $ ______________

CFSTI PRICE(S) $ ______________

Hard copy (HC) 3.00

Microfiche (MF) .65

ff 653 July 65

N68-35912

(ACCESSION NUMBER)	(THRU)
54	1
(PAGES)	(CODE)
CR-97101	07
(NASA CR OR TMX OR AD NUMBER)	(CATEGORY)

FACILITY FORM 602

Department of

ELECTRICAL ENGINEERING

UNIVERSITY OF NOTRE DAME, NOTRE DAME, INDIANA

ANALYSIS OF DECODERS FOR CONVOLUTIONAL CODES
BY STOCHASTIC SEQUENTIAL MACHINE METHODS

by

Thomas N. Morrissey, Jr.

May 1, 1968

Technical Report No. EE-682

Department of Electrical Engineering
University of Notre Dame
Notre Dame, Indiana 46556

This work was supported by the National Aeronautics and Space Administration (NASA Grant NGR 15-004-026) in liaison with the Flight Data Systems Branch of the Goddard Space Flight Center.

ABSTRACT

Previous analyses of the error probability for feedback decoding of convolutional codes have focused almost exclusively upon the syndrome portion of the decoder, the contents of which are statistically dependent upon the infinite past due to the feedback of previous decoding estimates. This viewpoint makes exact error probability calculations intractable. In this paper, however, the entire decoder is modeled as an autonomous stochastic sequential machine and finite Markov chain theory applied in order to obtain a precise expression for $P_{FD}(u)$, the probability of error associated with the feedback decoding of the u^{th} subblock of information digits, thus circumventing the problems imposed by the dependencies on the infinite past. The analysis technique developed here applies to any syndrome feedback decoder for a systematic, rate $R = \frac{K_o}{N_o}$ convolutional code of memory order m over GF(2), used for transmission over a binary symmetric channel.

The limit of $P_{FD}(u)$ as u tends to infinity, when the limit exists, is termed P_{FD}, the steady-state probability of error of feedback decoding. Sufficient conditions on decoders are given in order for P_{FD} to exist, and two classes of minimum distance decoders exhibited which meet these sufficient conditions.

P_{FD} is calculated for a particular simple example and found to satisfy $P_{FD} < P_{DD}$, $0 < p < \frac{1}{2}$, where P_{DD} is the probability of error associated with definite (i.e., feedback free) decoding of the same code, and p is the transition probability of the binary symmetric channel.

The stochastic sequential machine approach is also used in order to calculate the probability of error associated with semi-definite decoding,

a decoding technique intermediate to feedback and definite decoding. Sufficient conditions are given for $P_{SSD}(k)$, the probability or error of a k stage semi-definite decoder, to tend to P_{FD} as k tends to infinity. Also, examples are exhibited for which there exists a particular value of k, k_α, such that $P_{SSD}(k_\alpha)$ is strictly less than both P_{FD} and P_{DD}, indicating that semi-definite decoding may be of some practical value.

ANALYSIS OF DECODERS FOR CONVOLUTIONAL CODES

BY STOCHASTIC SEQUENTIAL MACHINE METHODS*

by

Thomas N. Morrissey, Jr.

I. Introduction

A diagram of a general, binary, systematic, rate $R = \frac{1}{2}$ convolutional coding and decoding scheme at time u + m is shown in Figure 1.

The digits i(t), p(t), e(t), ξ(t), and s(t) represent the information, parity, information error, parity error, and syndrome input digits respectively at time t, t = 0, 1, 2,...., $\underline{\sigma}(t) \triangleq (\sigma_0(t), \sigma_1(t)---, \sigma_{m-1}(t))$ is the m dimensional column vector representing the state of the syndrome register at time t, and $e^*(u) = f(\underline{\sigma}(u+m), s(u+m))$ is the estimate of e(u) formed at time u+m, where the decoding function f is designed to provide a "reasonable" estimate of e(u) based upon $\underline{\sigma}$(u+m) and s(u+m).

In the sequel column vectors will be denoted by parentheses () and row vectors by brackets [].

All digits are elements of GF(2), the finite field of two elements, and all operations are assumed to be carried out in this field. The plus and minus signs in Fig. 1 represent identical operations in GF(2), and are distinguished here to emphasize the operation of the decoding process.

* This work was supported by the National Aeronautics and Space Administration (NASA Grant NGR 15-004-026) in liaison with the Flight Data Systems Branch of the Goddard Space Flight Center, and forms part of a dissertation to be submitted in partial fulfillment of the requirements for the doctor's degree in electrical engineering at the University of Notre Dame, Notre Dame, Indiana.

The noise digits are assumed to be statistically independent, each having probability p of being a "1" and probability 1-p of being a "0". This noise source model is equivalent, with respect to the transmission of the information and parity digits, to a memoryless binary symmetric channel (BSC) with crossover probability p. Given that a "0" or "1" is transmitted over the BSC, the digit is received incorrectly with probability p, the probability that a noise source in Fig. 1 will generate a "1" and complement the corresponding transmitted binary digit.

Since the information sequence appears unaltered among the transmitted sequences, the code is said to be in systematic form. The remaining transmitted sequence is called the parity sequence.

The syndrome input digits are formed by passing the received information sequence through a replica of the discrete linear filter used to form the parity sequence at the encoder, and subtracting from the result the received parity sequence. Since the system is linear with respect to the formation of the syndrome input digits, this sequence is equal to the sum of the components at the input of the syndrome register resulting from the information and noise sources. However, the component due to the information source stream is formed by passing the information sequence through two identical discrete linear filters in parallel and subtracting the output of one from the other. Hence this component is zero, and the syndrome state and input are independent of the information stream and depend only on the noise digits.

Since the code is in systematic form $e^*(u)$ is subtracted from the output of the encoder replica in the decoder in order to form $i^*(u) = i(u) + (e(u) - e^*(u))$, the estimate at time u+m of the information digit at time u. Thus the associated probability of error $Pr(i^*(u) \neq i(u))$ is equal to $Pr(e^*(u) \neq e(u))$, and is independent of the information sequence since the syndrome state and input used to form $e^*(u)$ are functions only of the channel errors.

Three modes of decoder operation, depicted by the position of the switch, are shown in Fig. 1.

The position labeled DD, in which the syndrome register is not modified, corresponds to the decoding scheme known as definite decoding [1]. For definite decoding,

$$\begin{array}{llll}
\sigma_o(u+m) = s(u) = & g_m e(u-m) + \cdots + g_1 e(u-1) + g_o e(u) & & -\xi(u) \\
\sigma_1(u+m) = s(u+1) = & g_m e(u-m+1) + \cdots + g_1 e(u) + g_o e(u+1) & & -\xi(u+1) \\
\vdots & & & \vdots \\
\sigma_{m-1}(u+m) = s(u+m-1) = & g_m e(u-1) + g_{m-1} e(u) + \cdots + g_o e(u+m-1) & & -\xi(u+m-1) \\
s(u+m) = & g_m e(u) + \cdots + g_o e(u+m) & & -\xi(u+m)
\end{array} \quad (1$$

and thus a total of 3m+2 statistically independent noise bits affect the estimate $e^*(u)$.

The definite decoding probability of error is defined to be $P_{DD}(u) \triangleq Pr(e^*(u) \neq e(u)/_{DD\ mode})$. For $u \geq m$ the probability distribution for $(e(u-m),\cdots,e(u),\cdots,e(u+m),\ \xi(u),\cdots\xi(u+m))$, and therefore $P_{DD}(u)$, is independent of u. Hence the steady-state decoding probability of error for definite decoding, P_{DD}, is simply equal to $P_{DD}(m)$.

Since in definite decoding $e(u-m),\cdots,e(u-1)$ affect the estimate $e^*(u)$ but have already been estimated by the decoder, presumably with low probability of an incorrect decision, it would seem reasonable to use $e^*(u-m),\cdots,e^*(u-1)$ in an attempt to cancel the effect of $e(u-m),\cdots,e(u-1)$ in $(\underline{\sigma}(u+m),s(u+m))$. It is useful to assume that the true values of $e(u-m),\cdots e(u-$

have been estimated, in which case the syndrome state and input are functions only of 2m+2 statistically independent noise digits, a situation which intuitively is superior to definite decoding. This is done in the decoding position labeled GD, or "genie decoding" after Wozencraft and Jacobs [2], in which use is made of a genie who corrects the estimated error bits, if necessary, and feeds them back in order to form the modified syndrome equations

$$
\begin{aligned}
\sigma_o(u+m) &= g_o e(u) && -\xi(u) \\
\sigma_1(u+m) &= g_1 e(u) + g_o e(u+1) && -\xi(u+1) \\
&\;\;\vdots && \;\;\vdots \\
\sigma_{m-1}(u+m) &= g_{m-1} e(u) + \cdots + g_o e(u+m-1) && -\xi(u+m-1) \\
s(u+m) &= g_m e(u) + \cdots + g_1 e(u+m-1) + g_o e(u+m) && -\xi(u+m).
\end{aligned}
\tag{2}
$$

Since genies are rumored to be quite scarce, the genie decoding mode is used merely as a handy conceptual tool for calculating an upper bound on system performance. The probability of error associated with genie decoding, $P_{GD}(u)$, is defined to be $Pr(e^*(u) \neq e(u)/_{GD\ mode})$. Since this probability is independent of u for $u \geq 0$, the steady-state error probability for genie decoding, P_{GD}, is simple $P_{GD}(0)$.

In the mode of operation labeled FD for feedback decoding, error estimates are fed back to modify the syndrome without the services of a genie. The resulting equations are

$$
\begin{aligned}
\sigma_o(u+m) &= g_m(e(u-m)-e^*(u-m)) + \cdots + g_1(e(u-1)-e^*(u-1)) + g_o e(u) && -\xi(u) \\
\sigma_1(u+m) &= g_m(e(u-m+1)-e^*(u-m+1)) + \cdots + g_1 e(u) + g_o e(u+1) && -\xi(u+1) \\
&\;\;\vdots && \;\;\vdots \\
\sigma_{m-1}(u+m) &= g_m(e(u-1)-e^*(u-1)) + g_{m-1} e(u) + \cdots + g_o e(u+m-1) && -\xi(u+m-1) \\
s(u+m) &= g_m e(u) + \cdots + g_o e(u+m) && -\xi(u+m).
\end{aligned}
\tag{3}
$$

As long as no decoding mistakes have been made, feedback decoding coincides with genie decoding; however, decoding errors in feedback decoding affect the syndrome as would a pattern of channel errors, and could result in further decoding mistakes even in the absence of additional channel noise [3].

The probability of error associated with feedback decoding, $P_{FD}(u)$, is the quantity $Pr(e^*(u) \neq e(u)_{/FD\ mode})$. The quantity $\lim_{u\to\infty} P_{FD}(u)$, when it exists, is denoted as P_{FD}, the steady-state probability of error of feedback decoding.

$[e^*(u)]_{FD}$, $[e^*(u)]_{DD}$, and $[e^*(u)]_{GD}$ will be used to denote the estimates made by the feedback, definite, and genie decoding operations respectively when it is necessary to distinguish which decoding method is being used.

For both definite and genie decoding the syndrome state and input are functions only of statistically independent channel error digits over a finite span, and $Pr(e^*(u) \neq e(u))$ may be easily calculated in principle by summing the probabilities of those error patterns for which $e^*(u) \neq e(u)$.

However, for feedback decoding, the estimated digits $e^*(u-m)$,---,$e^*(u-1)$ affecting $\underline{\sigma}(u+m)$ are dependent among themselves and upon $e(u-m)$,---,$e(u)$,------,$e(u+m-1)$, $\xi(u)$,---$\xi(u+m-1)$, the channel noise digits affecting the syndrome state. Alternatively, the syndrome state and input are dependent on the entire past history of the error sequences, i.e., upon $\xi(u+m)$, $\xi(u+m-1)$,------,$\xi(0)$, $e(u+m)$, $e(u+m-1)$,---,$e(0)$. Thus if attention is focused exclusively upon the syndrome equations for the purpose of calculating the probability of error, the complexity of the calculation necessary in order to calculate $P_{FD}(u)$ <u>exactly</u> grows exponentially with increasing time u.

In this paper the decoder is modeled as an autonomous stochastic sequential machine [4,5] and finite Markov chain theory applied in order to calculate the feedback decoding probability of error exactly. This approach can best be illustrated by an example.

II EXAMPLE

Consider the special case ($m=1$, $g_o=1$, $g_1=1$, $f(\sigma_o(u),s(u))=\sigma_o(u)\cdot s(u)$) illustrated in Fig. 2 of the general systematic rate $R = \frac{1}{2}$, binary convolutional coding and decoding scheme in Fig. 1.

Since with syndrome decoding the probability of error is independent of the information stream, as was stated in the previous section, the all-zero information sequence may be assumed without loss of generality. Thus for the purpose of calculating $Pr(i^*(u) \neq i(u)) = Pr(e^*(u) \neq e(u))$, the encoder portion of Fig. 2 may be ignored and the system modeled as in Fig. 3.

When the switch is in the DD position

$$
\begin{aligned}
[e^*(u)]_{DD} &= s(u+1)\cdot\sigma_o(u+1) = s(u+1)\cdot s(u) \\
&= (e(u)+e(u+1)-\xi(u+1))\cdot(e(u-1)+e(u)-\xi(u)) \\
&= e(u)+e(u)[e(u-1)-\xi(u)+e(u+1)-\xi(u+1)] \\
&\quad +e(u+1)e(u-1)-e(u+1)\xi(u)-\xi(u+1)e(u-1)+\xi(u+1)\xi(u)
\end{aligned}
$$

and thus $P_{DD} = Pr([e^*(u)]_{DD} \neq e(u))$

$$
= Pr[e(u)\{e(u-1)-\xi(u)+e(u+1)-\xi(u+1)\}+e(u+1)e(u-1)-e(u+1)\xi(u)-\xi(u+1)e(u-1) +\xi(u+1)\xi(u)=1].
$$

Since the error digits in this expression are statistically independent, each having probability p of being a "1", P_{DD} may be easily calculated to be

$$
P_{DD} = 8p^2(1-p)^3 + 4p^3(1-p)^2 + 4p^4(1-p). \tag{4}
$$

When the switch is in the GD position

$$
\begin{aligned}
[e^*(u)]_{GD} &= s(u+1)\cdot\sigma_o(u+1) \\
&= (e(u)+e(u+1)-\xi(u+1))\cdot(e(u)-\xi(u)) \\
&= e(u)+e(u)[e(u+1)-\xi(u+1)-\xi(u)]-e(u+1)\xi(u)+\xi(u+1)\cdot\xi(u)
\end{aligned}
$$

and $P_{GD} = Pr([e^*(u)]_{GD} \neq e(u))$

$$
= Pr[e(u)\{e(u+1)-\xi(u+1)-\xi(u)\}-e(u+1)\xi(u)+\xi(u+1)\cdot\xi(u) = 1].
$$

The error digits here also are statistically independent, and P_{GD} may be computed to be

$$P_{GD} = 5p^2(1-p)^2 + 2p^3(1-p) + p^4. \qquad (5)$$

However, for feedback decoding

$$[e^*(u)]_{FD} = s(u+1)\cdot\sigma_o(u+1)$$
$$= (e(u)+e(u+1)-\xi(u+1))\cdot(e(u-1)-[e^*(u-1)]_{FD}+e(u)-\xi(u))$$
$$= e(u)+[e(u-1)-[e^*(u-1)]_{FD}][e(u)+e(u+1)-\xi(u+1)]$$
$$+e(u)e(u+1))-e(u)\xi(u+1))-e(u)\xi(u)-e(u+1)\xi(u)+\xi(u)\xi(u+1),$$

and $P_{FD}(u) = Pr([e^*(u)]_{FD} \neq e(u))$

$$= Pr[\{e(u-1)-[e^*(u-1)]_{FD}\}\{e(u)+e(u+1)-\xi(u+1)\}+e(u)e(u+1)-e(u)\xi(u+1)$$
$$-e(u)\xi(u)-e(u+1)\xi(u)+\xi(u)\xi(u+1)=1].$$

But $[e^*(u-1)]_{FD}$ depends upon $e(u)$, $\xi(u)$ and the previous error digits $e(u-1)$,---,$e(0)$ and $\xi(u-1)$,---$\xi(0)$. Therefore the procedure for calculating $Pr([e^*(u)]_{FD} \neq e(u))$ by means of summing the probabilities of the decoding-error-causing error patterns as was done for definite and genie decoding is extremely unwieldy and, for all practical purposes, impossible for large u.

However this problem of infinite past history dependence is circumvented by the representation of the (feedback) decoder of Fig. 3 as a four state sequential machine with two dimensional random input vector $(e(u+1),\xi(u+1))$ and single output $e^*(u)$ at time u+1. The four states are denoted as $q_o=00$, $q_1=01$, $q_2=10$, and $q_3=11$, where the first digit represents the content of the buffer and the second the content of the syndrome register.

The decoder then may be modelled as a four-state autonomous stochastic machine in which the states are the same as the original sequential machine, and the transition probabilities between pairs of states are the sums of the probabilities of those input noise vectors that take the machine from the first state to the second. The state-behavior of the four-state autonomous stochastic machine is thus described by the following Markov transition matrix:

$$\pi = \begin{bmatrix} P_{00} & P_{01} & P_{02} & P_{03} \\ P_{10} & P_{11} & P_{12} & P_{13} \\ P_{20} & P_{21} & P_{22} & P_{23} \\ P_{30} & P_{31} & P_{32} & P_{33} \end{bmatrix} = \begin{bmatrix} (1-p)^2 & p(1-p) & p^2 & p(1-p) \\ 1-p & 0 & p & 0 \\ p(1-p) & (1-p)^2 & p(1-p) & p^2 \\ 1-p & 0 & p & 0 \end{bmatrix},$$

where P_{ij} is the conditional probability that the system will be in state q_j at time u+1 given that it is in state q_i at time u.

