BOUNDS ON THE RATE OF CODES WHICH FORBID SPECIFIED DIFFERENCE SEQUENCES

Bruce E. Moision, Alon Orlitsky, and Paul H. Siegel, *Fellow, IEEE*

Department of Electrical and Computer Engineering
University of California at San Diego
La Jolla, California 92093-0407

Abstract— **Certain magnetic recording applications call for a large number of sequences whose differences do not include certain disallowed patterns. We show that the number of such sequences increases exponentially with their length and that the exponent, or capacity, is the logarithm of the joint spectral radius of an appropriately defined set of matrices. We derive new algorithms for determining the joint spectral radius of sets of nonnegative matrices and combine them with existing algorithms to determine the capacity of several sets of disallowed differences that arise in practice.**

I. Introduction

The error probability of many magnetic-recording systems may be characterized in terms of the differences between the sequences that may be recorded [1], [2], [3]. In fact, the bit-error-rate is often dominated by a small set of potential difference patterns. Recently, binary codes have been proposed which exploit this fact [4], [5], [6], [7], [8]. The codes are designed to avoid the most problematic difference patterns by constraining the set of allowed recorded sequences and have been shown to improve system performance.

In this paper we study the largest number of sequences whose differences exclude a given set of disallowed patterns. We show that the number of such sequences increases exponentially with their length and that the exponent, or capacity, is the logarithm of the joint spectral radius of an appropriately defined set of matrices. We derive new algorithms for determining the joint spectral radius of sets of non-negative matrices and combine them with existing algorithms to determine the capacity of several sets of disallowed differences that arise in practice.

The paper is organized as follows. In the next section we summarize known results showing that the error probability is determined by the differences between recorded sequences. In Section III we formally describe the resulting combinatorial problem and the notation used. Section IV derives the connection to the joint spectral radius. In Section V we describe some known algorithms for determining the spectral radius, and derive some new ones. Finally, in Section VI, we list the capacities of some simple sets of disallowed patterns. Due to space limitations, many proofs are omitted. Details, however, may be found in [9].

This work was supported in part by the UC Micro Project 98-140 in cooperation with Marvell Semiconductor Incorporated and NSF grant NCR-9612802.

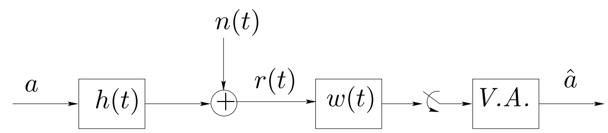

Fig. 1. Communications channel model

II. Motivation

Consider the binary communications channel in Figure 1 where a binary sequence $a = (\ldots, a_0, a_1, a_2, \ldots)$ passes through a linear channel with impulse response $h(t)$ and $n(t)$ is additive white Gaussian noise. The received signal is given by

$$r(t) = \sum_k a_k h(t - kT) + n(t),$$

where T is the bit-period. The receiving filter $w(t)$ is chosen such that the signal at the input to the Viterbi detector in the absence of noise approximates Xa, where X is the Toeplitz matrix corresponding to a finite target response. The Viterbi algorithm is then used to obtain the sequence $X\hat{a}$ closest in Euclidean distance to the sequence received at the input to the Viterbi detector.

At high signal-to-noise ratios, the probability of a bit error for this estimate is well approximated by

$$Pr(\text{bit error}) \approx \sum_{a-\hat{a} \in D} w(a - \hat{a}) Q\left(\frac{1}{2} d_{\mathit{eff}}(a - \hat{a})^2\right)$$

e.g. [10], where $w(a - \hat{a}) = Pr(a, \hat{a}) \sum |a_i - \hat{a}_i|$ is a weighting factor, $Q(\cdot)$ is the error function,

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2} dt,$$

$d_{\mathit{eff}}(a - \hat{a})$ denotes the *effective distance*[10] between the sequences $Xa, X\hat{a}$, and D is a set of *dominant* difference patterns, i.e. those patterns with small effective distance. At high signal-to-noise ratios, the performance of the sequence estimator is largely determined by a small set of such difference patterns.

