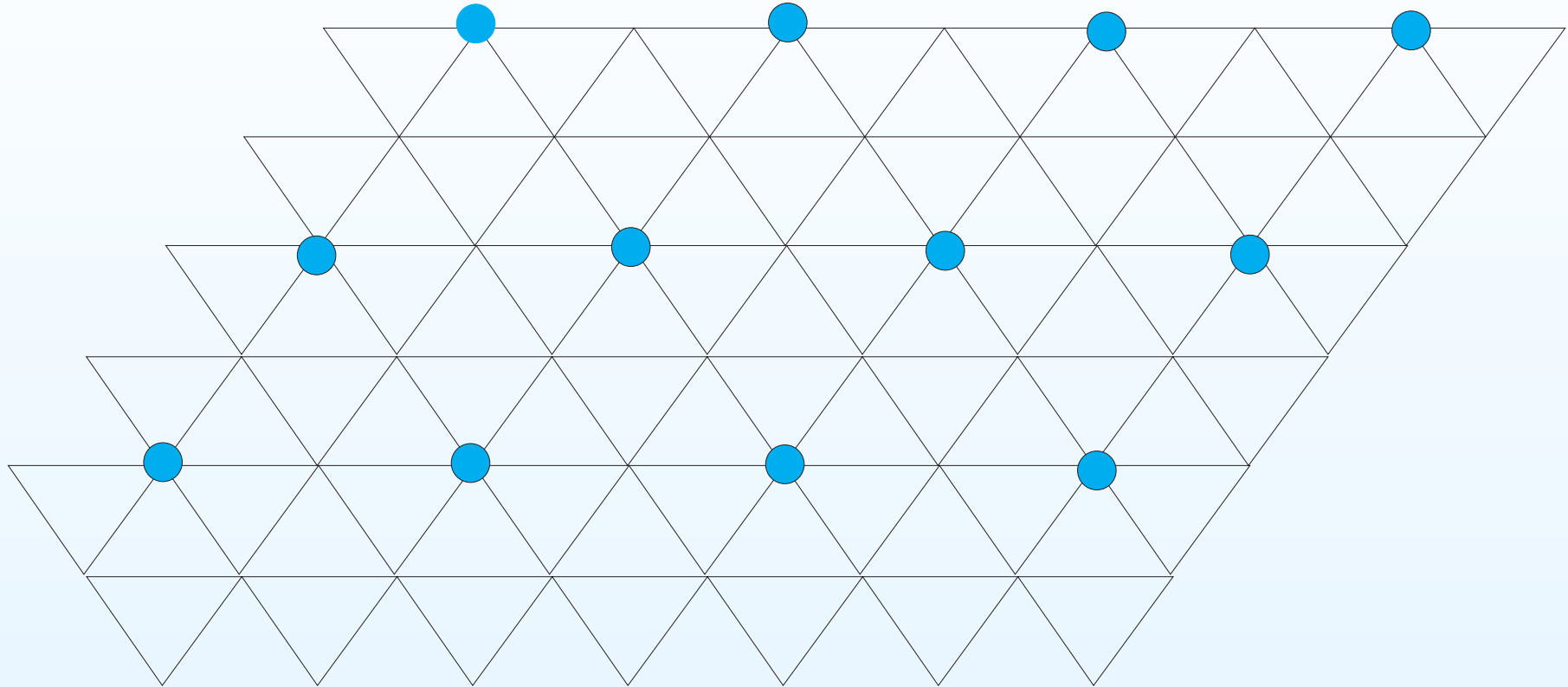


# *Codes for locating sets of vertices in a graph*

Tero Laihonen

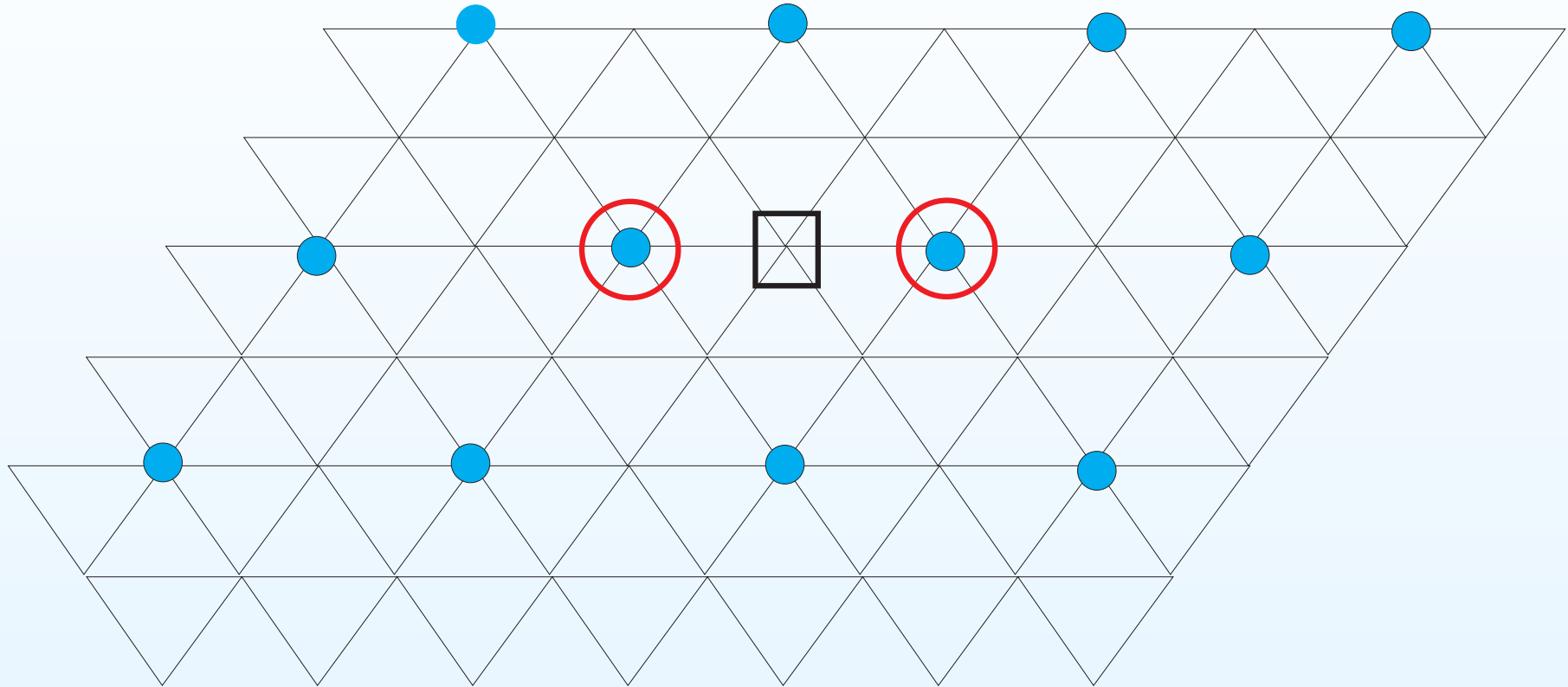
Department of Mathematics, University of Turku, Finland

# Locating many objects



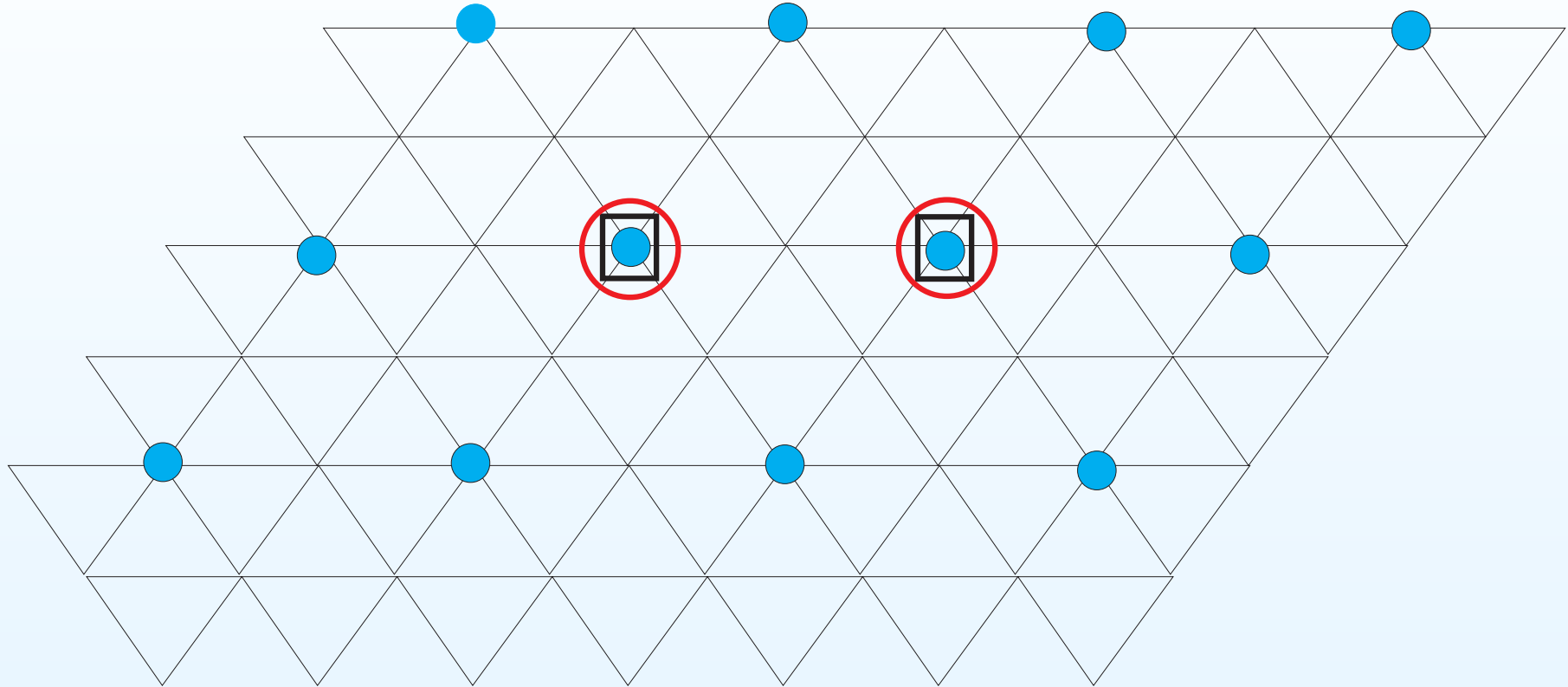
An optimal 1-identifying code

## Locating many objects



One object (fire, intruder, faulty processor) to be identified

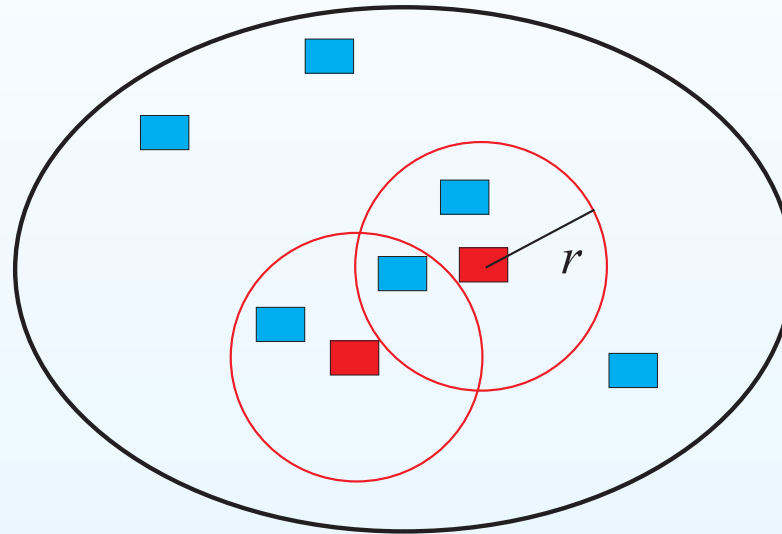
# Locating many objects



Two objects leads to a problem

## Identifying several objects :

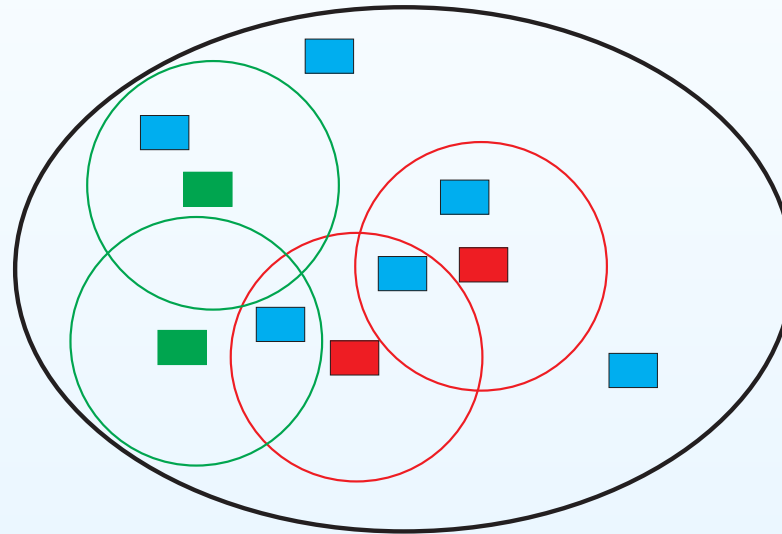
A more demanding class of codes introduced also by Karpovsky *et al.*



Consider codewords in a union of balls.

## Identifying several objects :

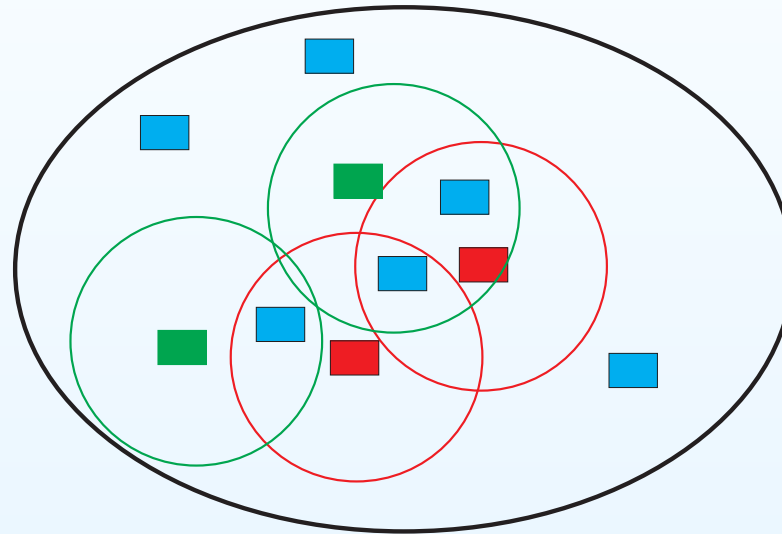
A more demanding class of codes introduced also by Karpovsky *et al.*



Again codewords in a union must be distinct.

