Available online at www.sciencedirect.com
ScienceDirect
Linear Algebra and its Applications 422 (2007) 442–454
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

# On codes with local joint constraints

Bruce Moision [a,*], Alon Orlitsky [b], Paul H. Siegel [c]

[a] *Jet Propulsion Laboratory, 4800 Oak Grove Drive, MS 238-420, Pasadena, CA 91109-8099, United States*
[b] *Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA 92093-0407, United States*
[c] *Center for Magnetic Recording Research, University of California at San Diego, La Jolla, CA 92093-0401, United States*

Received 22 December 2004; accepted 4 November 2006
Available online 22 January 2007
Submitted by H. Schneider

**Abstract**

We study the largest number of sequences with the property that any two sequences do not contain specified pairs of patterns. We show that this number increases exponentially with the length of the sequences and that the exponent, or *capacity*, is the logarithm of the joint spectral radius of an appropriately defined set of matrices. We illustrate a new heuristic for computing the joint spectral radius and use it to compute the capacity for several simple collections. The problem of computing the achievable rate region of a collection of codes is introduced and it is shown that the region may be computed via a similar analysis.


*Keywords:* Joint spectral radius; Generalized spectral radius; Information capacity; Forbidden words

## 1. Introduction

Collections of sequences with the property that the difference of any pair does not contain a pattern from a specified set have been used as the basis for codes in magnetic recording channels [2]. In [1], it was shown 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. In this paper we introduce two generalizations of this problem and provide a new heuristic for computing the joint spectral radius.

* Corresponding author.
*E-mail addresses:* bmoision@jpl.nasa.gov (B. Moision), alon@ece.ucsd.edu (A. Orlitsky), psiegel@ucsd.edu (P.H. Siegel).

0024-3795/$ - see front matter 
doi:10.1016/j.laa.2006.11.002

In the first generalization, we consider collections of sequences with the property that any pair does not contain a pair of patterns from a specified set. Extensions to disallow larger collections, e.g. triples, is straightforward. Sequences with this property would be better suited for channels with multiple-user or inter-track interference, e.g. [3,4], or channels whose performance is characterized by pairs or triples rather than differences, e.g. [5,6]. We show that the maximum growth rate of the number of such sequences is the joint spectral radius of a certain set of matrices.

In the second generalization, we consider the achievable rates of a pair of codes such that the two codes do not jointly contain pairs of patterns from a specified set. We show that an upper bound on the sum of the rates is similarly given by the joint spectral radius of an appropriately defined set, and illustrate an algorithm for computing a tight lower bound on the rate region.

Underlying the solutions for these various problems is the computation of the joint spectral radius. Computation of the joint spectral radius is NP-hard even for special cases [7] and the determination of a strict bound is undecidable [8]. We illustrate a new heuristic for computing the joint spectral radius, and use it to compute the capacity for several simple collections, giving new examples and extending prior results from [1].

The paper is organized as follows. In the next section we formally describe the first generalization. Section 3 states the connection to the joint spectral radius. In Section 4 we review known algorithms for determining the joint spectral radius and illustrate a new heuristic, computing the capacity for several collections of pairs. Finally, in Section 5 we introduce and discuss the rate pairs problem.

## 2. Notation and definitions

For simplicity, we assume sequences are binary. A *pattern* is a finite string of bits. A *joint pattern* is a set of two distinct equal-length patterns. Let $\mathscr{J}$ denote a collection of joint patterns. An $n$-bit *code* $\mathscr{C}$ is a collection of $n$-bit sequences, or *codewords*. $\mathscr{C}$ *avoids* $\mathscr{J}$ if for all $u, v \in \mathscr{C}$ and all $i \leqslant j$ in $[1, n]$,

$$\{u_{[i,j]}, v_{[i,j]}\} \notin \mathscr{J}, \tag{1}$$

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

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

and

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

We are interested in

$$\delta_n(\mathscr{J}) \stackrel{\text{def}}{=} \max\{|\mathscr{C}| : \mathscr{C} \text{ avoids } \mathscr{J}\},$$

the size of the largest $n$-bit code that avoids $\mathscr{J}$. It is easy to verify that $\delta_n(\mathscr{J})$ is *sub-multiplicative*, i.e.,

$$\delta_{n_1+n_2}(\mathscr{J}) \leqslant \delta_{n_1}(\mathscr{J}) \cdot \delta_{n_2}(\mathscr{J})$$

for all $n_1, n_2 > 0$. Hence, by the Sub-Additivity Lemma, e.g. [9], we can define the *capacity* of $\mathscr{J}$ as the limit

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

We would like to determine the capacities of various sets $\mathscr{J}$.

