Near BER optimal partial response codes

Rusek, Fredrik; Anderson, John B

Published in:
2005 IEEE International Symposium on Information Theory, Adelaide

DOI:
10.1109/ISIT.2005.1523393

2005

Near BER Optimal Partial Response Codes

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

***Abstract*— Partial Response Signaling (PRS) codes with maximal minimum Euclidean distance have previously been found by linear programming. These perform very well in the narrowband–high energy region, but they were not optimized for minimal Bit Error Rate (BER), so they are only optimal in the limit of infinite signal to noise ratio. Here we search for codes that perform better for more practical signal to noise ratios. The BER objective function is no longer linear, but it is still convex.**

I. Introduction

The CPM, TCM and partial response signaling (PRS) classes of coded modulation are well known. In the last class are intersymbol interference removal and codes based on filtering and real-number discrete-time convolution. These classes operate in different parts of the energy–bandwidth plane and PRS coding in particular is characterized by narrow bandwidth and high energy per bit. Here we synthesize PRS-type codes by severe time–discrete filtering of a data train.

Consider a binary partial response coded modulation generated by the linear filter $h_c[n]$. The time continuous baseband signal $s_u(t)$ transmitted over the channel is generated as

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

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

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

where $\mathbf{b}$ is a time discrete filter of length δ and $\psi(t)$ is an orthogonal pulse, here a root raised cosine pulse. This is illustrated in figure 1. By selecting $\mathbf{b}$ suitably, considerable bandwidth reduction compared to the full response case can be obtained with a small, controlled loss in minimum Euclidean 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 energy performance of the system are governed by $\mathbf{b}$.

Much research over the years has gone into finding a generator $\mathbf{b}$ that optimizes the performance of the system [1]–[4]. More recently Said [5] derived a linear formulation of Euclidean distance optimal generators at a given bandwidth. This class of optimal PRS (OPRS) codes has very good performance; generally it improves on TCM. Furthermore, the OPRS class can be improved significantly by concatenating an outer code [6].

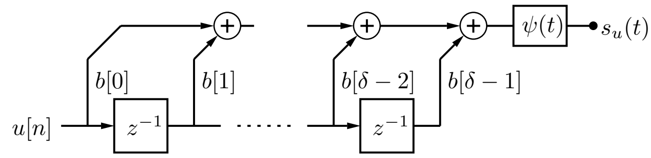

Fig. 1. A PRS coded modulation system.

The OPRS class is optimal in a Euclidean distance sense, but it may not be BER-optimal. In this paper we search for codes with better BER properties than the OPRS codes. To do this we need to modify the OPRS optimization enough so that it takes some account of BER while still being tractable. An exact and tractable optimization of BER is probably not possible, and the test of a good procedure has to be that it leads to codes with good measured receiver BER. Naturally our new codes have minimum distance less than OPRS. OPRS codes are asymptotically optimal in the limit of infinite signal to noise ratio, and our new codes can only outperform OPRS in a finite, but hopefully useful range.

The paper is organized in the following way. In section 2 we give some basic notation and give the linear formulation of OPRS, which is a basic building block. In Section 3 we give the optimization methods used for good BER codes. Numerical results and simulations are presented in section 4.

II. Properties of PRS

It can be shown that if the transmitted signal is generated according to (1) and (2), then samples each T seconds of a filter matched to $\psi(t)$ form a set of statistically independent sufficient statistics. Therefore we always work with the Euclidean distance equivalent time-discrete model $s_u[n] = \sum_{k=0}^{\infty} u[k]b[n-k]$ instead of the continuous version. The energy per bit, E_b, equals $\sum_{n=0}^{\delta-1} |b[n]|^2$.

A. OPRS

Here we follow the derivations of [5]. An error event is denoted by $\xi[n]$ and is the difference between two data sequences, i.e. $\xi[n] = u_1[n] - u_2[n]$. We begin by recapitulating Said's linear formulation of OPRS since this is a building block for us.

