Discrete Applied Mathematics 198 (2016) 264–273

Contents lists available at ScienceDirect

Discrete Applied Mathematics

journal homepage: www.elsevier.com/locate/dam

The channel capacity of read/write isolated memory[☆]

Chuanlong Wang[a], Xuerong Yong[b,*], Mordecai Golin[c]

[a] Department of Mathematics, Taiyuan Normal University, Taiyuan, Shanxi, China

[b] Department of Mathematical Sciences, The University of Puerto Rico, Mayaguez, PR 00681, USA

[c] Department of Computer Science and Engineering, Hong Kong University of Science & Technology, Clear Water Bay, Hong Kong

ARTICLE INFO

Article history:
Received 24 October 2013
Received in revised form 3 May 2015
Accepted 22 May 2015
Available online 18 July 2015

Keywords:
Constrained codes
Channel capacity
Transfer matrix
Spectral radius
Power method

ABSTRACT

A read/write isolated memory is a binary re-writable medium in which (i) two consecutive locations cannot both store 1's and also in which (ii) two consecutive locations cannot both be modified during the same rewriting pass. Its channel capacity C, in bits per symbol per rewrite, is defined as

$$C = \lim_{k,r\to\infty} \frac{\log_2 N(k,r)}{kr},$$

where k is the size of the memory in binary symbols, r is the lifetime of the memory in rewriting cycles, and $N(k,r)$ is the number of distinct sequences of r-characters that satisfy the constraints. This quantity was originally considered by Cohn (1995) who proved that $0.509\ldots \le C \le 0.560297\ldots$ and conjectured that $C = 0.537\ldots$. Subsequently, Golin et al. (2004) refined the bounds to $0.53500\ldots \le C \le 0.55209\ldots$ and conjectured that $C = 0.5350\ldots$.

In this paper, we develop a new technique for computing C as a particular type of constrained binary matrix and obtain that

$$C = 0.53501\ldots.$$

The methods introduced in this note are not specific to this particular problem but can also be used to consider various other computational counting problems.

1. Introduction

Binary (0, 1) sequences in a communication channel are often required to satisfy predefined specific constraints that guarantee reliable storage or transmission. The set of all permissible binary memory configurations of a given size can be viewed as a channel alphabet that obeys specific restrictions. A *read/write isolated memory* (RWIM) is a binary, linearly ordered, re-writable storage medium satisfying two restrictions. The first, the *read restriction* states that no two consecutive positions in the memory may both store 1's. The second, the *write restriction*, states that when the memory is rewritten no two consecutive positions in the memory are allowed to change. A fixed size memory can be viewed as a character sent over

☆ Research was supported partly by DIMACS (NSF center at Rutgers, The State University of New Jersey), Shanxi Beiren Jihua Projects of China, The University of Puerto Rico at Mayaguez, and Hong Kong RGC CERG grants.

* Corresponding author.
E-mail addresses: clwang218@126.net (C. Wang), xryong@dimacs.rutgers.edu (X. Yong), golin@cse.ust.hk (M. Golin).

http://dx.doi.org/10.1016/j.dam.2015.05.027
0166-218X/© 2015 Published by Elsevier B.V.

$$B_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 0 & 0 & \mathbf{1} & \mathbf{1} & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & \mathbf{0} & \mathbf{1} \\ 0 & 0 & 0 & 0 & \mathbf{1} & \mathbf{0} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Fig. 1. Matrix B_1 satisfies the RWIM constraints. Matrix B_2 does not for two reasons: the two consecutive bold 1's in the first row contradict the read restriction and the bold 2×2 submatrix in the fourth and fifth columns contradicts the write restriction.

a noiseless communication channel; the memory contents after rewriting pass i is the ith character sent over the channel. The write restrictions dictate which characters may follow which characters in the channel.

The *channel capacity* for a rewritable memory, measured in bits per character was studied in [20,21]. The capacity of memory, in bits per rewrite, is defined as

$$C_k = \lim_{r\to\infty} \frac{1}{r} \log_2 N(k, r),$$

where k is the size of the memory in binary symbols, r is the lifetime of the memory in rewriting cycles, and $N(k, r)$ is the number of distinct sequences of r-characters that satisfy the constraints. Note that $N(k, r)$ is equal to the number of distinct paths through a channel graph that describes permissible transitions among characters. Shannon showed that $C_k = \log_2 \lambda_k$, where λ_k is the spectral radius of the channel graph, *i.e.*, the largest eigenvalue of its adjacency matrix. (From the Perron–Frobenius Theorem in the theory of nonnegative matrices [1], λ_k is a positive number.) Shannon also proved that the capacity is an upper bound on the rate achievable by any coding scheme. A code is nearly *optimal* if the capacity obtained from computation is very close to its true value.

The channel capacity C of RWIM, in bits per symbol per rewrite, can be defined as [7,11]

$$C = \lim_{k,r\to\infty} \frac{\log_2 N(k, r)}{kr}. \tag{1}$$

The limit has been proven to exist but the exact value of C is difficult to obtain.

In applications the codes with the read and/or write restrictions above are typical of those that used magnetic recording and optical recording. To the best of our knowledge, Freiman and Wyner [8] were the first to consider read isolated memories, followed by Kautz [14]. Write isolated memories were originally examined by Robinson [18] and then Cohen [6]. This memory is used in the context of an asymmetric error-correcting ternary code and re-writable optical disc. They showed independently that the two different types of memories have the same capacity $\log_2 \varphi = 0.694\ldots$, in bits per symbol, in which φ is the larger root of the *Fibonacci recurrence*: $F_{n+2} = F_{n+1} + F_n$.

A (d, k)-*Runlength Limited (RLL)* code over the binary alphabet $\{0, 1\}$ has d and k being the minimum and maximum permitted numbers of 0's separating consecutive 1's in its each codeword, respectively. This class of codes has wide applications [13]. As examples, the $(1, 3)$-RLL constraint is often used in magnetic disc drivers and the $(2, 10)$-RLL is in compact audio discs. The $(1, \infty)$-RLL is the read isolated constraint.