## Identifying several objects :

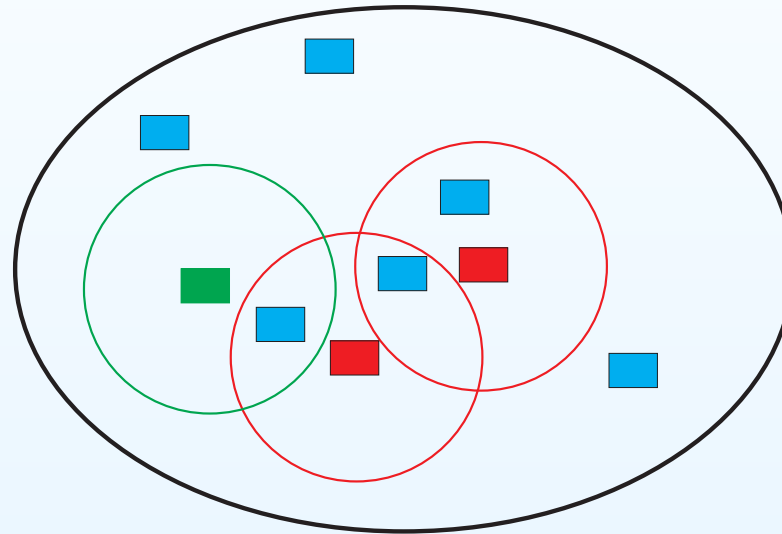
A more demanding class of codes introduced also by Karpovsky *et al.*



A violation of the property.

## Identifying several objects :

A more demanding class of codes introduced also by Karpovsky *et al.*



Also single vs a pair.



## Definitions:

- A code  $C \subseteq V$  (a set of sensors).
- For a set of objects  $X \subseteq V$ ,

$$I_r(X) = \bigcup_{x \in X} B_r(x) \cap C.$$

- A code is  $(r, \leq \ell)$ -*identifying* if

$$I_r(X) \neq I_r(Y)$$

for all  $X \neq Y$  where  $|X|, |Y| \leq \ell$ .

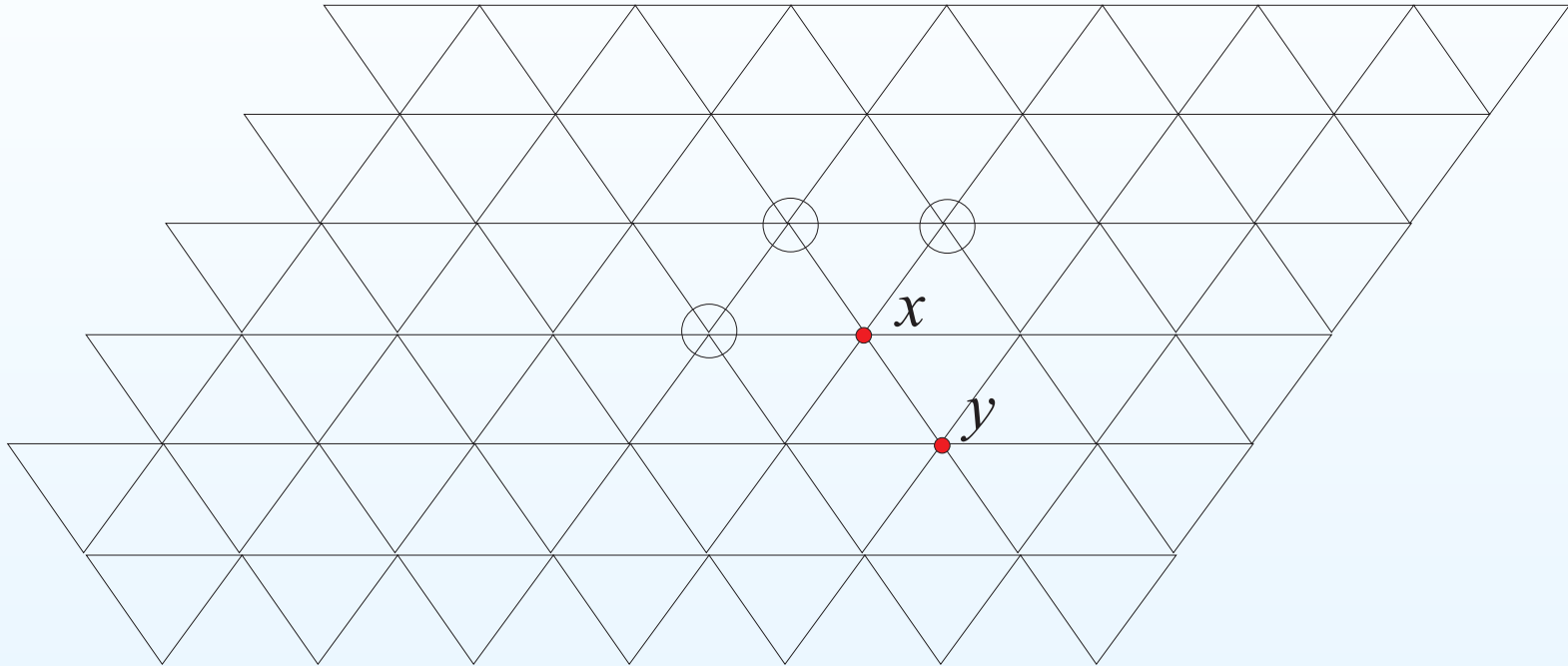
- The aim again is to find the smallest cardinality of such a code.

# Content

---

- Triangular and King grids for  $\ell = 2$ .
- Binary hypercubes for  $\ell \geq 2$ .
- Universally identifying codes
- Weakly identifying codes
- Two generalizations of locating-dominating sets

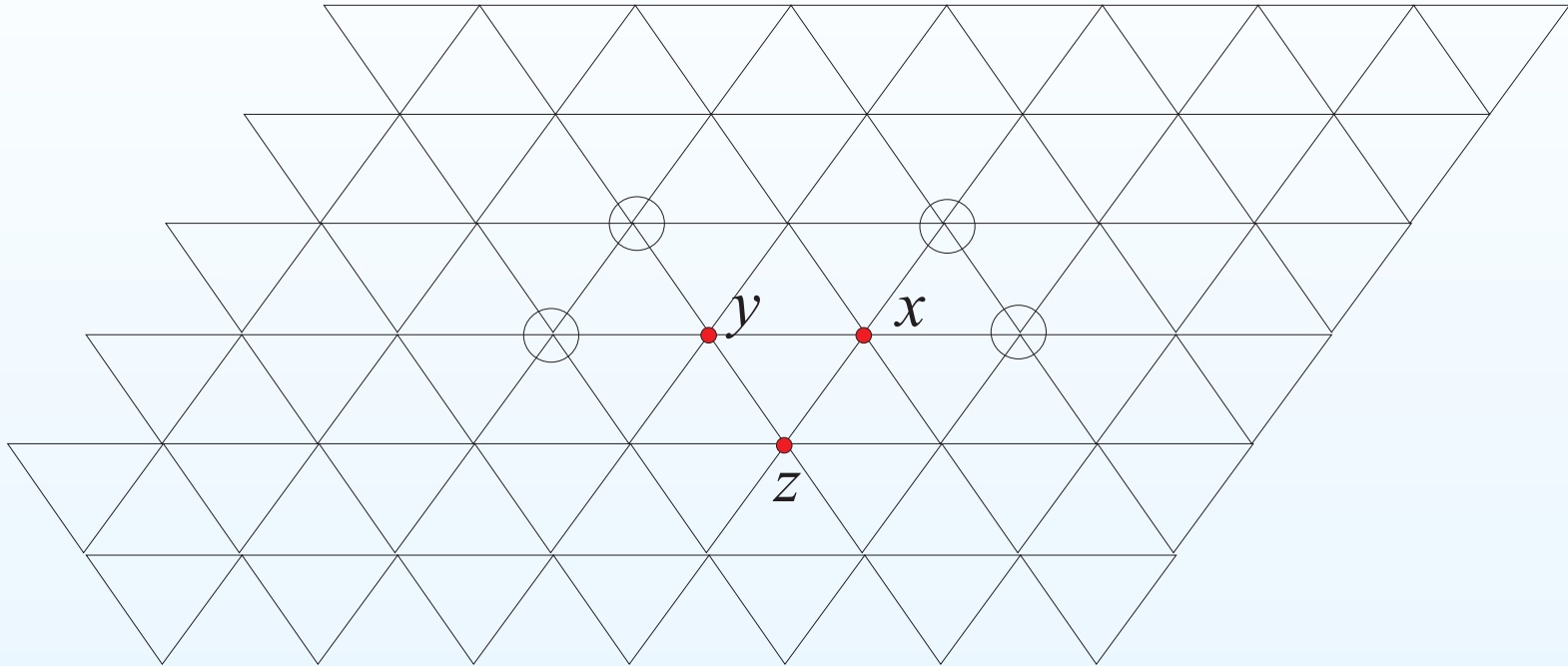
## Triangular grid and $\ell = 2$



Pattern (A): We must have  $I(y) \neq I(x, y)$ .

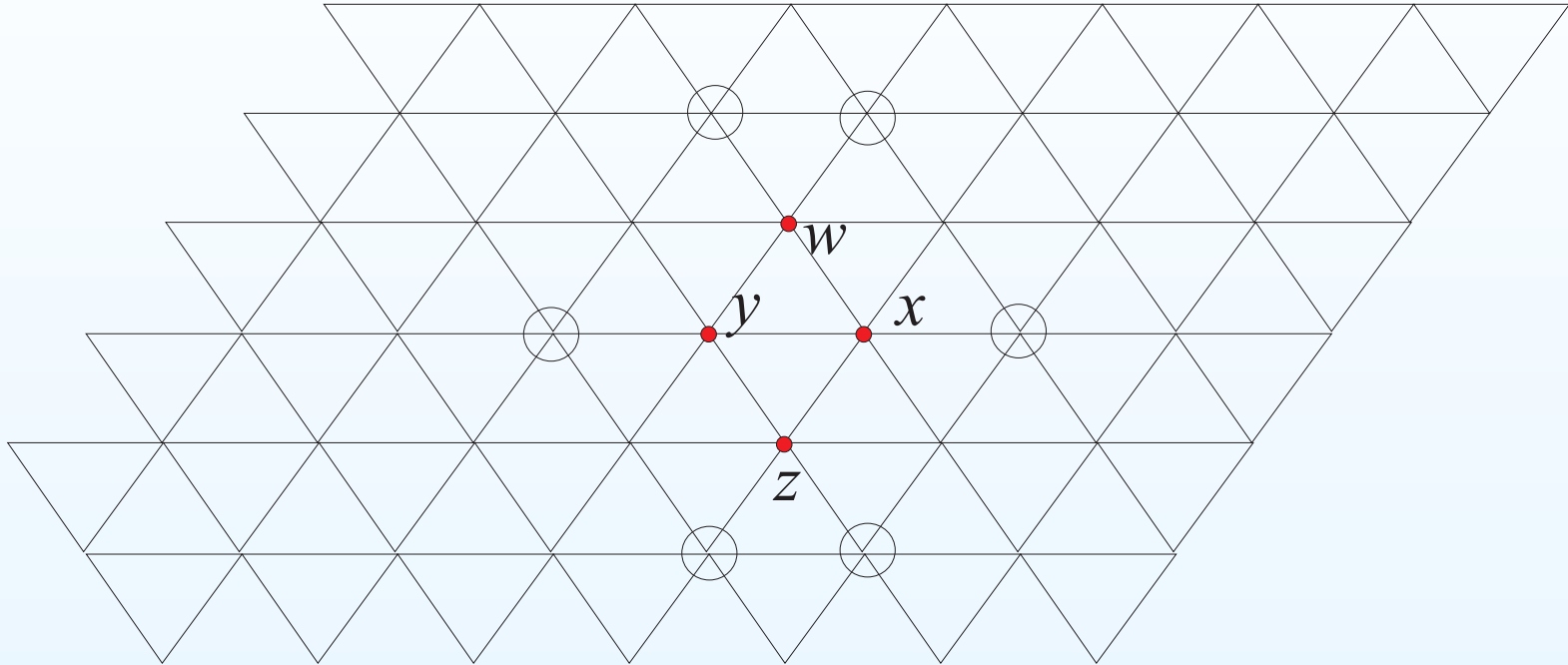
The symmetric difference of  $B(y)$  and  $B(x, y)$ .

## Triangular grid and $\ell = 2$



Pattern (B): We must have  $I(z, x) \neq I(z, y)$ .  
Together with the rotations and translations

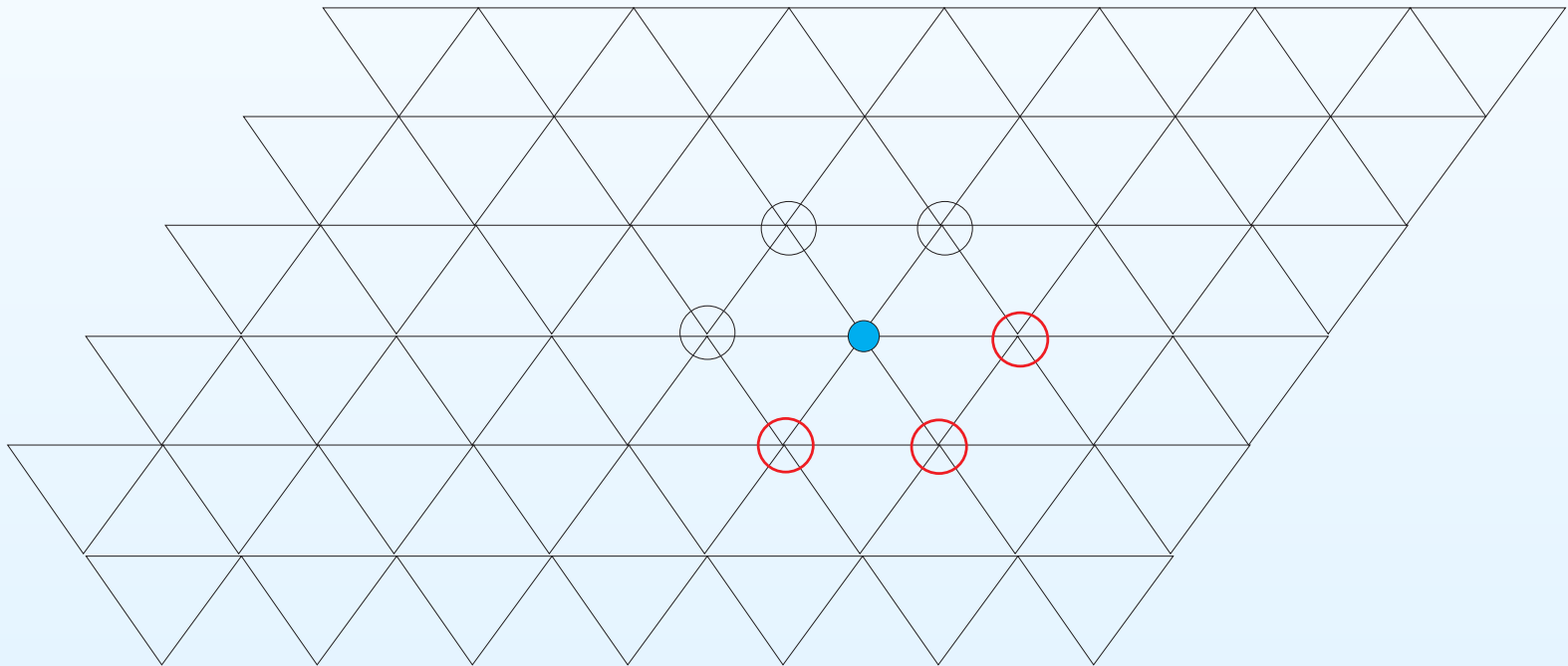
## Triangular grid and $\ell = 2$



Pattern (C): We must have  $I(x, y) \neq I(z, w)$ .

