New Bounds for Codes Identifying Vertices in Graphs

Gérard Cohen
cohen@inf.enst.fr

Iiro Honkala
honkala@utu.fi

Antoine Lobstein
lobstein@inf.enst.fr

Gilles Zémor
zemor@infres.enst.fr

Abstract

Let $G = (V, E)$ be an undirected graph. Let C be a subset of vertices that we shall call a code. For any vertex $v \in V$, the neighbouring set $N(v, C)$ is the set of vertices of C at distance at most one from v. We say that the code C *identifies* the vertices of G if the neighbouring sets $N(v, C), v \in V$, are all nonempty and different. What is the smallest size of an identifying code C ? We focus on the case when G is the two-dimensional square lattice and improve previous upper and lower bounds on the minimum size of such a code.

AMS subject classification: 05C70, 68R10, 94B99, 94C12.

Submitted: February 12, 1999; Accepted: March 15, 1999.

G. Cohen, A. Lobstein and G. Zémor are with ENST and CNRS URA 820, Computer Science and Network Dept., Paris, France, I. Honkala is with Turku University, Mathematics Dept., Turku, Finland

1 Introduction

In this paper, we investigate a problem initiated in [3]: given an undirected graph $G = (V, E)$, we define $B(v)$, the *ball* of radius one centered at a vertex $v \in V$, by

$$B(v) = \{x \in V : d(x, v) \leq 1\},$$

where $d(x, v)$ represents the number of edges in a shortest path between v and x. The vertex v is then said to *cover* all the elements of $B(v)$. We often refer to a distinguished subset C of V as a *code*, and to its elements as *codewords*.

A code C is called a *covering* if the sets $B(v) \cap C$, $v \in V$, are all nonempty; if furthermore they are all different, C is called an *identifying code*. The set of codewords covering a vertex v is called the *identifying set* (I-set) of v.

Now, what is the minimum cardinality of an identifying code ? This problem originates in [3] and is also taken up in [1].

Let us mention an application. A processor network can be modeled by an undirected graph $G = (V, E)$, where V is the set of processors and E the set of their links. A selected subset C of the processors constitutes the code. Its codewords report to a central controler the state of their neighbourhoods (typically, balls of radius one) by sending one bit of information (e.g., 1 if it does not contain a faulty processor, 0 otherwise). Based on these $|C|$ bits, the controler must locate the faulty processor. Common network architectures are the n-cube or the two-dimensional mesh or grid.

In this paper we focus on the case when G is a square grid drawn on a torus, that is G is the graph $\mathbb{T}_{nm}$ with vertex set $V = \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$ and edge set $E = \{\{u, v\} : u - v = (\pm 1, 0) \text{ or } u - v = (0, \pm 1)\}$. We shall also consider the limiting infinite case, i.e. when G is the graph $\mathbb{T}$ with vertex set $\mathbb{Z} \times \mathbb{Z}$. The *density* $D(C)$ of $C \subseteq V$ is defined as $|C|/|V|$ for $\mathbb{T}_{nm}$ and for the infinite graph $\mathbb{T}$ as

$$D(C) = \limsup_{n \to \infty} \frac{|C \cap Q_n|}{|Q_n|}$$

where Q_n is the set of vertices $(x, y) \in V$ such that $|x| \leq n$ and $|y| \leq n$.

An example of an identifying code of $\mathbb{T}$ is given in figure 1. It is taken from [3] and its density is 3/8. Our purpose is to determine the minimum density D of an identifying code of $\mathbb{T}$. It is proved in [3] that $1/3 \leq D \leq 3/8$. We shall improve this to

$$\frac{23}{66} \leq D \leq \frac{5}{14}.$$

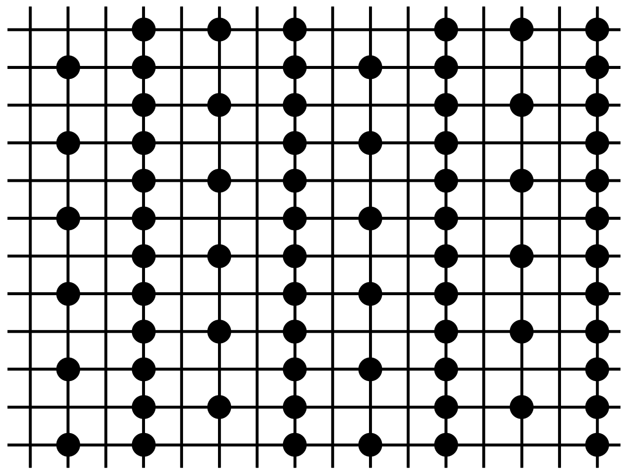


Figure 1: The pattern is periodic and extends to $\mathbb{Z}^2$ with density 3/8.

