Nested Codes for Constrained Memory and for Dirty Paper

Hans Georg Schaathun[1] and Gérard D. Cohen[2]

[1] Dept. Informatics, University of Bergen,
Pb. 7800, N-5020 Bergen, Norway
[2] Dept. Informatique et Reseaux,
Ecole Nationale Supérieure des Télécommunications,
46, rue Barrault, F-75634 Paris Cedex 13, France

Abstract. Dirty paper coding are relevant for wireless networks, multiuser channels, and digital watermarking. We show that the problem of dirty paper is essentially equivalent to some classes of constrained memories, and we explore the binary so-called nested codes, which are used for efficient coding and error-correction on such channels and memories.

Keywords: dirty paper, constrained memory, nested codes, covering codes.

The motivation of this paper is the dirty paper channel introduced by Costa [3]. This channel has received increased attention [4] in recent years, due to applications in wireless multiuser networks and digital fingerprinting [5].

We show that the dirty paper channel is practically equivalent to writing on reluctant memories, and we make a few improvements on the existing results for such channels. Our interest is mainly in the binary dirty paper channel (BDP).

1 Dirty Paper and Constrained Memory Coding

The dirty paper channel is depicted in Figure 1. There are two independent noise sources which are added to the transmitted signal to form the received signal. The first noise vector, which we will call the *state* of the channel is known to the sender but not to the receiver. The second noise vector, which we will refer to as *noise* is unknown to both.

The sender is subject to a power constraint $||\mathbf{x}|| \leq P$ on the transmitted signal. For a binary channel $||\cdot||$ is usually the Hamming norm; for a continuous channel it is usually the Euclidean norm.

Costa [3] introduced this channel with Gaussian sources for both the state and the noise. His surprising result was that the channel capacity depends only on the intensity of the noise; the intensity of the state does not change capacity. In more recent years, his results have been generalised to other source distributions. We will consider the binary dirty paper channel (BDP).

M. Fossorier et al. (Eds.): AAECC 2006, LNCS 3857, pp. 307–316, 2006.

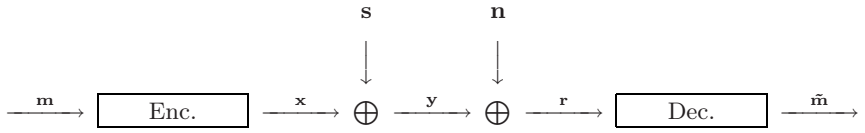

Fig. 1. The dirty paper channel

In a constrained memory, there are restrictions on writing to the memory, such that starting in one memory state, some states are reachable in one write operation and others are not. For each memory state, there is a feasible region of words which may be used to represent the next message. In this case the state is given by the previous message stored in memory.

Dirty paper coding and constrained memory coding are similar, in fact BDP channels are practically equivalent to WRM (write reluctant memories) with error-correction [2]. In WRM, one write operation cannot change more than a certain number P of bits. This corresponds to the power constraint in BDP; if $\mathbf{s}$ is the state (previous contents), $\mathbf{x}$ is the change, and $\mathbf{y} = \mathbf{s} + \mathbf{x}$ is the memory contents after writing, then $w(\mathbf{x}) \leq P$.

The state on dirty paper channels is externally given, whereas in constrained memories it is the old codeword (with possible errors). The state, together with power constraints, defines the feasible region of vectors $\mathbf{y}$ which can be generated. For BDP/WRM, the feasible region is a Hamming sphere around the state.

Remark 1. Occasionnally, in constrained memories, one assumes that $\mathbf{s}$ is a codeword with few errors, since nobody would write rubbish to the memory. We will not make this assumption, for two reasons. Primarily, it does not extend to BDP. Also, we know of no cases where results can be improved due to this assumption. Furthermore, by avoiding such assumption, the system can recover after an error pattern which could not be corrected.

Example 1. Another example of constrained memory is the Write Isolated Memory (WIM), where two consecutive memory bits cannot be changed in the same operation. In other words, the feasible region is the set $\{\mathbf{x}+\mathbf{s} : \mathbf{x} = (x_1, \ldots, x_n), x_i = 1 \Rightarrow x_{i-1} = x_{i+1} = 0\}$, where $\mathbf{s}$ is the memory state and $x_0 = x_{n+1} = 0$ by convention.

