On Decision Depths for Partial Response Codes

Fredrik Rusek and John B. Anderson
Dept. of Information Technology
Lund University
Lund, Sweden
Email: {fredrikr,anderson}@it.lth.se

***Abstract*— The decision depth L_d for finite and infinite Partial Response Signaling (PRS) schemes is investigated. New measures are proposed that take into account error event multiplicity. We conclude that small to moderate decoder decision depths achieve performance near MLSE even for near catastrophic codes. Under some constraints, the decision depth should be an increasing function of the signal to noise ratio.**

I. Introduction

This paper treats a simple but fundamental question: what is the decision depth (or observation window) of a decoder used for PRS coded modulations?

Consider a binary time–discrete partial response coded modulation generated by the linear filter $h[n]$ of length δ. The time continuous baseband signal $s(t)$ transmitted over the channel is generated as

$$s(t) = \sum_{n=-\infty}^{\infty} u[n]h_c(t - nT) \tag{1}$$

where $u[n]$ is a binary data sequence and $h_c(t)$ is a time continuous pulse which can in turn be represented as

$$h_c(t) = \sum_{n=0}^{\delta-1} h[n]\psi(t - nT) \tag{2}$$

where $\psi(t)$ is an orthogonal pulse, here a root raised cosine pulse. This is illustrated in figure 1. By selecting $h[n]$ suitably, considerable bandwidth reduction compared to the full response case can be obtained without any loss in minimum Euclidian distance. Normalized bandwidth, denoted by *nbw*, is used to measure bandwidth; this is the physical positive-frequency bandwidth of the pulse $h_c(t)$ multiplied by T. The AWGN channel is assumed. The bandwidth and performance of the system are governed by $h[n]$. Furthermore, we are most interested in the case where $h[n]$ is the minimum Euclidian distance optimal generator. Said [1] has compiled the PRS codes that simultaneously optimize distance and bandwidth.

Assume a Viterbi decoder or some reduced Viterbi decoder such as the M–algorithm where M is big enough to achiveve performance similar to the full Viterbi case. The decision depth L_{dec} of the *decoder* is the length of trellis path kept before the decoder makes a decision on a bit. It is a crucial storage parameter. It has been studied extensively for convolutional codes and CPM, see for example [2], [3]. For partial response codes little is known.

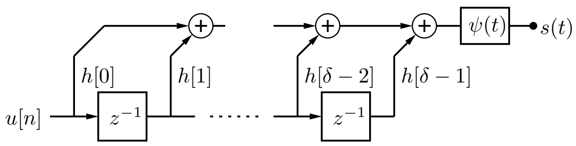

Fig. 1. A PRS coded modulation system.

An error event is the difference between two data sequences, $\xi[n] = u_1[n] - u_2[n]$. The standard way of thinking about decision depth is that it must be large enough that every possible error event, merged or unmerged, will build up a Euclidian distance larger or equal than the free distance of the code, d^2_{free}. This depth is property of the *code* and we denote it L'_d. Presumably, the decoder L_{dec} satisfies $L_{dec} \geq L'_d$. Wong [4] studied this problem for PRS codes up to δ=4. The outcome was that for most generators $L'_d \approx$ 10–15. However, in some cases L'_d grew unbounded meaning that those generators are infeasible from a decoder point of view.

By a catastrophic PRS coded modulation we mean one where an infinite error sequence generates finite Euclidian distance. All the Said optimal PRS codes are very close to catastrophic, which means that for near–MLSE performance they all seem to be infeasible. However, by simulations we quickly see that this is not true. Usually it is enough if L_{dec} is set to a rather small number, such as 20–50. In this paper we investigate the mimimum L_{dec} that actually achieve performance near MLSE.

II. Decision depths

First some notation. An error event is in general denoted by ξ and the set of all error events up to length m is denoted $\mathcal{A}_m$. Received signal to noise ratio per bit is denoted by γ_b. Furthermore, m_ξ and $d^2(\xi)$ denote the multiplicity and Euclidian distance of the error event ξ. By $d^2(\xi, n)$ we mean the accumulated Euclidian distance up to n trellis stages, i.e.

$$d^2(\xi, n) = \sum_{k=0}^{n} \left| \sum_{l=0}^{k} \xi[l]h[k-l] \right|^2. \tag{3}$$

The multiplicity of an error event is easy to derive [6]. Let $\nu(\xi)$ denote the number of non zero elements in the error event ξ; then $m_\xi = 2^{1-\nu(\xi)}$.