2 Lower bounds

For a given finite regular graph $G = (V, E)$, let $B = |B(v)|$ denote the size (independent of its centre) of a ball of radius one; let C be an identifying code. Since C is a covering of V, the *sphere-covering bound* holds:

$$|C| \cdot B \geq |V|.$$

But the identifying property implies a strictly better bound : let L_1 denote the set of vertices identified by singletons; now $|V| - |L_1|$ vertices have I-sets of size at least two. In other words, C is a double covering (see [2, Ch. 14]) of these vertices; thus, using the fact that $|L_1| \leq |C|$, we have:

$$|C| \cdot B \geq 2(|V| - |L_1|) + |L_1| = 2|V| - |L_1| \geq 2|V| - |C|.$$

We obtain, [3]

$$|C| \cdot \frac{B+1}{2} \geq |V|. \tag{2.1}$$

Bound (2.1) can be tight in some graphs, for example the triangular lattice, see [3].

2.1 The graphs $\mathbb{T}_{nm}$

Until the end of this section G will be a finite torus $\mathbb{T}_{nm}$ with $n, m \geq 30$, say. All balls of radius one have cardinality five. For $i = 1, 2, 3, 4, 5$, let L_i be the set of vertices identified by a set of exactly i codewords. Set $\ell_i = |L_i|$, $L_{\geq 3} = L_3 \cup L_4 \cup L_5$ and

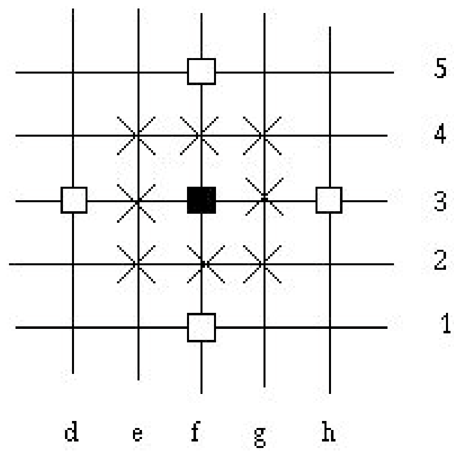


Figure 2: An element of C'.

$\ell_{\geq 3} = |L_{\geq 3}|$. Counting in two ways the number of couples (c, x) such that $c \in C$, $x \in V$ and $d(c, x) \leq 1$, we get:

$$5|C| = \sum_{1 \leq i \leq 5} i\ell_i. \tag{2.2}$$

From (2.2), we infer that $5|C| = \ell_1 + 2(|V| - \ell_1 - \ell_{\geq 3}) + 3\ell_{\geq 3} + \ell_4 + 2\ell_5$. Since $\ell_1 \leq |C|$, we obtain:

$$6|C| \geq 2|V| + \ell_{\geq 3} + \ell_4 + 2\ell_5. \tag{2.3}$$

If it were possible that $\ell_{\geq 3} = 0$ then the bound (2.3) would collapse to (2.1). But this is not the case for the square grids and for the rest of this section we shall bound $\ell_{\geq 3}$ from below as tightly as we can.

2.2 Partitioning C

We partition the code C into two subcodes C' and C'', with C'' consisting of all codewords belonging to at least one I-set of cardinality at least three. Thus, C' is the set of all codewords belonging only to I-sets of size one or two. Our strategy will be to bound $\ell_{\geq 3}$ from below by a function of $|C'|$. First, some facts about C' and C''.

In G, any vertex $c' \in C'$ has the neighbouring configuration of figure 2, where the black square represents c', a white square represents an element of C, and a cross represents a vertex not in C.