The (d, k)-RLL codes were sets of one dimensional strings. A 2-D code is a set of 2-D arrays/matrices that must each satisfy a set of constraints. The constraints are often given as horizontal (on the rows) and vertical (on the columns). Let $N(k, r)$ now be the number of $n \times r$ matrices that satisfy the given constraints. The capacity per bit of the 2-D code is given by

$$C = \lim_{k,r\to\infty} \frac{\log_2 N(k, r)}{kr}. \tag{2}$$

A 2-D (d, k)-RLL code is a matrix whose rows and columns all individually satisfy the 1-$D(d, k)$-RLL constraint. Over the last two decades the capacity of these codes have started to be studied. A relatively recent contribution is [9].

Notice that (2) is exactly in the same form as (1) so channel capacity looks very similar to 2-D code capacity. In fact, the RWIM channel can actually be rewritten as a 2-D code as follows. For a memory of size n with lifetime r create an $r \times k$ binary matrix B in which row i is the contents of the memory at time i. B must satisfy the following two constraints (see Fig. 1 for examples):

1. *Read restriction:* B does not contain any two horizontally consecutive ones, i.e., it does not contain any 1×2 submatrix $(1, 1)$.
2. *Write restriction:* B does not contain any 2×2 submatrix of the form $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ or $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Any B that satisfies these two constraints represents a legal set of rewrites for a r bit memory in the RWIM channel so $N(k, r)$ for the RWIM also represents the number of 2-D codewords that satisfy the read and write restrictions above and the RWIM channel capacity is the same as the capacity of the 2-D RWIM codes.

When Cohen originally investigated the capacity of the RWIM constraint in [7] he derived the following upper and lower bounds on the capacity C:

$$0.509\ldots \le C \le 0.560297\ldots.$$

On the basis of his computational results, he conjectured that the largest eigenvalue of the channel graph might satisfy a second-order ordinary linear difference equation, and the limiting capacity C would be about $\log_2 1.451 = 0.537\ldots$ in bits per symbol per rewrite. Subsequently, this problem was considered by the second and third authors [11] and the bounds were improved to

$$0.5350\ldots \le C \le 0.55209\ldots.$$

These bounds were obtained by modeling the problem as a constrained two-dimensional binary matrix problem and then by modifying the *transfer matrix* technique to bound the largest eigenvalues (the spectral radius) of the associated horizontal and vertical transfer matrices (defined in the next section). In that paper, the capacity C was conjectured to satisfy

$$0.5350\ldots \le C \le 0.5350\ldots. \tag{3}$$

We note that Roth [19] used a modified version of the cylindrical bounding technique introduced by Carkin and Wilf in [3] (which requires calculating the largest eigenvalues of a different type of matrix) to show that the capacity satisfies $C \le 0.535232$, strengthening the validity of the conjecture (3).

This paper further considers the RWIM capacity problem. Transfer matrix techniques define infinite sequences of larger and larger matrices. The spectral radius of each of these matrices gives bounds on capacity. The larger the original matrix, the better the bounds its spectral radius usually yields. Unfortunately, these matrices usually grow in size exponentially, severely restricting the number of matrices whose spectral radii can be calculated. Our major contribution[1] is to make use of the recursive properties of the vertical and horizontal transfer matrices, to compress the sizes of the matrices in ways that permit calculating the spectral radii. This permits showing that the capacity satisfies

$$C = 0.53501\ldots, \tag{4}$$

validating the conjecture. The compression techniques introduced are different than the ones that have appeared previously in the literature. More interestingly, the methods can be modified to consider some other counting problems in combinatorics.[2]

Throughout the paper, 't' stands for the transpose of a matrix or a vector, and '1' the all-ones column vector with an appropriate size. $\rho(A)$ denoted the spectral radius of a square matrix A.

In the next section we will describe the two classes of transfer matrices. Their recurrence properties are crucial in our consideration. In Section 3 we will derive our first improved upper bound on C by compressing the vertical transfer matrices through taking the Euclidean norm on their sub-matrices (blocks of the transfer matrices) and show that the upper bound goes to the true capacity as the size of matrix approached infinity. Our proof of the conjecture (4) will be given in Section 4 by peeling off the zero eigenvalues from the horizontal transfer matrices to halve their sizes and compute their largest eigenvalues. We will conclude the paper by proposing an open problem.

2. The vertical and horizontal transfer matrices

In the previous section we noted that a 2-D codeword is a 2-D array (matrix) that satisfies certain constraints. Let m be a fixed integer and V_m be the indexed set of all size m *vectors* (strings) that may appear as rows in some allowed $n \times m$ (0, 1) matrix in a 2-D code S with given constraints. We say that a pair of vectors (v_i, v_j) is *valid* if $v_i, v_j \in V_m$ and the $2 \times m$ array $\binom{v_i}{v_j}$ does not violate the given constraints. If n is a fixed integer we can set V_n to be all of the columns of size n columns that can appear in an allowed $m \times n$ matrix and say that (v_i, v_j) is *valid* if $v_i, v_j \in V_n$, if the $n \times 2$ array $(v_i v_j)$ does not violate the given constraints.

A *transfer matrix* $T_m = (t_{ij})$ corresponding to a given code satisfies that $t_{ij} = 1$ if (v_i, v_j) is *valid* and 0 otherwise.

Note that for a given constrained code for fixed m, (or n) the transfer matrix has size $|V_m| \times |V_m|$ ($|V_n| \times |V_n|$) with its structure being determined by the orderings of the vectors. In the literature the lexicographic order is often used. In two (or higher) dimensional constrained codes, the matrices are determined by the choice of direction chosen. For the 2-D RLL codes in which the horizontal and vertical restrictions are the same, the two sets of transfer matrices are the same. In other cases, such as the RWIM one, then transfer matrices in the different directions can be completely different [11]. We follow [11], in calling the transfer matrix A_k, introduced by Cohn [7], which fixes k the size of the memory, the *vertical transfer matrix* and call the transfer matrix $\bar{A}_r$ derived in [11] the *horizontal transfer matrix*, which is obtained by swapping the lifetime r and the memory size k and considering towards the r direction.