Forney [5] showed that the first event error probability can be upper bounded as

$$P_e \leq \sum_{\xi \in \mathcal{E}} m_\xi Q(\sqrt{d^2(\xi)\gamma_b}), \quad (4)$$

where $\mathcal{E}$ is a large set of error events. This set was significantly reduced in [6] giving a tighter upper bound. When γ_b grows large (4) is well approximated by only a few terms and in the limit of $\gamma_b \to \infty$ it depends on the event achieving the free distance alone.

A. The standard definition

We can now apply the standard approach to partial response codes. This gives the following definition of the decision depth of a code:

Definition 1:

$$L'_d \doteq \arg\min_n \quad d^2(\xi, n) \geq d^2_{free} \quad \forall \xi, \quad \xi[0] \neq 0 \quad (5)$$

The meaning of definition 1 is that a decoder with $L_{dec} \geq L'_d$ corrects all errors that lie $< d_{free}/2$ away from the correct signal path. It is straightforward to compute L'_d for different codes. Doing so for a collection of optimal binary codes gives the results in table I. The specific codes are listed in the appendix.

The decision depths are very large and simulations show that L'_d can be reduced significantly without any performance loss. The reason why L'_d is this big is that the optimal PRS codes are close to catastrophic. However, the near catastrophic events do not degrade the performance of the system due to their very small multiplicities. These decay exponentially with the length of the error events. For example the error event for code A that forces L'_d to be 882 has $m_\xi \approx 10^{-177}$.

B. A new definition

We now turn to a new definition on L_d that considers multiplicities as well. First some notation. Let $P_e(\xi)$ denote the overall probability of the error event ξ, i.e.

$$P_e(\xi) = m_\xi Q(\sqrt{d^2(\xi)\gamma_b}). \quad (6)$$

Finding the k largest $P_e(\xi)$ yields $P^k_{e,\max}$, the kth largest $P_e(\xi)$. Related to $P^k_{e,\max}$ is also the set $\mathcal{E}^k = \{\xi_1, \ldots, \xi_k\}$ consisting of the k events achieving $P^1_{e,\max}, \ldots, P^k_{e,\max}$. The new decision depth is now defined as

Definition 2:

$$L^k_d(\gamma_b) \doteq \arg\min_n \quad m_\xi Q(\sqrt{d^2(\xi, n)\gamma_b}) \leq P^\sigma_{e,\max}$$
$$\xi \in \mathcal{A}_n, \quad \xi[0] \neq 0, \quad (7)$$

where σ satisfies

$$\sigma = \begin{cases} j & \text{if } \xi = \xi_j \in \mathcal{E}^k \\ k & \text{otherwise.} \end{cases}$$

Note that the decision depth here depends on γ_b. Using this depth the k largest terms in (4) are kept intact. An immediate consequense of this definition is that $L^k_d(\gamma_b) \geq L^{k'}_d(\gamma_b)$ if $k > k'$. Furthermore, L^k_d is lower bounded as

$$L^k_d \geq \max_{\xi \in \mathcal{E}^k} \|\xi\| + \delta - 1, \quad (8)$$

where $\|\xi\|$ means the length of the error event ξ. This implies that we can only find decision depth for finite impluse responses. The new definition of decision depth is based upon (4); therefore we only use L^k_d where (4) makes sense, i.e. for moderate to high γ_b.