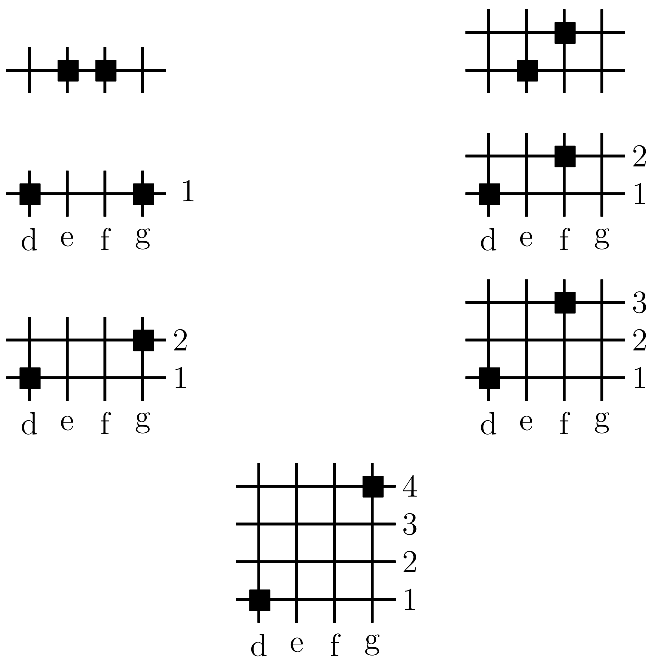


Figure 3: Forbidden configurations of two elements of C'.

Indeed, suppose that a codeword $c \in C$ is on $e3$; then, in order to give c' and c distinct I-sets, c' should belong to an I-set of size at least three. If $c \in C$ is on $e2$, then, in order to give $e3$ and $f2$ distinct I-sets, again c' must belong to an I-set of size at least three. This contradicts the definition of C'. Finally, $d3$, $f1$, $f5$ and $h3$ belong to C because $e3$, $f2$, $f4$ and $g3$ must have an I-set which is not reduced to $\{c'\}$. Actually, using similar arguments, it is easy to check (see figure 3) that two elements of C' cannot be at Euclidean distance 3 (e.g., on $d1$ and $g1$), $\sqrt{5}$ (on $d1$ and $f2$), $\sqrt{10}$ (on $d1$ and $g2$), $2\sqrt{2}$ (on $d1$ and $f3$), and even $3\sqrt{2}$ (on $d1$ and $g4$) from one another.

Obviously, we have $3\ell_3 + 4\ell_4 + 5\ell_5 \geq |C''|$, i.e.,

$$3\ell_{\geq 3} + \ell_4 + 2\ell_5 \geq |C''|. \tag{2.4}$$

Let $\ell_4 = \alpha\ell_{\geq 3}$, $\ell_5 = \beta\ell_{\geq 3}$ (with $\alpha, \beta, \alpha + \beta \in [0, 1]$). Then

$$\ell_{\geq 3} \geq \frac{|C''|}{3 + \alpha + 2\beta}.$$

Combining with (2.3), this leads to

$$6|C| \geq 2|V| + |C''|(1 - \frac{2}{3 + \alpha + 2\beta}).$$

The right hand side is smallest when $\alpha = \beta = 0$, hence

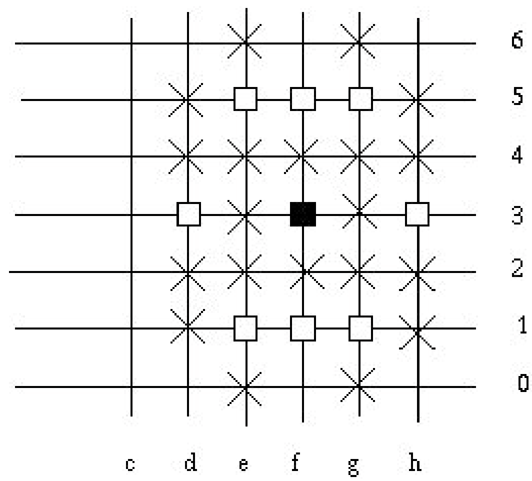


Figure 4: An element of C' with degree two in Γ.

Lemma 2.1 $6|C| \geq 2|V| + |C''|/3$. $\triangle$

2.3 An incidence relation between C' and $L_{\geq 3}$

For any vertex v, let $R(v)$ be the set of points at Euclidean distance either 2 or $\sqrt{5}$ from v. Now let us consider the bipartite graph Γ whose set of vertices is $C' \cup L_{\geq 3}$, and whose set of edges is included in $C' \times L_{\geq 3}$, with an edge between $c' \in C'$ and $x \in L_{\geq 3}$ if and only if $x \in C \cap R(c')$. We now study possible degrees in Γ.