The two different types of transfer matrices, which have very different recurrent properties, will be used differently in our analysis.

The first part of the following theorem is due to Cohn ([7], Theorem 1) and the second part was derived by the second and third authors, and P. Zhang and L. Sheng in ([11], Definition 1 and Lemma 5).

---

1 A preliminary version of these results appeared in DCC'02 [22].

2 For example, they can be applied to counting the number of structures in some special graphs, such as perfect matchings, spanning trees, and independent sets [5,15,16].

Theorem 1 *([7,11]). If we arrange the vectors of read/write isolated constraints in the lexicographic order, then the vertical, and horizontal transfer matrices A_k, and $\bar{A}_r$ are given respectively by*

$$A_0 = 1, \qquad A_1 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \qquad A_k = \begin{pmatrix} A_{k-1} & \hat{A}_{k-2} \\ \hat{A}^t_{k-2} & A_{k-2} \end{pmatrix},$$

where A_k is an $f_{k+2} \times f_{k+2}$ matrix,[3] *$f_{k+2} = f_{k+1} + f_k$, $f_0 = 0$, $f_1 = 1$, and $\hat{A}_{k-2} = \begin{pmatrix} A_{k-2} \\ 0 \end{pmatrix}$ is of size $f_{k+1} \times f_k$, and*

$$\bar{A}_r = \begin{pmatrix} \bar{A}_{r-1} & B_{r-1} \\ B^t_{r-1} & 0 \end{pmatrix}, \qquad B_{r-1} = \begin{pmatrix} \bar{A}_{r-2} & B_{r-2} \\ 0 & 0 \end{pmatrix}$$

with initial conditions: $\bar{A}_1 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, $B_1 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ where $\bar{A}_r$ is of size $2^r \times 2^r$.

An $n \times n$ nonnegative matrix is called *primitive* if there exists a positive integer p such that A^p is an entry-wise positive matrix. The Perron–Frobenius Theorem in nonnegative matrix theory states that the largest eigenvalue (in modulus) of a primitive matrix is positive and strictly dominants the modulus of all the other eigenvalues [1]. This eigenvalue therefore is equal to the *spectral radius* of the matrix. Noting that both A_k and $\bar{A}_r$ are symmetric and their first rows are all 1's, by straightforward calculations we see that their squares are entry-wise positive and so both of them are primitive. The following theorem is combination of Lemmas 1, 2, 6, and 7 from [11].

Theorem 2 *([11]). The channel capacity C of read/write isolated memory exists and satisfies:*

- $$C = \lim_{k,r\to\infty} \frac{\log_2 N(k,r)}{kr}, \quad \text{and} \quad N(k,r) = 1^t A_k^{r-1} 1 = 1^t \bar{A}_r^{k-1} 1.$$
- *if λ_k and μ_r are the largest eigenvalues of A_k and $\bar{A}_r$, respectively, then*
  $$\lim_{r\to\infty} \frac{\log_2 N(k,r)}{r} = \log_2 \lambda_k, \qquad \lim_{k\to\infty} \frac{\log_2 N(k,r)}{k} = \log_2 \mu_r.$$
- *for $\forall r, k \geq 1$,*
  $$\max\left\{\log_2 \frac{\mu_{2r}}{\mu_{2r-1}}, \log_2 \frac{\lambda_{2k}}{\lambda_{2k-1}}\right\} \leq C \leq \min\left\{\frac{\log_2 \mu_r}{r}, \frac{\log_2 \lambda_k}{k}\right\}.$$

Making use of Theorem 2 we now can prove the following lemma.

Lemma 3. *Let λ_k and μ_r be the largest eigenvalues of A_k and $\bar{A}_r$, respectively. Then we have*

$$\lim_{k\to\infty} \log_2 \frac{\lambda_k}{\lambda_{k-1}} = \lim_{r\to\infty} \log_2 \frac{\mu_r}{\mu_{r-1}} = \inf_{k\geq 1} \log_2 \lambda_k^{\frac{1}{k}} = \inf_{r\geq 1} \log_2 \mu_r^{\frac{1}{r}} = C.$$

Proof. Let $\mu_1, \mu_2, \ldots, \mu_{2^r}$ and $x_1, x_2, \ldots, x_{2^r}$ be the eigenvalues and their corresponding column orthonormal eigenvectors of $\bar{A}_r$, *i.e.*, $\bar{A}_r x_i = \mu_i x_i$ and $x_i^t x_i = 1$ for all i. Then using the spectral decomposition of $\bar{A}_r$, $\bar{A}_r = \sum_i \mu_i x_i x_i^t$, for all k we have that $\bar{A}_r^k = \sum_i \mu_i^k x_i x_i^t$ and therefore

$$N(k,r) = 1^t \bar{A}_r^{k-1} 1 = \sum_i \alpha_i \mu_i^{k-1},$$

in which μ_r is the largest eigenvalue of the matrix. Note that $\alpha_i = (1^t x_i)^2$, $i = 1, 2, \ldots, 2^r$, x_r is entry-wise positive [1], and so $\alpha_r > 0$. From Theorem 2

$$\begin{aligned} \lim_{k\to\infty} \frac{\lambda_k}{\lambda_{k-1}} &= \lim_{k\to\infty} \lim_{r\to\infty} \left(\frac{1^t A_k^{r-1} 1}{1^t A_{k-1}^{r-2} 1} \right)^{\frac{1}{r}} \\ &= \lim_{r\to\infty} \lim_{k\to\infty} \left(\frac{1^t \bar{A}_r^{k-1} 1}{1^t \bar{A}_r^{k-2} 1} \right)^{\frac{1}{r}} \end{aligned}$$

[3] f_i is the Fibonacci sequence.

$$= \lim_{r\to\infty} \lim_{k\to\infty} \left(\frac{\sum_i \alpha_i \mu_i^{k-1}}{\sum_i \alpha_i \mu_i^{k-2}} \right)^{\frac{1}{r}}$$

$$= \lim_{r\to\infty} \mu_r^{\frac{1}{r}} = 2^C.$$

The second equality follows similarly from Theorem 2. □

3. Approaching the capacity from above by making use of the vertical transfer matrices

In this section, by taking the Euclidean norm on the sub-matrices (blocks) in the vertical transfer matrices A_k, we obtain smaller *compressed matrices*. The largest eigenvalues of these smaller matrices, will permit deriving upper bounds on the capacity C that approach C as the size of these compressed matrices increases.