The normalized Euclidean distance built up from ξ is

$$\begin{aligned} d^2(\xi) &= \frac{1}{2E_b}\sum_{k=0}^{\infty}|s_\xi[k]|^2 \\ &= \frac{1}{2E_b}\sum_{k=0}^{\infty}\left|\sum_{m=0}^{\infty}\xi[m]b[k-m]\right|^2 \\ &= \frac{1}{2E_b}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\xi[m]g[n-m]\xi[n], \end{aligned} \quad (3)$$

where

$$g[n-m] = \sum_{k=-\infty}^{\infty} b[k+n-m]b[k] \quad (4)$$

and E_b denotes energy per bit. By some algebraic manipulation we bring (3) to the linear form

$$d^2(\xi) = \sum_{k=-\infty}^{\infty} g[k]r_\xi[k], \quad (5)$$

where

$$r_\xi[n] = \frac{1}{2E_b}\sum_{k=-\infty}^{\infty}\xi[k+n]\xi[k] = \frac{\xi[n]}{\sqrt{2}} * \frac{\xi[-n]}{\sqrt{2}} \quad (6)$$

if the energy of $\mathbf{b}$ is normalized to 1, i.e.

$$E_b = g[0] = 1. \quad (7)$$

The Fourier transform of the pulse $h_c(t)$ can be found to be

$$|H_c(f)|^2 = |\Psi(f)|^2 \sum_{n=1-\delta}^{\delta-1} g[n]e^{-j2\pi nfT}. \quad (8)$$

This gives that the concentration of the pulse in the frequency interval $[-W, W]$ equals

$$\int_{-W}^{W} |H_c(f)|^2 df = \sum_{n=1-\delta}^{\delta-1} g[n]\chi[n], \quad (9)$$

where

$$\chi[n] = \int_{-W}^{W} |\Psi(f)|^2 e^{j2\pi nfT}. \quad (10)$$

At this point we have the linear equations (5), (7) and (9) for Euclidean distance, energy normalization and bandwidth respectively. This means that we can solve for the optimal tap set $b[n]$ by using linear programming over $g[n]$. However, we must constrain $g[n]$ to be a valid autocorrelation sequence, meaning that there exists a tap set $b[n]$ having $g[n]$ as autocorrelation. The following infinite set of linear constraints on $g[n]$ ensures that a tap set $b[n]$ exists:

$$\sum_{n=1-\delta}^{\delta-1} g[n]e^{-j2\pi nf} = \sum_{n=1-\delta}^{\delta-1} g[n]\kappa_f[n] \geq 0, \quad \forall f \in [0,1). \quad (11)$$

These are called the admissibility constraints.

We now state the optimization problem. We replace the notation notation $g[n]$ by g:

$$\begin{aligned} d^2_{\text{min,opt}} &= \max_{x,g} \quad x \\ \text{s.t.} \quad & \begin{array}{lll} g^{\text{tr}} r_\xi & \geq x & \forall \xi \\ g[0] & = 1 & \\ g^{\text{tr}} \chi & = C & \\ g^{\text{tr}} \kappa_f & \geq 0 & \forall f \in [0,1) \end{array} \end{aligned} \quad (12)$$

The parameter C is called the spectral concentration and is set to $.999$ throughout the paper. This implies that the transmission has 99.9% of its power inside W Hz.

B. Error Probability

Forney [7] derived upper bounds to the first event error probability as well as the bit error probability for MLSE; these results were then revised by Foschini [8]. Here we are only interested in bounds on the BER. The standard Forney upper bound is

$$P_b \leq \sum_{\xi\in\mathcal{E}} d_H(\xi)m_\xi Q(\sqrt{d^2(\xi)\gamma_b}), \quad (13)$$

where $d_H(\xi)$ is the Hamming weight of the event ξ, m_ξ is the multiplicity of ξ and γ_b is the received signal to noise ratio. The set $\mathcal{E}$ is a huge set of error events. This set was reduced in [9]. The multiplicity m_ξ is easy to derive. Assume that ξ has length M and begins at time lT. Let $\hat{u}$ denote the data sequence $u[l], \ldots, u[l+M-1]$ and $\mathcal{U}_M$ be the the set of all possible data sequences of length M, i.e. $\mathcal{U}_M = \{+1, -1\}^M$. Then

$$\begin{aligned} m_\xi &= Pr\{\hat{u} + \xi \in \mathcal{U}_M\} \\ &= \prod_{k=0}^{M-1} Pr\{u[l+k] + \xi[k] \in \mathcal{U}_1\}. \end{aligned} \quad (14)$$