BDP (WRM) and WIM both fall into a class of channels, where the feasible regions are translation invariant, permitting some common techniques. By this we mean that if $F_{\mathbf{s}}$ is the feasible region from $\mathbf{s}$, then $F_{\mathbf{s}'} = F_{\mathbf{s}} - \mathbf{s} + \mathbf{s}'$. Let us call this class CCTIR (constrained channels with translation invariant regions).

2 Some Coding Theory

An $(n, M)_q$ code C is an M-set of n-tuples over a q-ary alphabet. When $q = 2$ we may suppress the subscript. The Hamming distance $d(\mathbf{x}, \mathbf{y})$ is the number of

positions where the two tuples differ. The minimum distance $d = d(C)$ of C is the least distance between two different codewords. We say that C is an $(n, M, d)_q$ code. The covering radius r of C is the largest distance between a vector $\mathbf{y} \in Q^n$ and the code.

The problem of covering codes amounts to finding codes minimising r given n and M, whereas the problem of error-correcting codes is about maximising d given n and M.

We also define normalised measures, which will be useful when moving to asymptotic codes. We define the rate $\log_q M/n$, the distance $\delta = d/n$, and the covering radius $\rho = r/n$.

3 Codes for CCTIR

In order to make a successful code for CCTIR, we need for every state $\mathbf{s}$ and every message $\mathbf{m}$, to have at least one codeword $\mathbf{x}$ corresponding to $\mathbf{m}$ in the feasible region of $\mathbf{s}$. Furthermore, we require any capability for error-correction that we may need. We will study e-error correcting CCTIR codes.

Lemma 1. *For CCTIR, if* $\mathbf{x} \in F_{\mathbf{y}}$ *then* $\mathbf{y} \in F_{\mathbf{x}}$.

Let B_i be the set of words corresponding to message i. We require that for any $\mathbf{s}$, $F_{\mathbf{s}} \cap B_i \neq \emptyset$. By the lemma above, this is equivalent to

$$\bigcup_{\mathbf{b} \in B_i} F_{\mathbf{b}} = \mathbb{F}^n, \tag{1}$$

i.e. that the feasible regions around the words of B_i cover the space. If the set of possible messages is $i = 1, \ldots, M$, then we define

$$C_F = \bigcup_{i=1}^{M} B_i.$$

When the feasible regions are spheres of radius ρ, this is to say that B_i must be a covering code of covering radius ρ or smaller. For other feasible regions it is a more general *covering by F-shapes*.

In order to correct up to e errors, we require that if $i \neq j$, then $d(B_i, B_j) > 2e$. It is sufficient to require that C_F has minimum distance $2e+1$ or more; i.e. that C_F is e-error correcting. Furthermore as a necessary condition, if there are two codewords with distance at most $2e$ apart, they must fall in the same set B_i.

In a sense, we try to pack the space with coverings B_i such that we maintain a minimum distance of $2e + 1$, a problem studied in [2].

We say that a CCTIR code $(B_1, \ldots, B_M)$ is linear if C_F is a linear e-error-correcting code, B_j is a subcode satisfying (1) for some j, and the B_i are cosets of B_j in C_F. Clearly by linearity, B_i satisfies (1) whenever B_j does.

Let $a_n = \#F_{\mathbf{0}}$. For CCTIR, all the feasible regions clearly have the same size.

Lemma 2. *For an (n, M) CCTIR code, we have*

$$M \leq a_n$$

Lemma 3. *For dirty paper codes, we have*

$$a_n = V(n, R) = \sum_{i=0}^{n} \binom{n}{i}.$$

In the case of WRM and dirty paper channels, a linear CCTIR code is also called a nested code. We call C_F the fine code and $C_C \subseteq C_F$ the coarse code. The nested code is the quotient $C = C_F/C_C$, and we say that C is an $[n, K; d_1, \rho]$ code, where $K = k_F - k_C$ is the dimension of C. The following lemma is well known.