The fact that a small set of difference patterns dominate the system performance has motivated the construction of codes designed to avoid the occurrence of the low-distance difference patterns [4]. The subsequent increase in the minimum effective distance is, however, offset by rate-loss of the codes. This led to the following question, which we address in this paper: what is the bound on the rate of a code which avoids a specified set of difference patterns?

III. Notation and definitions

The *difference* between two n-bit sequences $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n)$ is the sequence $u - v \stackrel{\text{def}}{=} (u_1 - v_1, \ldots, u_n - v_n) \in \{-1, 0, 1\}^n$ where subtraction is over the reals.

Given a set $D \subseteq \{-1, 0, 1\}^*$ of disallowed difference patterns and a sequence length n, we are interested in the largest number of n-bit sequences whose differences do not include any element of D. We are primarily interested in finite difference sets. Without loss of generality we therefore assume from here on that all patterns in D have the same length m.

An n-bit *code* $\mathcal{C}$ is a collection of n-bit sequences, or *codewords*, thought of as potential recorded sequences. $\mathcal{C}$ *avoids* D if for all $u, v \in \mathcal{C}$ and all $i \in [1, n']$,

$$u_{[i,i']} - v_{[i,i']} \notin D \tag{1}$$

where, for all $i \leq j$, we use the notation

$$[i, j] \stackrel{\text{def}}{=} \{i, \ldots, j\}$$

and

$$u_{[i,j]} \stackrel{\text{def}}{=} u_i, \ldots, u_j,$$

and, for i and n only, we let $i' \stackrel{\text{def}}{=} i + m - 1$ and $n' \stackrel{\text{def}}{=} n - m + 1$.

The largest number of sequences whose differences do not include any pattern in D is therefore

$$\delta_n(D) \stackrel{\text{def}}{=} \max\{|\mathcal{C}| : \mathcal{C} \text{ avoids } D\}.$$

We define the *capacity* of D as the limit

$$\mathsf{cap}(D) \stackrel{\text{def}}{=} \log \left[\lim_{n\to\infty} (\delta_n(D))^{1/n} \right]. \tag{2}$$

We would like to determine the capacities of various difference sets D and find codes that achieve them.

IV. From disallowed differences to joint spectral radius

A. Disallowed joint patterns

Represent an n-bit code $\mathcal{C}$ as an $|\mathcal{C}|$ by n array and for $i \in [1, n']$ let

$$M_i \stackrel{\text{def}}{=} \{0, 1\}^m - \{u_{[i,i']} : u \in \mathcal{C}\} \tag{3}$$

be the set of m-bit patterns *missing* from columns $[i, i']$.

A *joint pattern* is a set of two m-bit patterns. A joint pattern $\{p, p'\}$ is *disallowed* for a difference set D if

$$p - p' \in D \quad \text{or} \quad p' - p \in D.$$

Let $\mathcal{J}(D)$ denote the collection of all disallowed joint patterns.

B. Disallowed sets

A set $M \subseteq \{0, 1\}^m$ is a *representing set for* $\mathcal{J}(D)$ if it intersects every set in $\mathcal{J}(D)$ and no subset intersects every set in $\mathcal{J}(D)$. Let $\mathcal{M}(D)$ be the collection of all representing sets for $\mathcal{J}(D)$.

Let $M_1, \ldots, M_{n'} \subseteq \{0, 1\}^m$ be sets of m-bit patterns. An n-bit sequence $s_1, \ldots, s_n$ *avoids* the set sequence $M_1, \ldots, M_{n'}$ if $s_{[i,i']} \notin M_i$ for all $i \in [1, n']$. Let $\mu(M_1, \ldots, M_{n'})$ be the number of n-bit sequences that avoid $M_1, \ldots, M_{n'}$.