## Triangular grid and $\ell = 2$

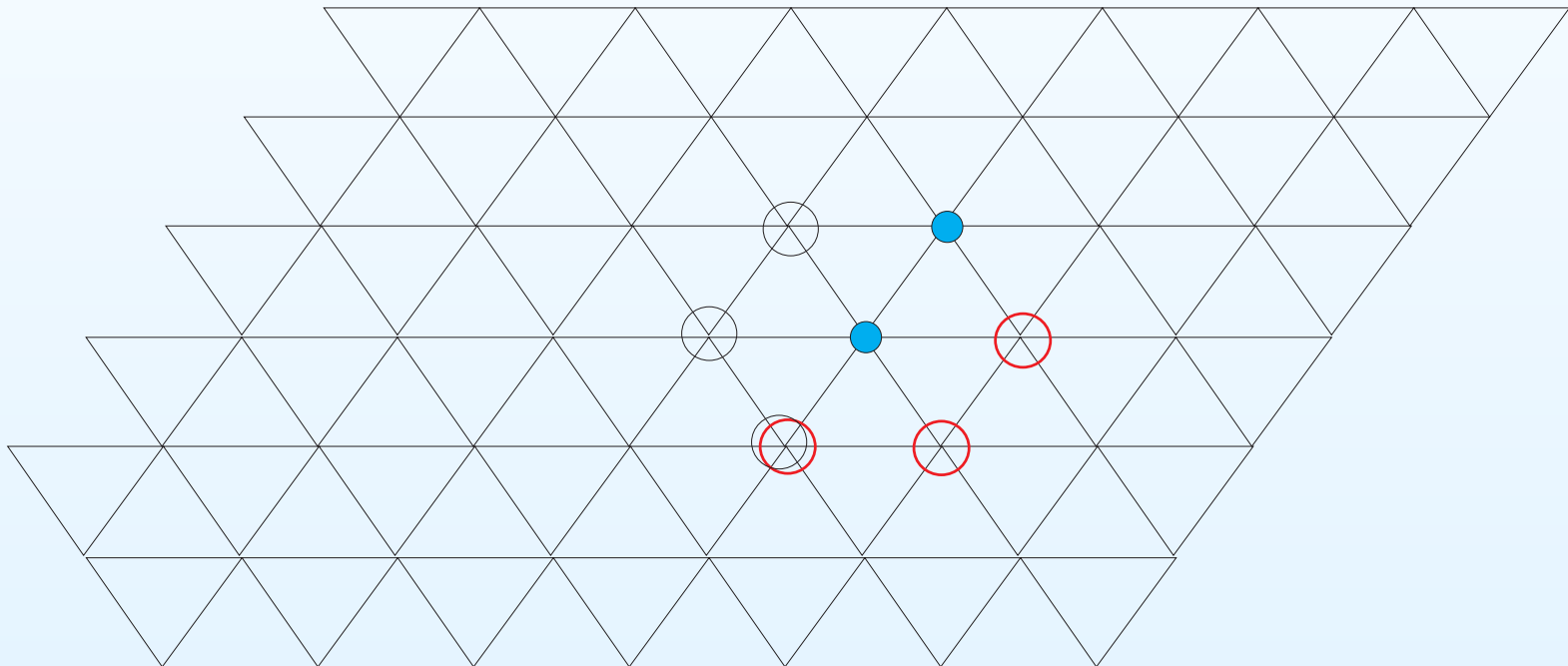
- A code  $C$  intersects every pattern (A), (B) and (C) if and only if  $C$  is  $(1, \leq 2)$ -identifying.
- The pattern (A) already gives useful information:



Pattern (A) gives  $|I(c)| \geq 3$ .

## Triangular grid and $\ell = 2$

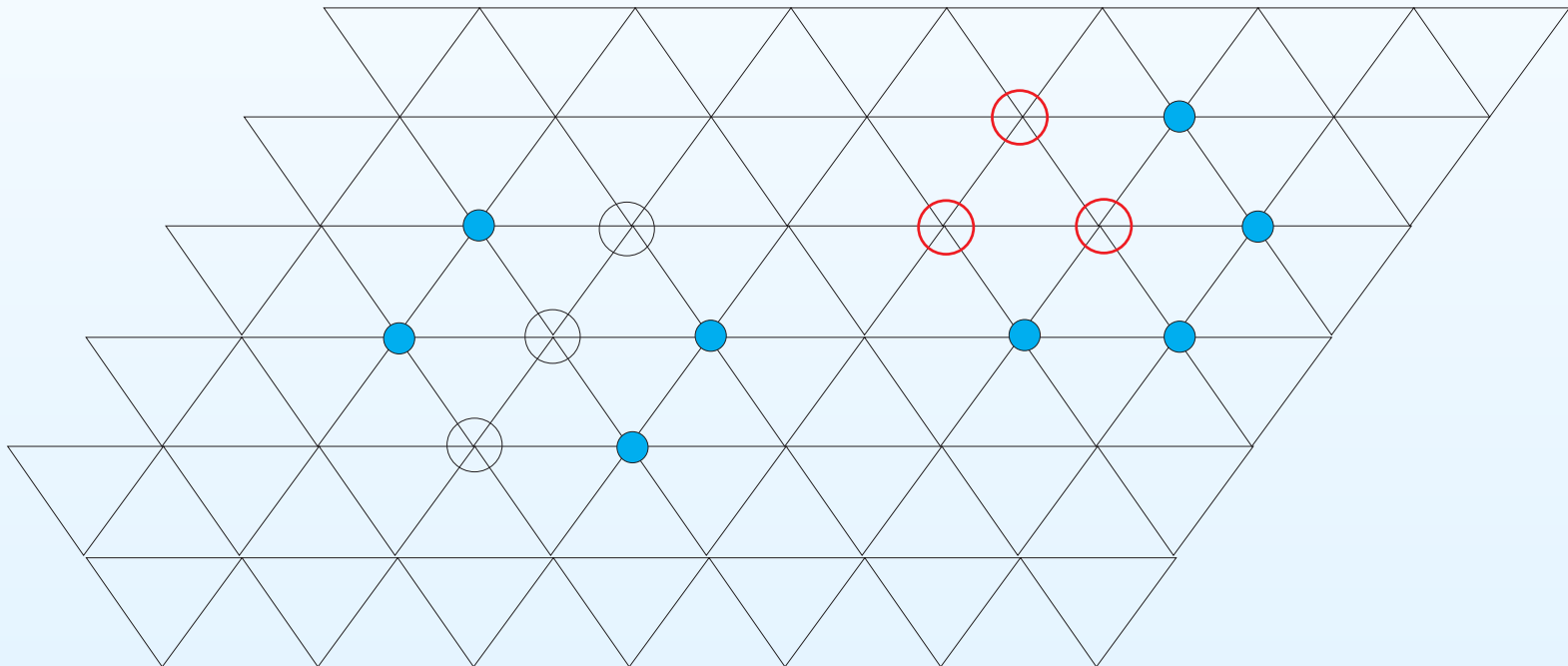
- A code  $C$  intersects every pattern (A), (B) and (C) if and only if  $C$  is  $(1, \leq 2)$ -identifying.
- The pattern (A) already gives useful information:



The only choice for  $|I(c)| = 3$ .

## Triangular grid and $\ell = 2$

- A code  $C$  intersects every pattern (A), (B) and (C) if and only if  $C$  is  $(1, \leq 2)$ -identifying.
- The pattern (A) already gives useful information:



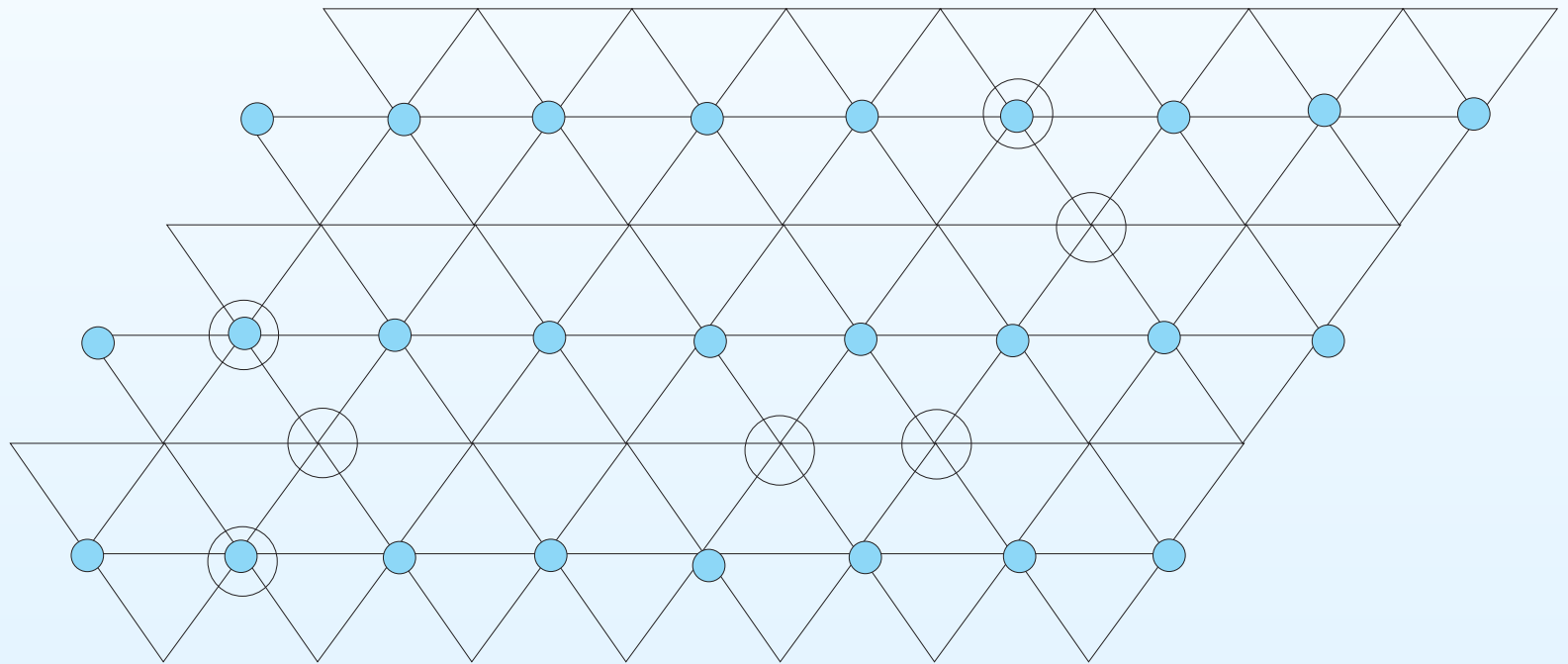
Two choices for  $|I(x)| = 4$  when  $x \notin C$ .



## Triangular grid and $\ell = 2$

- Patterns (A) and (B) would give density  $1/2$  where

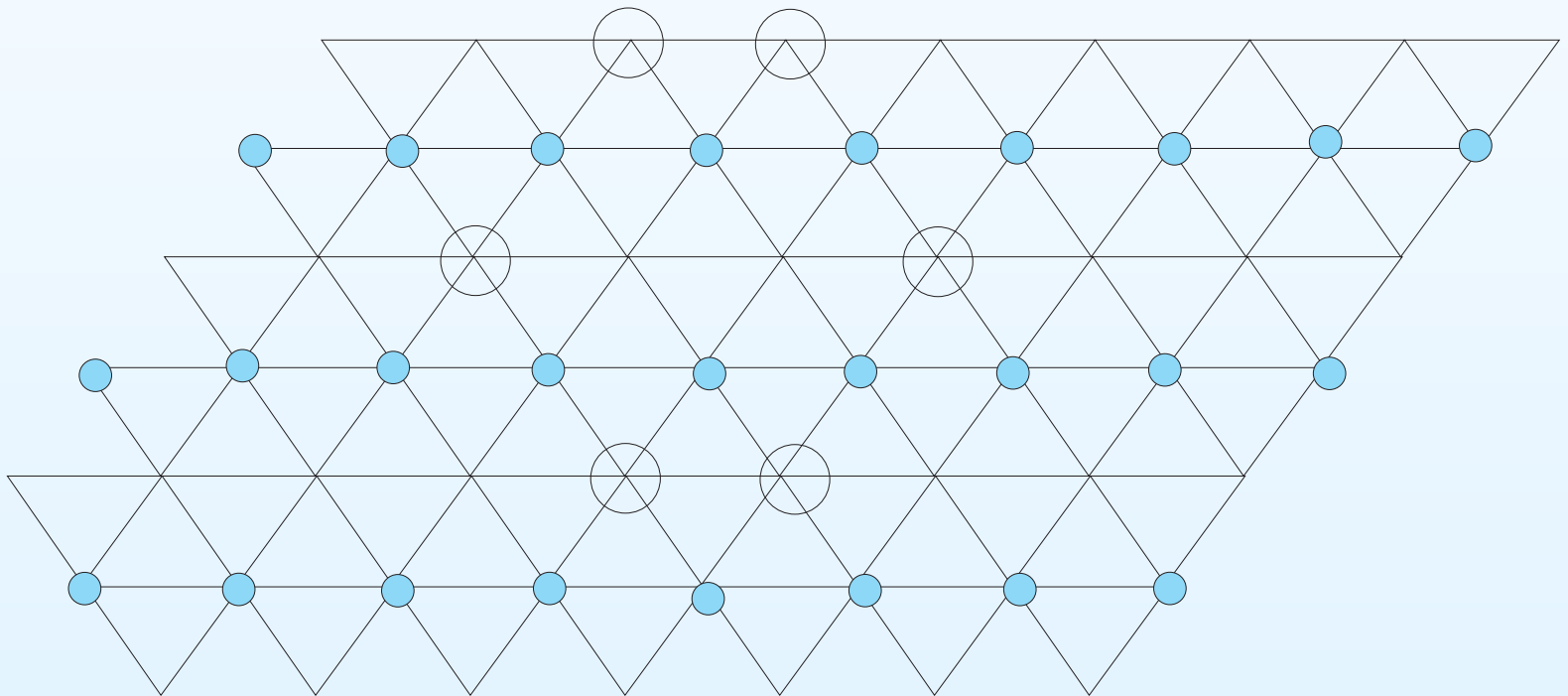
$$D = \limsup_{n \rightarrow \infty} \frac{|C \cap Q_n|}{|Q_n|}.$$