Lemma 4 (Supercode lemma). *For any $[n, K; d_1, \rho]$ nested code, we have $\rho \geq d_1$.*

4 Asymptotic Existence

Definition 1 (Entropy). *The (binary) entropy of a discrete stochastic variable X drawn from a set $\mathcal{X}$ is defined as*

$$H(X) = -\sum_{x \in \mathcal{X}} P(X = x) \log P(X = x).$$

The conditional entropy of X with respect to another discrete stochastic variable Y from $\mathcal{Y}$ is

$$H(X|Y) = -\sum_{y \in \mathcal{Y}} P(Y = y) \sum_{x \in \mathcal{X}} P(X = x|Y = y) \log P(X = x|Y = y).$$

The following general theorem appears in [2].

Theorem 1. *For n large enough, there are θn-error correcting codes for CCTIR with rate*

$$\kappa(\theta) \geq \kappa_0 - H(2\theta),$$

where κ_0 is the maximum rate for a non-error-correcting code for the same constrained channel.

The proof is by greedy techniques, as follows.

Proof. We write

$$S(B, i) = \bigcup_{\mathbf{b} \in B} \{\mathbf{x} : d(\mathbf{x}, \mathbf{b} \leq i\}.$$

First we make a code C_C of rate $1-\kappa$ without error-correction. Let $S_0 = S(C_C, 2\theta n - 1)$.

We start with $B = \{\mathbf{0}\}$, and construct a code C_C by the following greedy algorithm. In each step we take a random vector $\mathbf{y} \in S \backslash S(B + C_C, 2\theta n - 1)$, and update B to be the linear span of $\mathbf{y}$ and the vectors of B. We proceed until $S \backslash S(B + C_C, 2\theta n - 1)$ is empty. Since each word included in B excludes at most $\#S(C_C, 2\theta n - 1)$ elements from S_0, we get that

$$\#B \geq \frac{2^n}{\#C_C \#S(\{\mathbf{0}\}, 2\theta n - 1)} \geq \frac{2^{\kappa n}}{\#S(\{\mathbf{0}\}, 2\theta n - 1)}.$$

Assymptotically, we have $\#B \approx 2^{(\kappa - H(2\theta))n}$, Let $C_F = B + C_C$, so that $C = C_F / C_C \equiv B$. Clearly the rate of B and C is $\kappa - H(2\theta)$ as required.

In the case of dirty paper channel, $\kappa_0 = 1 - R_\rho$ where R_ρ is the minimum rate for a covering code with appropriate ρ.

Theorem 2. *For dirty paper codes with no error-correction, we can obtain rate* $\kappa_0 = H(\rho)$.

Observe that whenever $\rho > \delta$, we get asymptotic codes with non-zero rate from the above theorems. For $\rho = \delta$, however, the guaranteed rate is just zero.

Problem 1. *Are there asymptotic families of nested codes with* $R > 0$ *and* $\rho = d$*?*

5 Some Small Constructions

Lemma 5. *For any* $[n, K; 1, 1]$ *nested code with even* n*, we have* $K \leq \log n$.

Proof. For a $[n, k_C]1$ covering code, we have $k_C \leq n - \log n$ when n is even, and for an $[n, k_F, 1]$ code, we have $k_F \leq n$. Hence $K = k_F - k_C \leq \log n$.

Lemma 6. *There is a* $[2^K - 1, K; 1, 1]$ *nested code for any* K.

Proof. Let the coarse code be the $[2^K - 1, 2^K - 1 - K, 3]1$ Hamming code, and let the fine code be the $[2^K - 1, 2^K - 1, 1]$ code.