If $\mathcal{M}$ is a collection of sets in $\{0, 1\}^m$, we let

$$\mu_n(\mathcal{M}) \stackrel{\text{def}}{=} \max\{\mu(M_1, \ldots, M_{n'}) : M_i \in \mathcal{M}\ \forall i\}$$

be the largest number of n-bit sequences all avoiding a single sequence of sets in $\mathcal{M}$. The following Lemma converts the problem of finding $\delta_n(D)$ from a constraint on pairs of sequences to a constraint on individual sequences.

Lemma 1: For every n,

$$\delta_n(D) = \mu_n(\mathcal{M}(D)). \qquad \square$$

C. Bipartite and cascade graphs

In the previous subsection we reduced the difference constraint on pairs of sequences to a constraint on individual sequences. We now convert this problem to that of counting paths in graphs.

A bipartite graph (L, R, E) consists of a set L of *left vertices*, a set R of *right vertices*, and a set E of *edges*. Each edge $(l, r) \in E$ connects a left vertex $l \in L$ to a right vertex $r \in R$.

For $m \geq 2$ let G_m be the bipartite graph where $L = R = \{0, 1\}^{m-1}$ and $(l_1, \ldots, l_{m-1}) \in L$ is connected to $(r_1, \ldots, r_{m-1}) \in R$ if $l_i = r_{i-1}$ for all

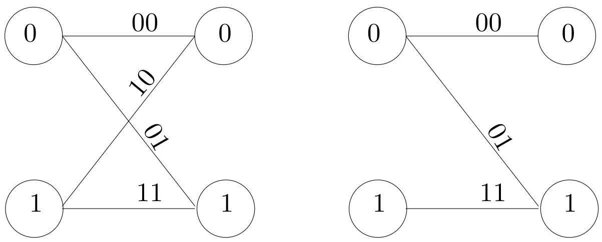

Fig. 2. G_2 and $G_{\{10\}}$

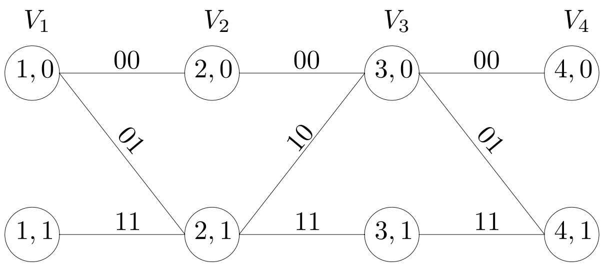

Fig. 3. $[G_{\{10\}}, G_{\{01\}}, G_{\{10\}}]$

$i = 2, \ldots, m-1$. We identify this edge with the m-bit sequence $l_1, l_2, \ldots, l_{m-1}, r_{m-1} = l_1, r_1, \ldots, r_{m-1}$.

For $M \subseteq \{0,1\}^m$, define G_M to be the bipartite graph obtained from G_m by removing the edges corresponding to elements of M. Figure 2 illustrates G_2 and $G_{\{10\}}$.

If $G_1, \ldots, G_{n'}$ are bipartite graphs with left vertex sets $L_1, \ldots, L_{n'}$ and right vertex sets $R_1, \ldots, R_{n'}$, respectively, such that $R_i = L_{i+1}$ for all $i \in [1, n'-1]$, we let

$$V_i \stackrel{\text{def}}{=} \begin{cases} \{1\} \times L_1 & \text{if } i = 1, \\ \{i\} \times R_{i-1} = \{i\} \times L_i & \text{if } 2 \le i \le n', \\ \{n'+1\} \times R_{n'} & \text{if } i = n'+1, \end{cases}$$

and define the *cascade* $[G_1, \ldots, G_{n'}]$ to be the graph whose vertex set is $V_1 \cup \ldots \cup V_{n'+1}$ and where for $i \in [1, n']$, the edges between V_i and V_{i+1} are the edges of G_i, and there are no other edges. Figure 3 illustrates the cascade $[G_{\{10\}}, G_{\{01\}}, G_{\{10\}}]$.

