Capacity of Hexagonal Checkerboard Codes
Zhun Deng, Jie Ding, Mohammad Noshad, and Vahid Tarokh

Abstract
In this paper, we propose a new method to bound the capacity of checkerboard codes on the hexagonal lattice. This produces rigorous bounds that are tighter than those commonly known.

Index Terms
Capacity, constrained codes, hexagonal lattice, transfer matrix.

I. INTRODUCTION
CHECKERBOARD codes are two-dimensional binary codes that are designed to satisfy specific constraints [1], [2]. An example is the two-dimensional rectangular binary arrays that satisfy the \((d, k)\) run-length constraint—there are at most \(k\) 0’s and the number of 0’s between any neighboring 1’s is at least \(d\) in each row and column [3–7]. It may naturally arise from data storage on a surface [8], [9]. An example of the two dimensional \((1, \infty)\) run-length constraint is shown in Fig. 1.

In practice, the distance between two data recorded points on a recording device should be no less than a given threshold due to the physical and fidelity constraints. It is of interest to study the checkerboard code on the hexagonal lattice, because it is known that the lattice arrangement of circles with the highest density in two dimensional Euclidean space is the hexagonal packing arrangement, in which the centers of the circles are arranged on the hexagonal lattice [10]. Specifically, we study the capacity of the checkerboard code shown in Fig. 2, where only 0’s can be arranged in the six neighbors of any 1, while both 0 and 1 can be arranged in the six neighbors of any 0 (which is referred to as the hexagonal constraint).

Let \(S_1 \subset S_2 \subset \cdots\) be a sequence of finite subsets of two dimensional hexagonal lattice \(A_2\) such that \(\bigcup_{i=1}^{\infty} S_i = A_2\). Let \(f(S_i)\) \((i \in \mathbb{N})\) be the total number of distinct arrangements of 0’s and 1’s satisfying the hexagonal constraint given a two-dimensional region \(S_i\), and \(|S_i|\) denotes the number of points inside \(S_i\). The capacity of the hexagonal checkerboard code is defined as

\[
C^* = \lim_{i \to \infty} \frac{\log_2 f(S_i)}{|S_i|},
\]

if the limit exists. The subsets \(S_i, i = 1, 2, \cdots\) considered in this work will mainly be the zigzag regions shown in Fig. 3. The existence of the capacity has been established in some prior papers [11], [12] using the subadditivity property of a double sequence.

Weeks et al. [11] studied a similar problem on the integer lattice. They provided rigorous lower and upper bounds on the capacity and produced high precision approximations of using Richardson extrapolation. A detailed discussion on their results is included in Section V. Wilf [13] proposed the transfer matrix method to study the problem of arranging kings in a checkerboard. Calkin et al. [14] further used the transfer matrix to compute the bounds on the capacity of the diamond constraint shown in Fig. 1. However, their method cannot be directly used in calculating the capacity of hexagonal lattice. Inspired by Calkin et al.’s work, we propose a modified transfer matrix method that provides rigorous bounds on the capacity of the hexagonal constraint. The bounds presented in this work are tighter than those calculated by Weeks et al. [11], under the same computational complexity.

II. TRANSFER MATRIX FOR HEXAGONAL LATTICE ARRAYS
A. Notations
Let \(\{F_i\}_{i=0}^{\infty}\) denote the Fibonacci numbers, i.e., \(F_0 = 1, F_1 = 1, \text{ and } F_i = F_{i-1} + F_{i-2}\). For any real-valued matrix \(A\), let \(\sigma_{\text{max}}(A)\), \(A^T\), and \(\text{Trace}(A)\) respectively denote the largest absolute value of the eigenvalues, the transpose, and the sum of all the diagonal elements of \(A\). We represent the inner product of any two real-valued vectors \(x\) and \(y\) of the same size by \(\langle x, y \rangle\). The notations \(1_{s \times r}\) and \(0_{s \times r}\) respectively denote the \(s \times r\) matrices of...
all ones and all zeros. We drop the subscripts $s \times r$ when there is no ambiguity. We note that for any real-valued symmetric $r \times r$ matrix $A$ and positive integer $n$, the following inequality holds

$$\left(\sigma_{\text{max}}(A)\right)^n \geq \frac{\langle 1_{r \times 1}, A^n 1_{r \times 1} \rangle}{\langle 1_{r \times 1}, 1_{r \times 1} \rangle}.$$  \hfill (2)

We provide the following lemma for the future convenience. It follows immediately from the Perron-Frobenius Theorem and Proposition 4.2.1 in [15].