We are primarily interested in finite collections. Without loss of generality we can assume from here on that all patterns in $\mathscr{J}$ have the same length $m$. Otherwise, let $m$ be the length of the longest pattern in $\mathscr{J}$ and replace every pair of length $m' < m$ by its $2^{m-m'}$ extensions of length $m$.

With this equal-length assumption, we restate constraint (1): Let $n' \stackrel{\text{def}}{=} n - m + 1$ and require that for all $u, v \in \mathscr{C}$ and all $i \in [1, n']$,

$$\{u_{[i,i']}, v_{[i,i']}\} \notin \mathscr{J},$$

where

$$i' \stackrel{\text{def}}{=} i + m - 1.$$

Note also that we use the term *pattern* to refer to strings of length $m$ and *sequence* for strings of length $n$.

A generalization of our results with $\mathscr{J}$ a collection of sets of arbitrary sizes is straightforward. In this paper, we address the case of collections of pairs to simplify the presentation.

## 3. From disallowed pairs to joint spectral radius

In [1], we consider the more restricted case where $\mathscr{J}$ is the collection of all pairs of patterns with difference belonging to a set of difference patterns $D$. For example, with $D = \{\{-11\}\}$, a single difference pattern from the ternary alphabet $\{-1, 0, 1\}$, we have $\mathscr{J} = \{\{01, 10\}\}$, since this pair uniquely yields the difference pattern. We show that the capacity of the collection is the joint spectral radius of a certain collection of matrices. The result, however, does not depend on $\mathscr{J}$ being derived from a set of difference patterns. It is a straightforward generalization to allow an arbitrary collection of pairs. In this section we state this generalization. We omit details of the proof which may be found in [1].

### 3.1. Representing sets

A set $M \subseteq \{0, 1\}^m$ *represents* or is a *representing set for* $\mathscr{J}$ if it intersects every set in $\mathscr{J}$. It is *minimal* if, in addition, none of its strict subsets represents $\mathscr{J}$. Clearly, every representing set contains a minimal one. Let $\mathscr{M}(\mathscr{J})$ be the collection of all minimal representing sets for $\mathscr{J}$.

### 3.2. Bipartite graph presenting $M \in \mathscr{M}(\mathscr{J})$

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 $r \in R$. Though we do not draw their direction explicitly, we think of the edges as directed from left to right.

For $m \geqslant 2$, 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 $i = 2, \ldots, m-1$. We identify this edge with the $m$-bit sequence $l_1, l_2, \ldots, l_{m-1}, r_{m-1} = l_1, r_1, \ldots, r_{m-1}$. Fig. 1 illustrates $G_2$ and $G_3$.

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 $M$. Fig. 2 illustrates $G_{\{10\}}$ and $G_{\{101\}}$.

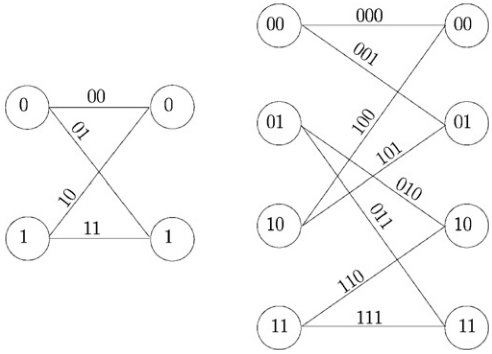


Fig. 1. $G_2$ and $G_3$.

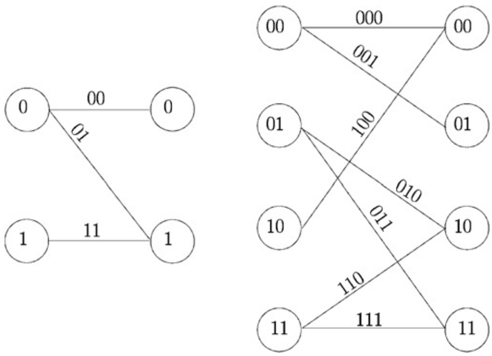


Fig. 2. $G_{\{10\}}$ and $G_{\{101\}}$.

### 3.3. Collection of adjacency matrices representing $\mathscr{M}(\mathscr{J})$

Identifying the elements of $L$ and $R$ of a bipartite graph $G = (L, R, E)$ with the intervals $[1, |L|]$ and $[1, |R|]$, respectively, we let the *adjacency matrix* $A_G$ be the $|L| \times |R|$ matrix whose $(l, r)$th element is 1 if $(l, r) \in E$, and 0 otherwise.

Let

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

denote the set of adjacency matrices corresponding to the collection $\mathscr{M}(\mathscr{J})$ of minimal representing sets for the disallowed joint patterns $\mathscr{J}$ and let

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

denote the set of products of $n$ matrices in $\Sigma$.