We have found values for L^k_d by computer search. A branch and bound algortihm was used. The bound works in the following way: assume that $m_\xi Q(\sqrt{d^2(\xi, n)\gamma_b}) < P^k_{e,\max}$ where $\|\xi\| - 1 = n$, then $m_{\xi_s} Q(\sqrt{d^2(\xi_s, n)\gamma_b}) < P^k_{e,\max}$ for all ξ_s having ξ as its first $n+1$ components. This means that they can all be eliminated from the search. By using this bound the computer search is a very easy task. A few examples of L^k_d are given in table I. The new values are considerably smaller. For nbw=.3 we used γ_b = 12 dB and for nbw=.2 we used γ_b = 16 dB. Figure 2 shows L^1_d as a function of bandwidth for different values of δ. It can be seen that no dramatic change in L^1_d occurs. The value of γ_b was chosen such that $P_{e,\max} < 10^{-5}$.

nbw = .3 Hz-s		nbw = .2 Hz-s	
Code A	Code B	Code C	Code D
δ = 8	δ = 12	δ = 8	δ = 12
L'_d = 882	L'_d = 492	L'_d = 250	L'_d = 277
L^1_d = 12	L^1_d = 14	L^1_d = 12	L^1_d = 17
L^3_d = 13	L^3_d = 15	L^3_d = 17	L^3_d = 23

TABLE I

DECISION DEPTH FOR SOME OPTIMAL PRS CODES LISTED IN TABLE IV.

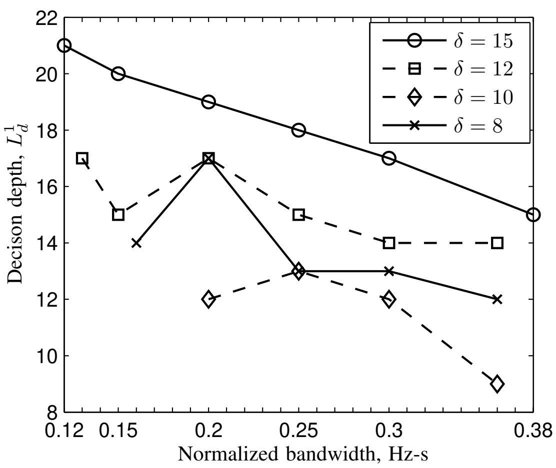

Fig. 2. L^1_d for a number of optimal PRS codes, as a function of the normalized bandwidth.

We close this subsection with two simple but interesting theorems.

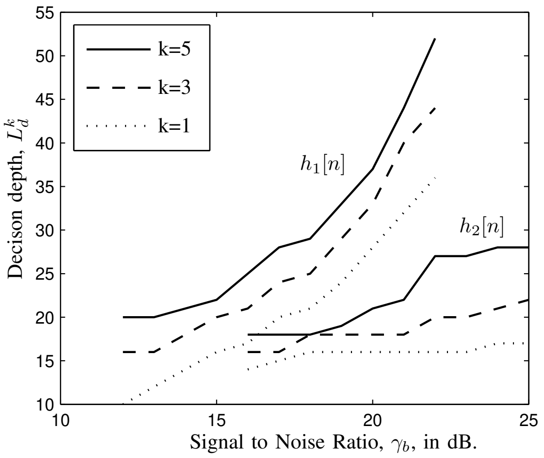

Fig. 3. Decision depth L_d^k as a function of the signal to noise ratio γ_b, for two optimal PRS codes.

Theorem 1:

$$L_d^1(\gamma_b) \to L'_d, \quad \text{as} \quad \gamma_b \to \infty \tag{9}$$

Proof When $\gamma_b \to \infty$ the distance inside the Q function determines everything. So we are back in definition 1. ■

The theorem is mainly of theoretical interest; in order to approach L'_d it is necessary to use a huge γ_b giving vanishing bit error rates. The only interesting case when $\gamma_b \to \infty$ is $k = 1$.

Theorem 2: Assume that $L_d^k(\gamma_b^1) < L'_d$ and $L_d^k(\gamma_b)$ depends only on a single error event ξ_s for $\gamma_b > \gamma_b^1$. Then $L_d^k(\gamma_b)$ is an increasing function for $\gamma_b > \gamma_b^1$ as long as $L_d^k(\gamma_b) < L'_d$.