Since

$$Pr\{u[l+k] + \xi[k] \in \mathcal{U}_1\} = \begin{cases} 1/2 & \xi[k] = \pm 2 \\ 1 & \xi[k] = 0, \end{cases} \quad (15)$$

we get

$$m_\xi = 2^{-d_H(\xi)} \quad (16)$$

By noting that $d^2(\xi) = d^2(-\xi)$, $d_H(\xi) = d_H(-\xi)$ and $m_\xi = m_{-\xi}$, we can incorporate the event $-\xi$ in the multiplicity for ξ and only sum over those $\xi \in \mathcal{E}$ which have $\xi[0] = 2$. Then we get the final expression for m_ξ as

$$m_\xi = 2^{1-d_H(\xi)} \quad (17)$$

Define further notation as follows. Let $P_e(\xi)$ denote the term in (13) related to the error event ξ, i.e.

$$P(\xi) = d_H(\xi)m_\xi Q(\sqrt{d^2(\xi)\gamma_b}). \quad (18)$$

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

III. GOOD BER CODES

We now turn to PRS codes with better BER than the OPRS codes. The problem with OPRS is that the event $\xi_{\min}$ that leads to $d^2_{\min}$ does not always lead to $P^1_{\max}$. In fact, sometimes $P(\xi_{\min}) \ll P^1_{\max}$. In view of the linear program (12) this means that the k dominating terms in (13) could possibly have their distances $d^2(\xi_1), \ldots, d^2(\xi_k)$ increased at the expense of $d^2_{\min}$. So instead of having $\max_\xi \; d^2(\xi)$ as objective function we take the upper bound to BER (13) as our new objective function. This function is not linear so we have to give up linear programming as the solution method. It is difficult to work with (13) when many terms are present, therefore we limit the number of terms to N_t. Now (13) only estimates the bound.

We arrive at the optimization problem stated next, where ξ^b_k denotes the kth worst error event for the code generator $\mathbf{b}$:

$$\begin{array}{rcl} \mathbf{b}_{\text{opt}} & = & \arg\min_{\mathbf{b}} \sum_{k=1}^{N_t} d_H(\xi^b_k) m_{\xi^b_k} Q(\sqrt{d^2(\xi^b_k)\gamma_b}) \\ \text{s.t.} & & \int_{-\infty}^{\infty} |h_c(t)|^2 dt \;=\; 1 \\ & & \int_{-W}^{W} |H_c(f)|^2 df \;=\; C \end{array} \qquad (19)$$

We now go in two directions: (i) assume $N_t = 2$ and (ii) assume not. The first case is a rather easy problem to solve, and we do this next.

Assume that the upper bound is dominated by only two terms. Furthermore, assume these two error events ξ^b_1 and ξ^b_2 are the same for all $\mathbf{b}$ under examination. (If not, two new events can be determined as needed and the algorithm continued). For the initial ξ^b_1 and ξ^b_2 we choose the two worst events for the OPRS code with same δ and W and we denote them by ξ_1 and ξ_2 in the sequal.

The set of autocorrelation functions satisfying the bandwidth, energy and admissibility constraints in (12) is a closed polyhedral set, denoted $\mathcal{G}$.

Definition 1: Let r_1 and r_2 be the autocorrelations of ξ_1 and ξ_2. The linear operator $R : \mathcal{G} \to \mathbb{R}^2$ is defined as

$$R(g) = (r_1^{\text{tr}} g, r_2^{\text{tr}} g).$$

The image set of R is denoted by $\mathcal{D}$ and is called the set of achievable distances.

Lemma 1: $\mathcal{D}$ is a compact and convex set.