### 3.4. Matrix norms and spectral radius

Let $\|\cdot\|$ denote a matrix norm. By sub-multiplicativity of matrix norms, the limit

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

exists, and is independent of the norm $\|\cdot\|$. Let

$$\check{\rho}(A) \stackrel{\text{def}}{=} \max\{|\lambda| : \lambda \text{ an eigenvalue of } A\}.$$

For any matrix norm and $A \in \mathbb{C}^{m\times m}$ we have, e.g. [10, Theorem 5.6.9],

$$\check{\rho}(A) \leqslant \|A\|. \tag{3}$$

It is also well known, e.g. [10, Corollary 5.6.14], that

$$\check{\rho}(A) = \hat{\rho}(A).$$

This quantity is called the *spectral radius* of $A$ and denoted by $\rho(A)$.

### 3.5. Joint spectral radius

The quantities $\hat{\rho}$ and $\check{\rho}$ can be generalized to sets of matrices. We begin with $\hat{\rho}$. Letting

$$\hat{\rho}_n(\Sigma, \|\cdot\|) \stackrel{\text{def}}{=} \sup\{\|A\| : A \in \Sigma^n\}$$

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

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

which is independent of the norm $\|\cdot\|$.

Daubechies and Lagarias [12] defined the *generalized spectral radius* of $\Sigma$ to be

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

where

$$\check{\rho}_n(\Sigma) \stackrel{\text{def}}{=} \sup\{\check{\rho}(A) : A \in \Sigma^n\}.$$

It follows from (3) that

$$\check{\rho}_n(\Sigma) \leqslant \hat{\rho}_n(\Sigma, \|\cdot\|)$$

for every $n$. Moreover, the joint and generalized spectral radius have been shown to be equal for all finite $\Sigma$ [13]. We denote this quantity by $\rho(\Sigma) = \check{\rho}(\Sigma) = \hat{\rho}(\Sigma)$, and refer to it as the joint spectral radius.

Substituting $\mathscr{J}$ for $\mathscr{J}(D)$ in the model and results of [1] we obtain:

**Theorem 1.** *For every finite* $\mathscr{J}$,

$$\mathsf{cap}(\mathscr{J}) = \log(\rho(\Sigma(\mathscr{J}))).$$

Namely, the capacity is the logarithm of the joint spectral radius of $\Sigma(\mathscr{J})$.

This equality generalizes known results on *constrained systems* where, instead of joint patterns, individual patterns are disallowed, and it is well known, e.g. [14, Theorem 3.9], that the growth rate of the number of sequences, or *Shannon capacity* of the constraint, is $\log(\rho(A))$, the logarithm of the spectral radius of a corresponding adjacency matrix $A$.

The joint spectral radius measures the maximum growth rate of the norm of a product of matrices drawn from the set $\Sigma$. Rota and Strang introduced this concept in [11], and it has been used to study convergence of infinite products of matrices, e.g. [15], with applications to wavelets [12]. The concept is also related to the stability properties of discrete linear inclusions, e.g. [16,17], wherein the logarithm of the joint spectral radius is referred to as the *Lyapunov indicator*.

In the next section we describe several existing algorithms for computing the joint spectral radius and introduce a heuristic for choosing a good norm in certain algorithms.

## 4. Algorithms for computation of the joint spectral radius

Computation of the joint spectral radius is NP-hard even for special cases [7] and the problem of determining whether $\rho(\Sigma) \leqslant 1$ is undecidable [8]. In this section we illustrate a new heuristic for computing the joint spectral radius by extending the branch-and-bound algorithm of [18], and use it to compute the capacity for several simple collections, giving new examples and extending prior results from [1]. We first discuss some known algorithms.

### 4.1. Branch-and-bound algorithm

Because of the sub-multiplicativity of $\hat{\rho}_n(\Sigma, \|\cdot\|)$,

$$\rho(\Sigma) = \hat{\rho}(\Sigma) \leqslant \hat{\rho}_n(\Sigma, \|\cdot\|)^{1/n}$$

for every $n$. Furthermore as $n$ increases, this upper bound generally better approximates the joint spectral radius in the sense that for every $n$, there exists an $n' > n$ such that

$$\hat{\rho}_{n'}(\Sigma, \|\cdot\|)^{1/n'} \leqslant \hat{\rho}_n(\Sigma, \|\cdot\|)^{1/n}.$$

Similarly, every $\check{\rho}_n(\Sigma)$ lower bounds $\rho(\Sigma)$, and as $n$ increases, $\check{\rho}_n(\Sigma)$ generally better approximates the joint spectral radius from below, in the sense that, for every $n$,

$$\check{\rho}_n(\Sigma)^{1/n} \leqslant \check{\rho}_{nk}(\Sigma)^{1/kn}$$

for any $k \geqslant 1$.