Proof Since $L_d^k(\gamma_b)$ only depends on ξ_s and $L_d^k(\gamma_b^1) = \Delta < L'_d$, we conclude that $d^2(\xi_s, \Delta) < d_{free}^2$. Furthermore, when γ_b grows, $P_{e,\max}^k$ will decay faster since it has a larger distance; thus $L_d^k(\gamma_b^2)$ has to be set to $\Delta + 1$ for some $\gamma_b^2 > \gamma_b^1$ in order to build up more distance. This implies that the decision depth can never be decreased as long as $L_d^k(\gamma_b) < L'_d$. ■

Theorem 2 simply states, perhaps contradicting intuition, that the decision depth is an increasing function over a range of the signal to noise ratio whenever there is single error event that gives the decision depth for all the γ_b. For all considered optimal PRS codes and some other codes this is the case.

Figure 3 is an illustration of Theorem 2. $h_1[n]$ is the optimal generator for normalized bandwidth .20 Hz-s with $\delta = 10$, and $h_2[n]$ is the optimal generator for 99.9% normalized bandwidth[1] .13 Hz/s with $\delta = 12$. The rightmost region of figure 3, having $P_{e,\max}^k \approx 10^{-10} - 10^{-15}$, is of less practical interest.

[1]This means that at most 0.1 % of the signal power lies outside the nbw.

C. *The Proakis* b *and* c *channels*

So far we have been only interested in intentional ISI, i.e. PRS coded modulation schemes. Of course we could investigate decision depth for ISI channels as well. Here we look at the well known b and c channels taken from [7]. The b channel has impulse response $\{.407\ .815\ .407\}$ and the c channel has impulse response $\{.227\ .460\ .688\ .460\ .227\}$.

We start by examining the upper bound (4) for the b channel. The ten largest $P_e(\xi)$ at $\gamma_b = 10$ dB are tabulated in table II. By using L_d^5 we get a new upper bound on first event error probability, which begins with the five largest terms in table II. The value of L_d^5 at 10 dB is found to be 11.

k	$P_{e,\max}^k(\xi)$	ξ
1	$.6690 \cdot 10^{-4}$	2,-2
2	$.3345 \cdot 10^{-4}$	2,-2,2
3	$.1673 \cdot 10^{-4}$	2,-2,2,-2
4	$.8362 \cdot 10^{-5}$	2,-2,2,-2,2
5	$.4181 \cdot 10^{-5}$	2,-2,2,-2,2,-2
6	$.4058 \cdot 10^{-5}$	2
7	$.2091 \cdot 10^{-5}$	2,-2,2,-2,2,-2,2
8	$.1045 \cdot 10^{-5}$	2,-2,2,-2,2,-2,2,-2
9	$.5226 \cdot 10^{-6}$	2,-2,2,-2,2,-2,2,-2,2
10	$.4986 \cdot 10^{-6}$	2,-2,0,2,-2

TABLE II

THE 10 LARGEST $P_e(\xi)$ AT $\gamma_b = 10$ AND CORRESPONDING ERROR EVENTS; THE EVENTS SHOULD BE DIVIDED BY $\sqrt{2}$. PROAKIS ISI CHANNEL a.

It is easy to verify that the error sequence $2, -2, 2, -2, 2, \ldots$ forms a nearly catastrophic error sequence for this channel. This means that L'_d is huge and not feasible to use. A common way to estimate a practical decoder L_{dec} is to base the estimation on the memory length and use simple rules such as five times the memory length. In this case such a computation gives $L_d = 10$ which matches L_d^5 very well. However, if the transmission has a higher γ_b this is no longer true. For example if $\gamma_b = 14$ dB, L_d^5 now equals 18. From figure 2 we suspect that such simple static rules do not hold since code decision depths seem to be dynamic, depending on bandwidth and signal to noise ratio. Similar results hold for the c channel as well.