A *path* in a cascade $[G_1, \ldots, G_{n'}]$ is a sequence $v_1, \ldots, v_{n'+1}$ of vertices where each $v_i \in V_i$, and v_i is connected to v_{i+1} for all $i \in [1, n']$. We let $\psi([G_1, \ldots, G_{n'}])$ be the total number of paths in the cascade.

For $n \ge m$, there is a bijection between n-bit sequences that avoid $M_1, \ldots, M_{n'}$ and paths in the cascade $[G_{M_1}, \ldots, G_{M_{n'}}]$, hence, letting

$$\psi_n(\mathcal{M}) \stackrel{\text{def}}{=} \max\{\psi([G_{M_1}, \ldots, G_{M_{n'}}]) : M_i \in \mathcal{M}\ \forall i\}$$

we obtain

Lemma 2:

$$\mu_n(\mathcal{M}(D)) = \psi_n(\mathcal{M}(D)). \qquad \square$$

D. Adjacency matrices

Let the *adjacency matrix* A_G of $G = (L, R, E)$ be the $|L| \times |R|$ matrix whose (l, r)th element is 1 if $(l, r) \in E$, and 0 otherwise. The number of left-to-right paths from leftmost vertex l to rightmost vertex r in the cascade $[G_1, \ldots, G_{n'}]$ is the (l, r)-th element of the product $A_{G_1} A_{G_2} \cdot \ldots \cdot A_{G_{n'}}$.

Letting

$$\|A\|_1 = \sum_{l,r} |A_{l,r}| \tag{4}$$

denote the L_1 norm of the matrix A, it follows that for every $M_1, \ldots, M_{n'} \subseteq \{0,1\}^m$

$$\psi([G_{M_1}, \ldots, G_{M_{n'}}]) = \|A_{G_{M_1}} \cdot \ldots \cdot A_{G_{M_{n'}}}\|_1.$$

Putting

$$\Sigma(D) \stackrel{\text{def}}{=} \{A_{G_M} : M \in \mathcal{M}(D)\}$$

and

$$\hat{\rho}_n(\Sigma, \|\cdot\|_1) \stackrel{\text{def}}{=} \max\left\{\left\|\prod_{i=1}^n A_i\right\|_1 : A_i \in \Sigma\ \forall i\right\}. \tag{5}$$

we get

Lemma 3:

$$\psi_n(\mathcal{M}(D)) = \hat{\rho}_{n'}(\Sigma(D), \|\cdot\|_1). \qquad \square$$

This suggests looking for algebraic methods to determine the capacity.

E. Joint spectral radius

The *spectral radius* of a matrix $A \in \mathbb{C}^{m \times m}$ is the nonnegative real number

$$\hat{\rho}(A) \stackrel{\text{def}}{=} \lim_{n \to \infty} \|A^n\|^{1/n},$$

The quantity $\hat{\rho}$ can be generalized to sets of matrices. Letting

$$\hat{\rho}_n(\Sigma, \|\cdot\|) \stackrel{\text{def}}{=} \sup\left\{\left\|\prod_{i=1}^n A_i\right\| : A_i \in \Sigma\ \forall i\right\}$$

for an arbitrary matrix norm $\|\cdot\|$, Rota and Strang [11] defined the *joint spectral radius* of $\Sigma \subseteq \mathbb{C}^{m \times m}$ to be

$$\hat{\rho}(\Sigma) \stackrel{\text{def}}{=} \lim_{n \to \infty} \hat{\rho}_n(\Sigma, \|\cdot\|)^{1/n}$$

Combining Definition (2) and Lemmas 1 to 3, we obtain our main result:

Theorem 1: For every finite D,

$$\mathsf{cap}(D) = \log(\hat{\rho}(\Sigma(D))).$$

$\square$.