## Triangular grid and $\ell = 2$

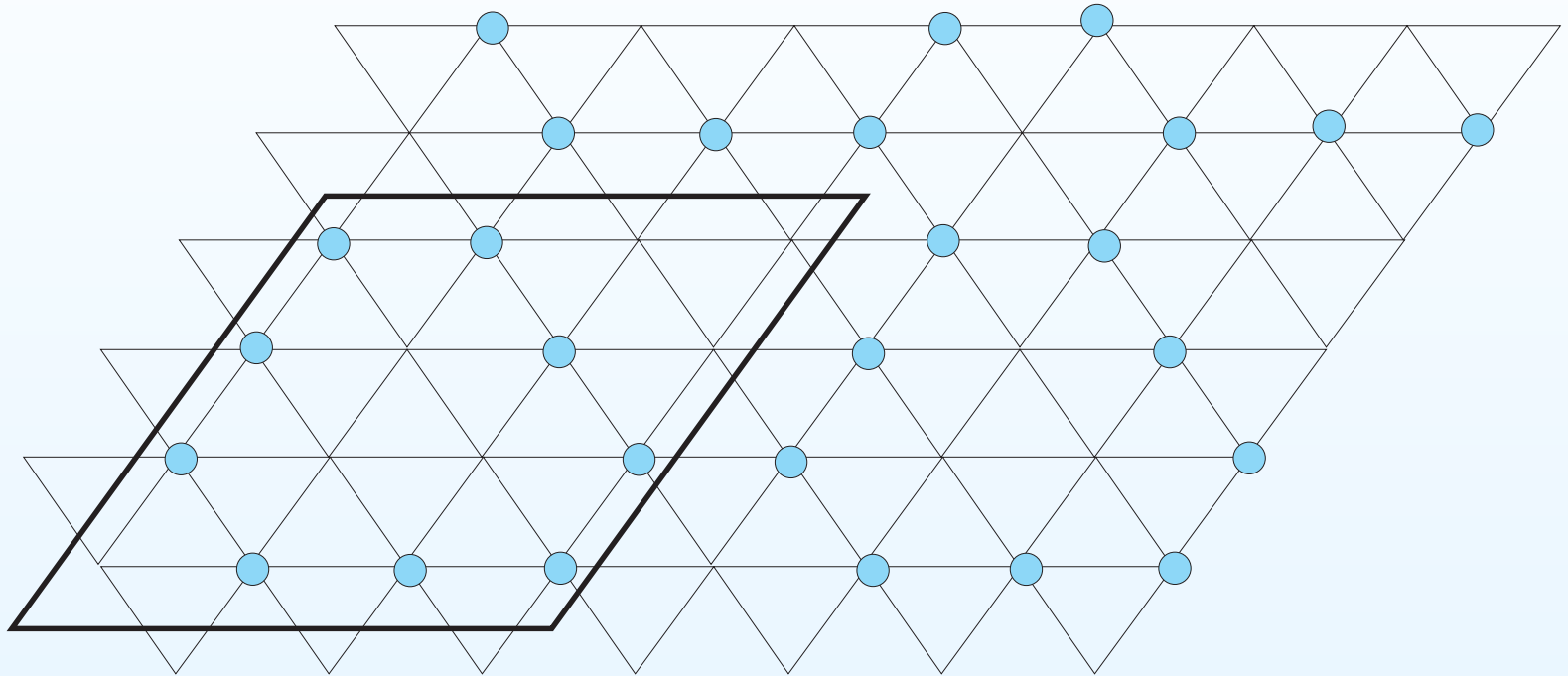
- Patterns (A) and (B) would give density  $1/2$  where

$$D = \limsup_{n \rightarrow \infty} \frac{|C \cap Q_n|}{|Q_n|}.$$



## Triangular grid and $\ell = 2$

- The Pattern (C) changes the game a lot!



## Triangular grid and $\ell = 2$

---

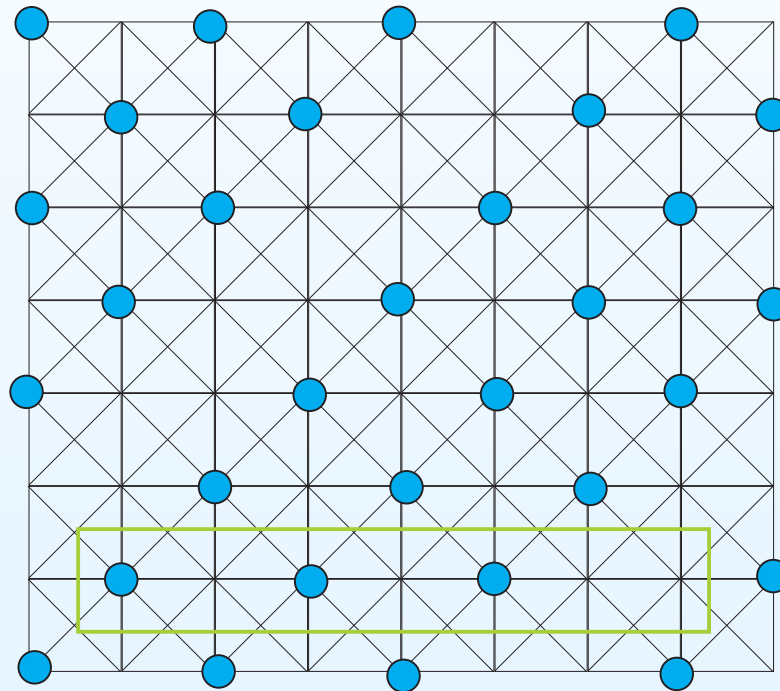
- Optimal density is  $9/16$  (L.- Honkala, 2004)
- The proof is based on showing that every vertex is covered in average by

$$\frac{63}{16} \approx 3.9375$$

codewords.

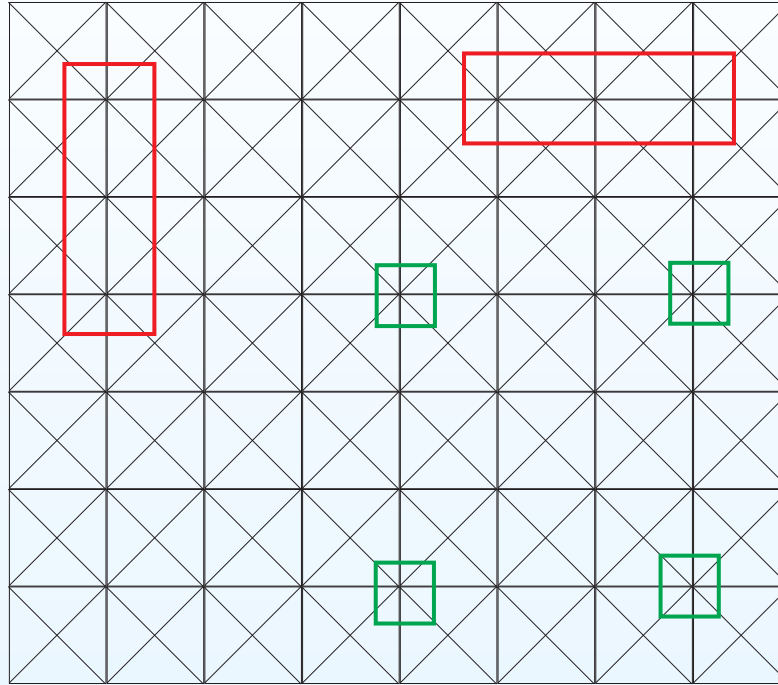
## King grid $\ell = 2$ .

- $V = \mathbb{Z}^2$  and adjacent if  $d(x, y) \leq \sqrt{2}$ .
- There is a code with density  $3/7 \approx 0.4286\dots$  (Honkala- L.).



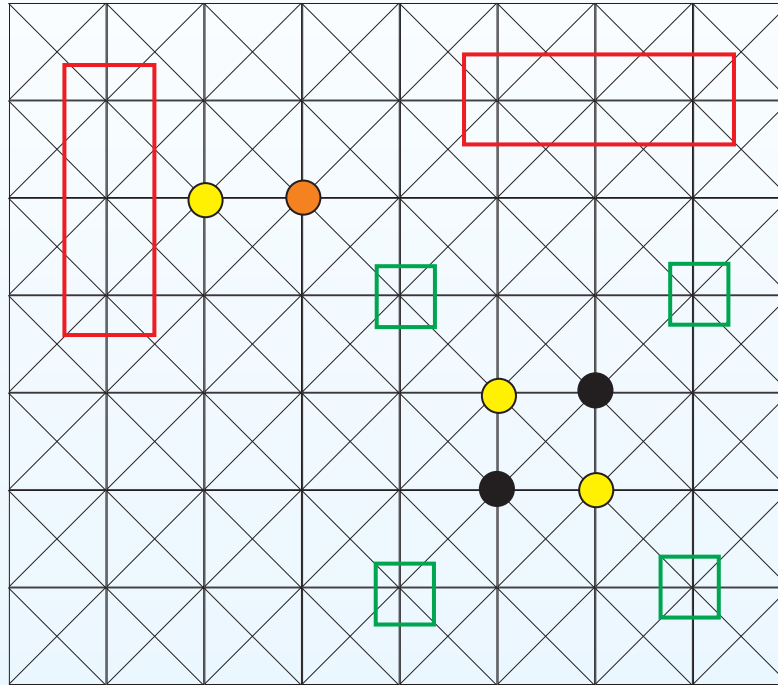
- A lower bound  $5/12 \approx 0.4166\dots$  by M. Peltó (2010).

## King grid $\ell = 2$ .



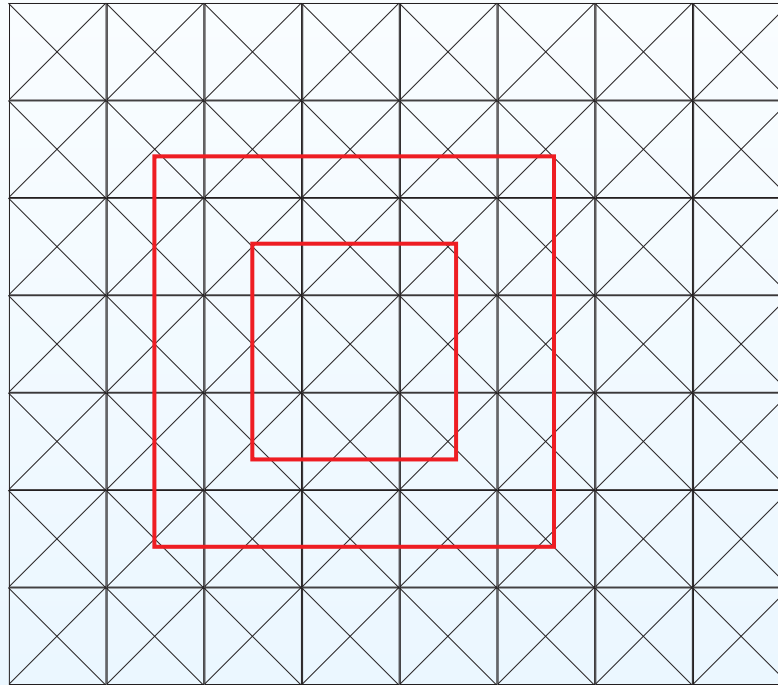
A code  $C$  is  $(1, \leq 2)$ -identifying if and only if the following patterns intersect with  $C$ .

# King grid $\ell = 2$ .



Where do they come from?

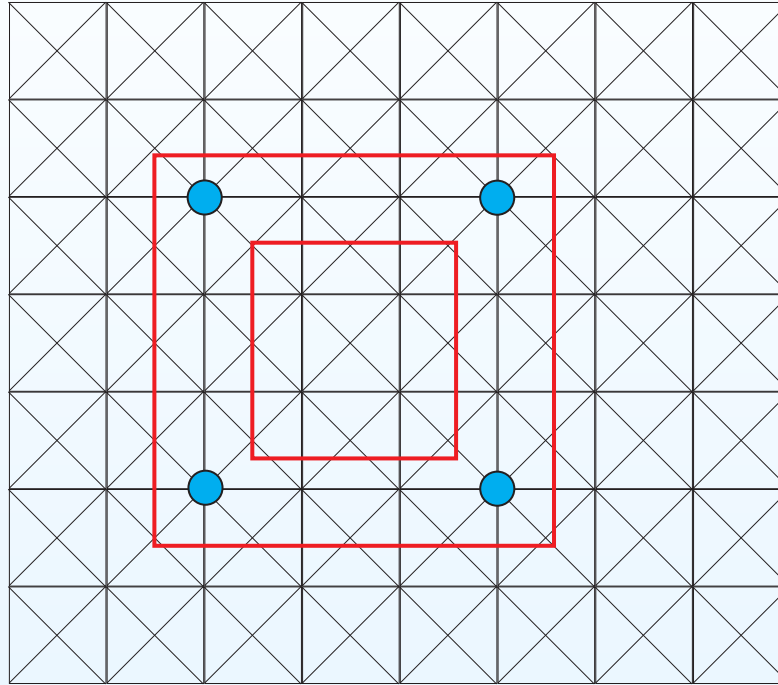
King grid  $\ell = 2$ .