As an example of how the P_{ij}'s were calculated for this decoder, assume that the system is in state $q_1=01$ at time u. Now if $e(u)=0$ and $\xi(u)=0$, or $e(u)=0$ and $\xi(u)=1$, the state at time u+1 will be $q_0=00$; and if $e(u)=1$ and $\xi(u)=0$, or $e(u)=1$ and $\xi(u)=1$, the state at time u+1 will be $q_2=10$.

Therefore, $P_{11}=0$, $P_{13}=0$,

$$P_{10}=Pr[e(u)=0,\xi(u)=0]+Pr[e(u)=0,\xi(u)=1]=(1-p)^2+p(1-p)=1-p,$$

and

$$P_{12}=Pr[e(u)=1,\xi(u)=0]+Pr[e(u)=1,\xi(u)=1]=p(1-p)+p^2=p.$$

The fact that $e(u)$ and $\xi(u)$ are statistically independent of one another and of the assumed state at time u makes the calculation of the P_{ij}'s straightforward. The actual state at time u accounts for all the past history of the error sequences that is relevant to future operation of the decoder.

The probabilistic transition diagram for this decoder is shown in Fig. 4.

Let $\underline{W}(u) \triangleq [W_0(u),W_1(u),W_2(u),W_3(u)]$ be the state probability vector at time u; i.e., $W_i(u)$ is the probability that the system is in state q_i at time u.

Now $\underline{W}(u) = \underline{W}(0)\pi^u$, with $\underline{W}(0) \triangleq [1000]$ since a feedback decoder is by definition in the all-zero state at the beginning of the decoding process.

With each state q_i is associated a probability of error Pq_i equal to the sum of the probabilities of those noise inputs $(e(u+m), \xi(u+m)) = (e(u+1), \xi(u+1))$ that cause $e^*(u) \neq e(u)$, given that the system is in state q_i at time $u+m=u+1$. For the example of Fig. 3, these quantities are readily found to be

$$Pq_0 = 0$$
$$Pq_1 = 2p(1-p)$$
$$Pq_2 = 1$$
$$Pq_3 = 2p(1-p).$$

As an example of the above calculations, if the system is in state $q_3=11$ at time $u+1$, and $e(u+1)=0$ and $\xi(u+1)=0$, or $e(u+1)=1$ and $\xi(u+1)=1$, then $e^*(u)=e(u)=1$. On the other hand, if $e(u+1)=0$ and $\xi(u+1)=1$, or $e(u+1)=1$ and $\xi(u+1)=0$, then $0=e^*(u) \neq e(u) = 1$, and an incorrect decision is made. Thus $Pq_3 = Pr[e(u+1)=0,\ \xi(u+1)=1] + Pr[e(u+1)=1,\ \xi(u+1)=0] = 2p(1-p)$.

$P_{FD}(u)$ may thus be written as

$$P_{FD}(u) = \sum_{i=0}^{3} W_i(u+1)Pq_i. \tag{6}$$

A <u>sufficient</u> condition for P_{FD} to exist is that $\lim_{u\to\infty} \underline{W}(u)=\underline{W}\overset{\Delta}{=}[W_0,W_1,W_2,W_3]$, the steady-state decoder state probability vector, with W_i equal to the steady-state probability of state q_i. If this is the case then $P_{FD}\overset{\Delta}{=}\lim_{u\to\infty} P_{FD}(u) = \lim_{u\to\infty} \sum_{i=0}^{3} W_i(u+1)Pq_i = \sum_{i=0}^{3} W_iPq_i$. $\underline{W}$ does in fact exist for this example (see Theorem 1 in the next section) and may be calculated by solving $\underline{W}\pi = \underline{W}$ for $\underline{W}$ with the added constraint $\sum_{i=0}^{3} W_i = 1$. The desired solution is

$$W_0 = \frac{1-2p+3p^2-2p^3}{1+3p^2-2p^3} \qquad W_2 = \frac{3p^2-2p^3}{1+3p^2-2p^3}$$
$$W_1 = \frac{p-3p^3+2p^4}{1+3p^2-2p^3} \qquad W_3 = \frac{p-3p^2+5p^3-2p^4}{1+3p^2-2p^3} \tag{7}$$

and thus $P_{FD} = \sum_{i=0}^{3} W_i P q_i = \frac{7p^2-12p^3+10p^4-4p^5}{1+3p^2-2p^3}$. (8)

P_{GD}, P_{DD}, and P_{FD} for this example are sketched in Fig. 5 as a function of p. Note that for small p, the asymptotic values $5p^2$, $7p^2$, and $8p^2$ are approached by P_{GD}, P_{FD}, and P_{DD} respectively.

It can be shown that $P_{GD}<P_{FD}<P_{DD}$ for $0<p<\frac{1}{2}$, which illustrates the global (i.e., all $p<\frac{1}{2}$) superiority of feedback decoding over definite decoding for this example. This is the first instance known that P_{FD} has been calculated exactly and compared with P_{DD}, although intuitively one suspects that in general $P_{FD}<P_{DD}$ for sufficiently small p.

III. GENERAL STOCHASTIC MODEL APPROACH

More generally, consider the arbitrary syndrome feedback decoder shown in Fig. 6 at time u+m for an $R = \frac{K_o}{N_o}$ binary systematic convolutional code with memory order m. For full details of the general syndrome feedback decoder, see the literature [6].

The decoder is described as follows: $\underline{S}(u+m) = G'\underline{E}(u+m)$, where $\underline{S}(u+m) \triangleq (S^{(K_o+1)}(u+m), ---, S^{(N_o)}(u+m))$ is the syndrome input at time u+m. And $\underline{E}(u+m) \triangleq (\underline{e}(u):--:\underline{e}(u+m))$, where $\underline{e}(i) \triangleq (e^{(1)}(i), ---, e^{(N_o)}(i))$, is the N_o dimensional random input column vector at time i with components which are statistically independent and have probability p of being a "1", $o<p<\frac{1}{2}$. The superscripts above the error and syndrome digits refer to the particular input sequence and syndrome register respectively of these digits. The matrix G' is given by

$$G' \triangleq [G'_m : G'_{m-1} :---: G'_o]$$

where

$$G'_i \triangleq [G_i : I_{N_o-K_o}], \; i = o$$

$$G'_i \triangleq [G_i : \tilde{0}\,], \; i = 1, 2, --, m.$$

and

$$G_i \triangleq \begin{bmatrix} g_{i(1)}^{(K_o+1)} & ---- & g_{i(K_o)}^{(K_o+1)} \\ | & & | \\ g_{i(1)}^{(N_o)} & ------ & g_{i(K_o)}^{(N_o)} \end{bmatrix} \quad i = 0, 1, -- , m, \qquad (9)$$

$I_{N_o-K_o}$ is the $(N_o-K_o) \times (N_o-K_o)$ identity matrix, $\tilde{0}$ is the $(N_o-K_o) \times (N_o-K_o)$ all zero matrix, and $g_{i(j)}^{(k)}$ is the component of the response on the k^{th} output line of the encoder at time i due to an excitation (a "1") on the j^{th} input line at time 0, $j = 1, 2, ---, K_o$, $k = K_o+1, K_o+2, ---, N_o$.

Now the operation of the syndrome register is described by

$$\underline{\sigma}(u+1) = A\,\underline{\sigma}(u) + B\underline{S}(u) + G_{FB}\,\underline{e}^*_I(u-m). \qquad (10)$$

$$u = o, 1, 2, ---, \quad \underline{\sigma}(0) \triangleq \underline{0},$$

where $\underline{\sigma}(u) \triangleq (\underline{\sigma}_o(u) : \underline{\sigma}_1(u) :---: \underline{\sigma}_{m-1}(u))$,

with $\underline{\sigma}_i(u) \triangleq (\sigma_i^{(K_o+1)}(u), \sigma_i^{(K_o+2)}(u), ---, \sigma_i^{(N_o)}(u))$,

is the $m(N_o-K_o)$ dimensional column vector representing the syndrome state at time u, and

$$A \triangleq \begin{bmatrix} \tilde{0} & I_{N_o-K_o} & \tilde{0} & \cdots & \tilde{0} \\ \tilde{0} & \tilde{0} & I_{N_o-K_o} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \tilde{0} \\ \vdots & \vdots & & \ddots & I_{N_o-K_o} \\ \tilde{0} & \tilde{0} & \cdots & \cdots & \tilde{0} \end{bmatrix}, \tag{11}$$

$$B \triangleq \begin{bmatrix} \tilde{0} \\ \tilde{0} \\ \vdots \\ \tilde{0} \\ I_{N_o-K_o} \end{bmatrix}, \tag{12}$$

$$G_{FB} \triangleq \begin{bmatrix} G_1 \\ \cdot\cdot \\ G_2 \\ \vdots \\ G_m \end{bmatrix}. \tag{13}$$

$\underline{e}^*_I(u)$, the estimate at time u+m of the first K_o components (i.e., the information components) of $\underline{e}(u)$, is given by $\underline{e}^*_I(u) = \underline{f}(\underline{\sigma}(u+m) : \underline{S}(u+m))$, where $\underline{f}$ is a boolean function specified by the decoding algorithm. $\underline{\sigma}(u)$ may also be written as

$$\underline{\sigma}(u) = A^u\underline{\sigma}(0) + \sum_{i=0}^{u-1} A^i[B\underline{S}(u-1-i) + G_{FB}\underline{f}(\underline{\sigma}(u-1-i):\underline{S}(u-1-i))],$$

$$u = 1, 2, ---, \qquad \underline{\sigma}(0) \triangleq \underline{0}\ . \tag{14}$$

The estimates $\underline{e}^*_I(-m)$, $\underline{e}^*_I(-m+1)$, --, $\underline{e}^*_I(-1)$ above, although corresponding to no actual error digits since the first such digit is received at time u = 0, nevertheless must be considered in the syndrome state equations since the feedback loop of the decoder is assumed to be closed for all time; i.e., the feedback decoder is insensitive to the starting time.

Let $\underline{e}_I(u) \triangleq (e^{(1)}(u), e^{(2)}(u),---,e^{(K_o)}(u))$ denote the error pattern at time u in the information positions. Thus the vector $(\underline{e}_I(u) : \underline{e}_I(u+1) :---: \underline{e}_I(u+m-1) : \underline{\sigma}(u+m))$ represents a possible state of the decoder; and since there are 2^{mN_o} distinct vectors of this form there are 2^{mN_o} possible decoder states. However, since the registers in a feedback decoder are initially filled with zeroes, only the set of states Q_{RO} reachable from the all-zero state need be considered. Thus the system may be represented as an r state Markov chain ($r= \#Q_{RO} \leq 2^{mN_o}$) with transition probabilities from a given state determined by that state and the probability distribution of the noise input vector.

Associated with each state $q_j \epsilon Q_{RO}$ is a probability of error, P_{q_j}, which is the probability that $\underline{e}(u+m)$ is such that $\underline{e}^*_I(u) \neq \underline{e}_I(u)$, given that the system is in state q_j at time u+m.

If $W_j(u+m)$ is the probability that the decoder is in state q_j at time u+m, j = 0, 1,---, r-1, then $P_{FD}(u) \triangleq Pr(\underline{e}^*_I(u) \neq \underline{e}_I(u))$ is given by

$$P_{FD}(u) = \sum_{j=0}^{r-1} W_j(u+m) P_{q_j}, \tag{15}$$

where

$$[W_o(u+m), W_1(u+m),---,W_{r-1}(u+m)] \triangleq \underline{W}(u+m) = \underline{W}(0) \cdot \pi^{u+m}. \tag{16}$$

π is the rxr Markov transition matrix associated with the decoder, and $\underline{W}(0) \triangleq [W_0(0), W_1(0),---,W_{r-1}(0)] = [100---00]$, with q_o taken as the all-zero decoder state, is the initial probability distribution of the feedback decoder.

If steady-state probabilities exist for the Markov chain, then $\lim_{u \to \infty} \underline{W}(u) = \underline{W} \overset{\Delta}{=} [W_0, W_1, --, W_{r-1}]$, where W_j is the steady-state probability that the system is in state q_j. For this case,

$$P_{FD} = \lim_{u \to \infty} P_{FD}(u) = \sum_{j=0}^{r-1} W^{(j)} P_{q_j}, \tag{17}$$

the steady-state probability of error of feedback decoding.

Note that for large memory m and/or block length N_o this approach for calculating P_{FD} is impractical because the number of states which must be considered grows exponentially with mN_o. The Markov chains, however, are "loosely connected" since, there being only 2^{N_o} possible input vectors, each state can make a transition into at most 2^{N_o} other states.

The condition for existence of steady-state probabilities for a Markov chain is given by the following theorem [7].

<u>Theorem 1.</u> Let π be an r by r transition matrix associated with a finite Markov chain. Then steady-state probabilities exist if and only if there exists an integer $N, 1 \leq N \leq 2^{(r^2)}$, such that π^N has a positive column, that is, a column all of whose elements are > 0.

The above condition is equivalent to the existence of a state q_j and a positive integer N such that starting from any initial state, q_j can be reached (with nonzero probability) is exactly N steps. This condition when applied to feedback decoders results immediately in the following corollary.

<u>Corollary 1.1</u> For a syndrome feedback decoder for an $R = \frac{K_o}{N_o}$ systematic convolutional code with memory order m, a sufficient condition for steady-state state occupancy probabilities to exist is that there exists some positive integer N_{q_o} such that the all-zero state q_o is reachable from any initial state $q_j \in Q_{RO}$ in exactly N_{q_o} steps.

Now assume p = 0 and that the decoder is in the all-zero state q_o at time u+m. Now p = 0 => $\underline{S}(u+m) = \underline{0}$ and $\underline{E}(u+m) = \underline{0}$; $\underline{W}(u+m) = (10---0) => \underline{\sigma}(u+m) = \underline{\ }$; and if correct estimates are made by a decoder for this noiseless case (all decoders in this paper are assumed to have this property), $\underline{e}^*_I(u) =$ $= \underline{f}(\underline{\sigma}(u+m) : \underline{S}(u+m)) = \underline{f}(\underline{0}) = \underline{0}$, and thus a transition from the all-zero state into itself results from the all-zero noise input vector. Thus if N_{q_o} is the maximum number of transitions, over all $q_j \varepsilon Q_{RO}$, required in order to reach q_o, and if state q_k can be driven to q_o in $N'_{q_o} \leq N_{q_o}$ steps, then q_k can be driven to q_o in exactly N_{q_o} steps by first taking q_k to q_o in N'_{q_o} steps and then driving q_o into itself $N_{q_o} - N'_{q_o}$ times. Hence the following corollary is obtained.

Corollary 1.2 For a syndrome feedback decoder for a rate $R = \frac{K_o}{N_o}$ systematic convolutional code with memory order m, which decodes correctly in the absence of channel errors, a sufficient condition for steady-state state occupancy probabilities to exist is that the all-zero state q_o be reachable from any initial state $q_j \varepsilon Q_{RO}$.

Since the binary symmetric channel is assumed for this development, the random input vector $\underline{e}(u)$ assumes each of its 2^{N_o} possible values with nonzero probability at each time unit u. Thus if a sequence of inputs is found to drive the decoder from q_j to q_o, that sequence of inputs has nonzero probability.

By inspection, the K_o buffer registers which form the encoder replica may be driven to the all-zero state from any other buffer state by the shifting of m successive zeroes into each. And the inputs to the N_o-K_o syndrome registers may be controlled, independently of the states and inputs of the buffer registers, by proper choice of the N_o-K_o parity error sequences. The following theorem may thus be stated:

Theorem 2. For a syndrome feedback decoder for an $R = \frac{K_o}{N_o}$ systematic convolutional code with memory order m, which decodes correctly in the absence of channel errors, steady-state state occupancy probabilities exist if the nonlinear feedback shift register composed of the $N_o - K_o$ syndrome registers can be driven from any syndrome state $\underline{\sigma}_i$ reachable from $\underline{\sigma}_o = \underline{0}$ back to the $\underline{\sigma}_o$ state by means of a suitable choice of syndrome inputs.

A shift register is said to be "driven stable" [8] if any state reachable from $\underline{\sigma}_o = \underline{0}$ can be driven autonomously back into the $\underline{\sigma}_o$ state. With this definition the following corollary to Theorem 2 is immediate:

Corollary 2.1 If the syndrome register portion of a syndrome feedback decoder for an $R = \frac{K_o}{N_o}$, systematic convolutional code with memory order m, which decodes correctly in the absence of channel errors, is "driven stable", then steady-state state occupancy probabilities exist for the decoder.

The converse to this corollary is not necessarily true.

It might be noted that if a transition from any state to any other state is possible due to a malfunction in circuitry, no matter how small the probability of such a malfunction, then steady-state probabilities exist for the decoder.

The sufficient condition that the syndrome portion of the decoder be capable of being driven from any reachable state back to $\underline{\sigma}_o = \underline{0}$ is not necessary in order for $\underline{W}$ and P_{FD} to exist, as is shown by the following example.