Proof Since $\mathcal{G}$ is closed so must $\mathcal{D}$ be. Furthermore, from (3) we see that $\mathcal{D}$ must be bounded. This shows that $\mathcal{D}$ is compact. Since $\mathcal{G}$ is a polyhedral set it is convex; this and the linearity of R implies that for $x_1, x_2 \in \mathcal{D}$ we have

$$\begin{array}{rcl} \lambda x_1 + (1-\lambda)x_2 & = & \lambda R(g_1) + (1-\lambda)R(g_2) \\ & = & R(\lambda g_1 + (1-\lambda) g_2) \\ & = & R(\tilde{g}),\; \tilde{g} \in \mathcal{G} \\ & \Rightarrow & \lambda x_1 + (1-\lambda)x_2 \in \mathcal{D} \blacksquare \end{array} \qquad (20)$$

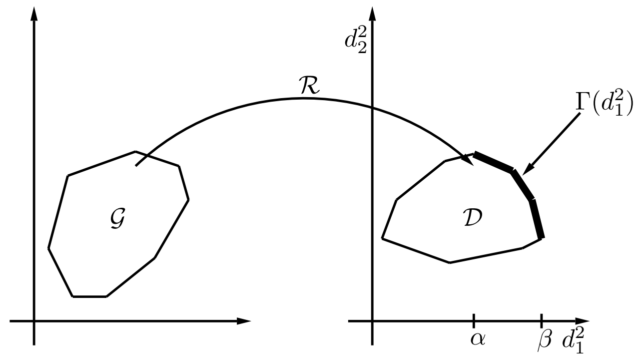

Fig. 2. The $\mathcal{G}$ and $\mathcal{D}$ domains and the linear operation R linking them. $\Gamma(d_1^2)$ is shown as well as the interval $[\alpha, \beta]$.

We can now express (19) as an optimization over $\mathcal{D}$ with the objective function

$$f(d_1^2, d_2^2) = C_1 Q(\sqrt{\gamma_b d_1^2}) + C_2 Q(\sqrt{\gamma_b d_2^2}), \qquad (21)$$

where $C_i = d_H(\xi_i) m_{\xi_i}$. Since f is continuous over $\mathcal{D}$ and $\mathcal{D}$ is compact we know that there exists a minimum of f on $\mathcal{D}$. This mimimum must occur somewhere on the northeast boundary of $\mathcal{D}$, see figure 2. At this boundary we can express d_2^2 as a function of d_1^2, i.e.

$$d_2^2 = \Gamma(d_1^2) \qquad (22)$$

The (interesting) values that d_1^2 can take can be found via linear search. The maximal value that d_1^2 can take is denoted by β and is computed by setting the objective function in (12) to $r_1^{\text{tr}} g$ and omitting the distance constraints. In a similar fashion the minimal interesting value, α, is computed by $\alpha = r_1^{\text{tr}} \tilde{g}$, where $\tilde{g}$ is the autocorrelation maximizing (12) when the objective function is replaced by $r_2^{\text{tr}} g$ and the distance constraints are omitted. Both α and β are indicated in figure 2.

If we define the function

$$y(x) = f(x, \Gamma(x)) = C_1 Q(\sqrt{\gamma_b x}) + C_2 Q(\sqrt{\gamma_b \Gamma(x)}), \qquad (23)$$

we can state the optimization problem in a compact way as

$$\min_{x \in [\alpha, \beta]} y(x). \qquad (24)$$

Lemma 2: $y(x)$ is a convex $\cup$ function on $[\alpha, \beta]$.

Proof If $y(x) \in \mathcal{C}^2([\alpha, \beta])$ then $y(x)$ is convex on $[\alpha, \beta]$ if and only if $y(x)'' \geq 0 \; \forall x \in [\alpha, \beta]$. Since $\mathcal{D}$ is a polyhedral set $\Gamma(x)$ has jumps in its derivative on the extrema points of $\mathcal{D}$. This can be avoided in the following way: assume that $\Gamma(x)$ has a corner point at x_0. Then connect the points $\Gamma(x_0 - \epsilon)$ and $\Gamma(x_0 + \epsilon)$ with a smooth arc. This does not affect the convexity of $\mathcal{D}$. If $\epsilon \to 0$ we do not affect $y(x)$ either. Now let $\tilde{Q}(x) = Q(\sqrt{\gamma_b x})$; then we have by the chain rule

$$y'(x) = \tilde{Q}'(x) + \Gamma'(x)\tilde{Q}'(\Gamma(x)), \qquad (25)$$

and

$$y''(x) = \tilde{Q}''(x) + \Gamma''(x)\tilde{Q}'(\Gamma(x)) + (\Gamma'(x)^2)\tilde{Q}''(\Gamma(x)). \qquad (26)$$