The frame with 12 vertices

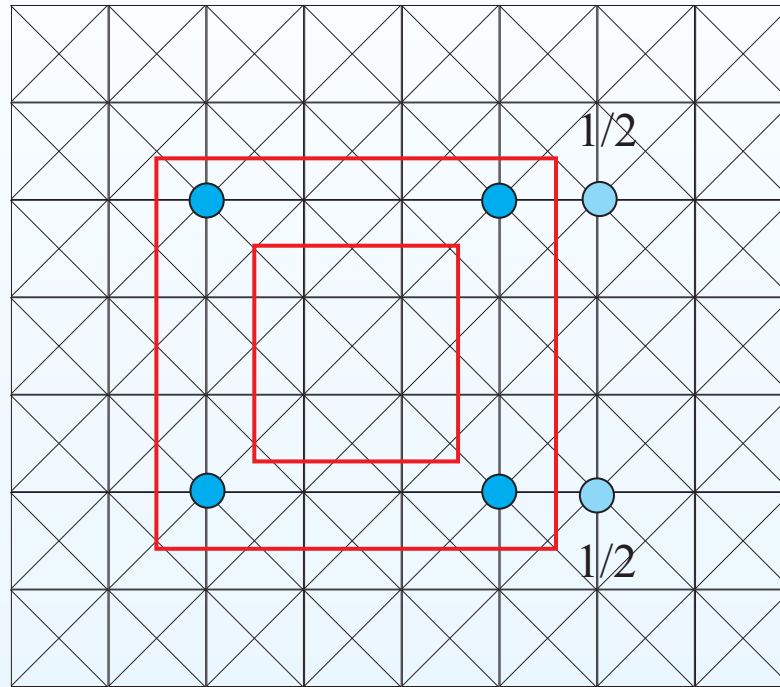


## King grid $\ell = 2$ .



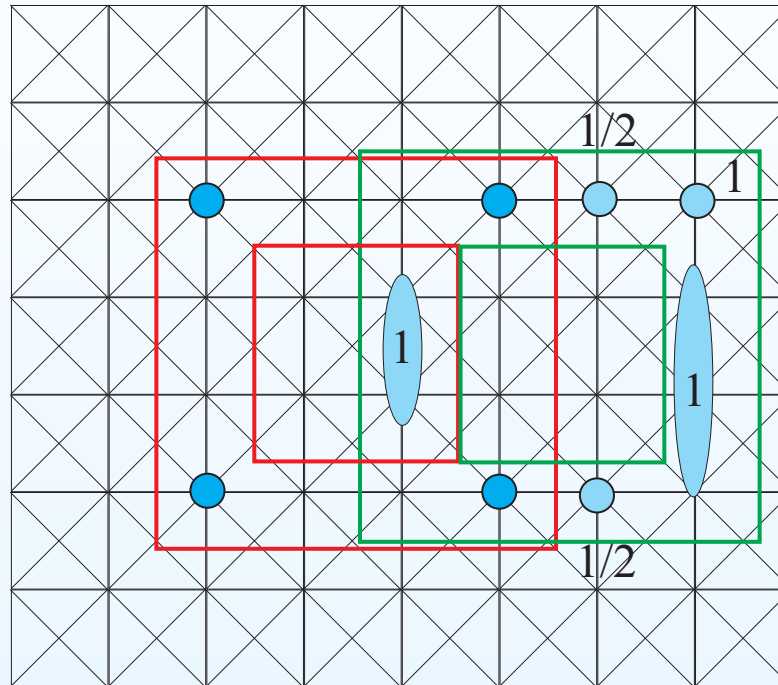
The only way to have exactly 4 codewords

# King grid $\ell = 2$ .



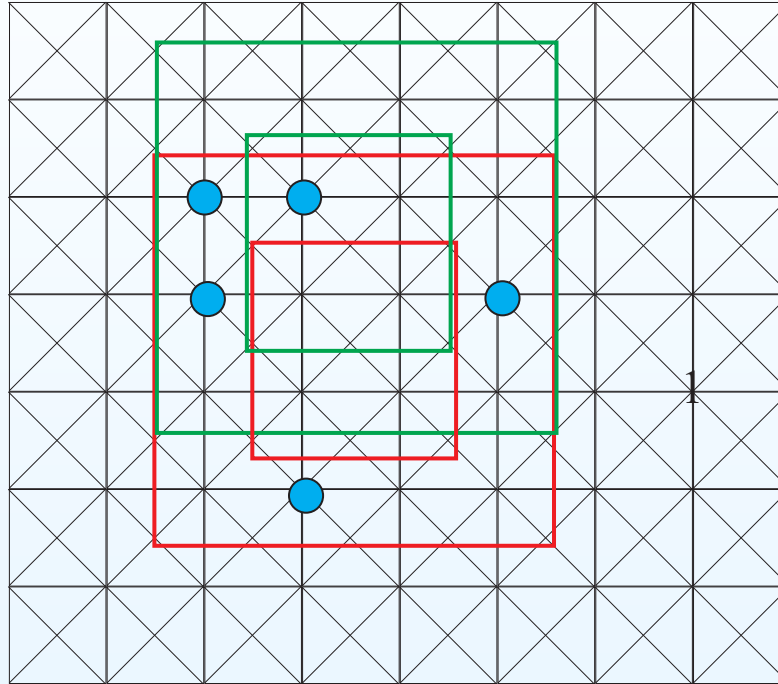
Codewords nearby

# King grid $\ell = 2$ .



A neighbouring  $6^+$  -frame

# King grid $\ell = 2$ .



For 5-frame also a neighbouring  $6^+$  -frame

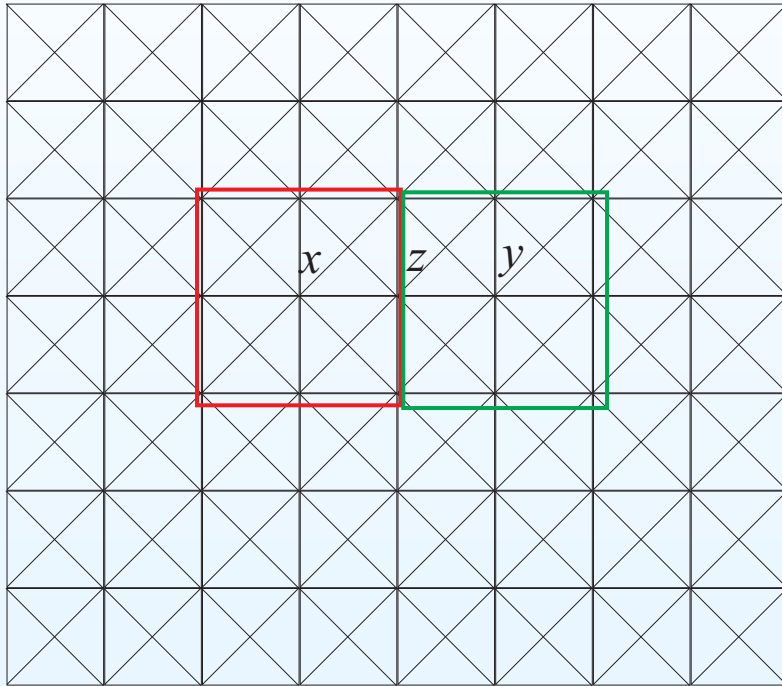
## King grid $\ell = 2$ .

- In fact, we can show that there is in average  $5 + 3/37$  codewords among a 12 frame.
- A  $(1, \leq 2)$ -identifying code in the king grid has density at least  $47/111 \sim 0.423\dots$
- Joint work with Foucaud and Parreau (2011)

$\ell = 3$  is not possible in king and triangular grid

---

It is not possible to distinguish between the sets  $I(x, y)$  and  $I(x, y, z)$ .

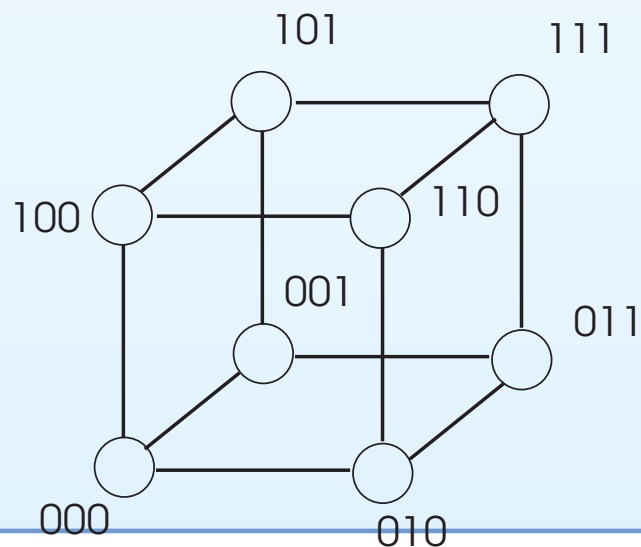


## Binary $n$ -cube

- If  $a \neq b$ , then

$$|B_1(a) \cap B_1(b)| = \begin{cases} 2 & \text{if } d(a, b) = 1 \text{ or } 2 \\ 0 & \text{otherwise.} \end{cases}$$

- If the intersection of three Hamming spheres of radius one is nonempty, then it consists of a unique point.
- no triangles



## Optimal codes when $\ell \geq 3$

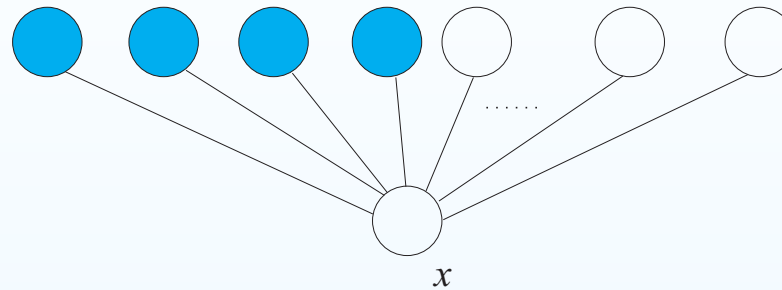
- A subset of vertices  $C$  is called a  **$\mu$ -fold covering**, if  $|I(x)| \geq \mu$  for all  $x \in V$ .
- Let  $\ell \geq 3$ . A code is  $(1, \leq \ell)$ -identifying  $\Leftrightarrow$  it is a  $(2\ell - 1)$ -fold covering.
- (L. 2002) Let  $\ell \geq 3$ . For an infinite family of  $n$

$$M_1^{(\leq \ell)}(n) = (2\ell - 1) \frac{2^n}{n + 1}.$$



Outline of the proof ( $\ell = 3$ ), so  $2\ell - 1 = 5$ :

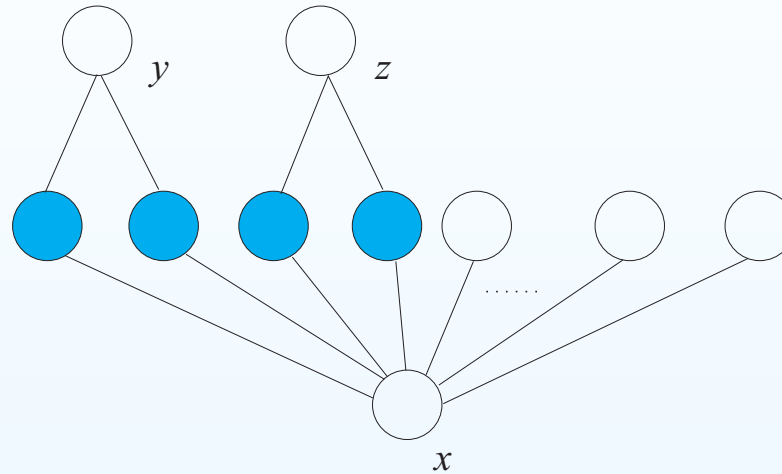
---



A word covered by  $2\ell - 2$  codewords.

Outline of the proof ( $\ell = 3$ ), so  $2\ell - 1 = 5$ :

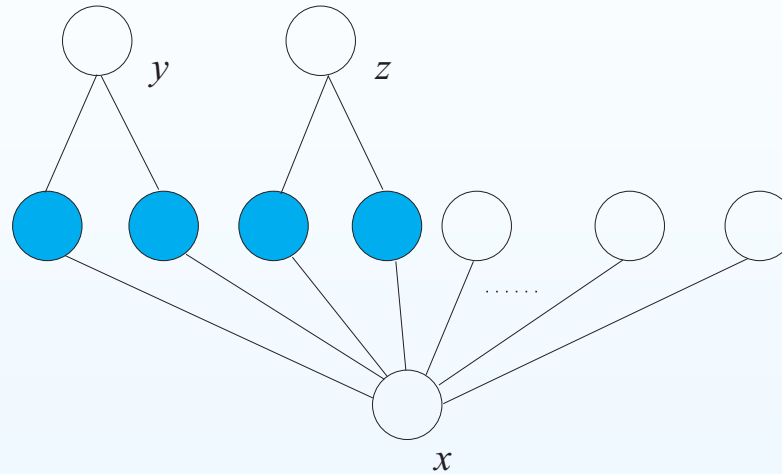
---



Other words at distance two.

Outline of the proof ( $\ell = 3$ ), so  $2\ell - 1 = 5$ :

---



A violation of the property.

$$I(y, z) = I(x, y, z)$$

## Outline of the proof: The other direction.

---

- Suppose  $C$  is  $(2\ell - 1)$ -fold covering.
- We need to show that  $I(X) \neq I(Y)$  for any distinct  $X$  and  $Y$  of size  $|X|, |Y| \leq \ell$ .
- Assume that  $x \in X \setminus Y$ .
- If  $|Y| \leq \ell - 1$ , then from  $2(\ell - 1) < 2\ell - 1 \leq |I(x)|$ , it follows that is enough to look at the case  $|X| = |Y| = \ell$ .

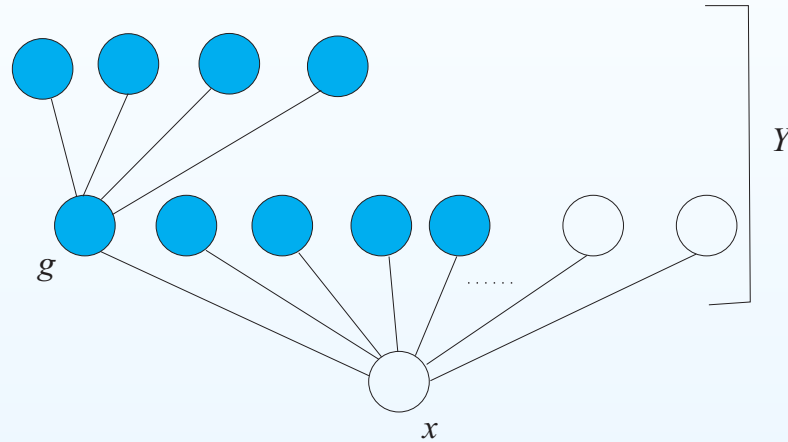
## Outline of the proof:

---

- Now there exists (besides  $x$ ) also  $g \in Y \setminus X$ .
- In fact,  $1 \leq d(x, a) \leq 2$  for all  $a \in Y$  and  $1 \leq d(g, b) \leq 2$  for all  $b \in X$ .
- In particular,  $1 \leq d(x, g) \leq 2$ .
- We may assume  $x = 000 \dots 0$ .
- Consider the case  $d(x, g) = 1$ .