The following lemma plays an important role in our discussion. Its proof is straightforward and can be found in [10].

Lemma 4. *Let T be an $n \times n$ matrix and n_i positive integers such that $\sum_{i=1}^{k} n_i = n$. Decompose*

$$T = \begin{pmatrix} T_{11} & T_{12} & \cdots & T_{1k} \\ T_{21} & T_{22} & \cdots & T_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ T_{k1} & T_{k2} & \cdots & T_{kk} \end{pmatrix} = (T_{ij}),$$

where, for $i, j = 1, 2, \ldots, k$, the T_{ij} are $n_i \times n_j$ matrices (called sub-matrices). For each T_{ij}, define a matrix norm by

$$\|T_{ij}\|_{n_i,n_j} = \sup_{x_j \neq 0} \frac{\|T_{ij}x_j\|_{n_i}}{\|x_j\|_{n_j}}, \quad i, j = 1, 2, \ldots, k,$$

and let $M = (m_{ij})$, $m_{ij} = \|T_{ij}\|_{n_i,n_j}$, $i, j = 1, 2, \ldots, k$. Then

$$\rho(T) \le \rho(M),$$

where $\rho(X)$ is the spectral radius of matrix X. The matrix M is called a compressed matrix of T.

From Theorem 1 we see that the transfer matrices A_k are real symmetric, $k = 1, 2, \ldots$, and so taking the Euclidean norm $\|\cdot\|_2$ on the matrices A_k, we have $\|A_k\|_2 = \sup_{x\neq 0} \frac{\|A_k x\|_2}{\|x\|_2} = \lambda_k$, the spectral radius of A_k, and

$$\|A_{k-1}\|_2 = \lambda_{k-1}, \qquad \|A_{k-2}\|_2 = \lambda_{k-2}, \qquad \|\hat{A}_{k-2}\|_2 = \lambda_{k-2}.$$

(See e.g., [4], or [17].) This implies that, if we let

$$M_1 = \begin{pmatrix} \|A_{k-1}\|_2 & \|\hat{A}_{k-2}\|_2 \\ \|\hat{A}^t_{k-2}\|_2 & \|A_{k-2}\|_2 \end{pmatrix} = \begin{pmatrix} \lambda_{k-1} & \lambda_{k-2} \\ \lambda_{k-2} & \lambda_{k-2} \end{pmatrix},$$

then from Lemma 4, $\lambda_k \le \rho(M_1)$, and therefore

$$\frac{\lambda_k}{\lambda_{k-1}} \le \frac{\rho(M_1)}{\lambda_{k-1}} = \rho\left(\frac{M_1}{\lambda_{k-1}}\right).$$

Now let $Q_0(x) = 1$ and $Q_1(x) = \begin{pmatrix} 1 & x \\ x & x \end{pmatrix}$. Since each eigenvalue of a matrix is a continuous function of its elements ([2], p. 153), taking the limit as k approaches to infinity on both sides of the above inequalities and applying Lemma 3 yield $2^C \le \rho(Q_1(2^{-C}))$. Similarly, noting that A_k also has the following recurrence relation

$$A_k = \begin{pmatrix} A_{k-2} & \hat{A}_{k-3} & A_{k-2} \\ \hat{A}^t_{k-3} & A_{k-3} & 0 \\ A_{k-2} & 0 & A_{k-2} \end{pmatrix},$$

we have

$$M_2 \stackrel{\text{def}}{=} \begin{pmatrix} \|A_{k-2}\|_2 & \|\hat{A}_{k-3}\|_2 & \|\hat{A}_{k-2}\|_2 \\ \|\hat{A}^t_{k-3}\|_2 & \|A_{k-3}\|_2 & 0 \\ \|A_{k-2}\|_2 & 0 & \|A_{k-2}\|_2 \end{pmatrix} = \begin{pmatrix} \lambda_{k-2} & \lambda_{k-3} & \lambda_{k-2} \\ \lambda_{k-3} & \lambda_{k-3} & 0 \\ \lambda_{k-2} & 0 & \lambda_{k-2} \end{pmatrix}.$$

Let

$$Q_2(x) \stackrel{\text{def}}{=} \begin{pmatrix} 1 & x & 1 \\ x & x & 0 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} Q_1(x) & \hat{Q}_0(x) \\ \hat{Q}_0(x)^t & 1 \end{pmatrix}, \qquad \hat{Q}_0(x) = \begin{pmatrix} Q_0(x) \\ 0 \end{pmatrix}.$$

Then, again from Lemma 4 we have $\lambda_k \le \rho(M_2)$, and therefore

$$\frac{\lambda_k}{\lambda_{k-1}} \cdot \frac{\lambda_{k-1}}{\lambda_{k-2}} = \frac{\lambda_k}{\lambda_{k-2}} \le \frac{\rho(M_2)}{\lambda_{k-2}} = \rho\left(\frac{M_2}{\lambda_{k-2}}\right).$$

Similarly, letting k go to infinity implies that $2^{2C} \le \rho(Q_2(2^{-C}))$. Continuing the same argument, for the following $F_{m+2} \times F_{m+2}$, $m = 3, 4, \ldots$, matrices

$$Q_m(x) = \begin{pmatrix} Q_{m-1}(x) & \hat{Q}_{m-2}(x) \\ \hat{Q}_{m-2}(x)^t & Q_{m-2}(x) \end{pmatrix}, \qquad \hat{Q}_{m-2}(x) = \begin{pmatrix} Q_{m-2}(x) \\ 0 \end{pmatrix}, \tag{5}$$

we infer that $2^{mC} \le \rho(Q_m(2^{-C}))$. Note that if we assume that $\lambda_0 = 1$, and that $A_0 = 1$ then, for $m = k$, $\lambda_k = \rho(Q_m(1)) = \rho(A_k)$.