**Lemma 1.** Let $A$ be a real-valued $r \times r$ matrix that is nonnegative and irreducible. Then $A$ has a positive eigenvector $v = [v[1], v[2], \cdots, v[r]]^T$ with corresponding eigenvalue $\sigma_{\text{max}}(A)$. Besides this, the following inequality holds for any positive integer $n$

$$\frac{\epsilon}{d} \left(\sigma_{\text{max}}(A)\right)^n \leq \langle 1, A^n 1 \rangle \leq \frac{rd}{\epsilon} \left(\sigma_{\text{max}}(A)\right)^n$$  \hfill (3)

where $\epsilon = \min\{v[1], v[2], \cdots, v[r]\}$ and $d = \max\{v[1], v[2], \cdots, v[r]\}$.

**B. Transfer Matrix**

Let $G_{m,n}$ denote the zigzag shaped array on the hexagonal lattice shown in Fig. 3 where $m$ is the number of points in each row and $n$ is the number of rows. The row vectors are categorized into two different types: the ones whose first point lies on the left of its upper row (or lower row) are termed as “L” (left) type, while the others belong to “R” (right) type. Assume that the lowest row of $G_{m,n}$ is of “R” type. We emphasize that the vector “100001” of “L” type is different from “100001” of “R” type. Clearly, an “L” type vector cannot be attached to an “L” type vector, while an “R” type vector can be attached to an “L” type vector–either upwards or downwards (see Fig. 3).
Let \( f(m, n) \) denote the total number of valid arrangements of 0’s and 1’s in \( G_{m,n} \) satisfying the hexagonal constraint. The value of (1) for the given set of of regions therefore becomes

\[
C = \lim_{m,n \to \infty} \frac{\log_2 f(m, n)}{mn}.
\]

We will provide lower and upper bounds on \( C \) in the rest part of the paper.

A row vector is called “valid” if none of its two neighboring elements are both ones. Let \( a_m \) denote the number of valid row vectors of length \( m \). Then \( a_m = F_{m+1} \). In fact, if a valid row starts with element zero, the remaining \( m-1 \) elements have \( a_{m-1} \) valid arrangements; if the row starts with one, the second element must be zero and the remaining \( m-2 \) elements have \( a_{m-2} \) valid arrangements. So, we obtain \( a_m = a_{m-1} + a_{m-2} \). Besides, it is easy to see that \( a_1 = 2, a_2 = 3 \), which results in \( a_m = F_{m+1} \). Let \( \mathcal{A}_m = \{ u_i \}_{i=1}^{2a_m} \) represent the collection of all the “L” type and “R” type valid row vectors of length \( m \) (without loss of generality, the “R” type are arranged before the “L” type within \( \mathcal{A}_m \), i.e., \( \{ u_i \}_{i=1}^{a_m} \) belong to “R” type). Clearly, each row of \( G_{m,n} \) belongs to \( \mathcal{A}_m \).

Next, we define the transfer matrix \( T_m \) to be a \( 2a_m \times 2a_m \) binary matrix whose \((i, j)\)th entry is

\[
T_m[i, j] = \begin{cases} 
1 & \text{if } u_j \text{ can be attached to } u_i \text{ (as is depicted in Fig. 3)}, \\
0 & \text{otherwise}
\end{cases}, \quad 1 \leq i, j \leq 2a_m.
\]

Note that \( T_m \) is a symmetric matrix due to its construction. Let \( f(m, n, i) \) denote the number of valid arrangements in \( G_{m,n} \) whose top row is \( u_i \). Then we obtain

\[
f(m, n + 1, j) = \sum_{i=1}^{2a_m} f(m, n, i) T_m[i, j].
\]

Let \( f_{m,n} = [f(m, n, 1), f(m, n, 2), \ldots, f(m, n, 2a_m)] \). Therefore, (6) can be rewritten in the following form

\[
f_{m,n+1} = T_m f_{m,n}, \quad \text{and } f_{m,1} = [1, 1, \ldots, 1, 0, 0, \ldots, 0]^T \quad (a_m 1’s),
\]

from which we get

\[
f(m, n) = \langle 1, T_m^{n-1} f_{m,1} \rangle.
\]

Due to symmetry, \( f(m, n) \) remains the same if the lowest row of \( G_{m,n} \) is “L” type instead of “R” type, i.e.,

\[
f(m, n) = \langle 1, T_m^{n-1} (1 - f_{m,1}) \rangle,
\]

Combining (7) and (8) yields

\[
2f(m, n) = \langle 1, T_m^{n-1} 1 \rangle,
\]

which leads to the following lemma.