Lemma 2.2 *Any element of C' has degree at least two in Γ.*

Proof. Consider again figure 2. To identify $e4$, we can assume, without loss of generality, that there is a codeword in $e5$. Since $e5$ and $f5$ must have distinct I-sets, at least one of them must have at least a third element in its I-set. The same is true for $f1$ and $g1$, or $h2$ and $h3$, according to which place you choose for covering $g2$. Actually, the only way for $c' \in C'$ to have degree exactly two is given by figure 4 (or its rotation). $\triangle$

Lemma 2.3 *Any element of $L_{\geq 3}$ has degree at most three in Γ.*

Proof. Assume that a codeword x in $L_{\geq 3}$ has degree four: four distinct codewords c'_1, c'_2, c'_3, and c'_4 of C' are adjacent to x in Γ. For each i, $c'_i \in R(x)$, because $x \in R(c'_i)$, and figure 5 shows, with black squares, the twelve possible locations for the four c'_i's around x; figure 5 also gives the two possible ways of identifying the vertex x on $f3$

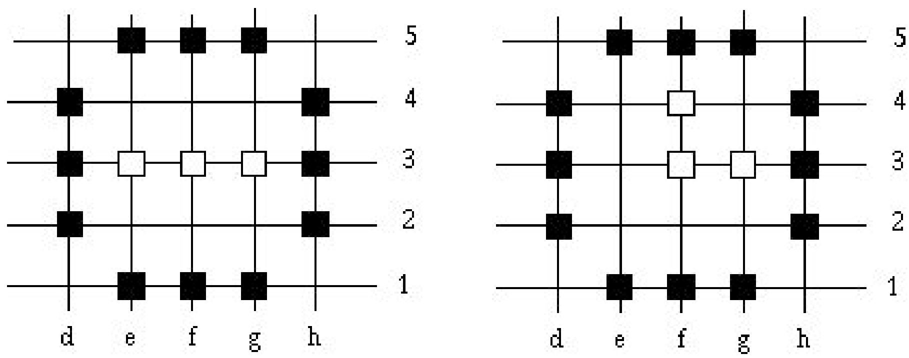


Figure 5: $R(x)$, the set of possible locations for elements of C'.

with three codewords, represented as white squares (more elements in the I-set of x would only mean more restrictions on the c_i''s). Now, keeping in mind figure 2 and the forbidden configurations of figure 3 it is not difficult to check that choosing four c_i''s among these twelve positions is impossible, and furthermore that figure 6 gives the only possible configurations with three elements of C' in $R(x)$ (this will help in proving our following lemma). △

Lemma 2.4 *If an element of $L_{\geq 3}$ has degree three in Γ, then at least two of its neighbours in Γ have degree at least four.*

Proof. Let us consider Configuration (b) of figure 6. There is necessarily a codeword on $f7$, in order to identify $f6$. The points $f2$ and $f4$ have different I-sets, so there is a codeword on $e2$. So in Γ we have the edges $(e5, f7)$, $(e5, f3)$, $(e5, e3)$; $(g5, f7)$, $(g5, f3)$; $(g1, f3)$, $(g1, e2)$. Now in order to cover $d6$ and $d4$, we must increase the degree of $e5$, and this will do nothing for the covering of $h6$, $h4$, $h2$, $f0$ and $h0$. For $h4$ and $h6$ we have two possibilities. Either we do not take $h3$ as a codeword: this allows the degree of $g5$ to increase by one only (if we take $i4$ and $i6$ as codewords). But then the covering of $h2$, $f0$ and $h0$ requires an increase of the degree of $g1$ of at least two, and in the best case we end up with degrees four, three and four for $e5, g5$ and $g1$, respectively. Or we take $h3$ in C: now $g3$ is in $L_{\geq 3} \cap C$ and the degrees of $g5$ and $g1$ both increase. The covering of $h6$, $f0$ and $h0$ will necessarily lead to another increase, and we end up with degrees at least four in Γ.