## Outline of the proof.

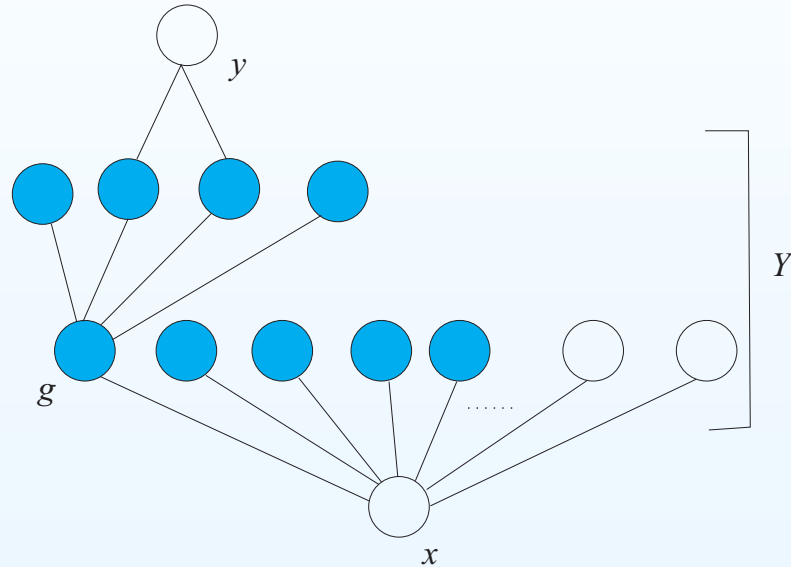
Let  $\ell = 3$ ,  $X = \{x, y, z\}$  and  $Y = \{g, h, k\}$ .



There must be a word  $y$  covering exactly two of  $I(g)$ .

## Outline of the proof.

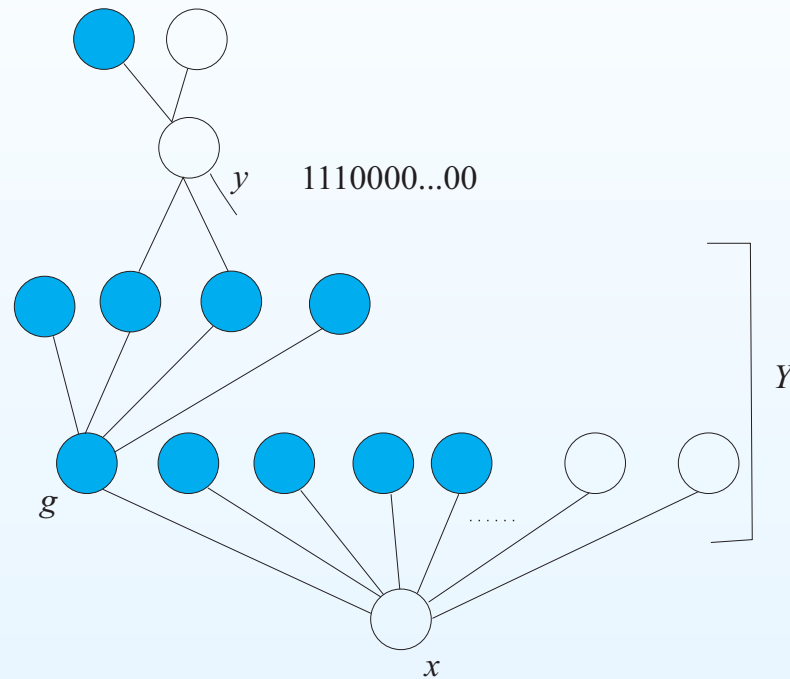
Let  $\ell = 3$ ,  $X = \{x, y, z\}$  and  $Y = \{g, h, k\}$ .



Distance must be three from  $x$ .

## Outline of the proof.

Let  $\ell = 3$ ,  $X = \{x, y, z\}$  and  $Y = \{g, h, k\}$ .

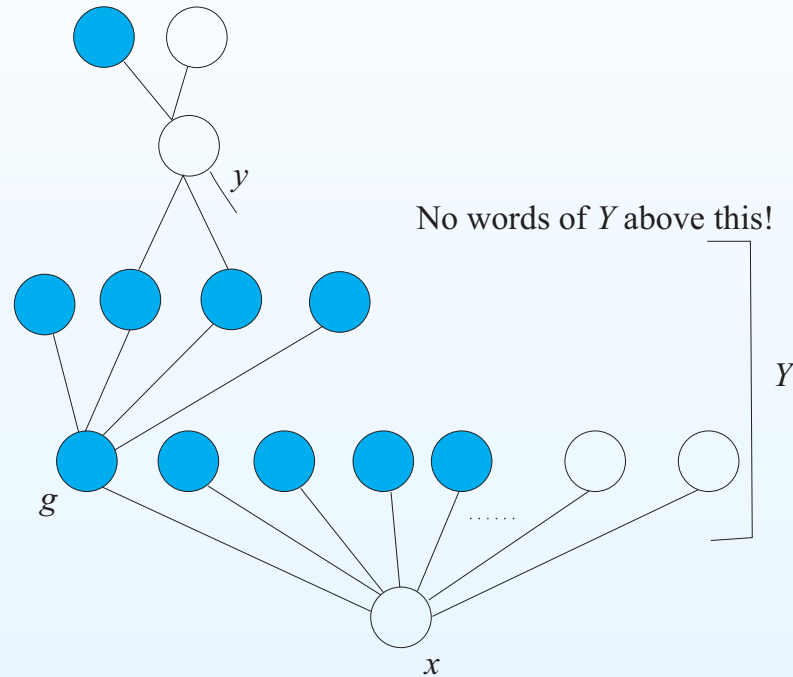


Codewords at distance four in  $I(X)$  .



# Outline of the proof.

Let  $\ell = 3$ ,  $X = \{x, y, z\}$  and  $Y = \{g, h, k\}$ .



Codewords at distance four in  $I(X)$  but not in  $I(Y)$ .

## Outline of the proof.

---

- Hence for  $\ell \geq 3$ , if  $C$  is  $(2\ell - 1)$ -fold covering, we always have  $I(X) = I(Y) \Rightarrow X = Y$ .
- If  $\ell = 2$ , it holds again that  $(1, \leq 2)$ -identifying code is 3-fold covering.
- However, the other direction is not true.

## The case $\ell = 2$

- Idea: Optimal initial codes  $C$  for  $n = 5$  and  $n = 7$  combined with construction

$$\Pi(C) = \{(\pi(u), u, u + c) \mid u \in F_2^n, c \in C\} \subseteq F_2^{2n+1}$$

where  $\pi(u)$  is the parity check bit:  $\pi(01100) = 0$  and  $\pi(10101) = 1$ .

- Two **optimal infinite families**:
- $M_1^{(\leq 2)}(3 \cdot 2^k - 1) = 2^{3 \cdot 2^k - k - 1}$  for  $k \geq 1$
- $M_1^{(\leq 2)}(2^k - 1) = 3 \cdot 2^{2^k - k - 1}$  for  $k \geq 3$

## Radius $r > 1$ , results on the rate when $\ell = 1$

---

- $M_r(n)$  denotes the smallest cardinality of an  $r$ -identifying code of length  $n$ .
- What is  $\alpha$  in  $M_r(n) = p(n) \cdot 2^{\alpha n}$ ?
- Honkala and Lobstein (2002)

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log_2 M_r(n) = 1 - h(\rho)$$

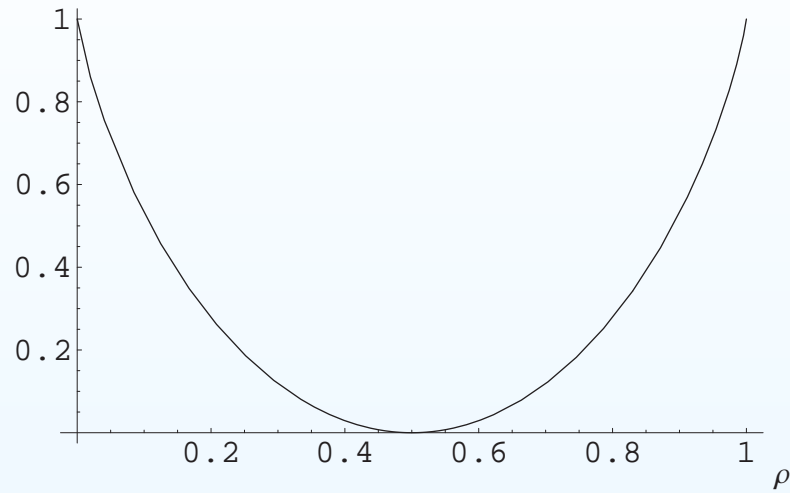
where  $r = \lfloor \rho n \rfloor$  and  $h(x)$  is the binary entropy function

$$h(x) = -x \log_2 x - (1 - x) \log_2(1 - x).$$

- Independently also found by R. Ahlswede in 2006.

## A figure of the optimal rate:

---



$$\lim_{n \rightarrow \infty} \frac{1}{n} \log_2 M_r(n) = 1 - h(\rho)$$

where  $r = \lfloor \rho n \rfloor$ .

## Polynomial size of codes:

---

Honkala and Lobstein showed that for  $n \geq 3$ ,

$$M_{\lfloor \frac{n}{2} \rfloor}(n) \leq \begin{cases} \frac{n^2 - n + 2}{2}, & n \text{ odd,} \\ \frac{n^2 - 4}{2}, & n \text{ even.} \end{cases}$$

A trivial lower bound:

$$M_r(n) \geq n + 1$$

for any radius  $r$ .

## Radius $r > 1$ and $\ell > 1$

---

- Let now  $\ell > 1$  be fixed.
- What is  $\alpha$  in  $M_r^{(\leq \ell)}(n) = p(n) \cdot 2^{\alpha n}$ ?
- Clearly,

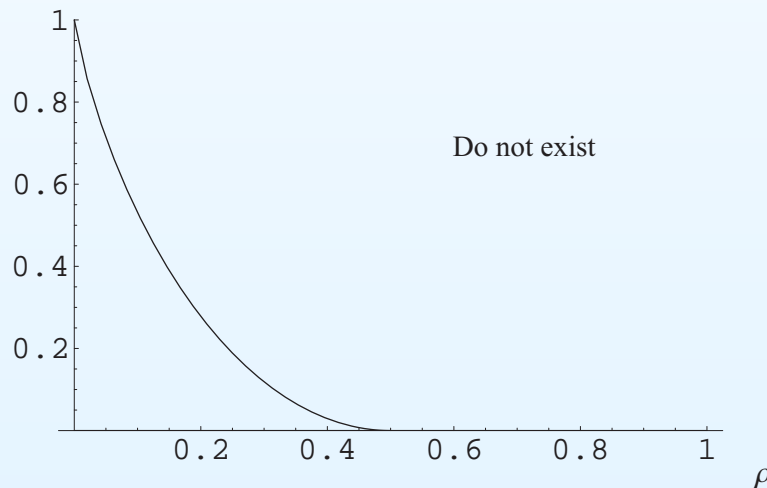
$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log_2 M_r^{(\leq \ell)}(n) \geq 1 - h(\rho).$$

## Radius $r > 1$ and $\ell > 1$

- Janson- L. 2009:
- For **any fixed**  $\ell \geq 1$  we obtain

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log_2 M_r^{(\leq \ell)}(n) = 1 - h(\rho)$$

when  $\rho \in (0, 1/2)$  and  $r/n \rightarrow \rho$ .

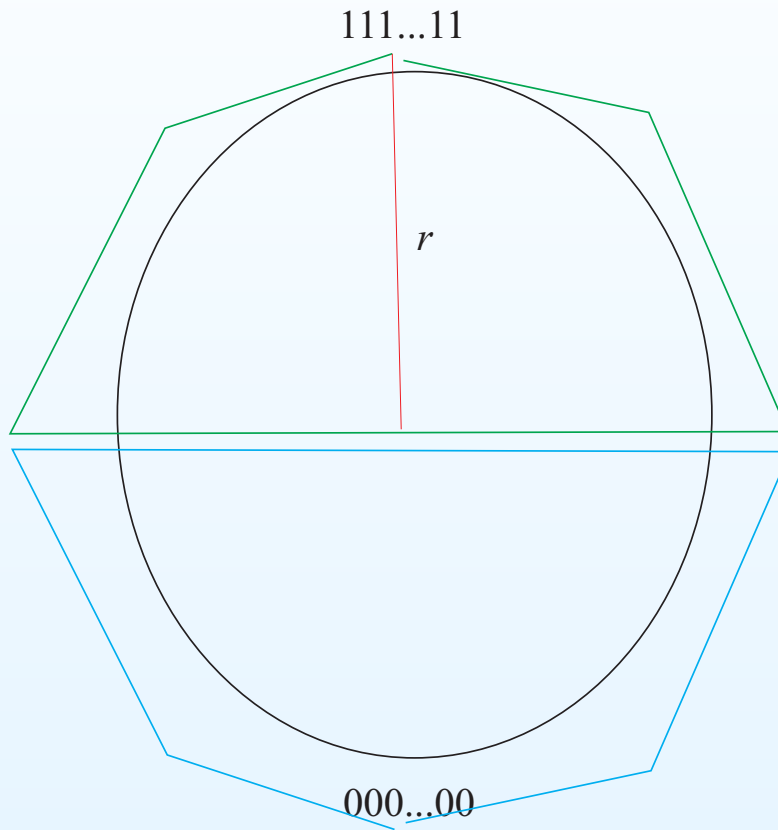




## Radius $r > 1$ and $\ell > 1$

---

- For  $r \geq n/2$ , there is no  $(1, \leq 2)$ -identifying code:



## About the proof:

---

A random code gives

$$M_r^{(\leq \ell)}(n) \leq 2n\ell \frac{2^n}{m_n(r, \ell)}$$

where

$$m_n(r, \ell) = \min_{X, Y \subseteq \mathbb{F}^n, |X|, |Y| \leq \ell} |B_r(X) \Delta B_r(Y)|$$

and  $\Delta$  denotes the symmetric difference.