**Lemma 2.** For any given positive integer \( m \), we obtain

\[
\lim_{n \to \infty} \frac{\log_2 f(m, n)}{mn} = \lim_{n \to \infty} \frac{\log_2 \left( f(m, n)^{\frac{1}{2}} \right)}{m} = \lim_{n \to \infty} \frac{\log_2 \left( \left( \langle 1, T_m^{n-1} 1 \rangle \right)^{\frac{1}{2}} \right)}{m} = \log_2 \left( \frac{\sigma_{\max}(T_m)}{m} \right).
\]

**Proof.** \( T_m \) is an irreducible matrix since the directed graph it represents is strongly-connected. In fact, we can build a path from one vertex to any other vertex by inserting vertices that represent the all-zero vectors. Furthermore, Lemma 1 implies that there are positive constants \( d \) and \( e \) that depend on \( m \) such that

\[
ed \left( \frac{\sigma_{\max}(T_m)}{m} \right)^{n-1} \leq \langle 1, T_m^{n-1} 1 \rangle \leq \frac{2a_m d}{e} \left( \frac{\sigma_{\max}(T_m)}{m} \right)^{n-1}.
\]

Raising both sides of (11) to the power of \( 1/(n-1) \) and letting \( n \to \infty \), we get

\[
\lim_{n \to \infty} \langle 1, T_m^{n-1} 1 \rangle^{1/(n-1)} = \sigma_{\max}(T_m)
\]

which implies the existence of the limit in (10).

\[\square\]

From Lemma 2 and since \( \lim_{m,n \to \infty} \frac{\log_2 f(m, n)}{mn} \) exists, the iterated limit \( \lim_{m \to \infty} \lim_{n \to \infty} \frac{\log_2 f(m, n)}{mn} \) also exists and they are equal. Therefore, we only need to compute bounds on the iterated limit.
Example 1. If \( m = 3 \), the valid rows are \([0, 0, 0]_R, [0, 0, 1]_R, [0, 1, 0]_R, [1, 0, 0]_R, [1, 0, 1]_R, [0, 0, 0]_L, [0, 0, 1]_L,\)
\([0, 1, 0]_L, [1, 0, 0]_L, [1, 0, 1]_L\), and the transfer matrix is
\[
T_3 = \begin{pmatrix} 0 & A^T \\ A & 0 \end{pmatrix}, \quad \text{where } A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.
\]

III. LOWER BOUNDS ON THE CAPACITY

We denote the shaded part of \( G_{m,n} \) that is formed by concatenating \( n \) column vectors by \( G_{m,n}^{\text{shaded}} \), as shown in Fig.4. Similar to the arguments in Subsection II-B, there are \( a_n = F_{n+1} \) valid column vectors, denoted by \( \{w_i\}_{i=1}^{a_n} \).

We define \( X_n \) to be a \( a_n \times a_n \) binary matrix whose \((i,j)\)th entry is
\[
X_n[i,j] = \begin{cases} 1 & \text{if } w_j \text{ can be attached to } w_i \\ 0 & \text{otherwise} \end{cases}, \quad 1 \leq i, j \leq a_n.
\]

Example 2. If \( n = 3 \), the valid columns are \([0, 0, 0], [0, 0, 1], [0, 1, 0], [1, 0, 0], [1, 0, 1]\), and the transfer matrix is
\[
X_3 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.
\]

One can observe that each row of of \( G_{m,n}^{\text{shaded}} \) is of length \( m - h_n \), where \( h_n = \lceil n/2 \rceil \) (\( \lceil u \rceil \) denotes the largest integer that is no larger than \( u \)). Similar to the derivation of (7), the number of distinct arrangements satisfying the hexagonal constraint in \( G_{m,n} \) is calculated to be \( \langle 1, X_n^{m-h_n-1} \rangle \). Besides this, the number of points in \( G_{m,n} \) but not in \( G_{m,n}^{\text{shaded}} \) is \( nm - n(m-h_n) = nh_n \). Thus, there exists a positive integer \( c_{m,n} \), which depends on \( m, n \) and upper-bounded by \( 2^{nh_n} \), such that
\[
\frac{1}{2} \langle 1, T_n^{m-1} \rangle = f(m, n) = c_{m,n} \frac{1}{2} \langle 1, X_n^{m-h_n-1} \rangle.
\]

For example, when \( n = 5 \) (Fig.4), we have \( h_5 = 2 \) and \( 1 \leq c_5 \leq 2^{10} \).