This equality generalizes known results on *constrained systems* where, instead of differences, certain patterns are disallowed, and it is well known, e.g., [12, Theorem 4.4.4], that the growth rate of the number of sequences, or *Shannon capacity* of the constraint, is $\log(\hat{\rho}(A))$, the logarithm of the spectral radius of a corresponding adjacency matrix A.

V. Computing the joint spectral radius

A. Existing algorithms

Daubechies and Lagarias [13] defined the *generalized spectral radius* of Σ to be

$$\check{\rho}(\Sigma) \stackrel{\text{def}}{=} \limsup_{n\to\infty}\{\check{\rho}_n(\Sigma)^{1/n}\}$$

where

$$\check{\rho}_n(\Sigma) \stackrel{\text{def}}{=} \max\left\{\hat{\rho}\left(\prod_{i=1}^{n} A_i\right) : A_i \in \Sigma\ \forall i\right\}.$$

They showed that for every Σ,

$$\sup_{n\geq 1}\check{\rho}_n(\Sigma)^{1/n} \leq \check{\rho}(\Sigma) \leq \hat{\rho}(\Sigma) \leq \inf_{n\geq 1}\hat{\rho}_n(\Sigma, \|\cdot\|)^{1/n}$$

and conjectured that all inequalities hold with equality, which was proven by Berger and Wang [14] for all finite Σ. This suggests approximating the joint spectral radius by computing the lower bounds $max_{1\leq k\leq n}\check{\rho}_k(\Sigma)^{1/k}$ and upper bounds $min_{1\leq k\leq n}\hat{\rho}_k(\Sigma, \|\cdot\|)^{1/k}$ for $n = 1, 2, \ldots$. However, the number of matrix operations increases as $|\Sigma|^n$.

Several steps have been taken to reduce the growth rate of the number of computations required to approximate $\hat{\rho}(\Sigma)$. Daubechies and Lagarias [13] proposed a recursive algorithm to upper bound $\hat{\rho}(\Sigma)$, examples of which may be found in [15], [16], [17].

The remainder of this section describes an algorithm which empirical results show has a computation time competitive with the algorithm in [17].

B. The pruning algorithm

The pruning algorithm may be used to bound $\hat{\rho}(\Sigma)$ when all the matrices in Σ are non-negative. The method replaces the search for the largest norm among all (exponentially many) products of n matrices with a search over a smaller set with the same largest norm.

We write $A \geq 0$ if every element of A is nonnegative and $A \geq B$ if every element of A is at least as large as the corresponding element of B.

A matrix A *dominates* matrix B with respect to the norm $\|\cdot\|$ if

$$\|AM\| \geq \|BM\|$$

for all $M \geq 0$. Let

$$\Sigma^n \stackrel{\text{def}}{=} \left\{\prod_{i=1}^{n} : A_i \in \Sigma\right\}$$

denote the set of products of n matrices in Σ. A subset S of Σ^n is *dominating* if every matrix in Σ^n is dominated by some matrix in S. Let Ψ_n be any dominating subset of Σ^n. By definition,

$$\hat{\rho}_n(\Sigma, \|\cdot\|) = \max\{\|A\| : A \in \Psi_n\}. \qquad (6)$$

Furthermore, it is easy to verify that if all matrices in Σ are non-negative then $\Psi_n\Sigma$ is a dominating subset of Σ^{n+1}.

Given a matrix norm one can construct a recursive algorithm which computes a dominant set Ψ_n from Ψ_{n-1} by 'pruning' those products in $\Psi_{n-1}\Sigma$ which are dominanted by another product. The subsequent growth rate of $|\Psi_n|$ will depend on the condition for domination. For example, the following Lemma provides a sufficient condition for domination with respect to the *spectral norm* $\|\cdot\|_s$.

Lemma 4: If $A^*A \geq B^*B$, then A dominates B with respect to $\|\cdot\|_s$. $\square$

An analogous algorithm may be used to construct a sequence of convergent lower bounds on $\check{\rho}(\Sigma)$.