This suggests approximating the joint spectral radius $\rho(\Sigma)$ by computing the lower bounds $\max_{1 \leqslant k \leqslant n} \check{\rho}_k(\Sigma)^{1/k}$ and upper bounds $\min_{1 \leqslant k \leqslant n} \hat{\rho}_k(\Sigma, \|\cdot\|)^{1/k}$ for $n = 1, 2, \ldots$. However, the number of matrix operations increases as $|\Sigma|^n$; consequently determining $\rho(\Sigma)$ with an arbitrary error may be computationally prohibitive.

Several algorithms have been introduced to more efficiently compute or bound the joint spectral radius. Maesumi [19] showed that the number of matrix operations required to compute $\check{\rho}_n(\Sigma)$ need be no greater than $|\Sigma|^n/n$. Daubechies and Lagarias [12] developed a recursive 'branch-and-bound' algorithm to upper bound $\rho(\Sigma)$, e.g. [12,20,21]. This was extended by Gripenberg [18] to include a sequence of lower bounds such that $\rho(\Sigma)$ may be specified to lie within an arbitrarily small interval.

### 4.2. Pruning algorithm

In [1] a pruning algorithm was presented for bounding $\rho(\Sigma)$ when all the matrices in $\Sigma$ 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. It can be applied to compute $\check{\rho}_n(\Sigma)$ and $\hat{\rho}_n(\Sigma, \|\cdot\|)$ for several norms. We briefly describe the algorithm here.

We write $A \geqslant 0$ if every element of $A$ is nonnegative and $A \geqslant B$ if every element of $A$ is at least as large as the corresponding element of $B$. It can be shown, e.g. [10, Theorem 8.1.18], that if $A \geqslant B \geqslant 0$ then

$$\rho(A) \geqslant \rho(B). \tag{4}$$

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

$$\|AM\| \geqslant \|BM\|$$

for all $M \geqslant 0$. A subset $S$ of $\Sigma^n$ is *dominating* if every matrix in $\Sigma^n$ is dominated by some matrix in $S$. Let $\Psi_n$ be any dominating subset of $\Sigma^n$. By definition,

$$\hat{\rho}_n(\Sigma, \|\cdot\|) = \max\{\|A\| : A \in \Psi_n\}. \tag{5}$$

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

Given a matrix norm one can therefore construct a recursive algorithm which computes a dominating set $\Psi_n$ from $\Psi_{n-1}$ by considering all products in $\Psi_{n-1}\Sigma$ and 'pruning' those that are dominated by another product. The subsequent growth rate of $|\Psi_n|$ will depend on the condition for domination. Sufficient conditions for domination for several norms are described in [1].

### *4.3. A new heuristic algorithm*

In all cases we have observed, the lower bounds provided by branch-and-bound or pruning based algorithms do not increase after a finite depth. This suggests that the joint spectral radius is achieved by a finite product (this finiteness conjecture was made in [22] but has since been disproven [23]). However, the upper bound may converge slowly. We propose the following heuristic for increasing the convergence rate of the upper bound. It is essentially a method of choosing a good norm for a branch-and-bound algorithm when we suspect a given matrix in $\Sigma^n$ achieves $\rho(\Sigma)$.

Suppose we observe that for some $A \in \Sigma^n$, $\check{\rho}_k(\Sigma) \leqslant \check{\rho}(A)^{1/n}$ for all computed values of $k$. We want to check whether $A \in \Sigma^n$ achieves the joint spectral radius, i.e., if $\rho(\Sigma) = \check{\rho}(A)^{1/n}$. Let $S$ be the nonsingular matrix such that $S^{-1}AS$ is in Jordan form. We conjecture that, if $A$ achieves the joint spectral radius, then $\hat{\rho}_k(S^{-1}\Sigma S, \|\cdot\|)$ will converge more rapidly than $\hat{\rho}_k(\Sigma, \|\cdot\|)$. The intuition behind the heuristic is that the growth rate of the norm of the product that achieves the joint spectral radius will be larger than the growth rate of any other product. Recall the *spectral norm* is given by

$$\|A\|_s \stackrel{\text{def}}{=} \max_{\|x\|_2=1} \|Ax\|_2,$$

where

$$\|x\|_2 \stackrel{\text{def}}{=} \left(\sum_{i=1}^{m} |x_i|^2\right)^{1/2}$$

is the *Euclidean norm* of the vector $x$, and the $L_1$ norm is given by

$$\|A\|_1 = \sum_{i,j} |A_{i,j}|.$$

Now, if $A$ is full rank, then $S^{-1}AS$ will be diagonal and if $S^{-1}A^kS$ achieves the upper bound $\hat{\rho}_{kn}(S^{-1}\Sigma S, \|\cdot\|)$ with either the $L_1$ or spectral norms, then the algorithm will terminate. We note that since the pruning algorithm as proposed operates under the assumption that all matrices are non-negative, which may not be the case after the similarity transformation, we will use the heuristic with branch-and-bound algorithms such as that proposed in [18].