Table 1. Some nested codes for $n = 3$

Parameters	Coarse code	Fine code
$[3, 1; 1, 1]$	$\begin{bmatrix}110\\101\end{bmatrix}$	$\begin{bmatrix}100\\010\\001\end{bmatrix}$
$[3, 1; 2, 2]$	$[110]$	$\begin{bmatrix}110\\011\end{bmatrix}$
$[3, 2; 1, 2]$	$[110]$	$\begin{bmatrix}100\\010\\001\end{bmatrix}$

Table 2. Some nested codes for $n = 4, 5, 6$

Parameters	Coarse code	Fine code
$[4, 2; 1, 1]$	$\begin{bmatrix}1110\\1001\end{bmatrix}$	$\begin{bmatrix}1000\\0100\\0010\\0001\end{bmatrix}$
$[4, 2; 2, 2]$	$\begin{bmatrix}1111\end{bmatrix}$	$\begin{bmatrix}1100\\1010\\1001\end{bmatrix}$
$[4, 3; 1, 2]$	$\begin{bmatrix}1111\end{bmatrix}$	$\begin{bmatrix}1000\\0100\\0010\\0001\end{bmatrix}$
$[5, 2; 1, 1]$	$\begin{bmatrix}11100\\10011\\00110\end{bmatrix}$	$[5, 5, 1]$ full code
$[5, 4; 1, 2]$	$\begin{bmatrix}11111\end{bmatrix}$	$[5, 5, 1]$ full code
$[5, 2; 2, 2]$	$\begin{bmatrix}11111\end{bmatrix}$	$\begin{bmatrix}11000\\10100\\10010\end{bmatrix}$
$[6, 2; 2, 2]$	$\begin{bmatrix}111000\\000111\end{bmatrix}$	$\begin{bmatrix}111000\\000111\\100100\\010010\end{bmatrix}$
$[6, 4; 1, 2]$	$\begin{bmatrix}111000\\000111\end{bmatrix}$	$[6, 6, 1]$ full code
$[6, 4; 2, 3]$	$\begin{bmatrix}111111\end{bmatrix}$	$[6, 5, 2]$ even weight

Lemma 7. *There are* $[2^K, K; 1, 1]$ *and* $[2^K - 1, K; 1, 1]$ *nested codes.*

Proof. The fine code is the $[n, n, 1]$ full space. The coarse codes are the Hamming codes, and the $[2^K, 2^K - K, 1]$ direct sum of a Hamming code padded with a zero column, and the code generated by a single word of weight one.

Lemma 8. *There are* $[2^K, K; 2^{K-1}, 2^{K-1}]$ *and* $[2^K - 1, K; 2^{K-1} - 1, 2^{K-1} - 1]$ *nested codes for any positive* K*.*

Proof. Let the coarse code be the $[2^K, 1, 2^K]$ repetition code, and let the fine code be the $[2^K, K, 2^{K-1}]$ Reed-Muller code. The second set of parameters comes from puncturing the above code.

Lemma 9. *If there is an* $[n, K; d, \rho]$ *code, then there is an* $[n - 1, K; d - 1, \rho]$ *code by puncturing and an* $[n - 1, K; d, \rho + 1]$ *code by shortening.*

Proof. This follows easily from the standard results on puncturing and shortening of error-correcting and covering codes.

Lemma 10 ([2]). *The* $[2^m - 1, 2^m - 1 - 2m, 5]$ *BCH code has* $\rho = 3$ *for* $m \geq 3$.