III. EXTENSION TO IIR IMPULSE RESPONSES

Our definition of decision depth led to the lower bound (8). Since this bound contains δ it follows that it is worthless to use our approach for IIR filters or for very long FIR filters. However, it is possible to modify definition 2 slightly to include the IIR case as well. We propose the following definition

Definition 3:

$$L_d^k(\gamma_b) \doteq \arg\min_n : m_\xi Q(\sqrt{d^2(\xi, n)\gamma_b}) \leq \beta P_{e,\max}^\sigma$$
$$\xi \in \mathcal{A}_n, \quad \xi[0] \neq 0, \quad \beta > 1 \tag{10}$$

where σ satisfies

$$\sigma = \begin{cases} j & \text{if } \xi = \xi_j \in \mathcal{E}^k \\ k & \text{otherwise.} \end{cases}$$

As β we take a small number such as 1.01. This definition doesn't lead to a lower bound based on δ, and should therefore be suitable for IIR impulse responses. Furthermore, it guarantees that the upper bound (4) will start with the true k largest terms multiplied by β. For this definition theorem 1 holds as well.

Other definitions are also possible; for example we could replace β in definition 3 by β_k and have variable losses for the different terms.

Consider the case where $h[n]$ is a time discrete Butterworth filter [8]. We create $h[n]$ by

```
[B,A]=butter(8,2fc);
h[n]=filter(B,A,[1 zeros(1,50)]);
```

in MATLAB, where f_c is the cutoff frequency. We consider 8–pole filters. Codewords are made up according to (1). By changing the filter cutoff frequency, different bandwidths and minimum distances are created. Strictly speaking, L'_d is always infinite since the impulse response is. However, in practice we always truncate the impulse response after 30–50 symbol intervals. It is easy to verify that $\xi = 2, -2, 2, -2, 2, \ldots$ forms a nearly catastrophic error sequence for these codes. Thus, even for a truncated version of $h[n]$ we get an infeasible L'_d.

Table III shows decision depths according to definition 3

$\beta = 1.01$		$\beta = 1.10$	
$f_c = .35$	$f_c = .20$	$f_c = .35$	$f_c = .20$
L_d^1 = 20	L_d^1 = 24	L_d^1 = 15	L_d^1 = 18
L_d^5 = 27	L_d^5 = 28	L_d^5 = 20	L_d^5 = 22

TABLE III

DECISION DEPTH FOR TWO BUTTERWORTH FILTERS.

for two Butterworth filters. We use γ_b large enough to achieve $P^1_{e,\max} = 10^{-5}$. No truncation of the impulse response was done when finding these decision depths.

IV. Receiver tests

We now verify that we get performance close to optimal when using L_d^k, both for the PRS codes in table IV and for the b and c channels. We use the M–algorithm with M = 32; this is big enough to make simulations independent of this parameter. We measure the first event error probability for $L_{dec} = L_d^1$ and L_d^5 as well as $L_{dec} = 200$. The simulations for codes A and C are shown in figure 4. The first event bound (4) is plotted for comparison. From this it is clear that near MLSE performance is obtained using L_d^5. Less than 1 dB is lost by using L_d^1. For code A, L_d^5 varies between 9 and 19, considerably smaller than the simple estimation $L_d = 5 \cdot 7 = 35$.

We also simulate the b and c channels from [7]. In this case a full Viterbi decoder was used. The simulations are shown in figure 5. Again it is clear that near MLSE performance is obtained using L_d^5.

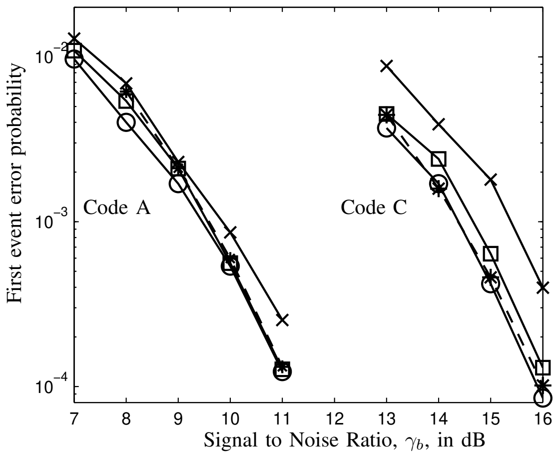

Fig. 4. Measurements of first event error probability. ○ denotes $L_{dec} = 200$, × denotes $L_{dec} = L_d^1$ and □ denotes $L_{dec} = L_d^5$. The dashed line marked with $*$ is the upper bound for the first error event probability (4)