**Example 1.** As an example of applying the heuristic, let

$$\Sigma = \left\{ \begin{bmatrix} 3/5 & 0 \\ 1/5 & 3/5 \end{bmatrix}, \begin{bmatrix} 3/5 & -3/5 \\ 0 & -1/5 \end{bmatrix} \right\},$$

which was considered in [21,18]. At a search depth of 243, taking into account finite precision, an application of the branch-and-bound algorithm in [18] yields

$$0.6596789\ldots \leqslant \rho(\Sigma) < 0.6596924,$$

but the algorithm did not seem to give a smaller interval. However, we observe that

$$A = \begin{bmatrix} 3/5 & -3/5 \\ 0 & -1/5 \end{bmatrix} \begin{bmatrix} 3/5 & 0 \\ 1/5 & 3/5 \end{bmatrix}^{12}$$

achieves the lower bound. Using the heuristic and exact arithmetic, we find $S$ which diagonalizes $A$ and compute bounds on $\rho(S^{-1}\Sigma S)$ using the spectral norm in the branch-and-bound algorithm in [18]. The upper and lower bounds agree at a depth of 34 yielding

$$\rho(\Sigma) = \frac{3^{12/13}(5 + 2\sqrt{7})^{1/13}}{5} = 0.6596789\ldots$$

### 4.4. Examples

In this section we illustrate computations of the capacity for several collections of pairs. When all matrices in $\Sigma$ are Hermitian, it follows, e.g. [10, 5.6.6], that

$$\check{\rho}_1(\Sigma) = \hat{\rho}_1(\Sigma, \|\cdot\|_s),$$

hence,

$$\rho(\Sigma) = \hat{\rho}_1(\Sigma, \|\cdot\|_s).$$

For example, this can be used to provide a simple calculation of $\mathsf{cap}(\{\{00, 11\}\})$.

**Example 2.** For $\mathscr{J} = \{\{00, 11\}\}$ we have

$$\Sigma(\mathscr{J}) = \left\{ \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \right\},$$

hence $\mathsf{cap}(\mathscr{J}) = \log((1 + \sqrt{5})/2)$.

**Example 3.** For $\mathscr{J} = \{\{010, 001\}, \{110, 101\}\}$, computation of bounds on $\Sigma(\mathscr{J})$ using the spectral norm in the branch-and-bound algorithm of [18] yields

$$\mathsf{cap}(\mathscr{J}) \in [0.6942, 0.7095)$$

at a depth of 13, with all candidate products of 13 or fewer matrices having been considered. At this depth, the algorithm has 12,388 candidate products of 13 matrices. Much tighter bounds are hindered by the growth rate of the number of candidate products. The pruning algorithm limits the growth rate of the number of candidates, allowing a search to a larger depth. Using the spectral norm in the pruning algorithm yields

$$\mathsf{cap}(\mathscr{J}) \in [0.6942, 0.6946)$$

at a depth of 375. Although the number of candidates remains small – there are 22 – the upper bound is converging slowly.
However, we observe that

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \in \Sigma(\mathscr{J}),$$

which is full-rank, achieves the lower bound. Applying the heuristic, we find $S$ which diagonalizes $A$ and use the branch-and-bound algorithm in [18] to bound $\rho(S^{-1}\Sigma(\mathscr{J})S)$. Using the spectral norm the upper and lower bounds agree at a depth of 6, yielding

$$\mathsf{cap}(\mathscr{J}) = \log_2\left(\frac{1+\sqrt{5}}{2}\right) = 0.6942\ldots$$

**Example 4.** For $\mathscr{J} = \{\{0101, 1010\}\}$, computation of $\Sigma(\mathscr{J})$ using the spectral norm and either a branch-and-bound [18] or pruning algorithm yields

$$\mathsf{cap}(\mathscr{J}) \in [0.9467, 0.9468).$$

Tighter bounds are hindered by the growth rate of the number of candidate products.
However, we observe that