Consider the $R = \frac{1}{2}$ systematic convolutional decoder shown in Fig. 7, with $f(\sigma_o'(u+1), S(u+1)) = \overline{S(u+1)}\cdot\sigma_o(u+1)$. With $q_o = 00$, $q_1 = 10$, $q_2 = 01$, $q_3 = 11$, $Q_{RO} = \{q_0 q_1 q_2 q_3\}$, the associated Markov matrix is

$$\pi = \begin{bmatrix} (1-p)^2 & p^2 & p(1-p) & p(1-p) \\ p(1-p) & p(1-p) & (1-p)^2 & p^2 \\ 0 & 0 & 1-p & p \\ 0 & 0 & 1-p & p \end{bmatrix}.$$

Since column 3 of π is nonzero, steady-state state occupancy probabilities exist from Theorem 1 with $N = 1$, and $\underline{W}$ and P_{FD} may be easily calculated to be $\underline{W} = [0\ 0\ 1\text{-}p\ p]$ and $P_{FD} = p^2 + (1-p)^2$. However, states q_2 and q_3, reachable from q_o, cannot be driven back to $\underline{\sigma}_o = 0$, and thus the sufficient condition given in Theorem 2 for $\underline{W}$ and thus P_{FD} to exist does not hold.

The state q_o is here a "transient state" in that it has zero steady-state probability ($W_o = 0$). In general, if steady-state probabilities exist for the decoder states and yet $q_o \triangleq 0$ is not reachable from $q_j \in Q_{RO}$, it follows that q_o has zero steady-state probability and hence $\lim_{p \to o} P_{FD} \geq \frac{1}{m}$ since at least one decoding error must be made each m time units or else the syndrome register would clear itself.

As an example of a decoder for which neither steady-state state occupancy probabilities nor P_{FD} exists, consider the $R = \frac{1}{2}$ systematic decoder with $f(\underline{\sigma}(u+2), S(u+2)) = S(u+2) \cdot \sigma_1(u+2) \cup \overline{S(u+2)} \cdot \sigma_o(u+2)$, where "$\cup$" denotes "inclusive or", shown in Fig. 8.

The state diagram of the syndrome register portion of the decoder is shown in Fig. 9, with the first digit of the 2-tuple state representation representing $\sigma_1(\ )$ and the second $\sigma_o(\ )$. S refers to the syndrome input which causes the syndrome state transition.

For this example there exists no syndrome state $\underline{\sigma}_i$ for which there exists an N such that every syndrome state reachable from $\underline{\sigma}_o = \underline{0}$ (including $\underline{\sigma}_i$) can be driven to $\underline{\sigma}_i$ in exactly N steps. (It is assumed that

$p \approx 0$, in which case $\underline{\sigma}_0$, $\underline{\sigma}_1$ and $\underline{\sigma}_2$ are all reachable from $\underline{\sigma}_0$). Clearly, therefore, there exists no decoder state $q_i \epsilon Q_{RO}$ to which every state ϵQ_{RO} can be driven in exactly N steps, and thus no power of the associated Markov matrix π has a nonzero column. Thus from Theorem 1 the conclusion is made that steady-state state occupancy probabilities do not exist for this decoder.

The existence of P_{FD} for this example must still be investigated, since the existence of $\underline{W}$ is merely a sufficient condition for the existence of P_{FD}.

The cyclic operation of the syndrome register is begun at time u*, where u* is the time instant such that $S(i) = 0$, $i = 0,1,\ldots,u^*-1$, and $S(u^*) = 1$.

Now the quantity $Pr(u^* \geq \alpha)$ is conservatively bounded as $Pr(u^* \geq \alpha) \leq (1-p)^\alpha$ $\alpha = 0,1,2,---,0<p<\frac{1}{2}$, which implies that $\lim_{\alpha \to \infty} Pr(u^* \geq \alpha) = 0$, and u* may be assumed to be finite.

For $u \geq u^* + 3$ and $\underline{\sigma}(u+2) = \underline{\sigma}_1 \stackrel{\Delta}{=} (10)$, $e^*(u) = S(u+2) = e(u) + e(u+2) + \xi(u+2)$ and $P_{FD}(u) = 2p(1-p)$. However, if $\underline{\sigma}(u+2) = \underline{\sigma}_2 \stackrel{\Delta}{=} (01)$ for $u \geq u^*+3$, then $e^*(u) = \overline{S(u+1)} = 1+e(u) + e(u+2) + \xi(u+2)$, and $P_{FD}(u) = p^2+(1-p)^2$. Therefore for $u \geq u^*+3$, $P_{FD}(u) = Pr(u^* \text{ is even})\, 2p(1-p) + Pr(u^* \text{ is odd}) [p^2+(1-p)^2]$, u even and $P_{FD}(u) = Pr(u^* \text{ is even}) [p^2+(1-p)^2] + Pr(u^* \text{ is odd}) \cdot 2p(1-p)$, u odd. Thus $\lim_{u \to \infty} P_{FD}(u) = P_{FD}$ if and only if $Pr(u^* \text{ is even}) = Pr(u^* \text{ is odd})$, which intuitively does not seem true for all p.

More specifically, for $p = 0.4$, $Pr(u^*=0) = 2p(1-p) = .48$ and $Pr(u^*=2) = 2p(1-p)^6 + 3p^3(1-p)^4 + 2p^5(1-p)^2 + p^7 = .0712$. Thus $Pr(u^* \text{ is even}) \geq Pr(u^*=0) + Pr(u^*=2) = .5512$, $Pr(u^* \text{ is even}) \neq Pr(u^* \text{ is odd})$, and P_{FD} does not exist for this example for $p = 0.4$.

However, in the limit as $p \to 0$, terms of order p^2 may be neglected and thus $\Pr(u^*=\alpha) = 2p(1-2p)^{\alpha}$. Therefore $\Pr(u^* \text{ is even}) = \sum_{i=0}^{\infty} 2p(1-2p)^{2i} = \frac{2p}{1-(1-2p)^2} = \frac{2p}{1-1+4p-4p^2} = \frac{2p}{4p} = \frac{1}{2} = \Pr(u^* \text{ is odd})$, and

$$P_{FD} = \frac{1}{2}[2p(1-p)] + \frac{1}{2}[p^2 + (1-p)^2] = \frac{1}{2}.$$

Thus in the limit as $p \to 0$, P_{FD} exists even though $\lim_{u\to\infty} \underline{W}(u)$ does not. (For $p = 0$, however, $\lim_{u\to\infty} \underline{W}(u) = \underline{W} = [100---0]$ and $P_{FD} = 0$ since the decoder never leaves the all-zero state).

IV EXISTENCE OF STEADY-STATE STATE OCCUPANCY PROBABILITIES FOR CERTAIN CLASSES OF DECODERS

Definition 1. A "Quasi-Maximum Likelihood Decoder" (QMLD) is a feedback decoder for a binary, rate $R = \frac{K_o}{N_o}$, systematic convolutional code with memory order m which operates in the following manner:

Conceptually, the QMLD first determines (one of) the most likely error pattern $\hat{\underline{E}}(u+m)$ consistent with the syndrome state and input at time u+m, assuming past decisions have been correctly made, and estimates $\underline{e}^*_I(u)$ as the first K_o components of the first subblock of $\hat{\underline{E}}(u+m)$. Ties among the most probable error patterns are resolved on the basis of the fewest number of errors in the leading positions of $\hat{\underline{E}}(u+m)$.

More precisely, let $\underline{Y}(u+m) \triangleq (\underline{\sigma}(u+m) : \underline{S}(u+m))$ be the $(N_o-K_o)(m+1)$ dimensional column vector representing the syndrome state and input at time u+m. Also define $E[\underline{Y}]$ as the set of error vectors consistent with $\underline{Y}$; i.e., $E[\underline{Y}(u+m)]$ is the set of error patterns $\underline{E}_i(u+m)$ which give $\underline{Y}(u+m) \triangleq (\underline{\sigma}(u+m) : \underline{S}(u+m))$ given that $\underline{e}^*_I(j) = \underline{e}_I(j)$, j = u-1, u-2,---, u-m. In other words $\underline{E}_i(u+m) \in E[\underline{Y}(u+m)] \Longleftrightarrow \underline{Y}(u+m) = H\underline{E}_i(u+m)$, where

$$H = \begin{bmatrix} G'_o & "0" & "0" & \cdots & "0" \\ G'_1 & G'_o & & & \vdots \\ \vdots & \ddots & \ddots & & \vdots \\ G'_{m-1} & & \ddots & \ddots & "0" \\ G'_m & G'_{m-1} & \cdots & G'_1 & G'_0 \end{bmatrix}, \tag{18}$$

the G'_i's are defined as in the previous section, "0" is the K_o X N_o all zero matrix, and $\underline{E}_i(u+m) \triangleq (\underline{e}_i(u) : \underline{e}_i(u+1) : \text{---} : \underline{e}_i(u+m))$, with $\underline{e}_i(j) \triangleq (e_i^{(1)}(j), e_i^{(2)}(j), \text{---}, e_i^{(N_o)}(j))$, is any $(m+1)N_o$ dimensional error vector.

Since H is an $(m+1)(N_o-K_o)$ X $(m+1)N_o$ matrix with rank equal to $(m+1)(N_o-K_o)$, it has a null space of dimension $K_o(m+1)$, and hence $\#E[\underline{Y}(u+m)] = 2^{(m+1)K_o}$ for every $\underline{Y}(u+m)$. Note though that $\underline{E}(u+m)$, the actual noise vector at time u+m, may not be an element of $\underline{E}[\underline{Y}(u+m)]$ if the previous m error estimates $\underline{e}^*_I(u-1),---,\underline{e}^*_I(u-m)$ have not all been made correctly.)

Now let $\hat{\underline{E}}(u+m) \triangleq (\hat{\underline{e}}_{u+m}(u) : \hat{\underline{e}}_{u+m}(u+1):---:\hat{\underline{e}}_{u+m}(u+m))$ with $\hat{\underline{e}}_{u+m}(i) \triangleq (\hat{e}^{(1)}_{u+m}(i), \hat{e}^{(2)}_{u+m}(i),---,\hat{e}^{(N_o)}_{u+m}(i))$, be the minimum weight error pattern which is an element of $E[\underline{Y}(u+m)]$. In the event more than one such minimum weight error patterns exist, $\hat{\underline{E}}(u+m)$ is determined as follows. Given $\underline{E}_A(u+m)$ and $\underline{E}_B(u+m)$, $\underline{E}_A(u+m) \neq \underline{E}_B(u+m)$, as two minimum weight error patterns consistent with $\underline{Y}(u+m)$, let j be the smallest integer for which $\underline{e}_A(u+j) \neq \underline{e}_B(u+j)$, $j\varepsilon\{0,1,---,m\}$, and q the smallest integer for which $e_A^{(q)}(u+j) \neq e_B^{(q)}(u+j)$, $q\ \varepsilon\{1,2,---,N_o\}$. Describe the case $e_A^{(q)}(u+j) = 1$, and $e_B^{(q)}(u+j) = 0$ by $\underline{E}_A(u+m) > \underline{E}_B(u+m)$, and the case $e_A^{(q)}(u+j) = 0$ and $e_B^{(q)}(u+j) = 1$ by $\underline{E}_B(u+m) > \underline{E}_A(u+m)$. Thus if $\underline{E}_i(u+m)$, $i = 1,2,---,p$ are all minimum weight error patterns consistent with $\underline{Y}(u+m)$, $\hat{\underline{E}}(u+m) \triangleq \underline{E}_\ell(u+m)$, where $\underline{E}_\ell(u+m) > \underline{E}_j(u+m)$, $j = 1,2,---,p$, $j \neq \ell$.

Therefore, given $\underline{Y}(u+m)$, $\hat{\underline{E}}(u+m)$ as defined above is determined by the decoder, and the desired estimate $\underline{e}^*_I(u)$ chosen as $\underline{e}^*_I(u) \triangleq (\hat{e}^{(1)}_{u+m}(u), \hat{e}^{(2)}_{u+m}(u),---,\hat{e}^{(K_o)}_{u+m}(u))$. This completes the description of the QMLD.

A QMLD correctly decodes in the absence of channel errors, since $\hat{\underline{E}}(u+m) = \underline{0}$ for $\underline{Y}(u+m) = \underline{0}$.

As an example of a QMLD, consider the decoder of Fig. 3. Here

$$H = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \end{bmatrix},$$

and $\underline{Y}(u+1) = \begin{bmatrix} \sigma_o(u+1) \\ s(u+1) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \hat{e}^{(1)}_{u+1}(u) \\ \hat{e}^{(2)}_{u+1}(u) \\ \hat{e}^{(1)}_{u+1}(u+1) \\ \hat{e}^{(2)}_{u+1}(u+1) \end{bmatrix}$.

For $\underline{Y}(u+1) = (00)$, $E[(00)] = \{(0000), (1110), (1101), (0011)\}$,

$\hat{\underline{E}}(u+1) = (0000)$, and $e^*(u) = \hat{e}^{(1)}_{u+1}(u) = 0$;

for $\underline{Y}(u+1) = (01)$, $E[(01)] = \{(0010), (0001), (1100), (1011)\}$,

$\hat{\underline{E}}(u+1) = (0010)$, and $e^*(u) = \hat{e}^{(1)}_{u+1}(u) = 0$;

(Note here that two consistent error patterns are of minimum weight, namely (0010) and (0001). However (0010) > (0001) by the ordering relation ">" defined above, and thus $\hat{\underline{E}}(u+1) = (0010)$.)

for $\underline{Y}(u+m) = (10)$, $E[(10)] = \{(1010), (1001), (0100), (0111)\}$,

$\hat{\underline{E}}(u+1) = (0100)$, and $e^*(u) = \hat{e}^{(1)}_{u+1}(u) = 0$; and for $\underline{Y}(u+1) = (11)$, $E[(11)] = \{(1000), (1011), (0110), (0101)\}$,

$\hat{\underline{E}}(u+1) = (1000)$, and $e^*(u) = \hat{e}^{(1)}_{u+1}(u) = 1$.

Contrasted to the QMLD, a "Maximum Likelihood Decoder" (MLD) is a feedback decoder which operates as follows. Given $\underline{Y}(u+m)$, let $e_I^{B_j}[\underline{Y}(u+m)]$ be defined as the subset of $E[\underline{Y}(u+m)]$ for which the vectors have $\underline{e}_I^{B_j}(u)$ as the first K_o components, $j = 1,2,\text{---},t \leq 2^{K_o}$. Now $Pr(e_I^{B_j}[\underline{Y}(u+m)])$, the probability of the subset $e_I^{B_j}[\underline{Y}(u+m)]$, is defined as

$$Pr(e_I^{B_j}[\underline{Y}(u+m)]) = \sum_{\underline{E}_i(u+m)\,\varepsilon\, e_I^{B_j}[\underline{Y}(u+m)]} Pr(\underline{E}_i(u+m) \qquad (19)$$

where Σ is over all and

$$\text{where } Pr(\underline{E}_i(u+m)) = p^{W(\underline{E}_i(u+m))}(1-p)^{N_o(m+1) - W(\underline{E}_i(u+m))} \qquad (20)$$

is the apriori probability of receiving $\underline{E}_i(u+m)$, $W(\underline{E}_i(u+m))$ is the (Hamming) weight of $\underline{E}_i(u+m)$, and p is the transition probability of the binary symmetric channel. Now $\underline{e}^*_I(u)$, the estimate corresponding to $\underline{Y}(u+m)$, is chosen to be an $\underline{e}_I^{B_i}(u)$ for which

$$Pr(e_I^{B_i}[\underline{Y}(u+m)]) \geq Pr(e_I^{B_j}[\underline{Y}(u+m)]), \quad j = 1, 2, ---, t$$
$$i \varepsilon \ \{1, 2, ---, t\}.$$

Thus an MLD chooses $\underline{e}_I^*(u)$ as (one of) the most likely $\underline{e}_I^{B_j}(u)$('s) corresponding to $\underline{Y}(u+m)$, while a QMLD determines (one of) the most likely error vector(s) consistent with $\underline{Y}(u+m)$, i.e., $\hat{\underline{E}}(u+m)$, and chooses $\underline{e}_I^*(u)$ as the first K_o components of $\hat{\underline{E}}(u+m)$.

Note that for the case of only one minimum weight error pattern consistent with $\underline{Y}(u+m)$, $\underline{e}_I^*(u)$ is the same for both an MLD and a QMLD for sufficiently small p.

It is also interesting to note that the decoder of Fig. 3 is an MLD as well as a QMLD for all $p < \frac{1}{2}$.

It should be pointed out that the decoding algorithm of an MLD is a true maximum likelihood decoding algorithm for a genie decoder since the algorithm is predicated upon perfect removal of $\underline{e}(u-1)$,---,$\underline{e}(u-m)$ from $\underline{\sigma}(u+m)$.

That an MLD may not always put out the true maximum likelihood estimate in the feedback decoding mode because of past decoding errors is demonstrated by the decoder of Fig. 3. For this decoder let p^* be defined as the value of the BSC transition probability such that

$$\begin{array}{lll} P_{FD} < p & , & 0<p<p^* \\ P_{FD} = p & , & p = p^* \\ P_{FD} > p & , & p^*<p<\frac{1}{2}. \end{array} \qquad (21)$$

(p^*, the real root of $1-5p+5p^2-2p^3 = 0$, satisfies $.25<p^*<.3$.) Now for $p<p^*$ it can be shown that the maximum likelihood error estimate conditioned on the syndrome state and input $\underline{Y}(u+1)$ is always put out by the MLD of Fig. 3 in steady-state. However for $p^*<p<\frac{1}{2}$, again in steady-state operation and with $f(\sigma_o(u+1), S(u+1)) = \sigma_o(u+1).S(u+1)$ as before, the most likely estimate is $e^*(u) = 0$ for any $\underline{Y}(u+1)$. However, since $f(11) = 1$ the MLD estimates $e^*(u) = 1$ whenever $\underline{Y}(u+1) = (11)$, so that this decoder does not always make the true maximum likelihood decision for this range of p.