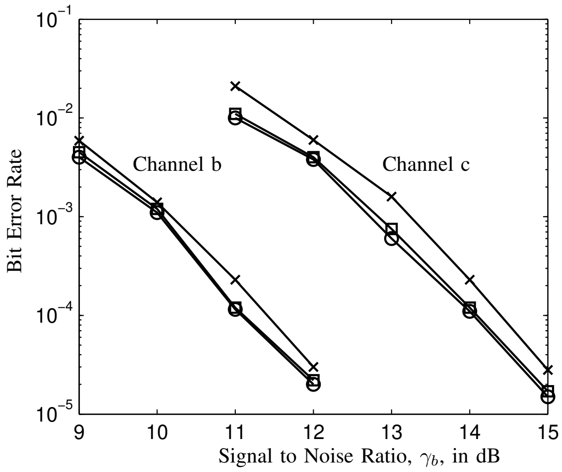

Fig. 5. BER for Proakis channels. ○ denotes $L_{dec} = 200$, × denotes $L_{dec} = L_d^1$ and □ denotes $L_{dec} = L_d^5$.

V. Conclusion

We have seen that the standard definiton of decision depth is not very useful when applied to PRS. By modifying the definition we get a new decision depth that more accurately predicts receiver error rate. The newly found depths are considerably smaller than the depths predicted from the standard definition on decision depths. We also saw that the decision depth depends on the signal to noise ratio and that in the limit the two definitions of L_d converge.

Furthermore, the optimal generators reported in [1] were found under constraints of maximal Euclidian minimum distance. However, to fully exploit the strength of these codes we should theoretically use L'_d and we have demonstarted that this is infeasible. Therefore it is clear that the Euclidian distance optimal codes may not be the optimal with a more practical decision depth. An alternative approach is to optimize the taps under constraint of minimal BER.

APPENDIX

We list $h[n]$ and d^2_{free} for Code A,B,C and D taken from [1].

Code	A	B	C	D
d^2_{free}	.99	1.13	.32	.45
h[0]	.281	.278	.122	.132
h[1]	.657	.660	.307	.330
h[2]	.642	.620	.488	.514
h[3]	.184	.141	.559	.545
h[4]	-.160	-.135	.480	.373
h[5]	-.093	.008	.302	.082
h[6]	.073	.092	.128	-.175
h[7]	.065	-.0850	.028	-.277
h[8]		-.148		-.221
h[9]		.035		-.104
h[10]		.147		-.020
h[11]		.071		.005

TABLE IV

OPTIMAL CODE GENERATORS AND CORRESPONDING FREE DISTANCES.

REFERENCES

[1] A. Said and J.B. Anderson, "Design of optimal signals for bandwidth-efficient linear coded modulation", *IEEE Trans. Information Theory*, vol. 44, pp. 701-713, March 1998.

[2] J.B. Anderson and K. Balachandran, "Decision depths of convolutional codes," *IEEE Trans. Information Theory*, IT–35, 455–459, March 1989.

[3] J.B. Anderson and A. Svensson, *Coded Modulation Systems*, Kluwer-Plenum, New York, 2003.

[4] C. W.-C. Wong and J.B. Anderson, "Optimal short impulse response channels for an MLSE receiver," *Conf. Rec.*, Int. Conf. Commun., Boston, Mass., 25.3.1–25.3.5, June 1979.

[5] G. D. Forney, Jr., "Maximum likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Trans. Information Theory*, IT-18, 363–378, May 1972.

[6] S. Verdu, "Maximum likelihood sequence detection for intersymbol interference channels: a new upper bound on error probability," *IEEE Trans. Information Theory*, vol. IT–33, pp. 62–68, Jan. 1987.

[7] J.G. Proakis, *Digital Communications*, 4th ed., McGraw-Hill, New York, 2001.

[8] F. Rusek and J.B. Anderson, "M–ary Coded Modulation by Butterworth Filtering," *Proc.*, 2003 Int. Symp. Inform. Theory, Yokohama, p. 184, June 2003.