Since

$$\tilde{Q}''(x) > 0 \quad (27)$$
$$\tilde{Q}'(x) < 0 \quad (28)$$
$$\Gamma''(x) \leq 0 \ (\mathcal{D}\,\text{convex}), \quad (29)$$

we finally get

$$y''(x) > 0,\ x \in [\alpha, \beta]. \quad \blacksquare \quad (30)$$

Even if we now know that $y(x)$ is convex we cannot use convex programming since we do not have full control over $y(x)$ (this requires knowledge of $\Gamma(x)$ which we do not have a priori). Therefore we use the golden section search (see [10]) over the interval $[\alpha, \beta]$. This method suits the problem well since $y(x)$ is convex and will therefore only have one minimum on $[\alpha, \beta]$; it also requires a low number of function evaluations which is attractive since an evaluation implies solving a linear program to find $\Gamma(d_1^2)$. Numerical results are in section IV.

We now turn to the case where we do not have the simplification $N_t = 2$. We still assume the error events to be invariant, up to reordering of the terms, over $\mathbf{b}$. This is no restriction if N_t is large enough. The case $N_t > 2$ is a more general view and we expect the results to be better than for the scalar search. However, the optimization is much tougher.

Analogous to the scalar case we express the general problem (19) in the form (21) over the variables $d_1^2, \ldots, d_{N_t}^2$. Then lemma 1 and 2 still hold. Lemma 2 can be proved by studying the Hessian of f; it is easy to see that the Hessian is positive semidefinite which implies that f is convex on the interiour of $\mathcal{D}$. But since we know that the optimum lies at the boundary of $\mathcal{D}$ we must extend the convexity to also include the boundary. Construct two sequences $y_{1,n} \to y_1,\ n \to \infty$ and $y_{2,n} \to y_2,\ n \to \infty$, where y_1 and y_2 are arbitrary points on the boundary. Then we have

$$f(\lambda y_{1,n} + (1-\lambda) y_{2,n}) \leq \lambda f(y_{1,n}) + (1-\lambda) f(y_{2,n}). \quad (31)$$

By letting $n \to \infty$ and using the continuity of f we obtain

$$f(\lambda y_1 + (1-\lambda) y_2) \leq \lambda f(y_1) + (1-\lambda) f(y_2), \quad (32)$$

and we have proved that f is convex on the whole $\mathcal{D}$.

Many different approaches to solve the optimization problem are possible. Since the function is convex we could in principle use $N_t - 1$ nested golden section searches. The complexity of this optimization grows exponentially with N_t. Instead we use the cyclic coordinates method [10]. The only departure from an ordinary cyclic coordinates method is that the domain is bounded. Since the objective function is convex the optimization converges fast.

IV. Numerical Results

We study good codes for PRS length $\delta = 6$ and 8. These codes depend on the SNR γ_b, see (23), but we only list codes for a single typical γ_b and use these at all SNRs. We perform enough iterations of the golden section search to get a region of uncertainty smaller than .001. The main reason for the scalar search is that compared to the multi dimensional search, it is extremely efficient.

The results are listed in tables I–IV. We show tests of the new codes in figures 3–4. In the figures are also simulations of OPRS codes with same parameters. It can be seen that the BER optimal codes perform .2–.4 dB better than the OPRS codes. All tests were done using the M-algorithm [12]. The decision depth was set to 50. Note that the tested codes are only optimized for a single value of γ_b.

nbw	.38	.35	.32	.30	.28
$d^2_{\min}$	1.380	1.122	1.004	.8549	.6375
d^2_{OPRS}	1.397	1.170	1.053	.8170	.6542
γ_b	10	12	12	13	13
b[0]	.4714	.3689	.2644	.2626	.2139
b[1]	.7473	.7631	.6745	.6276	.5390
b[2]	.0512	.5176	.6461	.6635	.6663
b[3]	-.4368	-.0117	.1595	.2889	.4511
b[4]	-.0489	-.1158	-.1558	-.0628	.1251
b[5]	.1532	-.0093	-.0895	-.0975	-.0245