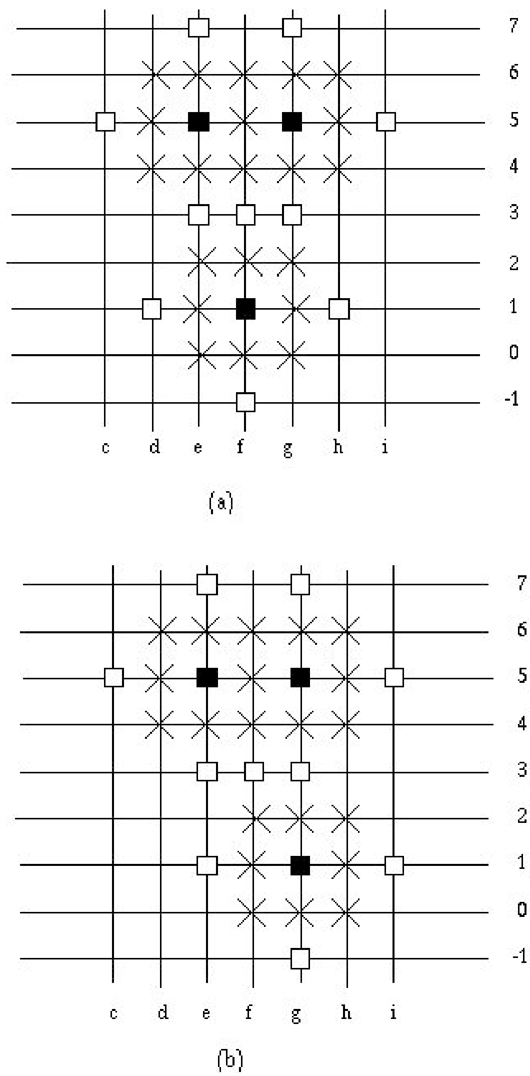


Figure 6: Possible locations for three elements of C' in $R(x)$.

In Configuration (a) of figure 6, there must also be a codeword on $f7$, so the two elements of C', $e5$ and $g5$, have $f7$ and $f3$ as neighbours in Γ. We now prove that $g5$ has at least two more edges in Γ; by symmetry, the same will be true for $e5$, proving our lemma.

Because $h6$ must be covered, $h7$ or $i6$ are in C. If $h7 \in C$, then the fact that $h4$ has to be covered gives the claim. Assume that $i6 \in C$. Since $h4$ must be covered, $h3$ or $i4$ belong to C. If $h3 \in C$, we are done. If $i4 \in C$ and $h3 \notin C$, then $i3 \in C$, because $h3$ and $g2$ must have distinct I-sets.

In all cases, $g5$ has degree at least four in Γ. △

Corollary 2.5 $\ell_{\geq 3} \geq |C'|$.

Proof. We partition $L_{\geq 3}$ into two sets, A and B: A is the set of vertices with degree exactly three in Γ and B is the set of vertices with degree at most two in Γ. We partition C' into two sets, X and Y: X contains the vertices having degree two or three in Γ and Y contains the vertices having degree at least four in Γ. Let a, b, c and d be the number of edges between X and A, X and B, Y and A, Y and B, respectively. Counting in different ways the edges of Γ, we obtain:

$$c + d \geq 4|Y|,\ a + b \geq 2|X|,\ a + c = 3|A|,\ b + d \leq 2|B|,$$

or

$$4|Y| - d \leq c = 3|A| - a \tag{2.5}$$

and

$$2|X| - a \leq b \leq 2|B| - d. \tag{2.6}$$

This leads to $4|C'| \leq 3|A| + 4|B| + a - d$. But Lemma 2.4 implies

$$a \leq |A|. \tag{2.7}$$

Therefore, $4|C'| \leq 4\ell_{\geq 3} - d \leq 4\ell_{\geq 3}$. △

We will now improve on this last result by showing that X and B cannot be both made up only of vertices of degree two in Γ.

2.4 A refined analysis of the degrees in Γ

Let us further partition the sets X and B: let C'_2 and C'_3 be the subsets of X with vertices of degree two and three in Γ, respectively; let B_0, B_1, and B_2 be the subsets of B containing vertices of degree zero, one, and two in Γ, respectively.

We study the elements of C'_2 and start from figure 4. Because $d2, d3$ and $d4$ must have distinct I-sets, we see that at least one of $c2$ and $c4$ must belong to C: we can assume, by symmetry, that $c4 \in C$. Then $c3$ or $c2$ are in C, and $c3 \in L_{\geq 3}$.