Table 3. Some nested codes for $n \geq 7$

Parameters	Coarse code	Fine code
$[7,3;1,1]$	$[7,4;3]$1 Hamming	$[7,7,1]$
$[7,3;3,3]$	$[7,1,7]$3 repetition	$[7,4,3]$ Hamming
$[8,3;1,1]$	$\begin{bmatrix}10001110\\01000110\\00101010\\00010110\\00000001\end{bmatrix}$	$[8,8,1]$
$[8,3;4,4]$	$[8,1,8]$4 repetition	$[8,4,4]$ ext. Hamming
$[15,4;3,3]$	$[15,7,5]$3 BCH(2)	$[15,11,3]$ Hamming
$[15,2;5,5]$	$[15,5,7]$5 BCH(3)	$[15,7,5]$3 BCH(2)
$[15,2;7,7]$	$[15,1,15]$7 repetition	$[15,3,7]$ BCH(3)
$[15,4;7,7]$	$[15,1,15]$7 repetition	$[15,5,7]$ punctured Reed-Muller
$[16,6;4,6]$	$[16,5,8]$6 RM$(1,4)$	$[16,11,4]$ RM$(2,4)$
$[16,4;8,8]$	$[16,1,16]$8 repetition	$[16,5,8]$ Reed-Muller
$[27,6;11,13]$	$[27,1,27]$13 repetition	$[27,7,11]$ [1]
$[28,6;12,14]$	$[28,1,28]$14 repetition	$[28,7,12]$ [1]
$[31,5;3,3]$	$[31,21,5]$3 BCH(2)	$[31,26,3]$ Hamming
$[31,5;5,5]$	$[31,16,7]$5 BCH(3)	$[31,21,5]$3 BCH(2)
$[31,5;7,7]$	$[31,11,11]$7 BCH(4)	$[31,16,7]$5 BCH(3)
$[31,5;11,11]$	$[31,6,15]$11 BCH(6)	$[31,11,11]$7 BCH(4)
$[31,5;15,15]$	$[31,1,31]$15 repetition	$[31,6,15]$ punctured Reed-Muller
$[32,5;2,2]$	$[32,26,4]$2 RM$(3,5)$	$[32,31,2]$ RM$(4,5)$
$[32,10;4,6]$	$[32,16,8]$6 RM$(2,5)$	$[32,26,4]$ RM$(3,5)$
$[32,10;8,12]$	$[32,6,16]$12 RM$(1,5)$	$[32,16,8]$ RM$(2,5)$
$[36,20;4,13]$	$[36,8,16]\rho$ $\rho \leq 13$	$[36,28,4]$ $C_C^{\perp}$
$[49,9;20,24]$	$[49,1,49]$24 repetition	$[49,10,20]$ [1]
$[63,6;1,1]$	$[63,57,3]$1 BCH(1)	$[63,63,1]$ full code
$[63,6;3,3]$	$[63,51,5]$3 BCH(2)	$[63,57,3]$1 BCH(1)
$[63,6;5,5]$	$[63,45,7]$5 BCH(3)	$[63,51,5]$3 BCH(2)
$[63,6;7,7]$	$[63,39,9]$7 BCH(4)	$[63,45,7]$5 BCH(3)
$[63,3;9,9]$	$[63,36,11]$9 BCH(5)	$[63,39,9]$7 BCH(4)
$[64,15;4,8]$	$(u,u+v)$ construction	
$[64,15;16,28]$	$[64,7,32]$28 RM$(1,6)$	$[64,22,16]$ RM$(2,6)$

Corollary 1. *There is a* $[2^m-1,m;3,3]$ *nested code for every* $m \geq 3$.

Proof. The coarse code is the $[2^m-1,2^m-1-2m,5]$3 BCH(2) code, and the fine code is the Hamming code.

Lemma 11 ([2]). *The* $[2^m-1,2^m-1-3m,7]$ *BCH code has* $\rho = 5$ *for* $m \geq 4$.

Corollary 2. *There is a* $[2^m-1,m;5,5]$ *nested code for every* $m \geq 4$.

Proof. The coarse code is the $[2^m-1, 2^m-1-3m, 7]5$ BCH(3) code, and the fine code is the $[2^m-1, 2^m-1-2m, 5]3$ BCH(2) code.

Lemma 12. *There are* $[2^{2m+1}-2^m, 2m+2; 2^{2m}-2^m, 2^{2m}]$ *and* $[2^{2m+1}-2^m-1, 2m+2; 2^{2m}-2^m-1, 2^{2m}-1]$ *nested codes.*

Proof. The coarse code is a repetition code. The fine code is a $[2^{2m+1}-2^m, 2m+3, 2^{2m}-2^m]$ code [1] or a punctured version of it.

Lemma 13. *There is no* $[6,4;2,2]$ *nested code, so the* $[6,3;2,2]$, $[6,4;1,2]$ *and* $[6,4;2,3]$ *codes are optimal.*

Proof. The smallest covering code of $\rho = 2$ and $n = 6$ has $k_C = 2$, so to get $K = 4$, we would need $k_F \geq 6$, which would give $d = 1$.