Summarizing the above derivations yields

Lemma 5. *Let $\{Q_m(x)\}$ be the sequence of matrices generated from (5). Then*

$$C \le \frac{\log_2 \rho(Q_m(2^{-C}))}{m}, \tag{6}$$

where $\rho(Q_m(x))$ is the spectral radius of $Q_m(x)$, $m = 1, 2, 3 \ldots$.

Below we show that the right hand side upper-bound in (6) converges to C as m approaches infinity. To see this, we apply the strictly monotonic property[4] of $\rho(Q_m(x))$ ($\forall x > 0$) to the matrices in Lemma 5. This implies that $\forall x \ge 2^{-C}$, $2^C \le \rho(Q_m(x))^{\frac{1}{m}}$. Since in [11] it is shown that $0.5350\ldots \le C$, we have

$$2^{-C} < 2^{-0.535} = 0.69015866\ldots < 1.$$

On the other hand, since $Q_m(1) = A_m$, from Theorem 2 and Lemma 3, $\forall x \in [0.69015866, 1)$,

$$\inf_{k\ge1} \rho(A_k)^{\frac{1}{k}} = 2^C \le \rho(Q_m(x))^{\frac{1}{m}} < \rho(A_m)^{\frac{1}{m}}.$$

If $x \ge 1$, then

$$\inf_{k\ge1} \rho(A_k)^{\frac{1}{k}} = 2^C \le \rho(Q_m(x))^{\frac{1}{m}} = x^{\frac{1}{m}} \rho(R_m(y))^{\frac{1}{m}},$$

where $y = \frac{1}{x} \le 1$, $R_1(y) = \begin{pmatrix} y & 1 \\ 1 & 1 \end{pmatrix}$, $R_2(y) = \begin{pmatrix} y & 1 & y \\ 1 & 1 & 0 \\ y & 0 & y \end{pmatrix}$, and for $m > 2$,

$$R_m(y) = \begin{pmatrix} R_{m-1}(y) & \hat{R}_{m-2}(y) \\ \hat{R}_{m-2}(y)^t & R_{m-2}(y) \end{pmatrix}, \qquad \hat{R}_{m-2}(y) = \begin{pmatrix} R_{m-2}(y) \\ 0 \end{pmatrix}.$$

Now $R_m(y) \le R_m(1) = A_m$, this gives rise to

$$2^C = \inf_{k\ge1} \rho(A_k)^{\frac{1}{k}} \le \rho(Q_m(x))^{\frac{1}{m}} = x^{\frac{1}{m}} \rho(R_m(y))^{\frac{1}{m}} \le x^{\frac{1}{m}} \rho(A_m)^{\frac{1}{m}}.$$

Combining these properties with Theorem 2 yields the following theorem.

Theorem 6. *Let m be any given positive integer. Then*

- *for $\forall x \in [0.69015866, 1)$, the compressed matrix $Q_m(x)$ satisfies*

$$\inf_{k\ge1} \rho(A_k)^{\frac{1}{k}} = 2^C \le \rho(Q_m(x))^{\frac{1}{m}} < \rho(A_m)^{\frac{1}{m}};$$

- *for $x \ge 1$*

$$\inf_{k\ge1} \rho(A_k)^{\frac{1}{k}} = 2^C \le \rho(Q_m(x))^{\frac{1}{m}} \le x^{\frac{1}{m}} \rho(A_m)^{\frac{1}{m}},$$

where $\rho(Q_m(x)) = \rho(A_m)$ if and only if $x = 1$.

Consequently, for any given number $x > 0.69015866$, we have (by Lemma 3)

$$2^C = \inf_{m\ge1} \rho(Q_m(x))^{\frac{1}{m}} = \lim_{m\to\infty} \rho(Q_m(x))^{\frac{1}{m}} = \lim_{m\to\infty} \rho(A_m)^{\frac{1}{m}} = \inf_{k\ge1} \rho(A_k)^{\frac{1}{k}}.$$

---

[4] The spectral radius of a nonnegative irreducible matrix is strictly increasing with its elements and its corresponding eigenvector is entrywise positive [1].

Table 1
F_m is the mth Fibonacci number. $\rho(Q_m(x_0))$ is the largest eigenvalue of the $F_{m+2} \times F_{m+2}$ compressed matrix $Q_m(x_0)$.

m	F_{m+2}	$\rho(Q_m(x_0))$	$\frac{\log_2(\rho(Q_m(x_0)))}{m}$
1	2	1.5524196988	0.6345186448
2	3	2.1732673267	0.5599328234
3	5	3.2543210622	0.5674521968
4	8	4.6152889202	0.5516052411
5	13	6.8116188717	0.5535995429
6	21	9.7408000844	0.5473400461
7	34	14.2668740681	0.5477996234
8	55	20.5065208518	0.5447513549
9	89	29.9058714421	0.5447063169
10	144	43.0634915800	0.5428393393
11	233	62.7213694907	0.5428077414
12	377	90.6138568737	0.5418049818
13	610	131.2975410982	0.5412843145
14	987	190.3410264771	0.5408887676
15	1597	276.1769813000	0.5406299513
16	2584	399.7392197852	0.5401822073
17	3181	579.6907368666	0.5399493895
18	5765	839.3943441870	0.5396224962