VI. Capacities of certain difference sets

Table I summarizes known values or ranges of $\mathsf{cap}(D)$ for all difference sets D consisting of a single pattern of length ≤ 3 and some patterns of larger length. Since the same number of sequences avoid a pattern p as its negation $-p$, we assume that the first nonzero element of p is $+$. Also, we do not list $\mathsf{cap}(D)$ if the (identical) capacity of the string obtained by reversing the order of p has already been addressed.

Next to the capacity, we list a constraint describing a sequence of codes, $\{\mathcal{C}_n\}$, such that each $\mathcal{C}_n$ is an n-bit code which avoids D and

$$\lim_{n\to\infty} \log|\mathcal{C}_n|^{1/n}$$

achieves the lower bound on the capacity, or $\mathsf{cap}(D)$ when it is known. In a notation similar to that used to describe shift spaces [12, Defn. 1.2.1], the constraint is defined by a list of forbidden patterns $\mathcal{O}$ and the codes $\mathcal{C}_n$ can be taken to be the largest n-bit codes satisfying the constraint. If no superscript is listed with a pattern, the pattern is forbidden from appearing in all columns of the code. If superscripts appear, then the patterns are periodic and the period is one more than the largest superscript. The superscript then represents the column indices (modulo the period) in which the pattern is disallowed. For example, 101, 010 means that these triples do not appear in any three consecutive columns, and $10^{(0)}$, $01^{(1)}$ means that 10 does not appear in columns $[i, i+1]$ for even i and 01 does not appear in columns $[i, i+1]$ for odd i.

Several of these constraints have appeared in the magnetic recording literature. $\mathcal{O} = \{00^{(1)}, 11^{(1)}\}$ is referred to as the *biphase* constraint [18], $\mathcal{O} = \{1010, 0101\}$ as the *MTR* constraint [5], and $\mathcal{O} = \{1010^{(1)}, 0101^{(1)}\}$ as the *TMTR* constraint [6].

m	D	$\mathsf{cap}(D)$	$\mathcal{O}$
$m \geq 1$	$0^{m-1}+$	0	$-$
2	$+-$	α	$10^{(0)}, 01^{(1)}$
	$++$	α	11
3	$0+-$	$[\alpha, .6948)$	$101, 010$
	$0+0$	.5	$00^{(1)}, 11^{(1)}$
	$0++$	$[\alpha, .6948)$	11
	$+0-$	α	$110^{(0)}, 001^{(0)}, 011^{(1)}, 100^{(1)}$
	$+0+$	α	$101, 111$
	$++-$	γ	$110^{(0)}, 110^{(1)}, 001^{(2)}, 001^{(3)}$
	$+++$	δ	111
	$+-+$	δ	$101^{(0)}, 010^{(1)}$
4	$0+-+$	$[\delta, .8797)$	$1010, 0101$
	$+-+-$	$[\eta, .9468)$	$1010^{(0)}, 0101^{(1)}$
5	$0+-+0$	$[\epsilon, .9165)$	$10101^{(1)}, 01010^{(1)}$

TABLE I
CAPACITY OF VARIOUS DIFFERENCE SETS D.

$$\alpha = \log_2((1+\sqrt{5})/2) = .6942...$$

$$\gamma = \log_2((2 + (\frac{25-3\sqrt{69}}{2})^{1/3} + (\frac{25+3\sqrt{69}}{2})^{1/3})/3) = .8113...$$

$$\delta = \log_2((1 + (19+3\sqrt{33})^{1/3} + (19-3\sqrt{33})^{1/3})/3) = .8791...,$$

$$\epsilon = \log_2(\sqrt{(3+\sqrt{17})/2}) = .9162...,$$

$$\eta = \log_2((3 + \sqrt{3\zeta} + \sqrt{99 - 3\zeta + 234\sqrt{3/\zeta}})/12) = .9467...,$$

WHERE $\zeta = 11 - 56\beta + 4/\beta$, AND $\beta = (2/(-65+3\sqrt{1689}))^{1/3}$.