Case A: $c3 \notin C$. It implies that $c2 \in C$ and $c3$ has degree zero in Γ.
Case B: $c3 \in C$. What degree can $c3$ have in Γ? There are only four possible places for elements of C' around $c3$: $a2$, $a3$, $a4$ and $c1$. Keeping in mind the forbidden distances between two elements of C', it is easy to check that there are three possibilities: 1) $c3$ has degree zero in Γ; 2) $c3$ has degree one in Γ, and any of these four places is possible; 3) $c3$ has degree two in Γ and necessarily $a4 \in C'$ (the other neighbour of $c3$ in Γ being $a2$ or $c1$).
Case B1: $c3$ has degree zero in Γ.
Case B2: $c3$ has degree one in Γ.
Case B3: $c3$ has degree two in Γ. This implies that $a4 \in C'$ (and $c1$ or $a2$ is in C').
Case B3a: $c5 \in C$. This implies that $c4 \in L_{\geq 3} \cap C$; moreover, $c4$ has degree one in Γ, $a4$ being its only neighbour.
Case B3b: $c5 \notin C$. This implies that $b6 \in C$ (to cover $b5$) and $d6 \in C$ (because $e4$ and $d5$ have distinct I-sets). The vertex $e6$ is not a codeword, and, since its I-set is different from that of $d5$, $e6 \in L_{\geq 3}$, with degree zero in Γ.

In these five cases, we have exhibited a vertex with degree zero or one in Γ. Of course, each time, a second one exists in a symmetric position, on column g or i.

Now we gather Cases A and B3b, which generated elements of $L_{\geq 3} \setminus C$ (of degree zero in Γ); and Cases B2 and B3a, which generated codewords of degree one in Γ. Case B1 has produced a codeword with degree zero in Γ. The point is to see how many elements of C'_2 could produce the **same** vertex. Then we can have an estimate on the number of elements which have degree zero or one in Γ, thus improving the inequality linking $|C'|$ and $\ell_{\geq 3}$.

We give a sketch only for Cases A and B3b. The other cases are very similar. The following remark will be useful: two elements of C'_2 cannot be at distance two from each other.

In Case A (resp., B3b), we produced an element of $L_{\geq 3} \setminus C$, $c3$ (resp., $e6$), at Euclidean distance 3 (resp., $\sqrt{10}$) from our starting point $f3 \in C'_2$. In Case A, apart from $f3$, the only possible location for an element of C'_2 at Euclidean distance 3 from $c3$ is $z3$. In Case B3b, apart from $f3$, the only possible locations for an element of C'_2 at Euclidean distance $\sqrt{10}$ from $e6$ are $d9$ and $f9$, but, using our preliminary remark, at most one is possible. One "crossing" between Case A and Case B3b can occur only when there is an element of C'_2 on $e9$, which excludes $d9$ and $f9$. So in this case, one vertex with degree zero in Γ is shared by at most two elements of C'_2.

In Cases B2 and B3a, one vertex with degree one is shared by at most two elements of C'_2. In case B1, at most two elements of C'_2 generate the same vertex of degree zero.

Since, by symmetry, one element in C'_2 produces two vertices with degree zero or one in Γ, we have shown:

Lemma 2.6 $|B_0| + |B_1| \geq |C'_2|$. △

Now, following (2.6), we have $2|C'_2|+3|C'_3|-a = b \leq 2|B_2|+|B_1|-d$, or $3|X|-|C'_2| \leq 2|B_2| + |B_1| + a - d$. By the previous lemma, this implies that

$$3|X| \leq 2|B_2| + 2|B_1| + |B_0| + a - d = 2|B| - |B_0| + a - d.$$

Thus

$$3|X| \leq 2|B| + a - d, \tag{2.8}$$

which improves on (2.6) and, together with (2.5) and (2.7), leads to

$$4|C'| \leq 3|A| + \frac{8}{3}|B| + \frac{1}{3}a - \frac{1}{3}d \leq \frac{10}{3}|A| + \frac{8}{3}|B| - \frac{1}{3}d \leq \frac{10}{3}\ell_{\geq 3},$$

and we have just proved:

Lemma 2.7 $\ell_{\geq 3} \geq 6|C'|/5$. △

Corollary 2.8 $6|C| \geq 2|V| + 6|C'|/5$.