We are able to show that

$$m_n(r, \ell) \geq c \binom{n}{r}.$$

## Polynomial size of codes:

---

For  $\ell \geq 2$  there exists codes with

$$M_r^{(\leq \ell)}(n) \leq An^{3/2}$$

where  $r \approx n/2$ .

## Polynomial size of codes:

---

For  $\ell \geq 2$  there exists codes with

$$M_r^{(\leq \ell)}(n) \leq An^{3/2}$$

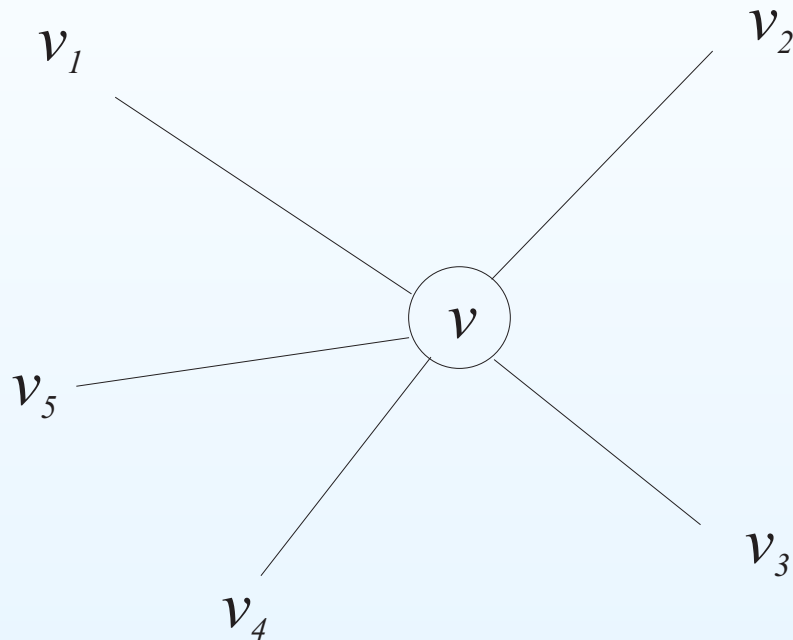
where  $r \approx n/2$ .

This is not far from the best possible, since

$$M_r^{(\leq \ell)}(n) \gtrsim \ell n.$$

## Small $\delta(G)$ do not allow too large $\ell$

- Suppose we have a graph, whose minimum degree equals  $\delta$ .



- Clearly,  $I(v_1, v_2, \dots, v_\delta) = I(v_1, v_2, \dots, v_\delta, v)$ , so the graph cannot admit a  $(1, \leq \ell)$ -identifying code for  $\ell > \delta$ .
- Therefore,  $\ell$  is modest for grids. For binary  $n$ -cube, we have  $\ell$  up to  $\lfloor n/2 \rfloor + 1$ .

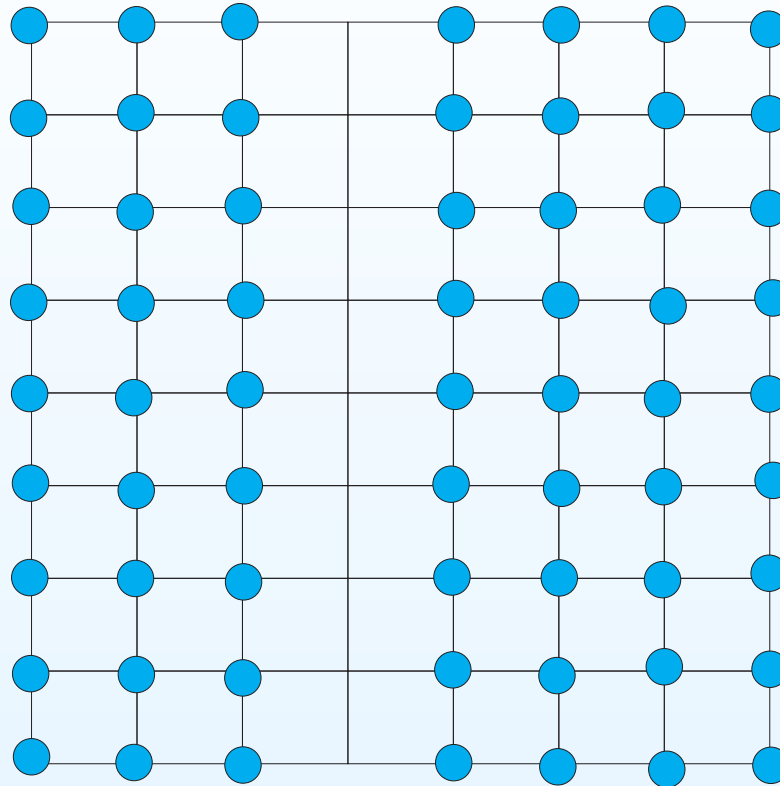
## Universally identifying codes

---

- Suppose we have a finite number of vertices to locate, but there is no a priori upper limit  $\ell$ .
- A dominating code  $C \subseteq V$  satisfying the following property is a **universally identifying code**: Suppose  $S = I(X)$  for finite  $X \subseteq V$ , then for every finite  $X' \subseteq V$  such that  $S = I(X')$  we have  $v_S \in \bigcap X'$ .
- Hence, we can determine at least one of the objects, fix it and continue the process.

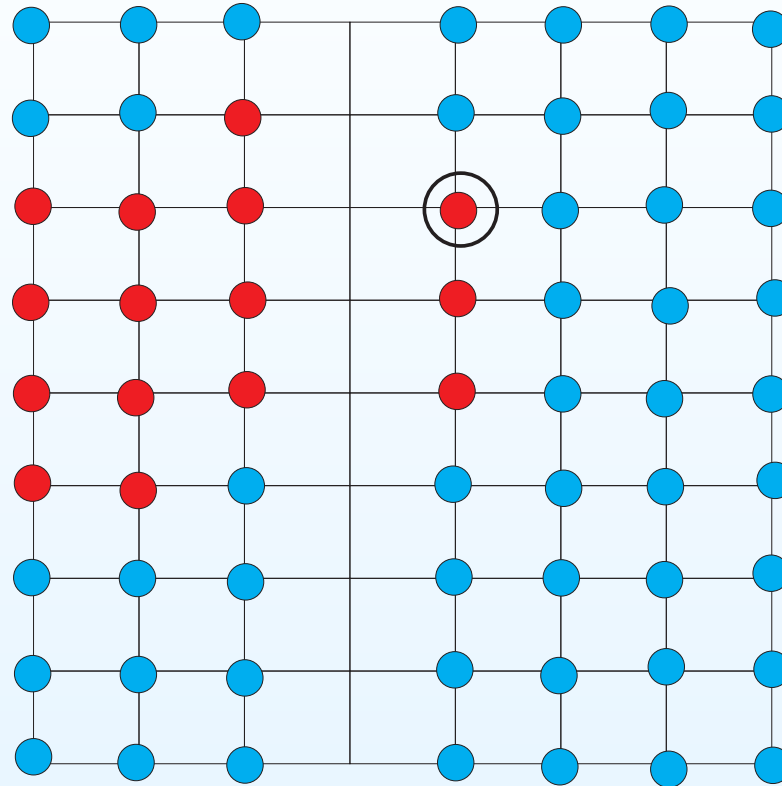
# Universally identifying codes

---



A universally identifying code in the square grid

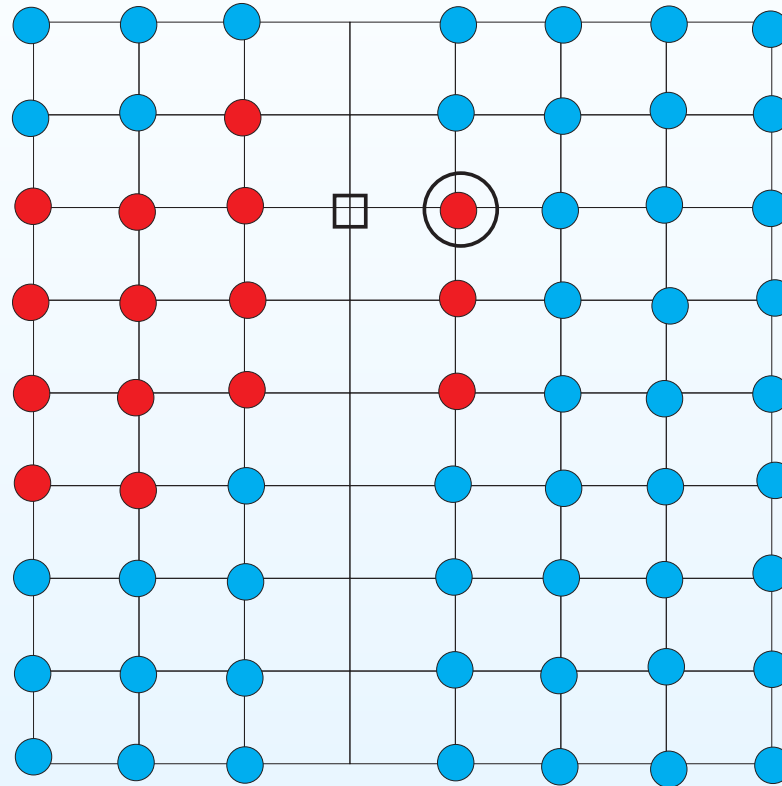
# Universally identifying codes



Take the point upper-rightmost point  $(i, j)$  of  $S$ .



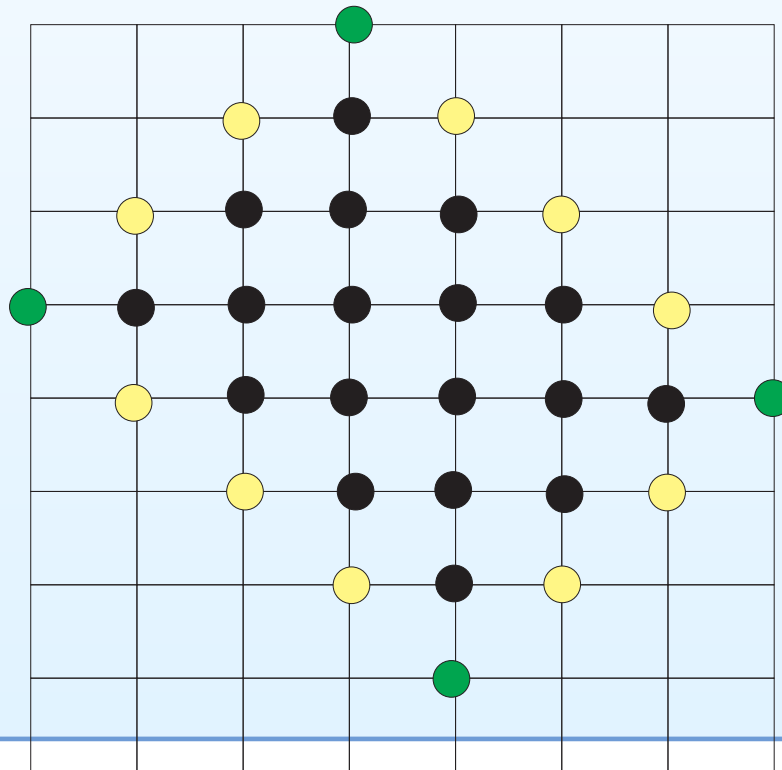
# Universally identifying codes



Conclude that  $(i - 1, j)$  is always in  $X$  when  $S = I(X)$ .

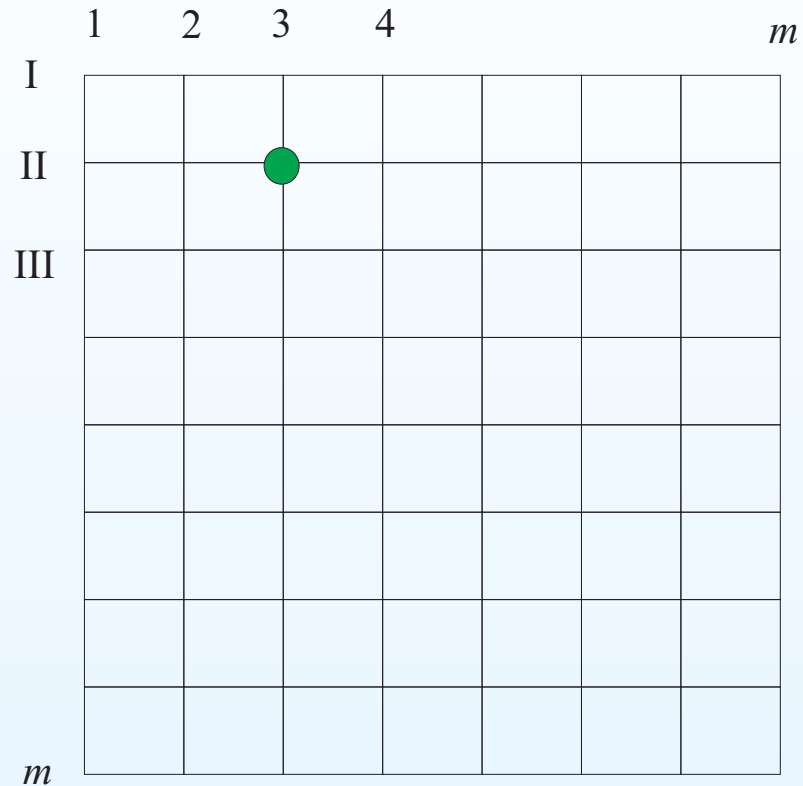
# Universally identifying codes

- We cannot do much better than this, density always 1.
- Non-codewords in green,  $B$  in black and  $Y$  in yellow. Now  $I(X) = (B \cup Y) \cap C$  for any  $X$  containing all but one of  $B$ .
- If there are noncodewords with positive density, then one of such patterns always appears.