References

[1] R. Karabed and N. Nazari, "Analysis of error sequences for PRML and EPRML signaling performed over Lorentzian channel," in *Proceedings IEEE Global Telecommunications Conference*, vol. 1, (London, UK), pp. 368–373, IEEE, Nov. 1996.

[2] S. A. Altekar, *Detection and coding techniques for magnetic recording channels*. PhD thesis, University of California, San Diego, Aug. 1997.

[3] B. E. Moision, P. Siegel, and E. Soljanin, "Error event characterization and coding for the equalized Lorentzian channel," in *Proceedings of IEEE International Symposium on Information Theory*, (Cambridge, MA, USA), p. 77, Aug. 1998.

[4] R. Karabed and P. H. Siegel, "Coding for higher order partial response channels," in *Proceedings of the SPIE* (M. R. Raghuveer, S. A. Dianat, S. W. McLaughlin, and M. Hassner, eds.), vol. 2605, (Philadelphia, PA, USA), pp. 92–102, Oct. 1995.

[5] J. Moon and B. Brickner, "Maximum transition run codes for data storage systems," *IEEE Transactions on Magnetics*, vol. 32, pp. 3992–3994, Sept. 1996.

[6] W. Bliss, "An 8/9 rate time-varying trellis code for high density magnetic recording," *IEEE Transactions on Magnetics*, vol. 33, pp. 2746–2748, Sept. 1997.

[7] K. K. Fitzpatrick and C. S. Modlin, "Time-varying MTR codes for high density magnetic recording," in *Proceedings IEEE Global Telecommunications Conference*, vol. 3, (Phoenix, AZ, USA), pp. 1250–1253, IEEE, Nov. 1997.

[8] R. Karabed, P. Siegel, and E. Soljanin, "Constrained coding for binary channels with high intersymbol interference," *IEEE Transactions on Information Theory*, vol. 45, pp. 1777–1797, Sept. 1999.

[9] B. E. Moision, A. Orlitsky, and P. H. Siegel, "On codes that avoid specified differences," *submitted to IEEE Transactions on Information Theory*, Nov. 1999.

[10] L. C. Barbosa, "Maximum-likelihood sequence estimators: a geometric view," *IEEE Transactions on Information Theory*, vol. 35, pp. 419–427, Mar. 1989.

[11] G. C. Rota and G. Strang, "A note on the joint spectral radius," *Indagationes Mathematicae*, vol. 22, pp. 379–381, 1960.

[12] D. Lind and B. Marcus, *An Introduction to Symbolic Dynamics and Coding*. New Jersey: Cambridge University Press, 1995.

[13] I. Daubechies and J. C. Lagarias, "Two-scale difference equations II. Local regularity, infinite products of matrices and fractals," *SIAM Journal Math. Anal.*, vol. 23, pp. 1031–1079, 1992.

[14] M. A. Berger and Y. Wang, "Bounded semi-groups of matrices," *Linear Algebra and its Applications*, vol. 166, pp. 21–27, 1992.

[15] M. Maesumi, "Spectral radius of sets of matrices," in *Proceedings of the SPIE*, vol. 2303, (San Diego, CA, USA), pp. 27–29, July 1994.

[16] D. Colella and C. Heil, "The characterization of continuous, four-coefficient scaling functions and wavelets," *IEEE Transactions on Information Theory*, vol. 38, pp. 876–881, Mar. 1992.

[17] G. Gripenberg, "Computing the joint spectral radius," *Linear Algebra and its Applications*, vol. 234, pp. 43–60, 1996.

[18] R. Karabed and P. H. Siegel, "Matched spectral-null codes for partial-response channels," *IEEE Transactions on Information Theory*, vol. 37, pp. 818–855, May 1991.