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} \in \Sigma^2(\mathscr{J})$$

achieves that lower bound. Applying the heuristic, we find $S$ such that $S^{-1}AS$ is in Jordan form, and compute $\rho(S^{-1}\Sigma(\mathscr{J})S)$ using a branch-and-bound algorithm [18] and the spectral norm, yielding

$$\begin{aligned}\mathsf{cap}(\mathscr{J}) &= \log_2\left(\left(3+\sqrt{3\zeta}+\sqrt{99-3\zeta+234\sqrt{3/\zeta}}\right)\Big/\ 12\right) \\ &= 0.9467\ldots,\end{aligned}$$

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

The following examples consider some classes of pairs $\mathscr{J}$ such that all matrices in $\Sigma(\mathscr{J})$ are full-rank. The initial motivation for investigating these classes was an intuition that the heuristic described in the prior section would perform well if the product that achieves the lower bound is full-rank. However, the capacities follow from more straightforward inductive arguments.

**Example 5.** Let $\mathscr{J}$ be the collection of all $m$-bit pairs with difference $0^{(m-2)}11$, $m \geqslant 2$. From Example 2, the case $m = 2$, we know $\mathsf{cap}(\mathscr{J}) \geqslant \log_2((1+\sqrt{5})/2)$. By inspection of the bipartite graphs $G_M$, $M \in \mathscr{M}(\mathscr{J})$, one can show via an inductive argument that $\delta_n(\mathscr{J}) \leqslant \delta_{n-1}(\mathscr{J}) + \delta_{n-2}(\mathscr{J})$, which implies $\mathsf{cap}(\mathscr{J}) \leqslant \log_2((1+\sqrt{5})/2)$. Hence $\mathsf{cap}(\mathscr{J}) = \log_2((1+\sqrt{5})/2)$.

**Example 6.** Let $\mathscr{J}$ be the collection of all $m$-bit pairs with difference $10^{(m-2)}1$, $m \geqslant 2$. By inspection of the bipartite graphs $G_M$, $M \in \mathscr{M}(\mathscr{J})$, one can show via an inductive argument that $\delta_n(\mathscr{J}) = \delta_{n-1}(\mathscr{J}) + \delta_{n-2}(\mathscr{J})$, hence $\mathsf{cap}(\mathscr{J}) = \log_2((1+\sqrt{5})/2)$.

## 5. Rate pairs

Consider the following scenario. We have two sources operating independently and transmitting over a channel wherein the two sources interfere with one another, e.g. inter-track interference in a magnetic recording channel or multi-user interference in a wireless channel. The performance of our system is enhanced if we can guarantee that the two users do not transmit a certain pair of patterns simultaneously. We would like to determine the achievable rate pairs for such a scheme.

This leads to the following modification of the problem. Let $\mathscr{J}$ be a collection of *ordered* pairs of possibly identical patterns. The $n$-bit codes $\mathscr{C}_1$ and $\mathscr{C}_2$ avoid $\mathscr{J}$ if, for all $u \in \mathscr{C}_1$, $v \in \mathscr{C}_2$ and all $i \leqslant j$ in $[1, n]$,

$$(u_{[i,j]}, v_{[i,j]}) \notin \mathscr{J}.$$

A rate pair $(R_1, R_2)$ is *achievable* if there exist codes $\mathscr{C}_1$ and $\mathscr{C}_2$ which avoid $\mathscr{J}$ and have rates greater than or equal to $R_1, R_2$ respectively. The *achievable rate region* is the set of all achievable rate pairs. Of particular interest is

$$\delta_n(\mathscr{J}) \stackrel{\text{def}}{=} \max\{|\mathscr{C}_1||\mathscr{C}_2| : \mathscr{C}_1, \mathscr{C}_2 \text{ avoid } \mathscr{J}\}$$

the largest product of the size of two $n$-bit codebooks. We similarly define the *capacity* of $\mathscr{J}$ as the limit

$$\mathsf{cap}(\mathscr{J}) \stackrel{\text{def}}{=} \log\left[\lim_{n\to\infty} (\delta_n(\mathscr{J}))^{1/n}\right]. \tag{6}$$

The capacity is an upper bound on the sum of the rates of the two codes. By translating $\mathscr{J}$ into a set of product trellises reflecting the pairs of paths simultaneously allowed in the two codes, one can show that $\mathsf{cap}(\mathscr{J})$ defined by (6) is the joint spectral radius of the corresponding set of adjacency matrices.