Furthermore, if we work a little bit on the matrix $A_m - Q_m(x)$, we have

Corollary 1. *For* $\forall x \in (-\infty, +\infty)$ *and an arbitrarily given number* m*, the compressed matrix* $Q_m(x)$ *satisfies* $\rho(A_m - Q_m(x)) = \rho(A_{m-2} \otimes (A_1 - Q_1(x))) = |1-x|\frac{1+\sqrt{5}}{2}\rho(A_{m-2})$*, and if* $x \neq 1$*, then*

$$2^C = \inf_{k\geq 1} \rho(A_k)^{\frac{1}{k}} = \inf_{m\geq 1} \rho(A_m - Q_m(x))^{\frac{1}{m-2}} < \rho(A_m)^{\frac{1}{m}},$$

where $\otimes$ *is the Kronecker product of matrices.*

Remark. Note from Lemma 4 that we can obtain different compressed matrices and so different bounds if we take different norms on the sub-matrices in A_k, or if we let M_i be divided by $\lambda_{k-j}, j \neq i$ to derive different matrices than $Q_m(x)$.

As an example, let us derive a new upper bound from the above discussion. Set $x_0 = 0.69015866$ and a positive integer m. For $i = 1$ to m compute $x_{i+1} = (\rho(Q_i(x_0)))^{\frac{1}{i}}$. Then, for every x_0, we can take x_{m+1} as an upper bound of 2^C, and from Theorem 6 our bound will be convergent to the true capacity C as m approaches infinity. In Table 1 we show the initial numerical results, where the numbers are computed by the Power Method [12] using MATLAB for $x_0 = 0.69015866$ and $m = 1, 2 \ldots, 18$. By Theorem 6 the results are better than the ones for the same sizes of transfer matrices. Note that $m = 18$ means that the read/write isolated memory size considered so far is more than 18 bits because we computed the spectral radii of the compressed matrices Q_m rather than those of the transfer matrices A_m. Theorem 6, implies that we would be able to get even better and better upper bounds on C if we continued the computation and the computer had enough memory space.

From Table 1 we see that the convergence speed is not slow, and it is strictly decreasing after $m = 8$. Compared with the previous results addressed in the Introduction, we obtain a new upper bound of C, which is $0.53962245\ldots$ (it is not better than Roth's bound $0.535232\ldots$.) We address that this is just an example of finding an upper bound of C by making use of our theoretical results obtained in this section. Our main interest is to update the bounds and approach the true capacity in the following section.

We also point out that Lemma 5, permits deriving new recursive and explicit relations on λ_k that improve the bound obtained in [7]. For example, for $k = 1, 2, \ldots$, we have

$$\lambda_k \leq \frac{1}{2}\left(\lambda_{k-1} + \lambda_{k-2} + \sqrt{(\lambda_{k-1} - \lambda_{k-2})^2 + 4\lambda_{k-2}^2}\right). \tag{7}$$

In fact, the right-hand side of the above inequality is the larger eigenvalue of matrix M_1 introduced in the proof of Lemma 5, which implies that we have

$$\lambda_k \leq \lambda_{k-1} + \lambda_{k-2}, \qquad \lambda_k \leq \frac{1}{2}\lambda_{k-1} + \frac{1+\sqrt{5}}{2}\lambda_{k-2}.$$

This is because, from *Perron–Frobenius Theorem* [1], we have that $0 < \lambda_{k-1} < \lambda_k$, so

$$\begin{aligned}\lambda_k &\leq \frac{1}{2}\left(\lambda_{k-1} + \lambda_{k-2} + \sqrt{(\lambda_{k-1} - \lambda_{k-2})^2 + 4\lambda_{k-2}^2}\right)\\ &\leq \frac{1}{2}(\lambda_{k-1} + \lambda_{k-2} + \lambda_{k-1} - \lambda_{k-2} + 2\lambda_{k-2})\\ &= \lambda_{k-1} + \lambda_{k-2}.\end{aligned}$$

Combining these inequalities yields

$$\begin{aligned}\lambda_k &\le \frac{1}{2}\left(\lambda_{k-1}+\lambda_{k-2}+\sqrt{(\lambda_{k-1}-\lambda_{k-2})^2+4\lambda_{k-2}^2}\right)\\ &\le \frac{1}{2}\left(\lambda_{k-1}+\lambda_{k-2}+\sqrt{\lambda_{k-3}^2+4\lambda_{k-2}^2}\right)\\ &\le \frac{1}{2}\left(\lambda_{k-1}+\lambda_{k-2}+\sqrt{5\lambda_{k-2}^2}\right)\\ &= \frac{1}{2}\lambda_{k-1}+\frac{1+\sqrt{5}}{2}\lambda_{k-2}.\end{aligned}$$

These inequalities are tighter than the ones derived in ([7], Theorem 7 and its Corollary).

4. Approaching the true capacity

In this section we address the conjectures addressed in the Introduction by calculating much better bounds on the true capacity C. To obtain the best possible approximation of the largest eigenvalue of the transfer matrices we reduce the sizes of the horizontal transfer matrices, $\bar{A}_r$, by peeling off the zero eigenvalues from the matrices through elementary operations. This permits calculating the same eigenvalues by doing the calculations on a much smaller matrix, reducing the work required, in turn permitting the calculation of eigenvalues of larger transfer matrices.