To show this more precisely, let $\underline{Y}(u+1) = (\sigma_o(u+1), S(u+1)) = (11)$, and assume that the system is in steady-state. Now

$$\Pr(e(u) = 1/\sigma_o(u+1) = 1, \; S(u+1) = 1)$$

$$= \Pr(e(u)=1, \sigma_o(u+1) = 0/\sigma_o(u+1) = 1, S(u+1) = 1) +$$

$$+ \Pr(e(u) = 1, \sigma_o(u+1) = 1/\sigma_o(u+1) = 1, S(u+1) = 1)$$

$$= 0 + \Pr(e(u) = 1, \sigma_o(u+1) = 1/\sigma_o(u+1) = 1, S(u+1) = 1)$$

$$= \frac{\Pr(\sigma_o(u+1)=1, S(u+1)=1/e(u)=1, \sigma_o(u+1)=1) \quad \Pr(e(u)=1, \sigma_o(u+1)=1)}{\Pr(\sigma_o(u+1) = 1, S(u+1) = 1)}$$

$$= \frac{[p^2 + (1-p)^2] \, P_{q_3}}{\Pr(\sigma_o(u+1) = 1, S(u+1) = 1)} = \frac{[p^2+(1-p)^2][p-3p^2+5p^3-2p^4]}{\Pr(\sigma_o(u+1) = 1, S(u+1) = 1)[1+3p^2-2p^3]} \tag{22}$$

And similarly,

$$\Pr(e(u) = 0/\sigma_o(u+1)=1, S(u+1)=1) = \frac{2p(1-p)[p-3p^3+2p^4]}{\Pr(\sigma_o(u+1)=1, S(u+1)=1)[1\pm 3p^2-2p^3]} \tag{23}$$

It can be readily shown that

$$\begin{aligned}
&\Pr(e(u) = 1/\sigma_o(u+1)=1, S(u+1)=1) > \Pr(e(u)=0/\sigma_o(u+1)=1, S(u+1)=1), \; p < p^* \\
&\Pr(e(u)=1/\sigma_o(u+1)=1, S(u+1)=1) = \Pr(e(u)=0/\sigma_o(u+1)=1, S(u+1)=1), \; p = p^* \\
&\Pr(e(u)=1/\sigma_o(u+1)=1, S(u+1)=1) < \Pr(e(u)=0/\sigma_o(u+1)=1, S(u+1)=1), \; p > p^*.
\end{aligned} \tag{24}$$

The following definition is prompted by this discussion.

<u>Definition 2:</u> A "True Maximum Likelihood Decoder" (TMLD) for a systematic, $R = \frac{K_o}{N_o}$ convolutional code of memory order m is a decoder of the general form of Fig. 6 for which the decoding function $\underline{f}(\underline{\sigma}(u):\underline{S}(u))$ is such that P_{FD} (exists and) is a minimum.

It can be shown (by exhaustion) that for a decoder of the form of Fig. 3, $f(\sigma_o(u),S(u))$ for a TMLD is given by

$$f(\sigma_o(u),S(u)) = \sigma_o(u).S(u), \qquad 0\le p<p^*$$

$$f(\sigma_o(u),S(u)) = \sigma_o(u).S(u) \text{ or } 0, \qquad p = p^*$$

$$f(\sigma_o(u),S(u)) = 0, \qquad p^*<p\le\tfrac{1}{2}$$

in which case

$$P_{FD} = \frac{7p^2-12p^3+10p^4-4p^5}{1+3p^2-2p^3}, \qquad 0\le p<p^* \tag{24}$$

$$P_{FD} = p, \qquad p^*\le p\le\tfrac{1}{2}.$$

For definite decoding or genie decoding, the probability distribution of $\underline{\sigma}(u)$ does not depend on the decoding function $\underline{f}$, and thus the determination of the $\underline{f}$ for which $\Pr(\underline{e}_I^*(u) \neq \underline{e}_I(u))$ is a minimum is relatively straightforward.

However, for feedback decoding the probability distribution of $\underline{\sigma}(u)$ is a function of $\underline{f}$, and thus the determination of $\underline{f}$ for a TMLD is more involved. Moreover, even if the maximum likelihood estimate $\underline{e}_I^*(u)$ for each $\underline{Y}(u+m)$ is always put out in steady-state for a feedback decoder with a specified $\underline{f}$, this is not sufficient to insure that the decoder is a TMLD, as the following example shows.

For the decoder of the form of Fig. 3 but with $e^*(u) = f(\sigma_o(u+1),S(u+1))=0$, it can be shown that in steady-state the state distribution is such that $e^*(u) = 0$ is the most likely error estimate for any syndrome state and input for $\hat{p}<p\le\frac{1}{2}$. $\hat{p}$, the real root of $1-6p+8p^2-4p^3 = 0$, satisfies $\hat{p} < p^*$.

However, as stated earlier, a feedback decoder of the form of Fig. 3 with $f(\sigma_o(u),S(u)) = \sigma_o(u).S(u)$ is a TMLD for $p < p^*$, which implies that the above feedback decoder with $f(\sigma_o(u),S(u)) = 0$ for $\hat{p}<p<p^*$ is not a TMLD, even though the maximum likelihood error estimate is put out for any $\underline{Y}(u+1)$ in steady-state.

Intuitively, decoders such as the one above for which the maximum likelihood estimate is a function of p are more difficult to analyze than maximum likelihood decoders for block codes, for which the maximum likelihood decoding function is not a function of p for $0<p<\frac{1}{2}$.

Lemma 1. For a syndrome feedback decoder for an $R = \frac{K_o}{N_o}$, binary, systematic convolutional code of memory order m, $N \geq m$, where N is the smallest power of the Markov transition matrix π associated with the decoder for which there exists a non-zero column.

Proof: Assume $N<m$. Then there exists a decoder state $q_j \in Q_{RO}$ which is reachable from every state $q_i \in Q_{RO}$ in exactly $N<m$ steps. Let $\underline{E}_j^{(1)} \triangleq [e_{j,m}^{(1)}, ---, e_{j,2}^{(1)}, e_{j,1}^{(1)}]$ be the row vector denoting the contents of the first buffer register when the decoder is in state q_j. But any state $q_k \in Q_{RO}$ with $\underline{E}_k^{(1)} \triangleq [e_{k,m}^{(1)}, ---, e_{k,1}^{(1)}] = [\overline{e_{j,1}^{(1)}}, \overline{e_{j,1}^{(1)}}, ---, \overline{e_{j,1}^{(1)}}]$, where $\overline{e_{j,1}^{(1)}}$ is the binary complement of $e_{j,1}^{(1)}$, cannot be driven to q_j in fewer than m steps, contradicting the assumption $N<m$. Hence $N \geq m$ and the lemma is proved.

Theorem 3. A QMLD for a systematic, rate $R = \frac{K_o}{N_o}$, binary convolutional code has steady-state state occupancy probabilities. Moreover $N=m$, where N is the smallest integer for which π^N has a positive column and π is the Markov transition matrix associated with the decoder.

Proof: From Lemma 1, $N \geq m$. Thus from Theorem 2 of the previous section, since a QMLD correctly decodes in the absence of channel errors, it is sufficient to demonstrate that any state of the syndrome register can be driven to the $\underline{\sigma}_o = \underline{0}$ state in a number of steps $N_{q_o} \leq m$. (The buffer registers can be cleared in at most m steps simultaneously.)

For $\hat{\underline{E}}(u+m) \triangleq (\hat{\underline{e}}_{u+m}(u) : \hat{\underline{e}}_{u+m}(u+1) : --- : \hat{\underline{e}}_{u+m}(u+m))$

associated with $\underline{Y}(u+m)$, let

$$\theta\hat{\underline{E}}(u+m) \triangleq (\hat{\underline{e}}_{u+m}(u+1) : \hat{\underline{e}}_{u+m}(u+2):---:\hat{\underline{e}}_{u+m}(u+m) : \underline{0}), \tag{25}$$

where $\underline{0}$ is the N_o dimensional all-zero column vector. Also let

$$\gamma\hat{\underline{E}}(u+m) \triangleq (\hat{\underline{e}}_{u+m-1}(u) : \hat{\underline{e}}_{u+m}(u): \hat{\underline{e}}_{u+m}(u+1) :---: \hat{\underline{e}}_{u+m}(u+m-1)), \tag{26}$$

where $\hat{\underline{e}}_{u+m-1}(u)$ is the first block of $\hat{\underline{E}}(u+m-1)$, the estimated error pattern corresponding to $\underline{Y}(u+m-1)$ at time u+m-1. Note that

$$\theta\gamma\hat{\underline{E}}(u+m) \triangleq \theta(\gamma\hat{\underline{E}}(u+m)) = (\hat{\underline{e}}_{u+m}(u):---:\hat{\underline{e}}_{u+m}(u+m-1) : \underline{0}). \tag{27}$$

Starting with $\underline{\sigma}(u+m)$ at time u+m, let the syndrome input vector $\underline{S}(u+m)$ be chosen such that $\underline{Y}(u+m) = H(\theta\hat{\underline{E}}(u+m-1))$, i.e., $\underline{S}(u+m) = G'\theta\hat{\underline{E}}(u+m-1)$, and assume that $\hat{\underline{E}}(u+m) \neq \theta\hat{\underline{E}}(u+m-1)$.

From the equations of operation of the syndrome (10), it can be shown that $\gamma\hat{\underline{E}}(u+m) \varepsilon E[\underline{Y}(u+m-1)]$; i.e., $\underline{Y}(u+m-1) = H(\gamma\hat{\underline{E}}(u+m))$. Now $W(\hat{\underline{E}}(u+m-1) \leq W(\gamma\hat{\underline{E}}(u+m)$ since $\hat{\underline{E}}(u+m-1)$ was the error estimate at time u+m-1. And since $\hat{\underline{E}}(u+m-1)$ and $\gamma\hat{\underline{E}}(u+m)$ have identical first blocks (by definition of the operator γ),

$$W(\theta\hat{\underline{E}}(u+m-1)) \leq W(\theta\gamma\hat{\underline{E}}(u+m)) \leq W(\hat{\underline{E}}(u+m)) \leq W(\theta\hat{\underline{E}}(u+m-1)) \tag{28}$$

$$\Rightarrow W(\theta\hat{\underline{E}}(u+m-1)) = W(\theta\gamma\hat{\underline{E}}(u+m)) = W(\hat{\underline{E}}(u+m)). \tag{29}$$

But

$$W(\theta\gamma\hat{\underline{E}}(u+m)) = W(\hat{\underline{E}}(u+m)) \Rightarrow \theta\gamma\hat{\underline{E}}(u+m) = \hat{\underline{E}}(u+m), \tag{30}$$

and

$$W(\theta\hat{\underline{E}}(u+m-1)) = W(\theta\gamma\hat{\underline{E}}(u+m)) \Rightarrow W(\hat{\underline{E}}(u+m-1)) = W(\gamma\hat{\underline{E}}(u+m)). \tag{31}$$

Now $W(\hat{\underline{E}}(u+m-1)) = W(\gamma\hat{\underline{E}}(u+m)) \Rightarrow$ the ordering relation "$>$" is defined between the two vectors, and $\hat{\underline{E}}(u+m-1) > \gamma\hat{\underline{E}}(u+m)$ since $\hat{\underline{E}}(u+m-1)$ was the estimate at time u+m-1. (Note that $\hat{\underline{E}}(u+m-1) = \gamma\hat{\underline{E}}(u+m)$ would contradict the assumption that $\hat{\underline{E}}(u+m) \neq \theta\hat{\underline{E}}(u+m-1)$, since $\theta\gamma\hat{\underline{E}}(u+m) = \hat{\underline{E}}(u+m)$.) But $\hat{\underline{E}}(u+m-1)$

$$> \gamma\hat{\underline{E}}(u+m) \Rightarrow \theta\hat{\underline{E}}(u+m-1) > \theta\gamma\hat{\underline{E}}(u+m) = \hat{\underline{E}}(u+m). \tag{32}$$

However, this is impossible since $\hat{\underline{E}}(u+m)$ is by definition the estimated error pattern at time u+m. Thus the only possible conclusion is that

$\underline{\hat{E}}(u+m) = \theta \underline{\hat{E}}(u+m-1)$.

By continuing to choose new syndrome input vectors which are consistent with the θ shift of the previously estimated error pattern (i.e., $\underline{S}(u+m+i) = G'(\theta^{i+1}\underline{\hat{E}}(u+m-1)$, $i = 1, 2, \text{---}, m)$, a $\underline{Y}(u+2m)$ can be obtained such that $\underline{Y}(u+2m) = H(\theta^{m+1}\underline{\hat{E}}(u+m-1))$, where $\theta^{m+1}\underline{\hat{E}}(u+m-1) \overset{\Delta}{=} \overbrace{\theta \cdots \theta}^{m+1} \underline{\hat{E}}(u+m-1)$

But $\theta^{m+1}\underline{\hat{E}}(u+m-1) = \underline{0} \Rightarrow \underline{Y}(u+2m) = \underline{0} \Rightarrow \underline{\sigma}(u+2m) = \underline{0}$. (33)

Thus the syndrome state can be driven to $\underline{\sigma}_o = \underline{0}$ in $N_{q_o} \leq m$ steps and the theorem is proved.

Definition 3. For a syndrome feedback decoder for a rate $R = \frac{K_o}{N_o}$, systematic convolutional code with syndrome state and input at time u+m given by $\underline{Y}(u+m)$, the error pattern $\underline{E}_\alpha(u+m) \overset{\Delta}{=} (\underline{e}_\alpha(u) : \underline{e}_\alpha(u+1) : \text{---} : \underline{e}_\alpha(u+m))$ is said to be a "fully correctable" error pattern if

i) $\underline{e}_I^*(u) = \underline{e}_{\alpha I}(u)$, and

ii) when the subsequent syndrome inputs are chosen as

$$\underline{S}(u+m+i) = G'(\theta^i \underline{E}_\alpha(u+m)),$$

then $\underline{e}_I^*(u+i) = \underline{e}_{\alpha I}(u+i)$, $\quad i = 1, 2, \text{---}, m$.

Note that whether $\underline{E}_\alpha(u+m)$ is fully correctable or not depends on both the code and the decoding rule. Also note that $\underline{E}_\alpha(u+m)$ need not be $\underline{E}(u+m)$, the actual error pattern at time u+m. However, if $\underline{E}(u+m)$ is "fully correctable", then choosing $\underline{S}(u+m+i) = G'(\theta^i \underline{E}(u+m))$, $i = 1, 2, \text{---}, m$, is equivalent to choosing $\underline{e}(u+m+j) = \underline{0}$, $j = 1, 2, \text{---}, m$, at the decoder input.

For $\underline{Y}(u+m)$ consistent with a correctable error pattern $\underline{E}_\alpha(u+m)$, the syndrome may be driven to $\underline{\sigma}(u+2m+1) = \underline{0}$ by the choice of $\underline{S}(u+m+i) =$ $= G'(\theta^i \underline{E}_\alpha(u+m))$, $i = 1, 2, \text{---}, m+1$, since for this case $\underline{Y}(u+2m+1) =$ $H(\theta^{m+1}\underline{E}_\alpha(u+m)) = \underline{0}$. For the QMLD, every $\underline{Y}(u+m)$ is consistent with a "fully correctable" error pattern, namely $\underline{\hat{E}}(u+m)$, and for this reason the syndrome portion of the decoder is capable of being driven to $\underline{\sigma}(u+2m+1) = \underline{0}$.

With the above as motivation, the following theorem is stated.

Theorem 4. For a syndrome feedback decoder for an $R = \frac{K_o}{N_o}$ systematic convolutional code, which correctly decodes in the absence of channel errors, steady-state state occupancy probabilities exist if every syndrome state $\underline{\sigma}_i(u+m)$ reachable from $\underline{\sigma}(0) \triangleq \underline{0}$ can be driven to a state $\underline{\sigma}_j(u+m+\delta_i)$ for which there exists vectors $\underline{S}(u+m+\delta_i)$ and $\underline{E}_j(u+m+\delta_i)$ such that $\underline{E}_j(u+m+\delta_i)$ is a "fully correctable" error pattern relative to $\underline{Y}(u+m+\delta_i)$.

Proof: Given the arbitrary syndrome state $\underline{\sigma}_i(u+m)$, let $\underline{\sigma}_i(u+m)$ be driven to $\underline{\sigma}_j(u+m+\delta_i)$ by means of a suitable choice of inputs. Then let

$$\underline{S}(u+m+\delta_i+k) = H(\theta^k \underline{E}_j(u+m+\delta_i)),\ k = 1, 2, ---, m+1, \tag{35}$$

be chosen as syndrome inputs in order to obtain $\underline{\sigma}(u+2m+1+\delta_i) = \underline{0}$.

Corollary 4.1 A syndrome feedback decoder for a systematic, $R = \frac{K_o}{N_o}$ convolutional code which correctly decodes any pattern of B or fewer errors over a $(m+1)N_o$ bit constraint length, and which puts out $\underline{e}_I^*(u) = \underline{0}$ as an estimate whenever $\underline{Y}(u+m)$ is consistent with no $\underline{E}_i(u+m)$ with $W(\underline{E}_i(u+m)) \leq B$, has steady-state state occupancy probabilities with $m \leq N \leq 2m+1$.