**Example 7.** For $\mathscr{J} = \{(11, 00)\}$, we have

$$\Sigma(\{(11,00)\}) = \left\{ \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \right\}$$

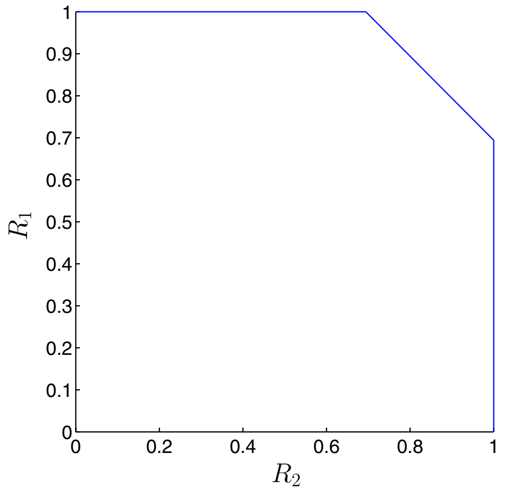


Fig. 3. Achievable rate region for $\mathscr{J} = \{(11, 00)\}$.

and $\mathsf{cap}(\mathscr{J}) = 1 + \log_2(\frac{1+\sqrt{5}}{2}) \approx 1.6942$. The capacity may be achieved by leaving one source unconstrained and disallowing 11 in the second source. The rate region, illustrated in Fig. 3, is achieved by time sharing.

Note that if the codes are allowed to cooperate, the problem reduces to the capacity of a source producing ordered binary pairs under the constraint that certain sequences of pairs are not allowed. This problem reduces to the computation of the spectral radius of an adjacency matrix, e.g. [14]. For example, if the two codes in Example 7 were allowed to cooperate, the problem reduces to the capacity of a source producing ordered binary pairs under the constraint that the pair $(1, 0)$ cannot be repeated. An adjacency matrix for this constraint is

$$\begin{bmatrix} 3 & 1 \\ 3 & 0 \end{bmatrix}$$

and the capacity is $\log_2(3 + \sqrt{21}/2) \approx 1.923$.

In general, the rate region is difficult to compute. We can, however, compute a tight lower bound by computing the rate pairs of all pairs of codes which avoid $\mathscr{J}$ and taking the convex hull of the resulting region, connecting outlying points by time-sharing.

The computation is simplified by applying to the tree search a pruning similar to that described in Section 4.2. Here, the leaves on the tree are pairs of products corresponding to the pair of codes. We say a pair $(A_1, A_2)$ dominates $(B_1, B_2)$ if $A_1 \geqslant B_1$ and $A_2 \geqslant B_2$. If $(A_1, A_2)$ dominates $(B_1, B_2)$, then the children of $(B_1, B_2)$ will fall within the rate region defined by the children of $(A_1, A_2)$.

**Example 8.** For $\mathscr{J} = \{(11, 10)\}$, we have $\mathsf{cap}(\mathscr{J}) = 1 + \log_2\left(\frac{1+\sqrt{5}}{2}\right)$. Fig. 4 illustrates upper and lower bounds on the boundary of the achievable rate region. The upper bound is defined by $R_1 \leqslant 1$, $R_2 \leqslant 1$, $R_1 + R_2 \leqslant \mathsf{cap}(\mathscr{J})$. The lower bound is obtained by the pruning technique described above.

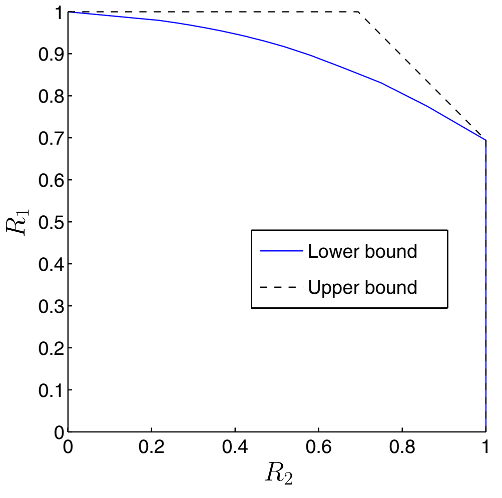


Fig. 4. Achievable rate region for $\mathscr{J} = \{(11, 10)\}$.

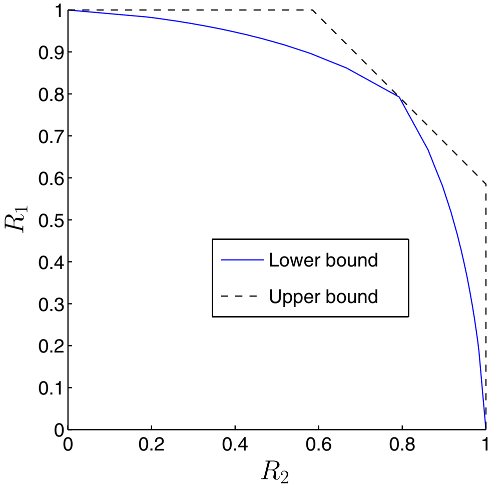


Fig. 5. Achievable rate region for $\mathscr{J} = \{(01, 10)\}$.