Lemma 7. *Let $p(x)=\det(xI-\bar{A}_r)$ be the characteristic polynomial of the $2^r\times 2^r$ horizontal transfer matrix $\bar{A}_r$. Then*

$$p(x)=\det(xI-\bar{A}_r)=x^{2^{r-1}}\det(xI-M_r),$$

where M_r is a nonnegative primitive matrix of size $2^{r-1}\times 2^{r-1}$ and is given by

$$M_r=\begin{pmatrix}\bar{A}_{r-4} & B_{r-4} & \bar{A}_{r-4} & B_{r-4} & \bar{A}_{r-4} & 2\bar{A}_{r-4} & 3\bar{A}_{r-4} & 6B_{r-4}\\ B^t_{r-4} & 0 & 0 & 0 & B^t_{r-4} & 2B^t_{r-4} & 0 & 0\\ \bar{A}_{r-4} & 0 & 0 & 0 & \bar{A}_{r-4} & 0 & 0 & 0\\ B^t_{r-4} & 0 & 0 & 0 & B^t_{r-4} & 0 & 0 & 0\\ \bar{A}_{r-4} & B_{r-4} & \bar{A}_{r-4} & B_{r-4} & 0 & 0 & 0 & 0\\ \bar{A}_{r-4} & B_{r-4} & 0 & 0 & 0 & 0 & 0 & 0\\ \bar{A}_{r-4} & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ B^t_{r-4} & 0 & 0 & 0 & 0 & 0 & 0 & 0\end{pmatrix}.$$

Proof. Using the recurrence relations of the horizontal transfer matrix $\bar{A}_r$ in Theorem 1 to get

$$\det(xI-\bar{A}_r)=\det\begin{pmatrix}xI-\bar{A}_{r-2} & -B_{r-2} & -\bar{A}_{r-2} & -B_{r-2}\\ -B^t_{r-2} & xI & 0 & 0\\ -\bar{A}_{r-2} & 0 & xI & 0\\ -B^t_{r-2} & 0 & 0 & xI\end{pmatrix}.$$

For the sake of simplicity, in the following we view the blocks (sub-matrices) in the matrix as elements and proceed elementary (block) operations. Adding (-1) times row 4 to row 2, and then adding column 2 to column 4, and then expanding the determinant by row (block) 2 (from linear algebra the operations do not change the value of determinant), we obtain smaller matrices as follows,

$$\det(xI-\bar{A}_r)=x^{2^{r-2}}\det\begin{pmatrix}xI-\bar{A}_{r-2} & -\bar{A}_{r-2} & -2B_{r-2}\\ -\bar{A}_{r-2} & xI & 0\\ -B^t_{r-2} & 0 & xI\end{pmatrix}.$$

Now using again the recurrence relations of $\bar{A}_{r-2}$ and B_{r-2} in Theorem 1, the determinant on the right hand side of the above determinant becomes

$$\det\begin{pmatrix}xI-\bar{A}_{r-3} & -B_{r-3} & -\bar{A}_{r-3} & -B_{r-3} & 2\bar{A}_{r-3} & -2B_{r-3}\\ -B^t_{r-3} & xI & -B^t_{r-3} & 0 & 0 & 0\\ -\bar{A}_{r-3} & -B_{r-3} & xI & 0 & 0 & 0\\ -B^t_{r-3} & 0 & 0 & xI & 0 & 0\\ -\bar{A}_{r-3} & 0 & 0 & 0 & xI & 0\\ -B^t_{r-3} & 0 & 0 & 0 & 0 & xI\end{pmatrix}.$$

For this determinant, adding (-1) times row 6 to row 4, then adding column 4 to column 6, and then expanding the resulting determinant by row 4, we see that $\det(xI - \bar{A}_r)$ is equal to

$$x^{2^{r-2}+2^{r-3}} \det \begin{pmatrix} xI - \bar{A}_{r-3} & -B_{r-3} & -\bar{A}_{r-3} & -2\bar{A}_{r-3} & -3B_{r-3} \\ -B^t_{r-3} & xI & -B^t_{r-3} & 0 & 0 \\ -\bar{A}_{r-3} & -B_{r-3} & xI & 0 & 0 \\ -\bar{A}_{r-3} & 0 & 0 & xI & 0 \\ -B^t_{r-3} & 0 & 0 & 0 & xI \end{pmatrix}.$$

Similarly, repeating the same procedure with the above determinant, one can see easily that $\det(xI - \bar{A}_r)$ is equal to $x^{2^{r-2}+2^{r-3}} \det(xI - M_1)$ where M_1 is given by

$$\begin{pmatrix} \bar{A}_{r-4} & B_{r-4} & \bar{A}_{r-4} & B_{r-4} & \bar{A}_{r-4} & B_{r-4} & 2\bar{A}_{r-4} & 2B_{r-4} & 3\bar{A}_{r-4} & 3B_{r-4} \\ B^t_{r-4} & 0 & 0 & 0 & B^t_{r-4} & 0 & 2B^t_{r-4} & 0 & 0 & 0 \\ \bar{A}_{r-4} & 0 & 0 & 0 & \bar{A}_{r-4} & 0 & 0 & 0 & 0 & 0 \\ B^t_{r-4} & 0 & 0 & 0 & B^t_{r-4} & 0 & 0 & 0 & 0 & 0 \\ \bar{A}_{r-4} & B_{r-4} & \bar{A}_{r-4} & B_{r-4} & 0 & 0 & 0 & 0 & 0 & 0 \\ B^t_{r-4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \bar{A}_{r-4} & B_{r-4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ B^t_{r-4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \bar{A}_{r-4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ B^t_{r-4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Now for $\det(xI - M_1)$, adding (-1) times row 10 to rows 8 and 6, respectively, then adding columns 8 and 6 to column 10, and then expanding the resulting determinant by row 8 and row 6, one sees that $\det(xI - \bar{A}_r)$ is equal to

$$x^{2^{r-2}+2^{r-3}+2^{r-4}+2^{r-4}} \det(xI - M_r) = x^{2^{r-1}} \det(xI - M_r)$$

where M_r is the matrix given in the lemma. The proof is thus completed. □

Smaller matrices than M_r can be obtained if we continue the same argument in the proof of the lemma.

Now let

$$D = \begin{pmatrix} I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sqrt{2}I & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{3}I & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{6}I \end{pmatrix}.$$

Then, since DM_rD^{-1} is symmetric and shares the same eigenvalues as M_r, it has the same nonzero eigenvalues as $\bar{A}_r$ does, we therefore can use the *Power Method* [12] to compute the largest eigenvalue of DM_rD^{-1}. Since the size of the matrix M_r is half of the size of the transfer matrix $\bar{A}_r$, in computation we reduced a lot of work.