Proof: From Lemma 1, $m \leq N$. Now given $\underline{\sigma}(u+m)$, let $\underline{S}(u+m+j) = \underline{0}$, $j = 0, 1, ---, i$, be chosen as syndrome inputs until a $\underline{Y}(u+m+i)$, $i \in \{0, 1, ---, m\}$, is obtained for which there exists an $\underline{E}_\alpha(u+m+i)$ such that $\underline{E}_\alpha(u+m+i) \in E[\underline{Y}(u+m+i)]$ and $W(\underline{E}_\alpha(u+m+i)) \leq B$, or until m+1 such zero syndrome inputs have been fed in. If such a $\underline{Y}(u+m+i)$ exists, then from Theorem 4 the syndrome registers may be driven to $\underline{\sigma}(u+2m+1+i) = \underline{0}$ since $\underline{E}_{\alpha}(u+m+i)$ is a "fully correctable" error pattern. However, if no such $\underline{Y}(u+m+i)$ exists, then $\underline{\sigma}(u+2m) = \underline{0}$ since the syndrome registers will have been driven autonomously for m consecutive time units while feeding back the all-zero vector as an estimate.

For $B = \left\lfloor \frac{d_{FD}\min-1}{2} \right\rfloor$, where $d_{FD}\min$ is the feedback decoding minimum distance and $\lfloor\ \rfloor$ denotes the "greatest integer less than or equal to", the above corollary gives a class of minimum distance feedback decoders for which steady-state state occupancy probabilities exist.

V. APPLICATION TO SEMI-DEFINITE DECODING

The concept of semi-definite decoding (SDD) was introduced by Massey and investigated experimentally by Frasco [9] as a decoding technique intermediate to feedback and definite decoding.

A diagram of a general semi-definite decoder of order k for a binary $R = \frac{1}{2}$ systematic convolutional code with memory order m is shown at time u+m in Fig. 10. The first intermediate decision relevant to the estimation of e(u), $e^{*1}(u-k+1)$, is made on the basis of the unmodified syndrome bits S(u-k+m+1),---,S(u-k+1) and is thus equivalent to a definite decoding decision. $e^{*1}(u-k+1)$ is then used ("fed back") in order to modify the syndrome bits S(u-k+m+2),---,S(u-k+2) and the second intermediate decision $e^{*2}(u-k+2)$ formed on the basis of these modified syndrome digits by means of the same decoding function f. In general, the i^{th} intermediate decision $e^{*i}(u-k+i)$, $i \in \{1,2,---,k\}$, is made on the basis of S(u-k+m+i), ---, S(u-k+i) (with the same f) after appropriate modification by $e^{*i-1}(u-k+i-1)$,---,$e^{*i-m}(u-k+i-m)$ if $i > m$, or by $e^{*i-1}(u-k+i-1)$,---$e^{*1}(u-k+1)$ if $i \leq m$. The desired estimate of e(u) is then chosen as $[e^*(u)]_{SDD} \triangleq e^{*k}(u)$. Hence $[e^*(u)]_{SDD}$ is the k^{th} intermediate estimate formed at time u+m, and is seen to be a function of only a finite number of syndrome bits, namely S(u+m), S(u+m-1), ---,S(u-k+1). At the next instant of time, u+m+1, k new intermediate estimates $e^{*1}(u+1-k+1)$, $e^{*2}(u+1-k+2)$,---,$e^{*k}(u+1)$ are formed and the estimate $[e^*(u+1)]_{SDD} \triangleq e^{*k}(u+1)$ of e(u+1) thus made from S(u+m+1),--- --,S(u-k+2) at time u+m+1 in exactly the same manner as was $[e^*(u)]_{SDD}$

from S(u+m),---,S(u-k+1) at time u+m.

More rigorously, in a semi-definite decoder of order k for a binary $R = \frac{K_o}{N_o}$ systematic convolutional code with memory order m, the estimate $[\underline{e}^*_I(u)]_{SDD}$ of $\underline{e}_I(u)$ is made at time u+m on the basis of $\underline{S}(u+m)$, $\underline{S}(u+m-1)$,---,$\underline{S}(u-k+1)$ in the following manner. Given the decoding function $\underline{f}$, the first intermediate decision $\underline{e}^{*1}_I(u-k+1)$ is made as

$$\underline{e}^{*1}_I(u-k+1) = \underline{f}\,(\underline{\sigma}_1(u+m-k+1) : \underline{S}(u+m-k+1)), \tag{36}$$

where

$$\underline{\sigma}_1(u+m-k+1) \triangleq (\underline{S}(u-k+1) : \underline{S}(u-k+2) :---: \underline{S}(u-k+m)). \tag{37}$$

The next k-1 intermediate decisions are then made as

$$\underline{e}^{*i}_I(u-k+i) = \underline{f}(\underline{\sigma}_i(u+m-k+i) : \underline{S}(u+m-k+i)), \tag{38}$$

where

$$\underline{\sigma}_i(u+m-k+i) = A\underline{\sigma}_{i-1}(u+m-k+i-1) + B\underline{S}(u+m-k+i-1) + G_{FB}\underline{e}^{*i-1}_I(u-k+i-1), \tag{39}$$

i = 2, 3,---,k, and where A, B, and G_{FB} are defined in equations (11), (12), and (13) in section III. The desired estimate $[\underline{e}^*_I(u)]_{SDD}$ is then chosen as

$$[\underline{e}^*_I(u)]_{SDD} \triangleq \underline{e}^{*k}_I(u). \tag{40}$$

This decoding method is similar to feedback decoding in that it utilizes some previous decoding decisions in arriving at an estimate. However, infinite error propagation is avoided as in definite decoding since each estimate is a function of only a finite number of syndrome bits, and hence a finite number of channel noise bits.

Let

$$P_{SDD}(u,k) \triangleq Pr([\underline{e}^*_I(u)]_{k\text{ stage SDD}} \neq \underline{e}_I(u)) = Pr(\underline{e}^{*k}_I(u) \neq \underline{e}_I(u)) \tag{41}$$

and

$$P_{SDD}(k) \triangleq P_{SDD}(u,k)/u \geq k+m-1. \tag{42}$$

For $u \geq k+m-1$, the probability distribution of $(\underline{e}(u-m-k+1) :---: \underline{e}(u+m)::\underline{S}(u-k+1):---:\underline{S}(u+m))$, and thus $P_{SDD}(u,k)$, is independent of u and the second definition is therefore meaningful as the steady-state error probability for a k-stage semi-definite decoder; i.e.

$$P_{SDD}(k) = \lim_{u\to\infty} P_{SDD}(u,k). \tag{43}$$

Note that although this latter limit always exists, $\lim_{k\to\infty} \lim_{u\to\infty} P_{SDD}(u,k)$ may not exist.

For k = 1, the semi-definite decoder reduces to a definite decoder since

$$[\underline{e}_I^*(u)]_{1\text{stage SDD}} = \underline{e}_I^{*1} = \underline{f}(\underline{S}(u) :---: \underline{S}(u+m)) = [\underline{e}_I^*(u)]_{DD}. \tag{44}$$

Writing the equations of semi-definite decoding operation in recursion relation form yields

$$\begin{aligned}\underline{\sigma}_i(u+m-k+i) &= A^{i-1}\,\underline{\sigma}_1(u+m-k+1)\\ &+ \sum_{j=0}^{i-2} A^j[B\underline{S}(u+m-k+i-1-j) + G_{FB}\underline{f}(\underline{\sigma}_{i-1-j}(u+m-k+i-1-j) : \underline{S}(u+m-k+i-1-j))]\\ i &= 2,3,---,k.\end{aligned} \tag{45}$$

These equations may also be written as

$$\begin{aligned}\underline{\sigma}_i(u+m-k+i) &= A^{i-\alpha}\underline{\sigma}_\alpha(u+m-k+\alpha)\\ &+ \sum_{j=0}^{i-1-\alpha} A^j[B\underline{S}(u+m-k+i-1-j) + G_{FB}\underline{f}(\underline{\sigma}_{i-1-j}(u+m-k+i-1-j):\underline{S}(u+m-k+i-1-j))]\\ i &= \alpha+1,\ \alpha + 2,---,k,\ \alpha \in(1,2,---,k-1).\end{aligned} \tag{46}$$

(For $\alpha = 1$, equations (46) reduce to equations (45).)

For the purpose of analyzing $P_{SDD}(u,k)$ it is appropriate to introduce the following definition.

Definition 4: For a semi-definite decoder for an $R = \frac{K_o}{N_o}$ systematic convolutional code with memory order m, the "equivalent feedback decoder" (EFD) is a syndrome feedback decoder for the same code with the same decoding function $\underline{f}$ as the semi-definite decoder.

With this definition, comparing the set of recursion equations (46) for a semi-definite decoder with a similar set for its corresponding "equivalent feedback decoder", namely

$$\underline{\sigma}(i) = A^{i-\alpha}\underline{\sigma}(\alpha) + \sum_{j=0}^{i-1-\alpha} A^{j}[B\underline{S}(i-1-j) + G_{FB}\underline{f}(\underline{\sigma}(i-1-j) : \underline{S}(i-1-j))]$$

$$i = \alpha+1, \quad \alpha+2, ---, \qquad (47)$$

it can be seen that if

$$\underline{\sigma}(u+m-k+\alpha) = \underline{\sigma}_{\alpha}(u+m-k+\alpha)$$

(where $\underline{\sigma}(i)$ refers to a state of the EFD and $\underline{\sigma}_i(u+m-k+i)$ an intermediate "state" of the semi-definite decoder), and if the same syndrome input vectors $\underline{S}(u+m-k+\alpha)$, $\underline{S}(u+m-k+\alpha+1)$,---,$\underline{S}(u+m)$ are input to both the semi-definite decoder and its "equivalent feedback decoder", then

$$\underline{\sigma}(u+m-k+\alpha+i) = \underline{\sigma}_{\alpha+i}(u+m-k+\alpha+i) \qquad (48)$$

and

$$[\underline{e}_I^*(u-k+\alpha+i)]_{EFD} = \underline{e}_I^{*\,\alpha+i}(u-k+\alpha+i) \qquad (49)$$

$$i = 0, 1, ---, k-\alpha,$$

which implies that $[\underline{e}_I^*(u)]_{EFD} = \underline{e}_I^{*k}(u) \triangleq [\underline{e}_I^*(u)]_{SDD}$. (50)

If $\alpha = k-u-m$ with $k>u+m$ for a semi-definite decoder, then

$$\underline{\sigma}_{\alpha}(u+m-k+\alpha) = \underline{\sigma}_{\alpha}(0) = A^{\alpha-1}\underline{\sigma}_1(u+m-k+1)$$
$$+ \sum_{j=0}^{\alpha-2} A^{j}[B\underline{S}(u+m-k+i-1-j) + G_{FB}\underline{f}(\underline{\sigma}_{i-1-j}(u+m-k+i-1-j):\underline{S}(u+m-k+i-1-j))] \qquad (51)$$

where

$$\underline{\sigma}_1(u+m-k+1) \triangleq (\underline{S}(u-k+1):---:\underline{S}(u-k+m)). \qquad (52)$$

Obviously $\underline{\sigma}_{\alpha}(0) = \underline{0}$, since $\underline{f}(\underline{0})=\underline{0}$ and $\underline{S}(B)=\underline{0}$ for $B<0$. Thus from the discussion in the previous paragraph, if the same syndrome vectors $\underline{S}(0)$, $\underline{S}(1)$,---,$\underline{S}(u+m)$ are input both to the semi-definite decoder and its "equivalent feedback decoder", and with $\underline{\sigma}(0) = \underline{0}$ for the EFD, $[\underline{e}_I^*(u)]_{EFD} = [\underline{e}_I^*(u)]_{SDD}$ for k>u+m, and for all such u the probability of error associated with a semi-definite decoder is equal to that of its "equivalent feedback decoder" under normal feedback decoding conditons (i.e., $\underline{\sigma}(0) = \underline{0}$). This may be stated mathematically as

$$\lim_{k\to\infty} P_{SDD}(u,k) = P_{FD}(u). \tag{53}$$

Note that $\lim_{u\to\infty} \lim_{k\to\infty} P_{SDD}(u,k) = \lim_{u\to\infty} P_{FD}(u)$ exists if and only if P_{FD} exists for the EFD.

Thus semi-definite decoding is seen to provide a spectrum of decoding techniques intermediate to definite deocding (k=1) and feedback decoding ($k\to\infty$ and $u<\infty$).

It is of more practical interest, however, to calculate the quantity $P_{SDD}(k) \triangleq P_{SDD}(u,k)/\underline{u>}k+m-1 = \lim_{u\to\infty} P_{SDD}(u,k)$ which is determined from the probability distribution of the vector $(\underline{e}(u-m-k+1):--- :\underline{e}(u+m):\underline{S}(u-k+1):---:\underline{S}(u+m))$, $\underline{u>}k+m-1$. The following lemma gives an expression for $P_{SDD}(k)$.

Lemma 2. $P_{SDD}(k)$, the "steady-state" probability of error associated with a semi-definite decoder of order k for a rate $R = \frac{K_o}{N_o}$, binary, systematic convolutional code of memory order m, is given by $P_{SDD}(k) = \Pr([\underline{e}_I^*(u)]_{EFD} \neq \underline{e}_I^*(u))$, $u \geq k+m-1$, where $[\underline{e}_I^*(u)]_{EFD}$ is the estimate formed by an "equivalent feedback decoder" at time u+m whose syndrome state is $\underline{\sigma}(u-k+m+1) \triangleq (\underline{S}(u-k+1):---:\underline{S}(u-k+m))$ at time u-k+m+1, and which is fed the syndrome inputs $\underline{S}(u-k+m+1), \underline{S}(u-k+m+2), ---, \underline{S}(u+m)$ at times u-k+m+1,--- --,u+m respectively.

Proof: From the previous discussion in which the operation of a semi-definite decoder was compared to that of its "equivalent feedback decoder" (Eqs. (48), (49), and (50), it is clear that $[\underline{e}_I^*(u)]_{EFD} = [\underline{e}_I^*(u)]_{SDD}$, and thus

$$P_{SDD}(k) \triangleq Pr([\underline{e}_I^*(u)]_{k \text{ stage SDD}} \neq \underline{e}_I(u)) = Pr([\underline{e}_I^*(u)]_{EFD} \neq \underline{e}_I(u)).$$

Equivalently, $P_{SDD}(k) = Pr([\underline{e}_I^*(u)]_{EFD} \neq \underline{e}_I^*(u))$, where $[\underline{e}_I^*(u)]_{EFD}$ is the estimate at time u+m of an "equivalent feedback decoder" which is in the decoder state given by $(\underline{e}_I(u-k+1):---:\underline{e}_I(u-k+m):\underline{S}(u-k+1):---:\underline{S}(u-k+m))$ at time u-k+m+1, and which is fed the error input vectors $\underline{e}(u-k+m+1)$,-----,$\underline{e}(u+m)$ at times u-k+m+1,---,u+m respectively.

It should be noted that each of the 2^{mN_o} possible values of $(\underline{e}_I(u-k+1):---:\underline{e}_I(u-k+m):\underline{S}(u-k+1):---:\underline{S}(u-k+m))$ occurs with non-zero probability, since each syndrome digit which is a part of the initial state is an additive function of a unique parity error digit. Thus all possible "initial" states must be considered for the "equivalent feedback decoder" in the calculation, even though all may not be reachable in normal feedback decoding operation (i.e., $\underline{\alpha}(0) = \underline{0}$).

The following formulation is helpful in calculating $P_{SDD}(k)$: Let Q be the set of the 2^{mN_o} possible states associated with the "equivalent feedback decoder" of a semi-definite decoder for a binary, $R = \frac{K_o}{N_o}$ systematic convolutional code with memory order m, and let Q be written as the union of its disconnected submachines. That is, $Q = Q_o \cup Q_1 \cup --- \cup Q_{n-1}$, where $Q_j = \{q_{j1}, q_{j2}, ---, q_{jr_j}\}$ and the q_{ij}'s represent the states of the decoder, $i = 1,2,---,r_j$, $j = 0, 1,---,n-1$. That Q_i and Q_j are disconnected, $i \neq j$, implies that q_{ik} cannot be driven into q_{jf}, $k = 1, 2, ---, r_i$, $f = 1, 2, ---, r_j$, $i,j = 0, 1, ---, n-1$, $i \neq j$.

Denote $\pi_o, \pi_1, ---, \pi_{n-1}$, as the Markov transition matrices associated with $Q_o, Q_1, ---, Q_{n-1}$ respectively. Also denote

$\underline{W}^{Q_j}(u) \triangleq [W_1^{Q_j}(u), W_2^{Q_j}(u), ---, W_{r_j}^{Q_j}(u)]$ as the state probability vector at time u, conditioned on the event that the "equivalent feedback decoder" began operation at time zero in a state $q_{jf} \in Q_j$, and

$P_o(Q_j)$ as the probability of this conditioning event, $f = 1, 2, ---, r_j$, $j = 0, 1, ---, n-1$.

With this formulation, $P_{SDD}(k)$ may be written as

$$P_{SDD}(k) = \sum_{j=0}^{n-1} P_o(Q_j) \sum_{i-1}^{r_j} W_i^{Q_j}(k-1)\, P_{q_{ji}}, \tag{54}$$

where

$$\underline{W}^{Q_j}(k-1) = \underline{W}^{Q_j}(0)\, \pi_j^{k-1}, \tag{55}$$

$P_{q_{ji}}$ is the probability of error associated with state q_{ji},

$$P_o(Q_j) = Pr((\underline{e}_I(u-k+1):---:\underline{e}_I(u-k+m):\underline{S}(u-k+1):---:\underline{S}(u-k+m)) \in Q_j), \tag{56}$$

and

$$W_i^{Q_j}(0)=Pr((\underline{e}_I(u-k+1):---:\underline{S}(u-k+m))=q_{ji}/(\underline{e}_I(u-k+1):---:\underline{S}(u-k+m)) \in Q_j) \tag{57}$$

$u \geq k+m-1$, $i = 1, 2, ---, r_j$, $j = 0, 1, ---, n-1$.