TABLE I

Best BER code generators with $\delta = 6$ found by the golden section search, γ_b in dB.

nbw	.30	.25	.20
γ_b	12	13	14
b[0]	.2620	.1722	.1065
b[1]	.6446	.4699	.2917
b[2]	.6512	.6484	.4783
b[3]	.2006	.5067	.5622
b[4]	-.1485	.1456	.4928
b[5]	-.0624	-.1355	.3145
b[6]	.1264	-.1695	.1288
b[7]	.098	-.0643	.0199

TABLE II

Best BER code generators with $\delta = 8$ found by the golden section search, γ_b in dB.

nbw	.42	.40	.35	.30
$d^2_{\min}$	1.70	1.50	1.112	.7949
d^2_{OPRS}	1.881	1.607	1.17	.8170
γ_b	10	10	12	13
b[0]	.5884	.5022	.3704	.2570
b[1]	.6270	.6831	.7631	.6255
b[2]	-.2821	-.1863	.5140	.6663
b[3]	-.1713	-.4401	-.0220	.2867
b[4]	.3633	.1428	-.1258	-.0752
b[5]	.1408	.1801	.0075	-.1040

TABLE III

Best BER code generators with $\delta = 6$ found by multidimensional search, γ_b in dB.

A different decoding approach is turbo equalization [11]. Turbo equalization is an iterative equalization and decoding technique that achieves impressive performance gains over ISI channels. Here we synthesize a coded modulation by including the ISI channel into the encoder. The PRS codes are concatenated with an interleaver and the (7,5) convolutional code, resulting in a scheme with twice the bandwidth consumption. The interleaver is MATLAB's `randintrlv` with blocksize 1024 information bits. Test results appear in figure 5. The

nbw	.40	.35	.33	.28	.20
$d^2_{\min}$	1.76	1.20	1.103	.797	.302
d^2_{OPRS}	1.793	1.255	1.162	.81	.322
γ_b	10	12	12	12	13
b[0]	.5224	.3897	.3227	.2163	.0941
b[1]	.6373	.7796	.7275	.5662	.2744
b[2]	-.2778	.4352	.5674	.6658	.4666
b[3]	-.4598	-.1365	.0214	.3331	.5624
b[4]	.1305	-.0934	-.1275	-.1039	.5052
b[5]	.0955	.1200	.0955	-.2388	.3285
b[6]	-.0156	-.0117	.1367	-.1027	.1358
b[7]	.0769	-.0955	.0163	.0051	.0176

TABLE IV

CODE GENERATORS WITH $\delta = 8$ FOUND BY MULTIDIMENSIONAL SEARCH.

coding gain over uncoded transmission is about 4 dB while at the same time the bandwidth falls from .7 to .56 Hz/bit/s for 40 % excess bandwidth ψ pulses.

V. CONCLUSIONS

The problem of finding optimal PRS codes for the AWGN channel has been attacked. Previously, Euclidean distance optimal codes were studied. This framework was extended and we showed that the BER objective function was convex. It turned out that a good code can be found by optimizing a real valued function of one variable. By optimizing over a multidimensional objective function better codes can be found. The optimal code depends on the signal to noise ratio. Tests show a .2–.4 dB gain in BER by using our modified generators.

REFERENCES

[1] S.A. Fredricsson, "Optimum transmitting filter in digital PAM systems with a Viterbi detector," *IEEE Trans. Information Theory*, IT–20, vol. 479–489, July 1974.

[2] G.J. Foschini, "Contrasting performance of faster binary signaling with QAM," *AT&T Bell Labs. Tech. J.*, vol. 63, pp 1419–1445, Oct. 1984.

[3] N. Chan, "Partial response signaling with a maximum likelihood sequence estimation receiver", M.Sc. thesis, Dept. Electrical and Computer Engin., McMaster Univ., Hamilton, Canada, Dec. 1980.