6 Some Upper Bounds on the Nested Code Dimension

Lemma 14. *For an* $[n, K; d, d]$ *nested code, we have*

$$2^K \leq \binom{n}{d} + 1.$$

Proof. Consider the points of C_C and the balls of radius $\rho = d$ around these points. Because ρ is the covering radius of C_C, these balls cover the space. Since C_F has minimum distance $d = \rho$, it can only contain points on the border of these balls, besides the points of C_C. Hence

$$\#C_F \leq \#C_C \cdot \left(\binom{n}{d} + 1 \right),$$

and hence

$$\#(C_F/C_C) \leq \left(\binom{n}{d} + 1 \right),$$

as required.

We have seen that this bound can be met with equality for $\rho = 1$. For $\rho > 1$ except $\rho = n = 2$ we have inequality; let's see this for $\rho = 2$ first.

Proposition 1. *For an* $[n, K; 2, 2]$ *nested code with* $n > 2$*, we have*

$$2^K < \binom{n}{2} + 1.$$

Proof. Suppose the bound were met with equality. Since C_C is a covering code of $\rho = 2$, we have

$$2^n \leq 2^{k_C} \left(1 + \binom{n}{1} + \binom{n}{2} \right) \leq 2^{k_C}(2^K + n) \leq 2^{k_F} + 2^{k_C} n.$$

For all $n > 2$, we have

$$n < 1 + \binom{n}{2},$$

which is equal to 2^K by assumption. This gives

$$2^n < 2^{k_F+1},$$

and clearly $n \geq k_F$, so we get $n = k_F$; but then $d = 1 < 2$, giving a contradiction.

We do have degenerate $[n, 1; n, n]$ nested codes for all n. They have only the zero word for C_C, an $[n, 1, n]$ repetition code for C_F.

Proposition 2. *For an $[n, K; d, d]$ nested code, we have*

$$2^K \leq A(n, d, d) + 1.$$

It is readily seen that this bound is stronger than Lemma 14 when $\rho > 2$.

Proof. We start as we did proving Lemma 14 with the balls of radius ρ around the points of C_C. The border of the ball around $\mathbf{x}$ are the points $\mathbf{x}+\mathbf{y}$ where $\mathbf{y}$ has weight ρ. Obeying the distance requirement, the $\mathbf{y}$ that we choose for C_F from this ball, will have to form a constant weight code of weight and distance $\rho = d$.

Generalising, we get the following proposition, for which we ommit the proof.

Proposition 3. *For an $[n, K; d, \rho]$ nested code, we have*

$$2^K \leq 1 + \sum_{w=d}^{\rho} A(n, d, w).$$

7 Some Constructions

Theorem 3. *Let $U = U_F/U_C$ and $V = V_F/V_C$ be $[n, K_U; d_U, \rho_U]$ and $[n, K_V; d_V, \rho_V]$ nested codes. Let $U_i \circ V_i$ denote the $(\mathbf{u}, \mathbf{u} + \mathbf{v})$ composition of U_I and U_V. Then we can form a nested code $C = U \circ V = (U_F \circ V_F)/(U_C \circ U_F)$, and C is a $[2n, K_U + K_V; d, \rho]$ nested code with $\rho \leq \rho_U + \rho_V$ and $d = \min\{2d_V, d_U\}$.*

The proof is obvious from fundamental results on the parameters of the component codes.

Acknowledgements

The authors are grateful for codes contributed by Carl Bracken of Dublin, and for the interest of and discussions with Wolfgang Willems of Magdeburg.

References

1. Carl Bracken. Private communication. 2004.
2. Gérard Cohen, Iiro Honkala, Simon Litsyn, and Antoine Lobstein. *Covering codes*, volume 54 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, 1997.
3. Max H. M. Costa. Writing on dirty paper. *IEEE Trans. Inform. Theory*, 29(3):439–441, 1983.
4. Frank Kschischang and David Tse, editors. *Proc. IEEE Intern. Symp. Inform. Theory*, pages 533–536. June 2004. A four-talk session on dirty paper coding.
5. M.L. Miller, G.J. Doerr, and I.J. Cox. Applying informed coding and embedding to design a robust high-capacity watermark. *IEEE Transactions on Image Processing*, 13(6):792–807, June 2004.