Proof. By (2.3), $6|C| \geq 2|V| + \ell_{\geq 3} \geq 2|V| + 6|C'|/5$. △
Since $|C'| + |C''| = |C|$, Lemma 2.1 and the above corollary yield:

$$66|C| \geq 23|V|. \tag{2.9}$$

By letting the two dimensions of $\mathbb{T}_{mn}$ grow to infinity, we obtain

Theorem 2.9 *The minimum density of an identifying code of the infinite square lattice* $\mathbb{T}$ *satisfies* $D \geq 23/66$. △

Remark : more detailed study of the possible degrees in Γ can lead to small improvements in the lower bound. For example, further refining the above argument can lead to the condition $d \geq a$ which gives $\ell_{\geq 3} \geq 4|C'|/3$ and $D \geq 15/43 \approx 23/66 + 0.00035$ (see [4]). But analysis of the above type tends to become more and more intricate and the improvements to the lower bound less and less significant.

3 A new construction

Consider the pattern of figure 7. This is an alternative construction to figure 1. One readily checks that it makes up an identifying code of density 3/8. Notice that it can be modified to yield the construction of figure 8 with the same density. But this identifying code is not optimal. Codewords can be deleted without losing the identifying property. We obtain the code of figure 9. Hence :

Theorem 3.1 *The minimum density of an identifying code of the infinite square lattice* $\mathbb{T}$ *satisfies* $D \leq 5/14$*.* $\triangle$

References

[1] U. Blass, I. Honkala and S. Litsyn: Bounds on identifying codes, *Discrete Math.*, to appear.

[2] G. D. Cohen, I. Honkala, S. Litsyn and A. Lobstein: *Covering Codes*, Elsevier, 1997.

[3] M. G. Karpovsky, K. Chakrabarty and L. B. Levitin: On a new class of codes for identifying vertices in graphs, *IEEE Trans. Inform. Th.*, vol. 44, pp. 599–611, 1998.

[4] http://www.infres.enst.fr/~lobstein/unpublished.html

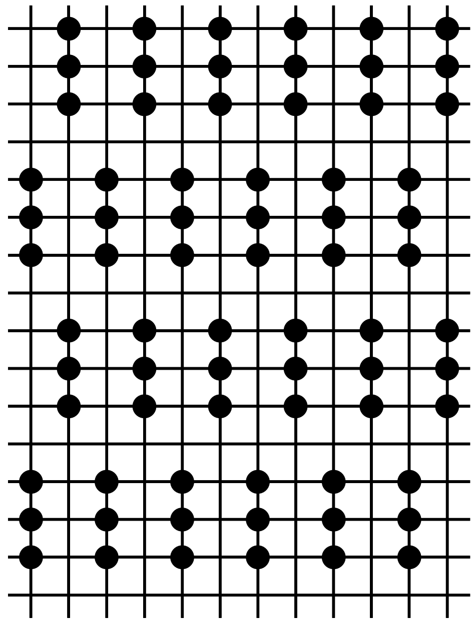


Figure 7: An alternative periodic identifying code of density 3/8.