Note that if steady-state probabilities exist for $\pi_o, \pi_1, ---, \pi_{n-1}$, then $\lim_{k \to \infty} P_{SDD}(k) = \lim_{k \to \infty} \lim_{u \to \infty} P_{SDD}(u,k)$ exists and will be denoted $P_{SDD}(\infty)$; i.e., one can speak of the error probability of "long" semi-definite decoders.

As an example of a calculation of $P_{SDD}(k)$ consider the semi-definite decoding scheme for the $R = \frac{1}{2}$ systematic binary convolutional code with $m = 1$, $g_{0(1)}^{(2)} = 1$, $g_{1(1)}^{(2)} = 1$, for which the decoding function is the "exclusive or" gate; i.e., $f(\sigma_o(u), S(u)) = \sigma_o(u) \oplus S(u)$. The "equivalent feedback decoder" is shown in Fig. 11, where the switch in the feedback loop is closed at time u-k+2 after S(u-k+1) enters the syndrome register, and for which $P_{SDD}(k) = P_{FD}(u)$.

For this example, $Q=Q_o \cup Q_1$, $Q_0=\{q_{01},q_{02}\}$, $Q_1=\{q_{11},q_{12}\}$, $n=2, r_1=2, r_2=2$, $q_{01}=00, q_{02}=10, q_{11}=01, q_{12}=11$, where the first digit represents the content of the buffer register and the second the content of the syndrome register. The error probabilities associated with the states are readily calculated to be

$$
\begin{aligned}
Pq_{01} &= 2p(1-p) \\
Pq_{02} &= 2p(1-p) \\
Pq_{11} &= p^2+(1-p)^2 \\
Pq_{12} &= p^2+(1-p)^2,
\end{aligned}
\tag{58}
$$

$$P_0(Q_0) = Pr(S(u-k+1) = 0) = (1-p)^3 + 3p^2(1-p), \tag{59}$$

$$P_0(Q_1) = Pr(S(u-k+1) = 1) = 3p(1-p)^2 + p^3, \tag{60}$$

$$\pi_0 = \begin{bmatrix} 1-p & p \\ 1-p & p \end{bmatrix}, \qquad \pi_1 = \begin{bmatrix} 1-p & p \\ 1-p & p \end{bmatrix}, \tag{61}$$

$$\underline{W}^{Q_0}(k-1) = \underline{W}^{Q_0}(0)\ \pi_0^{k-1} = [1-p,p] \begin{bmatrix} 1-p & p \\ 1-p & p \end{bmatrix}^{k-1} = [1-p,p], \tag{62}$$

$$\underline{W}^{Q_1}(k-1) = \underline{W}^{Q_1}(0)\ \pi_1^{k-1} = [1-p,p] \begin{bmatrix} 1-p & p \\ 1-p & p \end{bmatrix}^{k-1} = [1-p,p], \tag{63}$$

$$
\begin{aligned}
P_{SDD}(k) &= [(1-p)^3+3p^2(1-p)][(1-p)2p(1-p)+p\ 2p(1-p)] \\
&\quad + [3p(1-p)^2 + p^3][(1-p)\{p^2 + (1-p)^2\} + p\{p^2+(1-p)^2\}] \\
&= 5p - 20\ p^2 + 40\ p^3 - 40\ p^4 + 16\ p^5.
\end{aligned}
\tag{64}
$$

Note that for this example $P_{SDD}(k)$ is not a function of k, and hence $P_{SDD}(k) = P_{SDD}(1) = P_{DD}$, where P_{DD} is the probability of error of definite decoding.

On the other hand, when $P_{FD}(u)$ is calculated for this coding scheme for normal feedback decoding operation (i.e., the switch is closed for all time and $\underline{\sigma}(0) = \underline{0}$), only π_0 is considered since the decoder is initially in the all-zero state. (For this example, Q_0, the set of states connected with the all-zero state, is equal to Q_{RO}, the set of states reachable from the all-zero state, although this is not true in general.)

The feedback decoding probability of error is then

$$P_{FD}(u) = (1-p)2p(1-p) + p2p(1-p) = 2p(1-p). \qquad (65)$$

Here also no transient terms are present, and for all $u, P_{FD}(u) = P_{FD}$, the steady-state probability of error for feedback decoding.

Thus for this example, both $\lim_{k\to\infty} \lim_{u\to\infty} P_{SDD}(u,k)$ and $\lim_{u\to\infty} \lim_{k\to\infty} P_{SDD}(u,k)$ exist, and

$$P_{DD} = \lim_{k\to\infty} \lim_{u\to\infty} P_{SDD}(u,k) > \lim_{u\to\infty} \lim_{k\to\infty} P_{SDD}(u,k) = P_{FD}, <0<p\leq\frac{1}{2}. \qquad (66)$$

The question arises as to the conditions under which $\lim_{k\to\infty} P_{SDD}(k) \triangleq$ $= \lim_{k\to\infty} \lim_{u\to\infty} P_{SDD}(u,k) = \lim_{u\to\infty} \lim_{k\to\infty} P_{SDD}(u,k) \triangleq \lim_{u\to\infty} P_{FD}(u) = P_{FD}$

for a semi-definite decoder, where $P_{FD}(u)$ refers to the probability of error of its "equivalent feedback decoder" under normal feedback decoding conditions (i.e., $\underline{\alpha}(0) = \underline{0}$). Sufficient conditions for this occurance are given by the following theorem.

Theorem 5. For a semi-definite decoder and its "equivalent feedback decoder" for an $R = \frac{K_o}{N_o}$ binary systematic convolutional code with memory order m, if steady-state probabilities exist for π, the Markov matrix associated with Q, then

$$\lim_{k\to\infty} P_{SDD}(k) \triangleq \lim_{k\to\infty} \lim_{u\to\infty} P_{SDD}(u,k) = \lim_{u\to\infty} \lim_{k\to\infty} P_{SDD}(u,k) \triangleq \lim_{u\to\infty} P_{FD}(u) = P_{FD} \qquad (67)$$

Proof: Since steady-state probabilities exist for π, $Q=Q_o$ and $P_{SDD}(k) = \sum_{i=1}^{2^{mN_o}} W_i^{Q_o}(k-1)\, P_{q_{oi}}$ with

$$\underline{W}^{Q_o}(k-1) \triangleq [W_1^{Q_o}(k-1), W_2^{Q_o}(k-1), ---, W_{2^{mN_o}}^{Q_o}(k-1)]. \text{ Now}$$

$$\lim_{k\to\infty} \underline{W}^{Q_o}(k-1) = \lim \underline{W}^{Q_o}(0)\, \pi_o^{k-1} = \underline{W}^{Q_o} \triangleq [W_1^{Q_o}, \ldots, W_{2^{mN_o}}^{Q_o}] \qquad (68)$$

independently of the choice of $\underline{W}^{Q_o}(0)$, since steady-state probabilities exist for $\pi_o = \pi$, and thus $\lim_{k\to\infty} P_{SDD}(k) = \sum_{i=1}^{2^{mN_o}} W_i^{Q_o} P_{q_{oi}}$. But since $\lim_{k\to\infty} \underline{W}^{Q_o}(0)\, \pi_o^{k-1} = \underline{W}^{Q_o}$ independently of the choice of the initial probability distribution $\underline{W}^{Q_o}(0)$, if any state q_{oi} has a non-zero steady-state probability (i.e., $W_i^{Q_o} \neq 0$), that state is reachable from every state $q_{oj} \in Q_o$. In particular, every state with non-zero steady-state probability is reachable from the all-zero state; or, conversely,

$$q_{oi} \notin Q_{RO} \Rightarrow W_i^{Q_o} = 0. \qquad (69)$$

Thus

$$\lim_{k\to\infty} P_{SDD}(k) = \sum_{i=1}^{2^{mN_o}} W_i^{Q_o} Pq_{oi} = \sum_{i=1}^{r} W_i^{Q_{RO}} P_{q_{r_{o}i}} \triangleq P_{FD} \triangleq \lim_{u\to\infty} P_{FD}(u), \qquad (70)$$

and the theorem is proved.

A code which meets the conditions of Theorem 5 is that of the example in Section II, with the decoding function f given by $f(\sigma_o(u),S(u)) = \sigma_o(u)\cdot S(u)$ for both the feedback and semi-definite decoding schemes. The example in this section with $f(\sigma_o(u),S(u)) = \sigma_o(u) \oplus S(u)$ does not meet the condition of the theorem since $Q \neq Q_\sigma$, and for this case $\lim_{k\to\infty} P_{SDD}(k) \neq P_{FD}$.

An important point which might be raised is that of the shape of the $P_{SDD}(k)$ vs. k curve. That is, is there an intermediate value of k, say k_α, for which $P_{SDD}(k_\alpha) < P_{FD}$ and $P_{SDD}(k_\alpha) < P_{DD}$; i.e., is semi-definite decoding ever strictly superior to definite decoding and feedback decoding?

An example which illustrates this eventuality is that of section II, with $f(\sigma_o(u),S(u)) = \sigma_o(u)\cdot S(u)$ and $p = 0.4$.

For this case

$$\pi_o = \pi = \begin{bmatrix} .36 & .24 & .16 & .24 \\ .6 & 0 & .4 & 0 \\ .24 & .36 & .24 & .16 \\ .6 & 0 & .4 & 0 \end{bmatrix}, \qquad (71)$$

and

$$\pi^{k-1} \doteq \begin{bmatrix} .408 & .192 & .26 & .14 \\ .408 & .192 & .26 & .14 \\ .408 & .192 & .26 & .14 \\ .408 & .192 & .26 & .14 \end{bmatrix} + \begin{bmatrix} 0 & .4 & 0 & -.4 \\ 0 & .4 & 0 & -.4 \\ 0 & -.6 & 0 & .6 \\ 0 & -.6 & 0 & .6 \end{bmatrix} \delta(k-1) +$$

(72

$$+(.0967)^{k-1} \begin{bmatrix} .416 & -.416 & -.387 & .387 \\ -.0187 & .0187 & .0176 & -.0176 \\ -.627 & +.627 & .587 & -.587 \\ -.0187 & .0187 & .0176 & -.0176 \end{bmatrix} +(-.495)^{k-1} \begin{bmatrix} .176 & -.176 & .128 & -.128 \\ -.39 & .39 & -.278 & .278 \\ .218 & -.218 & .153 & -.153 \\ -.39 & .39 & -.278 & .278 \end{bmatrix}$$

where

$$\delta(k-1) = \begin{matrix} 1 & k = 1 \\ 0 & k \neq 1 \end{matrix} \qquad (73)$$

$$\underline{W}^{Q_o}(k-1) = \underline{W}^{Q_o}(0)\ \pi_o^{\ k-1} \qquad (74)$$

$$\underline{W}^{Q_o}(0) = [.312\ \ .288\ \ .192\ \ .208] \qquad (75)$$

and

$$P_{SDD}(k) = .48W_2^{\ Q_o}(k-1) + W_3^{\ Q_o}(k-1) + .48W_4^{\ Q_o}(k-1) \qquad (76)$$

$$\Rightarrow P_{SDD}(k) = .419 + .0107\ (-.495)^{k-1} - .00076\ (.0967)^{k-1}. \qquad (77)$$

Note that

$$\lim_{k\to\infty} P_{SDD}(k) = .419 = P_{FD}, \qquad (78)$$

$$P_{SDD}(1) = .429 = P_{DD}, \qquad (79)$$

and

$$\min_{\forall k} P_{SDD}(k) = P_{SDD}(2) = .414 \begin{matrix} < P_{FD} \\ < P_{DD} \end{matrix}. \qquad (80)$$

For normal feedback decoding operation, $\underline{W}(0) = [1000]$ and

$$P_{FD}(u) = .48\ W_2(u+1) + W_3(u+1) + .48W_4(u+1) \qquad (81)$$

$$= .419 + (.0967)^{u+1}\ (-.401) + (-.495)^{u+1}\ (-.018). \qquad (82)$$

Note that

$$\lim_{u\to\infty} P_{FD}(u) = .419 = P_{FD}, \qquad (83)$$

$$P_{FD}(-1) = 0, \qquad (84)$$

and

$$P_{FD}(0) = .389 = P_{GD}. \qquad (85)$$

$P_{SDD}(k)$ for the same code and decoder but with $p = .015$ was also calculated and found to be

$$P_{SDD}(k) = 1.54 \times 10^{-3} + 8.95 \times 10^{-4}(.1121)^{k-1} - 6.92 \times 10^{-4}(-.1271)^{k-1}. \qquad (86)$$

For this case

$$\min_{\forall k} P_{SDD}(k) = P_{SDD}(5) \begin{matrix} < P_{FD} \\ < P_{DD} \end{matrix}, \qquad (87)$$

although the difference $(P_{FD} - P_{SDD}(5))$ is negligible for all practical purposes.

The fact that it is possible for $P_{SDD}(k_\alpha)$ to be less than both P_{DD} and P_{FD} indicate that semi-definite decoding might be used to practical advantage. Moreover, semi-definite decoding has the further psychological advantage that infinite error propagation is impossible.

The results of this paper, although given for codes over GF(2), can easily be generalized to codes over GF(q) if the binary symmetric channel is replaced by a memoryless, time-invariant, additive GF(q) noise source. Generalizations can also be made to non-systematic codes and non-syndrome-type decoding, but the details shall not be carried out here.

ACKNOWLEDGEMENT

The author is greatly indebted to his thesis advisor, Professor James L. Massey of the Department of Electrical Engineering at the University of Notre Dame, for his many valuable suggestions pertaining both to the reported research and to the preparation of this paper.

It is a pleasure also to acknowledge the contribution of Professor Celso de Renna e Souza of the Electrical Engineering Department at Notre Dame, who first suggested the stochastic sequential machine approach for calculating P_{FD}, and who, along with Daniel J. Costello, spent much time proofreading this report.

REFERENCES

[1] J. P. Robinson, "Error Propagation and Definite Decoding of Recurrent Codes", Tech. Rept. No. 47, Princeton University, Princeton, N. J., June, 1965. p. 12.

[2] J. M. Wozencraft and I. M. Jacobs, Principles of Communication Engineering, New York: Wiley, 1965, p. 419.

[3] D. D. Sullivan, "Control of Error Propagation in Convolutional Codes", Tech. Rept. No. EE-667, University of Notre Dame, Notre Dame, Indiana., November, 1966.

[4] Carlyle, J. W., "Equivalent Stochastic Sequential Machines", Tech. Rept. Series 60, Issue 415, University of California at Berkeley, also as "Reduced Forms for Stochastic Sequential Machines", J. Math. Anal. and Appl., Vol. 7, 1963, pp. 167-175.

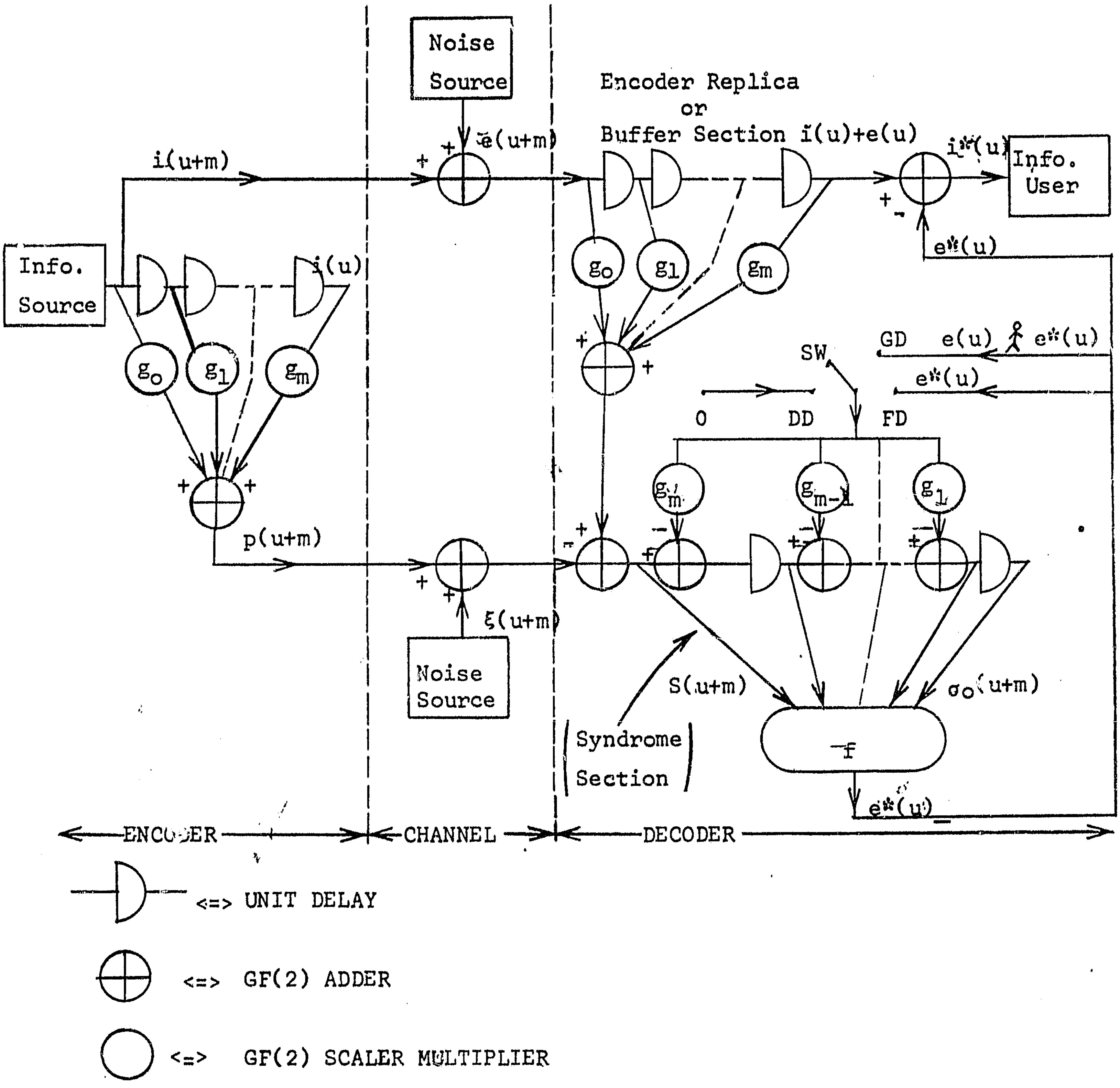
[5] Souza, C. R., and Leake, R. J., "Stochastic Finite-State Systems: Four Models and Their Relationships", Proc. 5th Annual Alerton Conference on Circuit & System Theory, Monticello, Illinois, Oct., 1967.