In Table 2 we show the numerical results of our computation of the largest eigenvalues of the compressed matrix DM_rD^{-1} by applying the Power Method and Matlab. The largest size of the matrix computed is 8192. In the table the initial 8 results are same as the ones obtained in [11].

Now combining Theorem 2, Lemma 3, and Table 2 proves the conjecture.

Theorem 8. *The channel capacity C of read/write isolated memory is*

$$C = 0.53501\ldots.$$

5. Conclusion and open problem

In this paper we developed a new technique to bound the channel capacity of read/write isolated memory as a particular type of constrained binary matrix. We derived an upper-bound function of C by compressing the *vertical transfer matrix* through taking the Euclidean norm on its 'elements' (sub-matrices) and showed that the upper bound function obtained

Table 2
μ_r is the largest eigenvalue of the $2^r \times 2^r$ transfer matrix $\bar{A}_r$.

r	2^r	μ_r	$\log_2 \frac{\mu_r}{\mu_{r-1}}$
1	2	1.618	
2	4	2.302775637	0.5091622465
3	8	3.346462191	0.5392628601
4	16	4.845619214	0.5340443229
5	32	7.021562462	0.5351110622
6	64	10.17359346	0.5349653450
7	128	14.74105370	0.5350103028
8	256	21.35908135	0.5350099454
9	512	30.948359597568	0.535010307223
10	1024	44.842824649376	0.535014214230
11	2048	64.975322373528	0.535014730251
12	4096	94.146459043118	0.535014947823
13	8192	136.414197132806	0.535015053352
14	16384	197.658348650048	0.535015093868

is convergent to the true capacity C. We then peeled off several zero eigenvalues from the horizontal transfer matrices to halve their sizes and showed that

$$C = 0.53501\ldots,$$

which indicates that the conjecture $C = 0.537\ldots$, posed by Cohn [7] and the conjecture $0.5350\ldots \le C \le 0.5350\ldots$ addressed in Conclusion of [11] are validated for the first two and four digits, respectively. An open question is:

Is it possible to derive an analytic function for the largest eigenvalue $\rho(A_k)$ of the transfer matrix A_k?

The techniques introduced in this paper are different than the ones that appear in the literature. More interestingly, our methods can be modified to consider many other counting problems in combinatorics. For example, combining the methods introduced in Sections 3 and 4, it is possible to consider counting the number of structures, such as perfect matchings, spanning trees, and independent sets in the special graphs discussed in [5,15,16] where transfer method techniques are already being used (they derived compressed matrix from combinatorics, which is completely different from ours).

Acknowledgments

The authors would like to thank the anonymous referees for their comments and suggestions towards improving the presentation of the article. Part of the research was completed when the second author studied at The Hong Kong University of Science and Technology. The second author would also like to thank Dr. Hailong Zhu for discussion of the techniques in computation.

References

[1] R.B. Bapat, T.E.S. Raghavan, Nonnegative Matrices and Applications, Cambridge University Press, 1997.
[2] R. Bhatia, Matrix analysis, in: Graduate Texts in Mathematics, Springer, 1997.
[3] N. Carkin, H. Wilf, The number of independent sets in a grid graph, SIAM J. Discrete Math. 11 (1) (1998) 54–60.
[4] F. Chatelin, Eigenvalues of Matrices, John Wiley & Sons, New York, 1993.
[5] M. Ciucu, An improved upper bounds for the 3-dimensional dimmer problem, Duke Math. J. 94 (1998) 1–11.
[6] G. Cohen, G. Zemor, Write-isolated memories, Discrete Math. 114 (1993) 105–113.
[7] M. Cohn, On the channel capacity of read/write isolated memory, Discrete Appl. Math. 56 (1995) 1–8.
[8] C.V. Freiman, A.D. Wyner, Optimum block codes for noiseless input-restricted channels, Inf. Control 7 (1964) 398–415.
[9] S. Friedland, P.H. Lundow, K. Markstrom, The 1-vertex transfer matrix and accurate estimation of channel capacity, IEEE Trans. Inf. Theory 56 (8) (2010) 3692–3699.
[10] J. Garloff, Block methods for the solution of linear interval equations, SIAM J. Matrix Anal. Appl. 11 (1) (1990) 89–106.
[11] M. Golin, X. Yong, Y. Zhang, L. Sheng, New upper and lower bounds on the channel capacity of read/write isolated memory, Discrete Appl. Math. 140 (2004) 35–48.
[12] G. Golub, C.van. Loan, Matrix Computations, third ed., The Johns Hopkins University Press, 1996.
[13] Kees A. Immink, Codes for Mass Data Storage Systems, second ed., Shannon Foundation Publishers, Eindhoven, The Netherlands, 2004.
[14] W.H. Kautz, Fibonacci codes for synchronization control, IEEE Trans. Inform. Theory 11 (1965) 284–292.
[15] P. Lundow, Compression of transfer matrices, Discrete Math. 231 (2001) 321–329.
[16] P. Lundow, K. Markstrom, Exact and approximate compression of transfer matrices for graph homomorphisms, London Mathematical Society, J. Comput. Math. 11 (2008) 1–14.
[17] B.N. Parlett, The Symmetric Eigenvalue Problem, Prentice-Hall, 1980.
[18] J.P. Robinson, An asymmetric error-correcting ternary code, IEEE Trans. Inform. Theory 24 (1978) 258–261.
[19] R. Roth, Private communication, 16 February, 2000.
[20] C.E. Shannon, The Mathematical Theory of Communication, University of Illinois Press, Urbana, IL, 1949.
[21] J.K. Wolf, A.D. Wyer, J. Ziv, J. Körner, Coding for write-once memory, AT&T Bell Laboratories Tech. J. 63 (6) (1984) 1089–1112.
[22] X. Yong, M. Golin, New techniques for bounding the channel capacity of read/write isolated memory, in: Poster at the 2002 IEEE Data Compression Conference (DCC2002), 2002, p. 482.