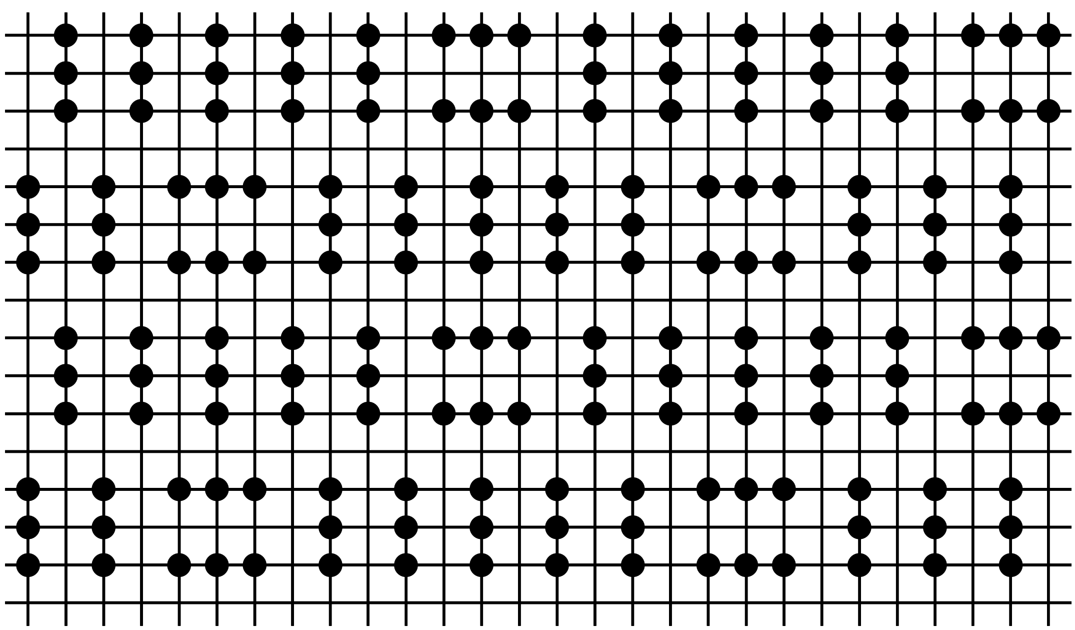
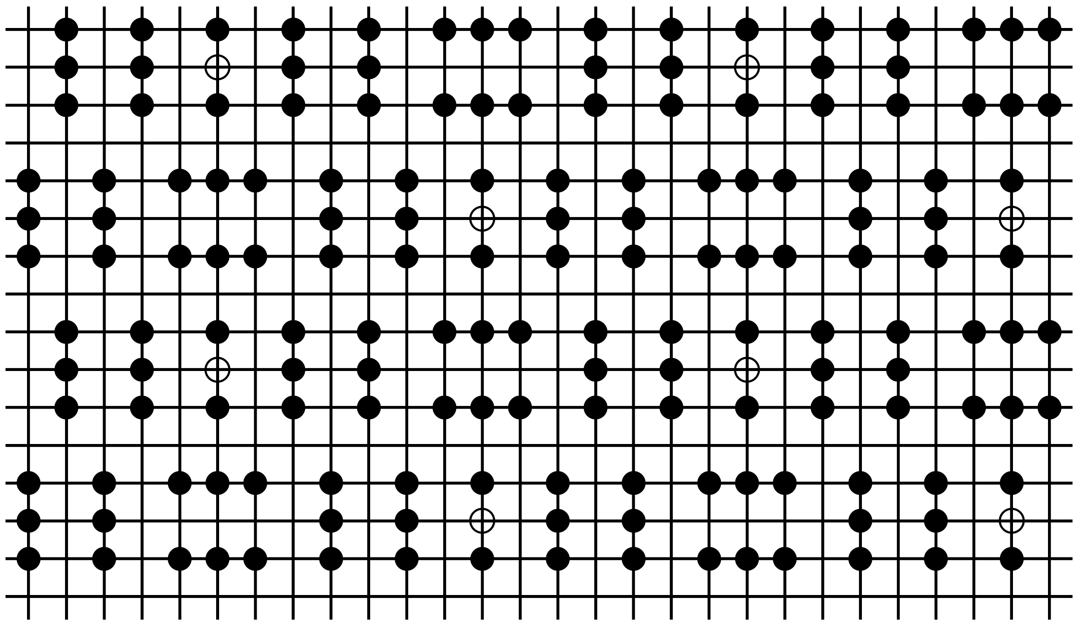


Figure 8: Another periodic identifying code of density 3/8.



The eight white codewords in the picture can be deleted without losing the identifying property. We obtain a periodic tiling of $\mathbb{Z}^2$ by the tile below.

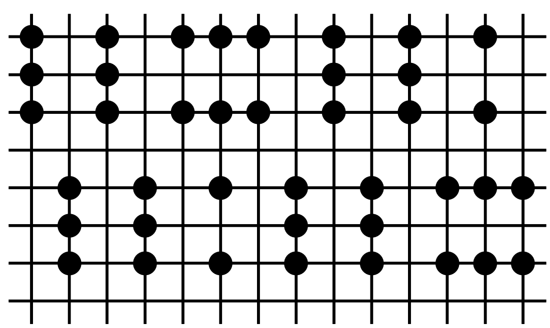


Figure 9: The improved identifying code : the tile is of size 112 and contains 40 codewords. Hence the density $40/112 = 5/14$.