[6] J. L. Massey, Threshold Decoding, The Technology Press of M.I.T., Cambridge, Mass., 1963, p. 55.

[7] R. Ash, Information Theory, New York: Wiley, 1965, p. 176.

[8] J. L. Massey and R. W. Liu, "Application of Lyapunov's Direct Method to the Error-Propagation Effect in Convolutional Codes," IEEE Transactions of the Professional Technical Group on Information Theory, Vol. IT-10, Num. 3, July, 1964.

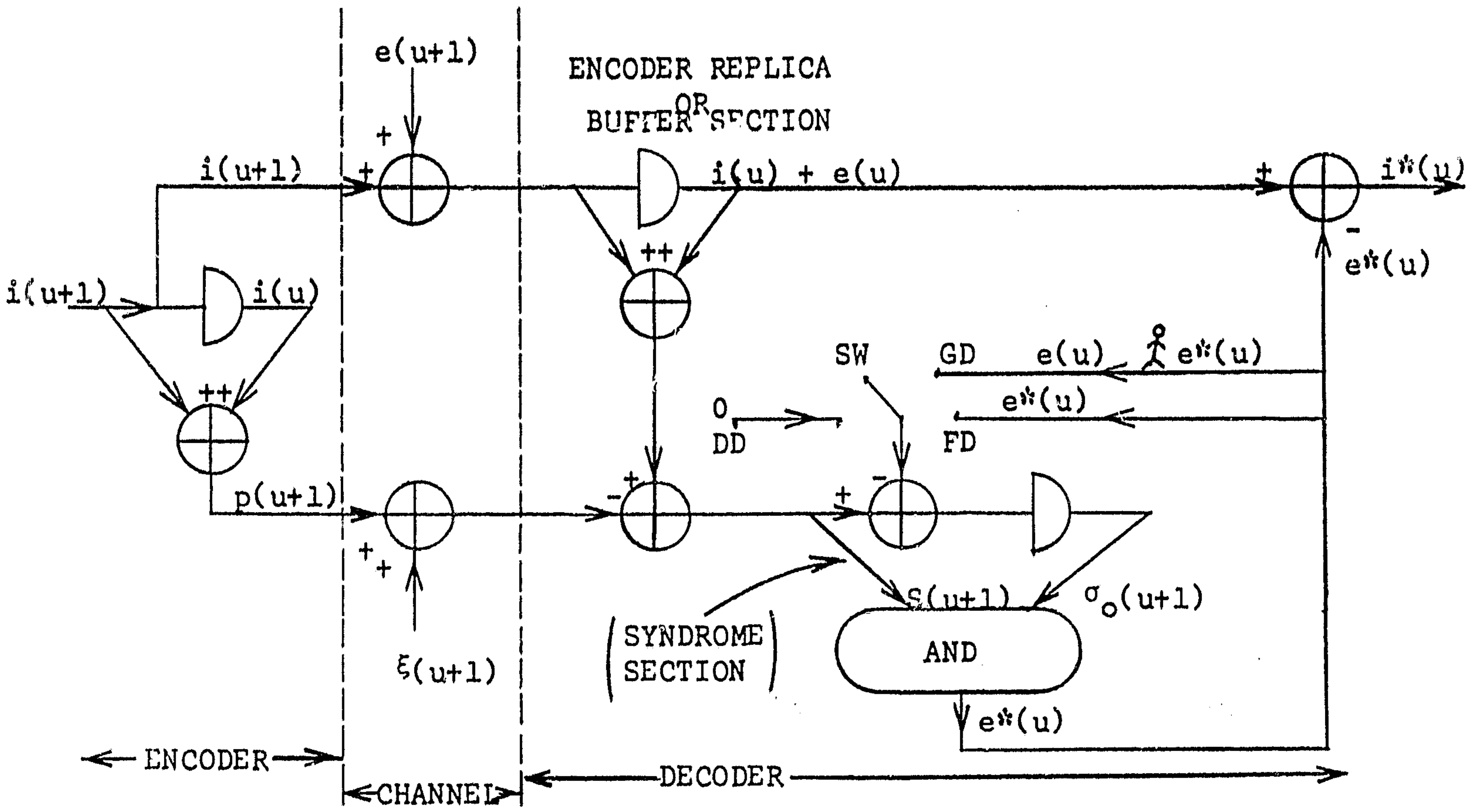
[9] L. A. Frasco, "Analysis and Simulation of Semi-Definite Decoding", S. M. Thesis, M.I.T., Course VI, June, 1967.



$g_i \in \{0,1\}$

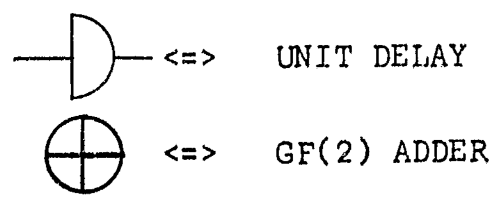
GENERAL BINARY, $R = \frac{1}{2}$ ENCODING & DECODING SYSTEM

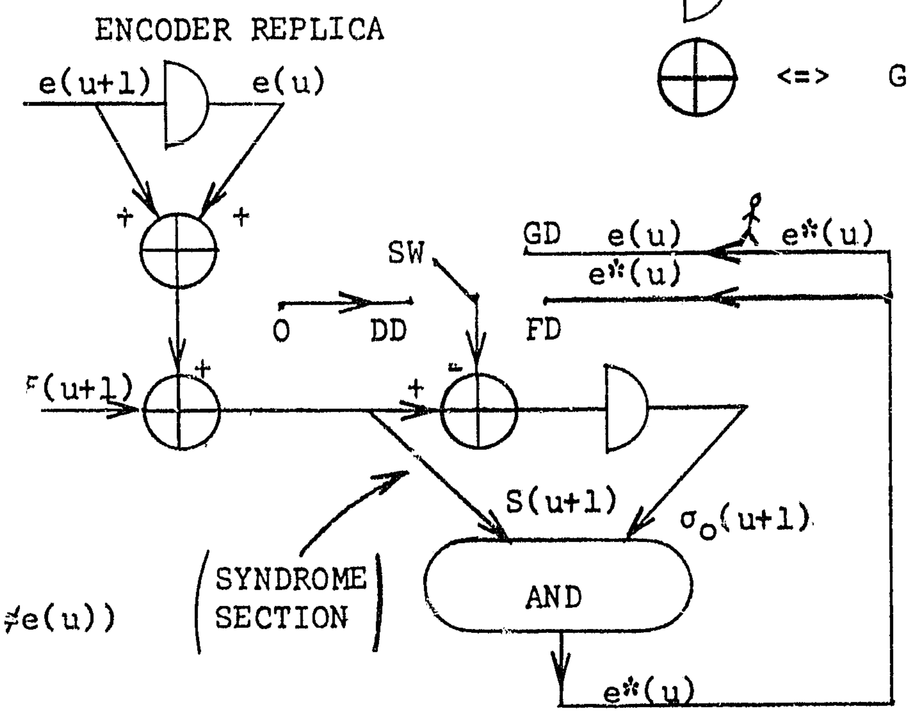
Fig. 1



BINARY, R = $\frac{1}{2}$ ENCODING and DECODING SYSTEM - SPECIAL CASE

Fig. 2

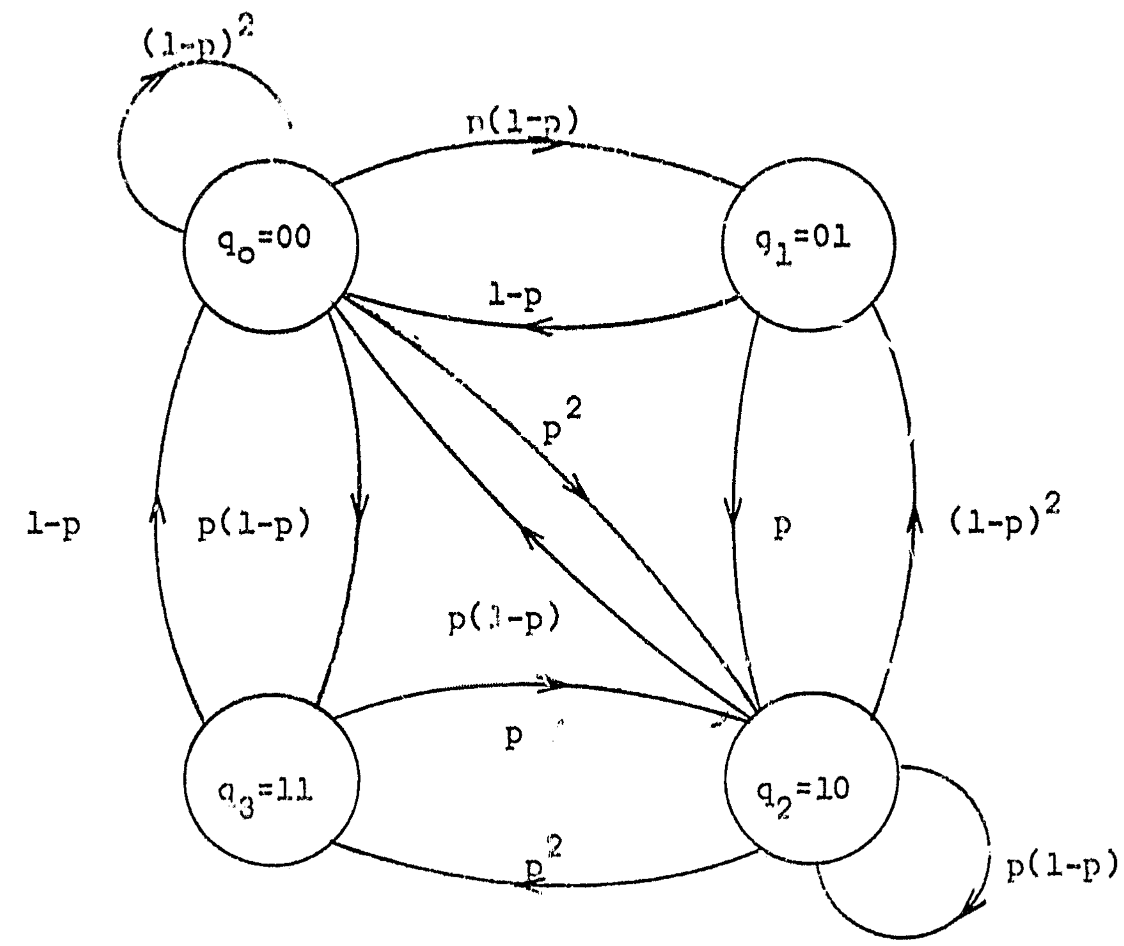




Pr(ERROR) = Pr(e*(u)≠e(u))