![Fig. 4: Illustration of the shaded part of \( G_{m,n} \)](image)

Because of (14) and the fact that \( T_m \) is symmetric, for any positive integer \( q \) we obtain
\[
(\sigma_{\text{max}}(T_m))^n \geq \frac{\langle T_q^{m+2q} \rangle \langle T_q^{m+2q} \rangle}{\langle T_q^{m+2q} \rangle} = \frac{\langle 1, T_q^{m+2q} \rangle \langle 1, T_q^{m+2q} \rangle}{\langle 1, T_q^{m+2q} \rangle} = \frac{c_{m,n+2q+1}}{c_{m,2q+1}} \frac{\langle 1, X_n^{m-h_n+2q+1-1} \rangle \langle 1, X_n^{m-h_n+2q+1-1} \rangle}{\langle 1, X_n^{m-h_n+2q+1-1} \rangle}.
\]

As a specific case, we choose \( n = 1, q = 8 \) and obtain
\[
\sigma_{\text{max}}(T_m) \geq \frac{1}{2^{17h_{18}}} \frac{\langle 1, X_n^{m-h_n-1} \rangle}{\langle 1, X_n^{m-h_n-1} \rangle}.
\]
The largest eigenvalues of the transfer matrix $X$ where $X$

Following a similar argument as in the proof of Lemma [2], $X_n$ is irreducible and the following equality holds for any fixed $n > 0$

$$
\sigma_{\text{max}}(X_n) = \lim_{m \to \infty} \langle 1, X_n^m 1 \rangle^{1/m} = \lim_{m \to \infty} \langle e_{m,n}, 1, X_n^{m-h_n-1} 1 \rangle^{\frac{1}{m}}. \quad (16)
$$

We thus obtain

$$
C = \lim_{m \to \infty} \frac{\log_2(\sigma_{\text{max}}(T_m))}{m} \geq \log_2 \lim_{m \to \infty} \left( \frac{1}{2^{1/n_1}} \langle 1, X_{18}^{m-h_{18}-1} 1 \rangle \right)^{\frac{1}{m}} = \log_2 \left( \frac{\sigma_{\text{max}}(X_{18})}{\sigma_{\text{max}}(X_{17})} \right) = 0.4807676144.
$$

The largest eigenvalues of the transfer matrix $X_n$ are listed in Table I for $n = 1, 2, \cdots, 18$.

### IV. Upper Bounds on the Capacity

Since $T_m$ is a real and symmetric matrix, the following inequality holds

$$
\sigma_{\text{max}}(T_m) \leq \left( \text{Trace}(T_m) \right)^{2/n}, \quad (17)
$$

where

$$
\text{Trace}(T_m) = \sum_{1 \leq x_0, x_1, \cdots, x_{2n-1} \leq 2a_m} T_m[x_0, x_1] T_m[x_1, x_2] \cdots T_m[x_{2n-1}, x_0]. \quad (18)
$$

Clearly, the right-hand side of (18) is the total number of valid arrangements of the row vectors $(u_{x_0}, u_{x_1}, \cdots, u_{x_{2n-1}}, u_{x_0})$ in a zigzag way from bottom to top, as shown in Fig. 5. We calculate the sum in (18) by considering the zigzag array as a concatenation of $(2n+1) \times 1$ valid column vectors (Fig. 5). The top element should be the same as the bottom one in each column, so that the first row and the last row become the same. We denote the collection of valid column vectors as $\mathcal{B}_n = \{v_i\}_{i=1}^{b_n}$.

TABLE I: The largest eigenvalue of $X_n$

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\lambda_{\text{max}}$</th>
<th>$n$</th>
<th>$\lambda_{\text{max}}$</th>
<th>$n$</th>
<th>$\lambda_{\text{max}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.618</td>
<td>7</td>
<td>11.526</td>
<td>13</td>
<td>83.112</td>
</tr>
<tr>
<td>2</td>
<td>2.147</td>
<td>8</td>
<td>16.082</td>
<td>14</td>
<td>118.773</td>
</tr>
<tr>
<td>3</td>
<td>3.054</td>
<td>9</td>
<td>22.443</td>
<td>15</td>
<td>165.747</td>
</tr>
<tr>
<td>4</td>
<td>4.233</td>
<td>10</td>
<td>31.319</td>
<td>16</td>
<td>231.297</td>
</tr>
<tr>
<td>5</td>
<td>5.922</td>
<td>11</td>
<td>43.706</td>
<td>17</td>
<td>322.773</td>
</tr>
<tr>
<td>6</td>
<td>8.256</td>
<td>12</td>
<td>60.991</td>
<td>18</td>
<td>450.425</td>
</tr>
</tbody>
</table>