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

[5] 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.

[6] F. Rusek and J.B. Anderson, "Coded Optimal Response Signaling," *Proc.*, 2004 Intern. Symp. Information Theory, Chicago, p. 263, June 2004.

[7] 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.

[8] G.J. Foschini, "Performance bound for maximum-likelihood reception of digital data," *IEEE Trans. Information Theory*, vol. IT–21, pp 47–50, Jan. 1975.

[9] 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.

[10] R. Fletcher, *Practical Methods of Optimization*, 2nd ed., Wiley, New York, 2000.

[11] C. Douillard *et al.*, "Iterative correction of intersymbol interference: Turbo equalization," *Eur. Trans. Telecomm.*, vol. 6, pp. 507–511, Sept./Oct. 1995.

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

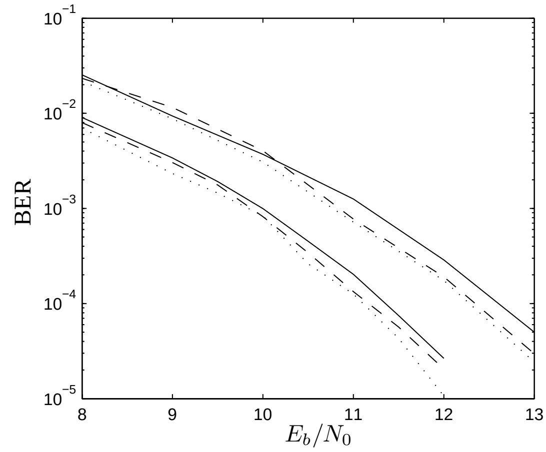

Fig. 3. Measured BER of 6 tap PRS-codes. Lower set of curves has nbw=.35 Hz/bit/s; upper has nbw=.30 Hz/bit/s. Solid curves refer to OPRS codes, dashed to codes optimized by scalar search, dotted to codes optimized by multidimensional search.

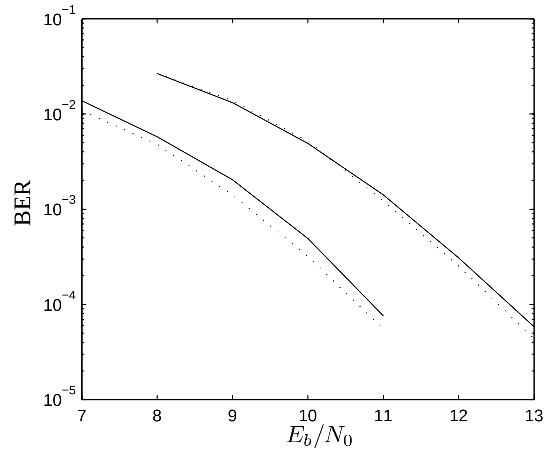

Fig. 4. BER of 8 tap PRS-codes. Lower set of curves has nbw=.35 Hz/bit/s; upper has nbw=.28 Hz/bit/s. Solid curves refer to OPRS codes, dotted to codes optimized by multidimensional search.

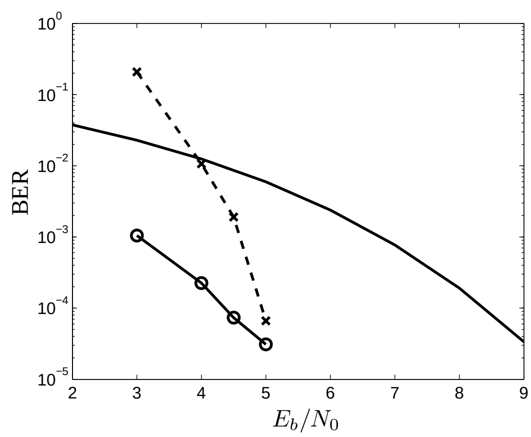

Fig. 5. Turbo equalization with 12 iterations for the (7,5) convolutional code concatenated with 6 tap BER optimal PRS code (solid) and 6 tap OPRS code (dashed), both designed for .28 Hz/bit/s. Uppermost curve is uncoded BPSK. Basic pulse is 40 % root RC pulse. $\mathrm{nbw} = .56$.