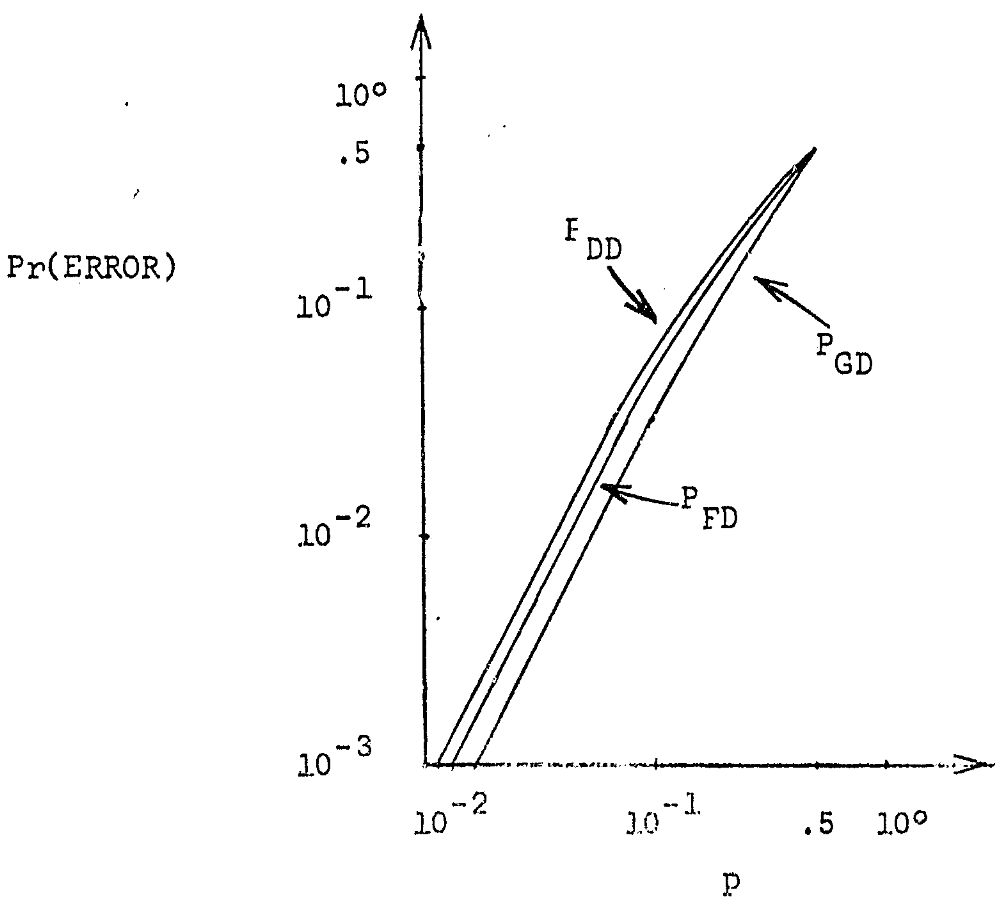
DECODER PORTION OF Fig. 2

Fig. 3



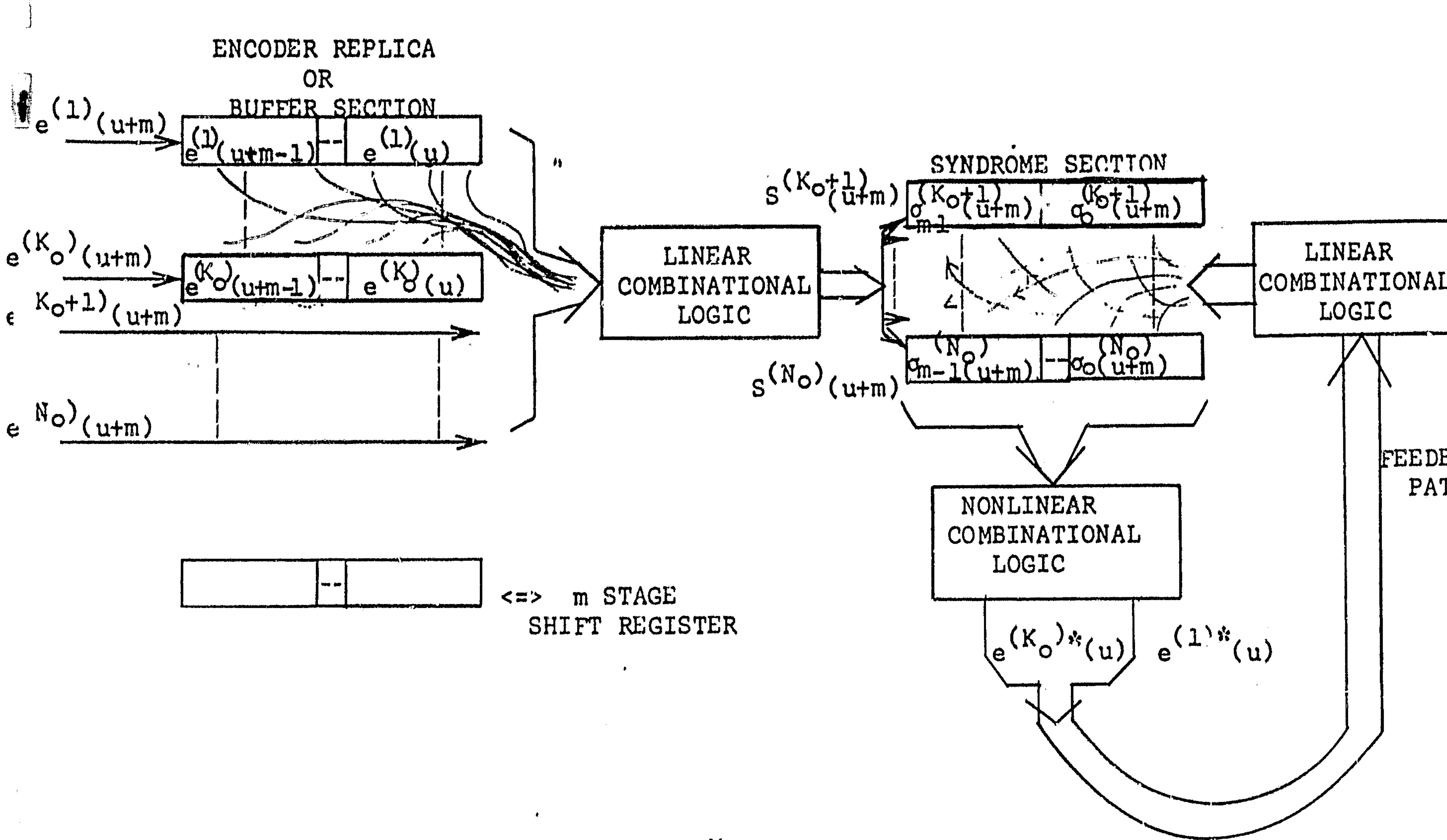
PROBABILISTIC STATE TRANSITION DIAGRAM FOR DECODER OF Fig. 3

Fig. 4



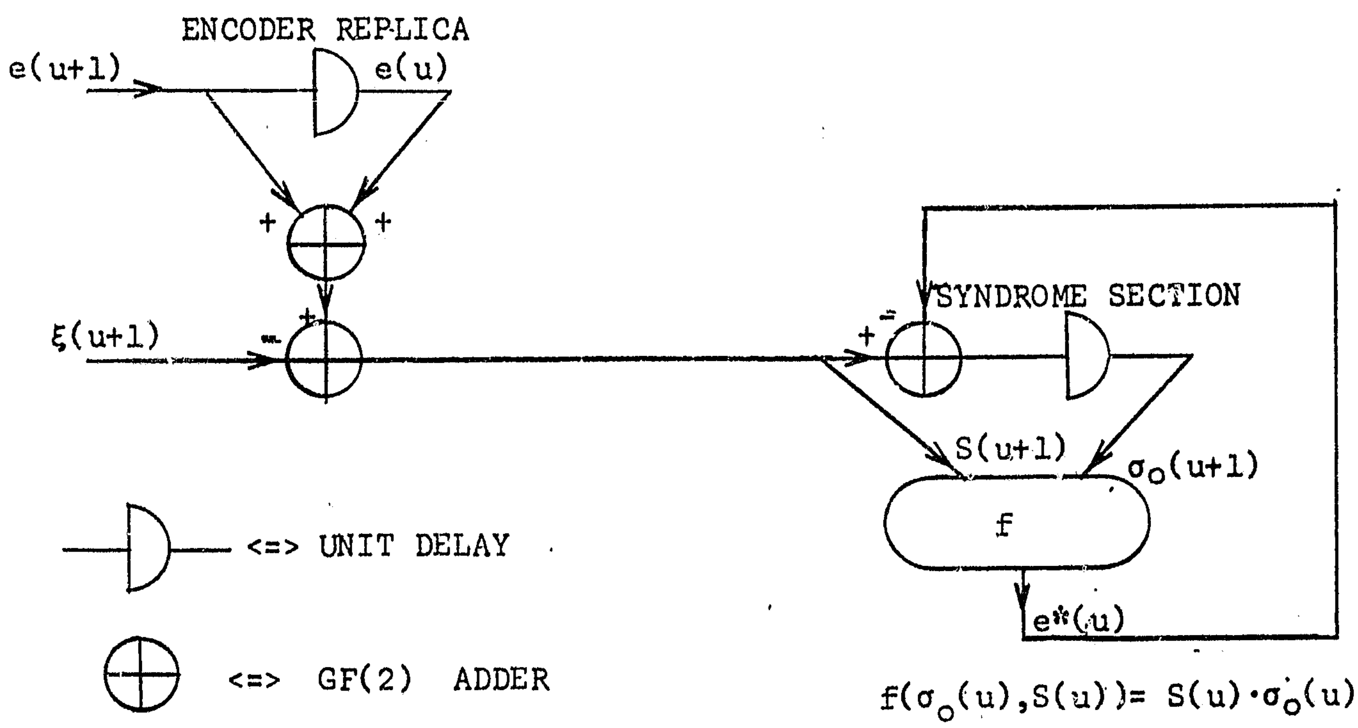
P_{GD}, P_{FD}, and P_{DD} vs. p FOR DECODER OF Fig. 3

Fig. 5



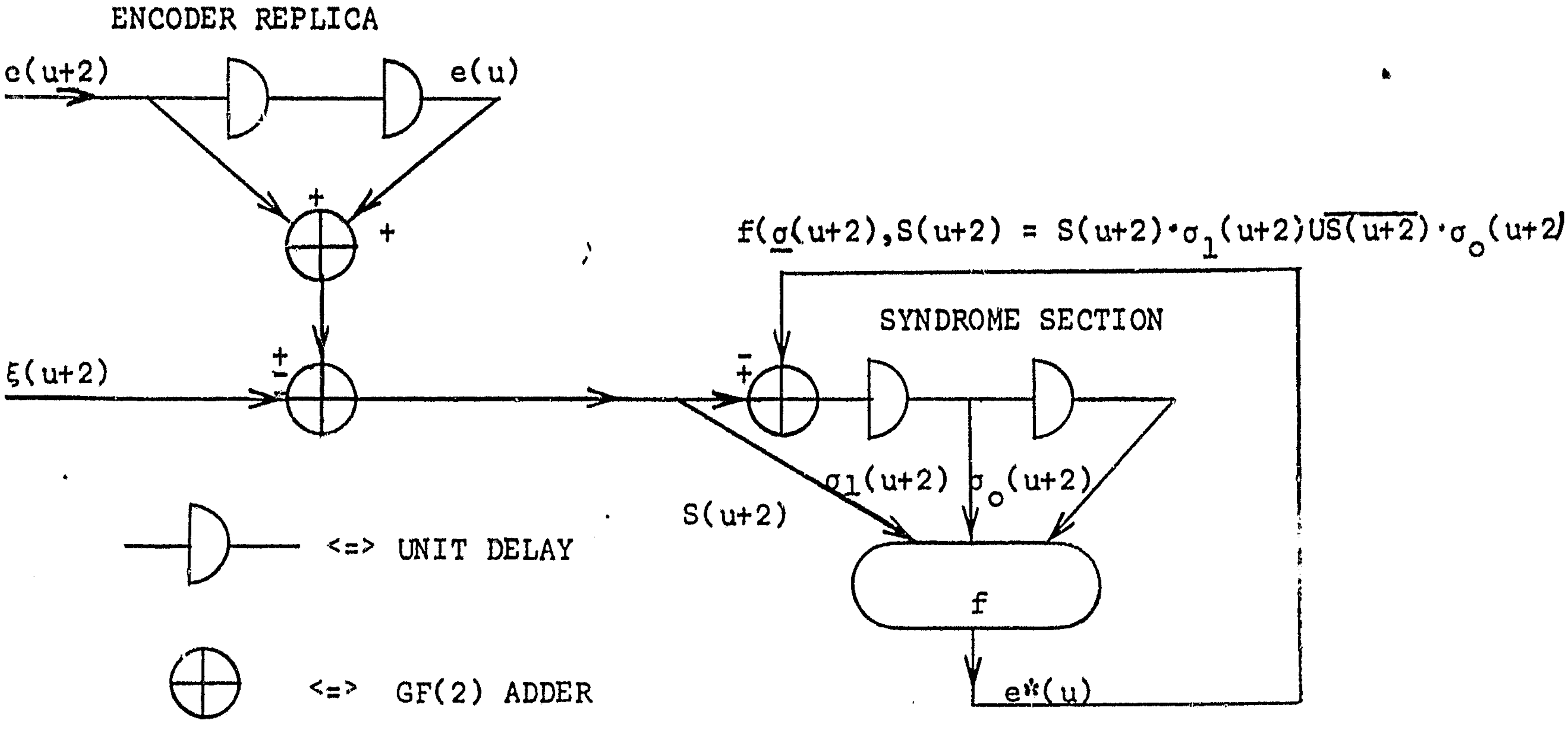
GENERAL, BINARY, $R = \frac{K_o}{N_o}$ SYNDROME FEEDBACK DECODER

Fig. 6



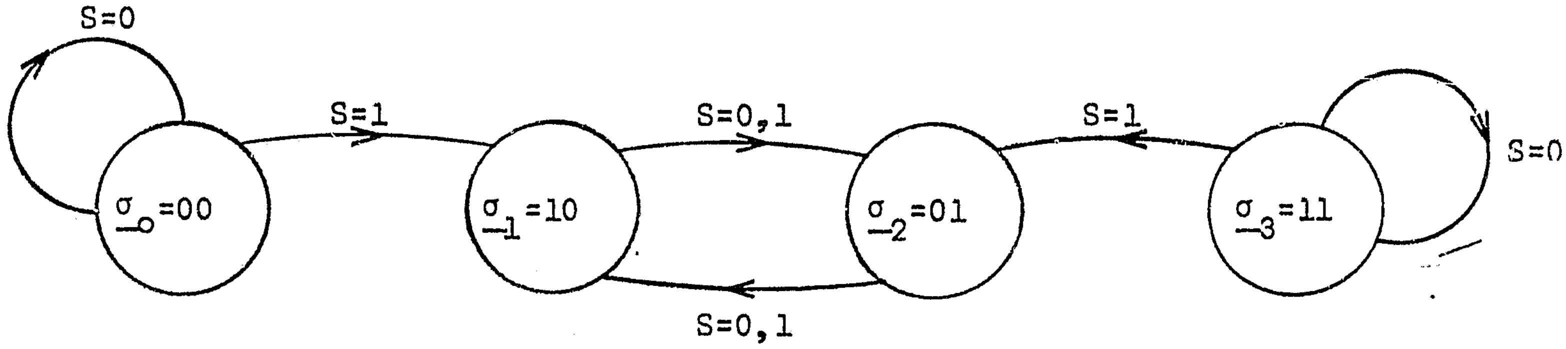
DECODER FOR WHICH W EXISTS BUT $W_o = 0$

Fig. 7



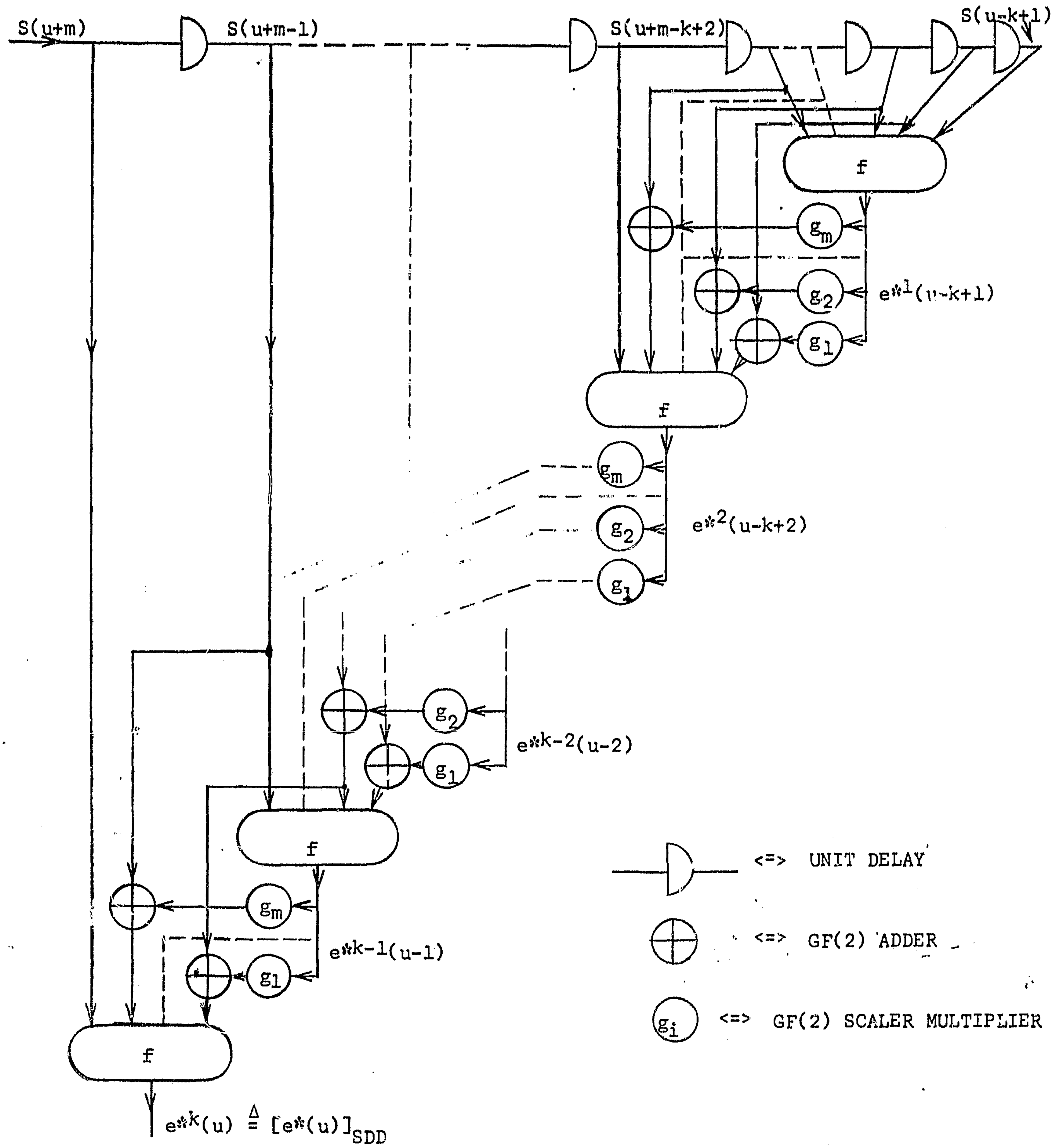
DECODER FOR WHICH NEITHER $\underline{W}$ NOR P_{FD} EXISTS

Fig. 8



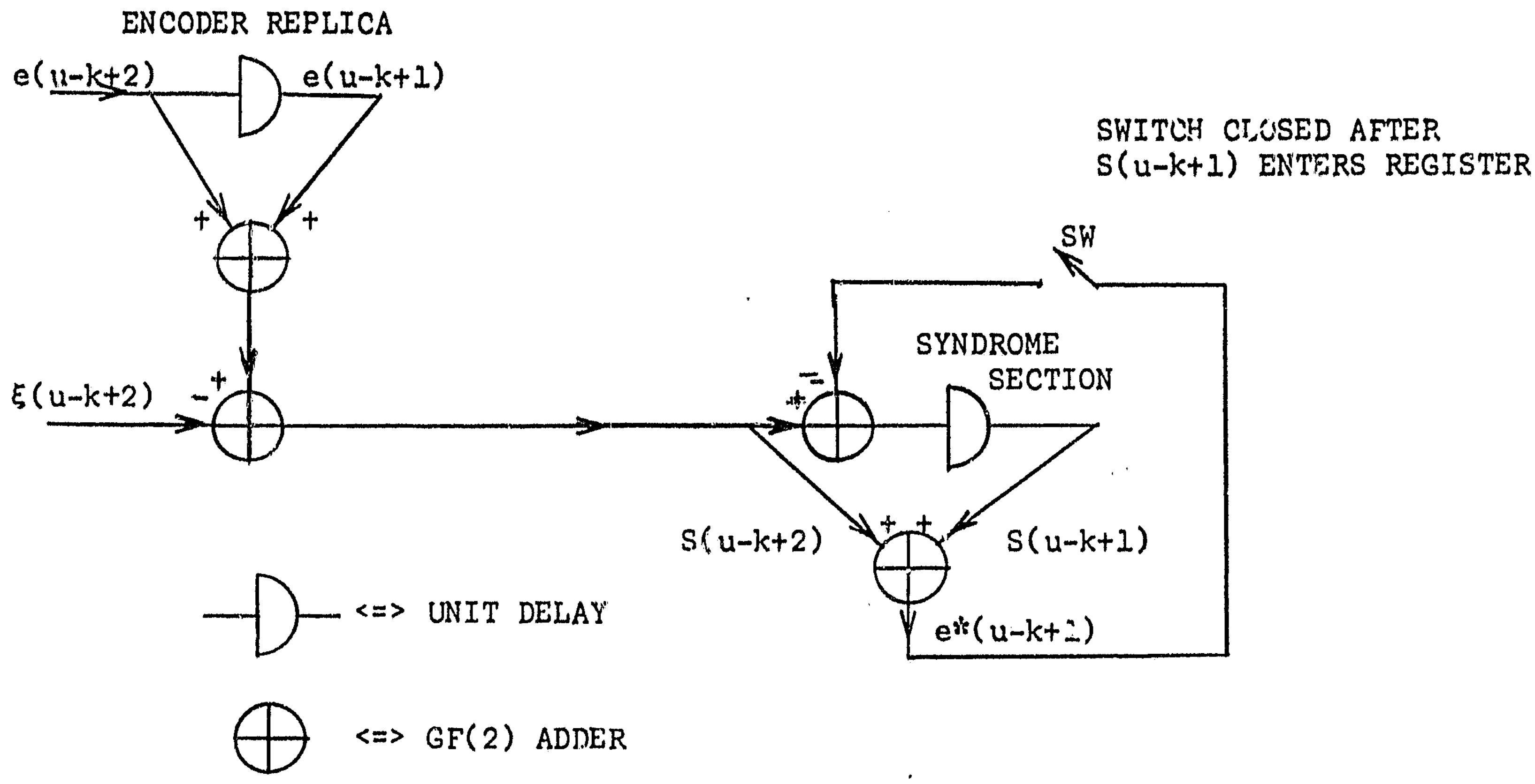
SYNDROME STATE DIAGRAM FOR DECODER OF Fig. 8

Fig. 9



GENERAL, SYSTEMATIC, BINARY, $R = \frac{1}{2}$ SEMI-DEFINITE DECODER OF ORDER k

Fig. 10



EQUIVALENT FEEDBACK DECODER

Fig. 11