Using an equation similar to (6), we get Trace$(T_m) = \langle 1, B_n^{-1} 1 \rangle$, which leads to

$$
C = \lim_{m \to \infty} \frac{\log_2(\sigma_{\text{max}}(T_m))}{m} \leq \log_2 \left( \lim_{m \to \infty} \left( \text{Trace}(T_m) \right)^{\frac{1}{2n}} \right) = \log_2 \left( \lim_{m \to \infty} \langle 1, B_n^{-1} 1 \rangle^{\frac{1}{2n}} \right) = \frac{\log_2(\sigma_{\text{max}}(B_n))}{2n}. \quad (20)
$$

Fig. 5: Illustration of $B_n$
Theorem 1. The following two limits exist and are equal
\[ \lim_{n \to \infty} \frac{\log_2 \sigma_{\max}(X_n)}{n} = \lim_{m \to \infty} \frac{\log_2 \sigma_{\max}(T_m)}{m}. \]  
\[ (22) \]

Proof. Because of the existence of
\[ 2f(m, n) = \lim_{m, n \to \infty} \frac{\log_2 \langle 1, T_m^{-1}1 \rangle}{mn} = \lim_{m, n \to \infty} \frac{\log_2 \langle c_{m, n} \langle 1, X_n^{m-h_n-1}1 \rangle \rangle}{mn} \]  
and (12), (16), the following iterated limits exist and are both equal to the double limits in (23)
\[ \lim_{m \to \infty} \frac{\log_2 \sigma_{\max}(T_m)}{m} = \lim_{m \to \infty} \frac{\log_2 \langle 1, T_m^{-1}1 \rangle}{mn} \]  
\[ \lim_{n \to \infty} \frac{\log_2 \sigma_{\max}(X_n)}{n} = \lim_{n \to \infty} \frac{\log_2 \langle c_{m, n} \langle 1, X_n^{m-h_n-1}1 \rangle \rangle}{mn}. \]  
\[ (24), (25) \]

They proposed a recursive method to compute \( X_n \) for \( n = 1, 2, \cdots \), and used
\[ \frac{n}{n+1} \log_2 \sigma_{\max}(X_n) \leq C \leq \frac{\log_2 \sigma_{\max}(X_n)}{n} \]  
\[ (26) \] to bound the capacity. However, the convergence becomes slow since the size of the transfer matrix increases as fast as the Fibonacci sequence with \( n \). As the size of the matrix increases, calculating its largest eigenvalue becomes computationally demanding. On the other hand, the approximation is not accurate for small \( n \). Applying \( X_{18} \) to (26) gives the bound \( 0.4640 \leq C \leq 0.4897 \) (which requires the eigenvalues of a matrix of size \( F_{19} = 6765 \)), while our calculated bound (which also involves a matrix of size \( F_{19} \)), \( 0.4807676144 \leq C \leq 0.4813 \), is much tighter. Fig. 6 shows the accuracy of the bounds vs. computational complexity, which is in terms of the size of the largest matrix whose eigenvalues need to be calculated, for both the method used by Weeks and the one derived in this paper.
Fig. 6: A comparison between the two methods: upper and lower bounds on the capacity vs. complexity (the middle two lines are the bounds obtained by our method)

Weeks et al. \[1\] also obtained 0.482644 as an estimate of the capacity using Richardson extrapolation and conjectured that it is correct up to the sixth decimal digit. Unfortunately, that conjecture is not true as this estimate is greater than our rigorous upper bound 0.4813. Halevy et al. \[7, 16\] presented a different estimation which up to nine decimal is 0.480767622. That approximation is consistent with our lower and upper bounds. Besides this, the value given by Halevy et al. and our lower bound (0.480767614) are the same up to 7 decimals.

VI. CONCLUSION

A new method was proposed to derive bounds on the capacity of constrained two-dimensional code on the hexagonal lattice. The methods rely on the zigzag patterns that we chose on the hexagonal lattice in order to obtain nonnegative and irreducible transfer matrices.

REFERENCES

Zhun Deng is a Ph.D. candidate in the School of Engineering and Applied Sciences, Harvard University. His current research are in statistical signal processing, information theory and fuzzy mathematics.

Jie Ding is a Ph.D. candidate in the School of Engineering and Applied Sciences, Harvard University. His current research are in cyclic difference sets, time series, information theory.

Mohammad Noshad is a postdoctoral fellow in the Electrical Engineering Department at Harvard University. He received his PhD in Electrical Engineering from the University of Virginia. He is a recipient of the “Best Paper Award” at IEEE Globecom 2012. His research interests include information and coding theory, statistical machine learning, free-space optical communications, and visible light communications.

Vahid Tarokh is a professor of applied mathematics in the School of Engineering and Applied Sciences, Harvard University. His current research interests are in data analysis, network security, optical surveillance, and radar theory.