# Universally identifying codes

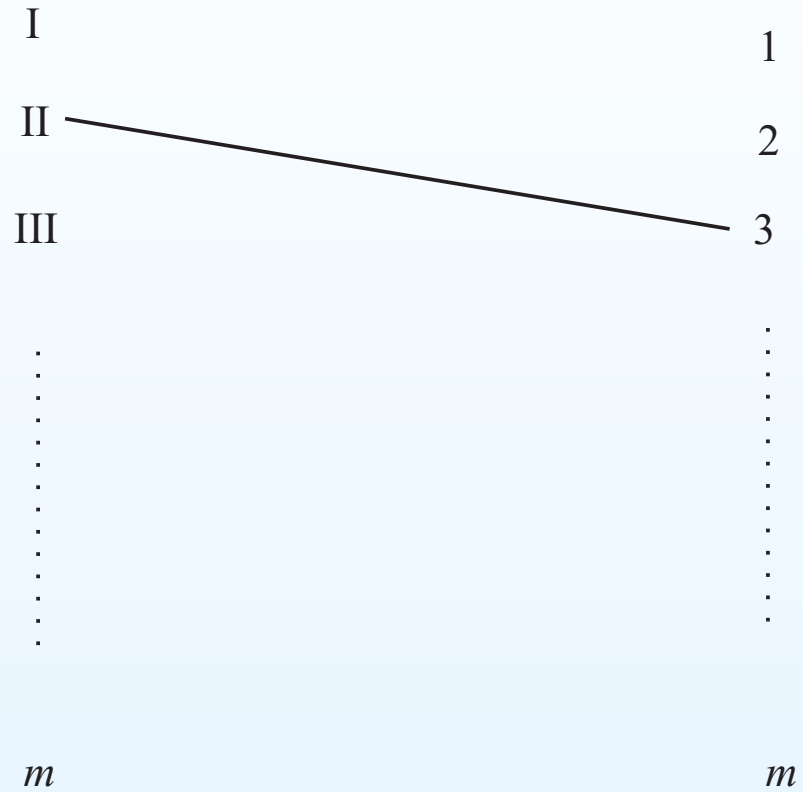
---



Back to the square grid (small cases avoided!)

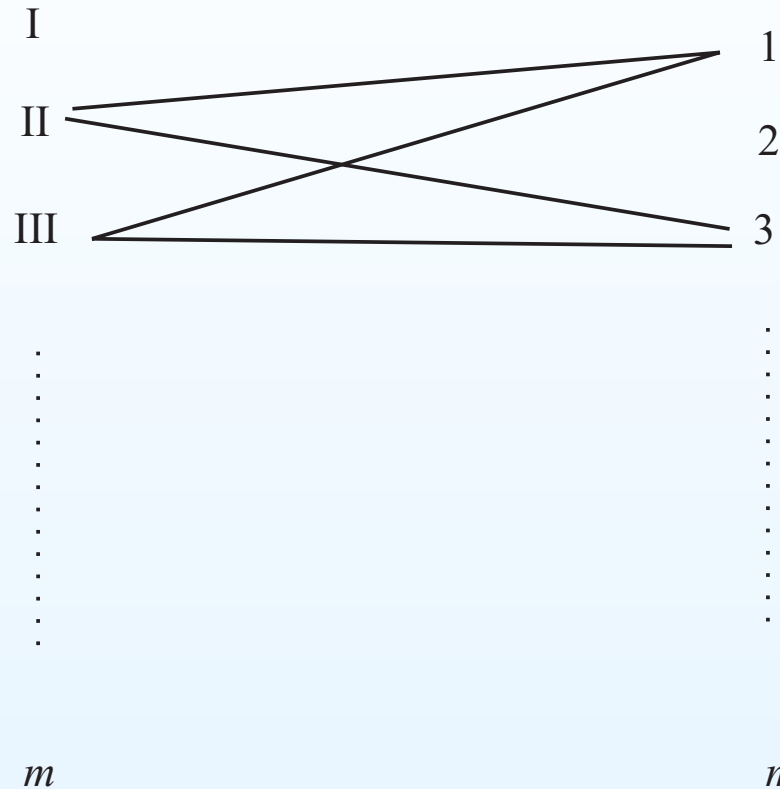
# Universally identifying codes

---



Construct a bipartite graph.

# Universally identifying codes



Number of edges in a graph with no 4-cycle  $O(K^{3/2})$ . So  $|N| \sim n^{3/2}$  and density equals zero.

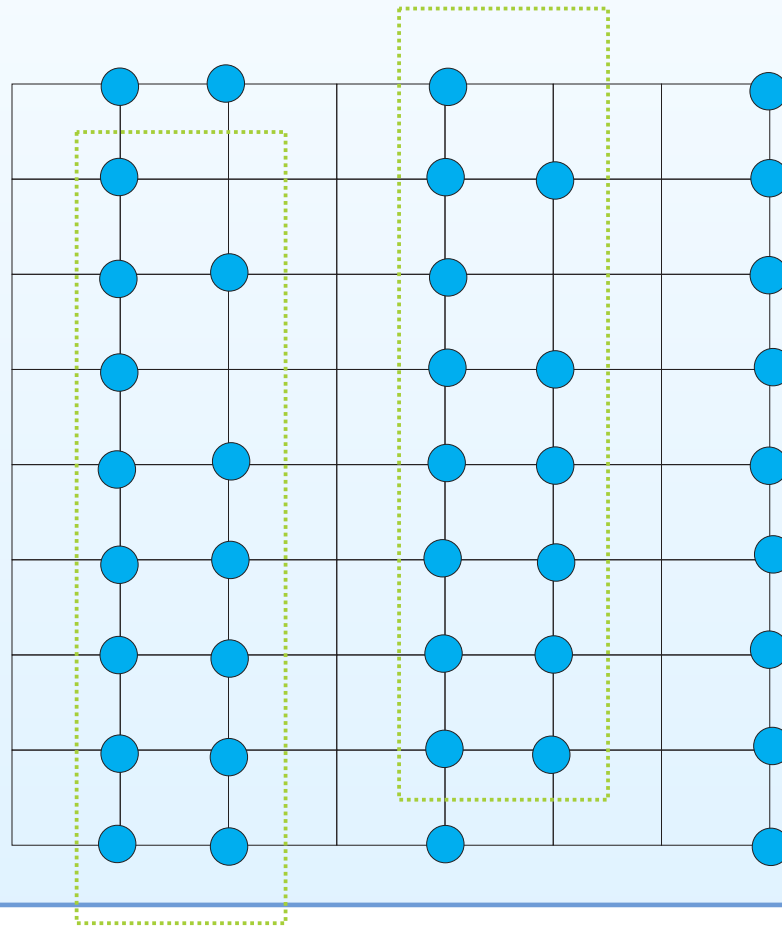
## Weakly identifying codes

---

- We know now an upper limit  $\ell$  and use the process, so that we can determine one by one  $F$  from  $I(F)$ .
- A dominating set  $C \subseteq V$  is **weakly  $(1, \leq \ell)$ -identifying** if it has the property: if  $S = I(X)$  for some  $X \subseteq V$  such that  $|X| \leq \ell$ , then for every  $X'$ ,  $|X'| \leq \ell$  with  $S = I(X')$  we have  $v_S \in \bigcap X'$ .
- There is **no**  $(1, \leq \ell)$ -identifying code in the square grid when  $\ell \geq 4$ .
- There is a **weakly**  $(1, \leq \ell)$ -identifying code in the square grid with density  $\ell/(\ell + 1)$  for any  $\ell$ .

## Weakly identifying codes

- Optimal density for usual  $(1, \leq 3)$ -identifying code in the square grid equals 1.
- A weakly  $(1, \leq 3)$ -identifying code of density  $7/12$ .



## From ID to LD

---

- Locating-dominating sets (P. Slater) are closely related to identifying codes.
- $C \subseteq V$  is a  $r$ -locating-dominating set if  $I_r(x)$  is nonempty for all  $x \in V \setminus C$  and

$$I_r(x) \neq I_r(y)$$

for any  $x \neq y$  where  $x, y \in V \setminus C$ .

- How to generalize the concept to higher  $\ell$ ?



## Generalizing locating-dominating sets

---

- $C \subseteq V$  is  $(r, \leq \ell)$ -LDB if

$$I_r(X) \neq I_r(Y)$$

for all  $X, Y \subseteq V \setminus C$ ,  $X \neq Y$ ,  $|X|, |Y| \leq \ell$ .

## Generalizing locating-dominating sets

---

- $C \subseteq V$  is  $(r, \leq \ell)$ -LDA if

$$I_r(X) \neq I_r(Y)$$

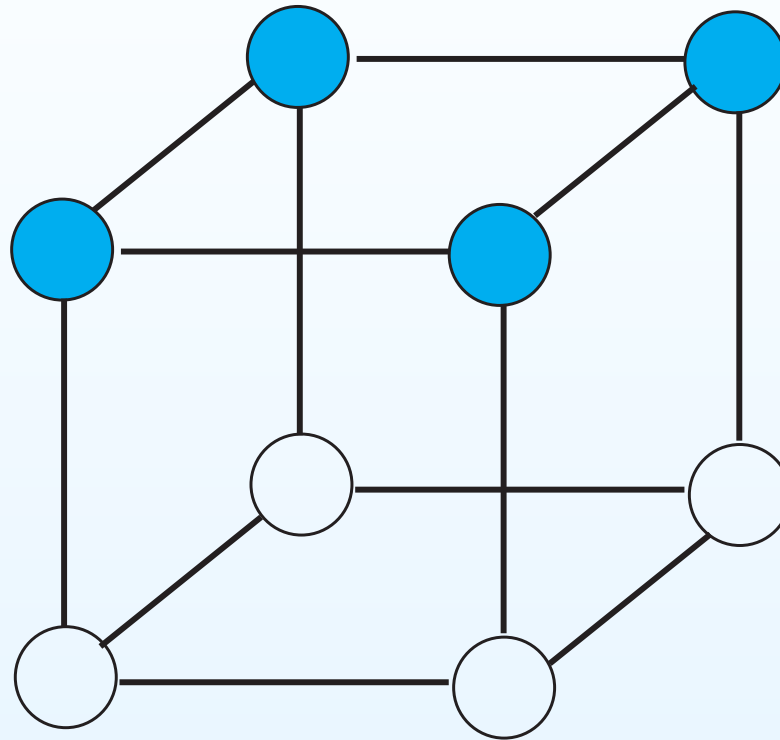
or

$$X \cap C \neq Y \cap C$$

for all  $X, Y \subseteq V$ ,  $X \neq Y$ ,  $|X|, |Y| \leq \ell$ .

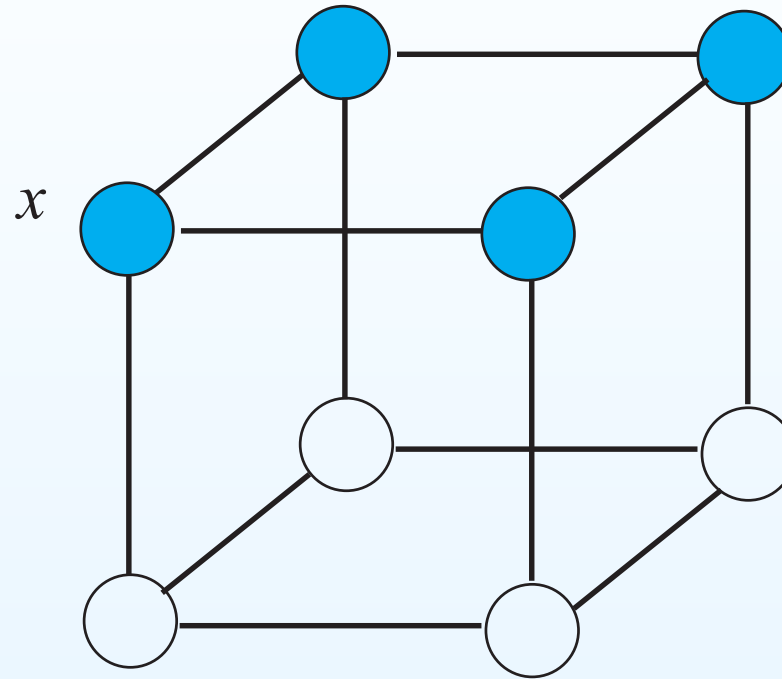
- Both are LD-sets if  $\ell = 1$ .

## LDA and LDB



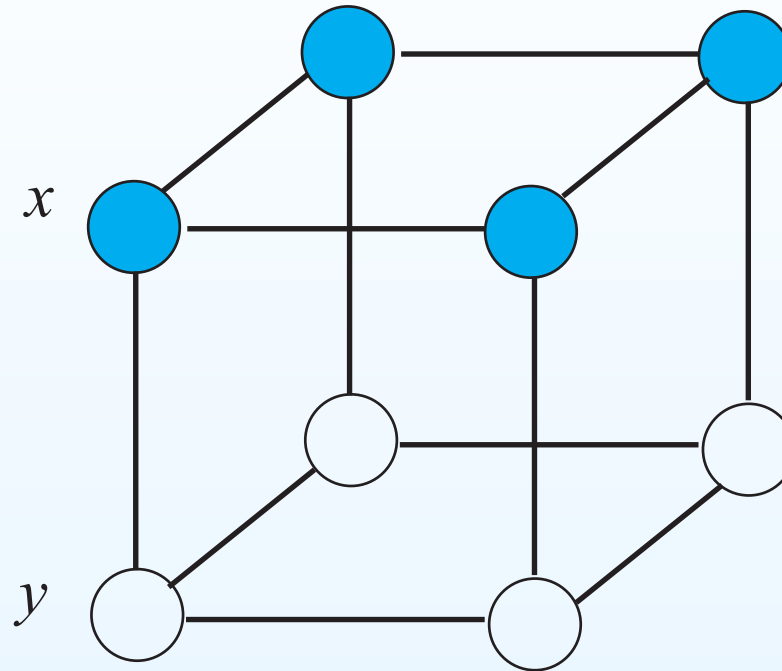
$(1, \leq 4)$ -LDB set in a hypercube

# LDA and LDB



not a  $(1, \leq 2)$ -LDA set

## LDA and LDB



Now  $I(x) = I(x, y)$  and  $\{x\} \cap C = \{x, y\} \cap C$

## LDA and LDB

---

- $(1, \leq \ell)$ -LDA is not  $(1, \leq \ell)$ -identifying, but some similarities exist (Honkala-L.-Ranto 2004):
- In a hypercube any non-codeword is covered by  $(2\ell - 1)$  vertices of  $(1, \leq \ell)$ -LDA, i.e. the optimal size

$$M \sim \frac{2\ell - 1}{n} 2^n.$$

- For  $(1, \leq \ell)$ -LDB the optimal size is

$$M = 2^{n-1}$$

for  $\ell \geq (\sqrt{2^n + 2n + 1} - 1)/2$ .

## LDA and LDB

- Consider cycles (Exoo-Junnila-L.,2010):
- There is only trivial  $(r, \leq \ell)$ -LDA set when  $\ell \geq 3$
- The size of optimal  $(r, \leq 2)$ -LDA is

$$\gcd(r, n) \left\lfloor \frac{n}{2 \gcd(r, n)} \right\rfloor$$

- The size of optimal  $(r, \leq \ell)$ -LDB is

$$\left\lfloor \frac{rn}{r+1} \right\rfloor$$

for  $\ell \geq 3$ ,  $n \geq 2r + 2$  and  $n \not\equiv r + 1 \pmod{2r + 2}$ .

Thank you!