**Example 9.** For $\mathscr{J} = \{(01, 10)\}$, using the heuristic we can show $\mathsf{cap}(\mathscr{J}) = \log_2(3) = 1.58496\ldots$ We note that we were unable to compute the capacity exactly via straightforward computations using the branch-and-bound or pruning algorithms. Fig. 5 illustrates upper and lower bounds on the boundary of the achievable rate region. The upper bound is defined by $R_1 \leqslant 1$, $R_2 \leqslant 1$, $R_1 + R_2 \leqslant \mathsf{cap}(\mathscr{J})$. The lower bound is obtained by the pruning technique described above.

## References

[1] B.E. Moision, A. Orlitsky, P.H. Siegel, On codes that avoid specified differences, IEEE Trans. Inform. Theory 47 (1) (2001) 433–441.

[2] K.A.S. Immink, P.H. Siegel, J.K. Wolf, Codes for digital recorders, IEEE Trans. Inform. Theory 44 (6) (1998) 2260–2299.
[3] E. Soljanin, C.N. Georghiades, Coding for two-head recording systems, IEEE Trans. Inform. Theory 41 (1995) 747–755.
[4] E. Kurtas, J.G. Proakis, M. Salehi, Coding for multitrack magnetic recording systems, IEEE Trans. Inform. Theory 43 (6) (1997) 2020–2023.
[5] J. Caroselli, J.K. Wolf, Error event characterization of partial response systems in magnetic recording systems with media noise, Proceedings IEEE Global Telecommunications Conference, vol. 5, IEEE, Sydney, Australia, 1998, pp. 2724–2729.
[6] E. Soljanin, On-track and off-track distance properties of class 4 partial response channels, in: M.R. Raghuveer, S.A. Dianat, S.W. McLaughlin, M. Hassner (Eds.), Proceedings of the SPIE, vol. 2605, Philadelphia, PA, USA, 1995, pp. 92–102.
[7] J.N. Tsitsiklis, V.D. Blondel, The Lyapunov exponent and joint spectral radius of pairs of matrices are hard—when not impossible—to compute and to approximate, Math. Control Signals Syst. 10 (1997) 31–40.
[8] V.D. Blondel, J.N. Tsitsiklis, The boundedness of all products of a pair of matrices is undecidable, Syst. Control Lett. 41 (2) (2000) 135–140.
[9] R.L. Graham, B.L. Rothschild, J.H. Spencer, Ramsey Theory, John Wiley & Sons, 1990.
[10] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, New Jersey, 1988.
[11] G.C. Rota, G. Strang, A note on the joint spectral radius, Indagationes Mathematicae 22 (1960) 379–381.
[12] I. Daubechies, J.C. Lagarias, Two-scale difference equations. II. Local regularity, infinite products of matrices and fractals, SIAM J. Math. Anal. 23 (1992) 1031–1079.
[13] M.A. Berger, Y. Wang, Bounded semi-groups of matrices, Linear Algebra Appl. 166 (1992) 21–27.
[14] B.H. Marcus, R.M. Roth, P.H. Siegel, Constrained systems and coding for recording channels, in: V.S. Pless, W.C. Huffman (Eds.), Handbook of Coding Theory, Elsevier Science, 1998 (Chapter 20).
[15] I. Daubechies, J.C. Lagarias, Sets of matrices all infinite products of which converge, Linear Algebra Appl. 161 (1992) 227–263.
[16] L. Gurvits, Stability of discrete linear inclusion, Linear Algebra Appl. 231 (1995) 47–85.
[17] N.E. Barabanov, Lyapunov indicators of discrete inclusions. I, Automat. Remote Control 49 (2) (1988) 152–157.
[18] G. Gripenberg, Computing the joint spectral radius, Linear Algebra Appl. 234 (1996) 43–60.
[19] M. Maesumi, An efficient lower bound for the generalized spectral radius of a set of matrices, Linear Algebra Appl. 240 (1996) 1–7.
[20] M. Maesumi, Spectral radius of sets of matrices, in: Proceedings of the SPIE, vol. 2303, San Diego, CA, USA, 1994, pp. 27–29.
[21] D. Colella, C. Heil, The characterization of continuous, four-coefficient scaling functions and wavelets, IEEE Trans. Inform. Theory 38 (2) (1992) 876–881.
[22] J. Lagarias, Y. Wang, The finiteness conjecture for the generalized spectral radius of a set of matrices, Linear Algebra Appl. 214 (1995) 17–42.
[23] T. Bousch, J. Mairesse, Asymptotic height optimization for topical IFS, Tetris heaps and the finiteness conjecture, J. Amer. Math. Soc. (15) (2002) 77–111.