



HAL
open science

Minimum Sizes of Identifying Codes in Graphs Differing by One Edge or One Vertex

Irène Charon, Iiro Honkala, Olivier Hudry, Antoine Lobstein

► **To cite this version:**

Irène Charon, Iiro Honkala, Olivier Hudry, Antoine Lobstein. Minimum Sizes of Identifying Codes in Graphs Differing by One Edge or One Vertex. 2012. hal-00731811

HAL Id: hal-00731811

<https://hal.science/hal-00731811>

Preprint submitted on 13 Sep 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Minimum Sizes of Identifying Codes in Graphs Differing by One Edge or One Vertex

Irène Charon

Institut Télécom - Télécom ParisTech & CNRS - LTCI UMR 5141
46, rue Barrault, 75634 Paris Cedex 13 - France
charon@telecom-paristech.fr

Iiro Honkala

University of Turku, Department of Mathematics and Statistics
20014 Turku, Finland
honkala@utu.fi

Olivier Hudry

Institut Télécom - Télécom ParisTech & CNRS - LTCI UMR 5141
46, rue Barrault, 75634 Paris Cedex 13 - France
hudry@telecom-paristech.fr

Antoine Lobstein

CNRS - LTCI UMR 5141 & Institut Télécom - Télécom ParisTech
46, rue Barrault, 75634 Paris Cedex 13 - France
lobstein@telecom-paristech.fr

Abstract

Let G be a simple, undirected graph with vertex set V . For $v \in V$ and $r \geq 1$, we denote by $B_{G,r}(v)$ the ball of radius r and centre v . A set $\mathcal{C} \subseteq V$ is said to be an r -*identifying code* in G if the sets $B_{G,r}(v) \cap \mathcal{C}$, $v \in V$, are all nonempty and distinct. A graph G admitting an r -identifying code is called r -*twin-free*, and in this case the size of a smallest r -identifying code in G is denoted by $\gamma_r(G)$.

We study the following structural problem: let G be an r -twin-free graph, and G^* be a graph obtained from G by adding or deleting a vertex, or by adding or deleting an edge. If G^* is still r -twin-free, we compare the behaviours of $\gamma_r(G)$ and $\gamma_r(G^*)$, establishing results on their possible differences and ratios.

Key Words: Graph Theory, Twin-Free Graphs, Identifiable Graphs, Identifying Codes.

1 Foreword

This preprint is a combination of the two submitted articles [5] and [6] which deal with closely related topics. The Introduction and Bibliography are put in common, but Part I, devoted to the addition or deletion of one vertex, and Part II, about the addition or deletion of one edge, have been made so that they can be read independently and in any order. Part I starts at page 6 and Part II at page 22.

2 Introduction

We introduce basic definitions and notation for graphs, for which we refer to, e.g., [1] and [12], and for identifying codes (see [18] and the bibliography at [21]).

We shall denote by $G = (V, E)$ a simple, undirected graph with vertex set V and edge set E , where an *edge* between $x \in V$ and $y \in V$ is indifferently denoted by $\{x, y\}$, $\{y, x\}$, xy or yx . The *order* of a graph is its number of vertices $|V|$.

A *path* $P_n = x_1x_2 \dots x_n$ is a sequence of n distinct vertices x_i , $1 \leq i \leq n$, such that x_ix_{i+1} is an edge for $i \in \{1, 2, \dots, n-1\}$. The *length* of P_n is its number of edges, $n-1$. A *cycle* $C_n = x_1x_2 \dots x_n$ is a sequence of n distinct vertices x_i , $1 \leq i \leq n$, where x_ix_{i+1} is an edge for $i \in \{1, 2, \dots, n-1\}$, and x_nx_1 is also an edge; its length is n .

A graph G is called *connected* if for any two vertices x and y , there is a path between them. It is called *disconnected* otherwise. In a connected graph G , we can define the *distance* between any two vertices x and y , denoted by $d_G(x, y)$, as the length of any shortest path between x and y , since such a path exists. This definition can be extended to disconnected graphs, using the convention that $d_G(x, y) = +\infty$ if there is no path between x and y .

For any vertex $v \in V$ and integer $r \geq 1$, the *ball of radius r and centre v* , denoted by $B_{G,r}(v)$, is the set of vertices within distance r from v :

$$B_{G,r}(v) = \{x \in V : d_G(v, x) \leq r\}.$$

Two vertices x and y such that $B_{G,r}(x) = B_{G,r}(y)$ are called (G, r) -*twins*; if G has no (G, r) -twins, that is, if

$$\forall x, y \in V \text{ with } x \neq y, \quad B_{G,r}(x) \neq B_{G,r}(y),$$

then we say that G is r -*twin-free*.

Whenever two vertices x and y are within distance r from each other in G , i.e., $x \in B_{G,r}(y)$ and $y \in B_{G,r}(x)$, we say that x and y r -*cover* each other. When three vertices x, y, z are such that $x \in B_{G,r}(z)$ and $y \notin B_{G,r}(z)$, we say that z r -*separates* x and y in G . A set is said to r -*separate* x and y in G if it contains at least one vertex which does.

A *code* \mathcal{C} is simply a subset of V , and its elements are called *codewords*. For each vertex $v \in V$, the *r -identifying set* of v , with respect to \mathcal{C} , is the set of codewords r -covering v , and is denoted by $I_{G,\mathcal{C},r}(v)$:

$$I_{G,\mathcal{C},r}(v) = B_{G,r}(v) \cap \mathcal{C}.$$

We say that \mathcal{C} is an *r -identifying code* [18] if all the sets $I_{G,\mathcal{C},r}(v)$, $v \in V$, are nonempty and distinct: in other words, every vertex is r -covered by at least one codeword, and every pair of vertices is r -separated by at least one codeword.

It is quite easy to observe that a graph G admits an r -identifying code if and only if G is r -twin-free; this is why r -twin-free graphs are also sometimes called *r -identifiable*.

When G is r -twin-free, we denote by $\gamma_r(G)$ the cardinality of a smallest r -identifying code in G . The search for the smallest r -identifying code in given graphs or families of graphs is an important part of the studies devoted to identifying codes.

In this preprint and the forthcoming [5] and [6], we are interested in the following issue: let G be an r -twin-free graph, and G^* be a graph obtained from G by adding or deleting one vertex, or by adding or deleting one edge. Now, if G^* is still r -twin-free, what can be said about $\gamma_r(G)$ compared to $\gamma_r(G^*)$? More specifically, we shall study their difference and, when appropriate, their ratio,

$$\gamma_r(G) - \gamma_r(G^*) \text{ and } \frac{\gamma_r(G)}{\gamma_r(G^*)},$$

as functions of the order of the graph G , and r .

Note that a partial answer to the issue of knowing the conditions for which an r -twin-free graph remains so when one vertex is removed was given in [4] and [8]: any 1-twin-free graph with at least four vertices always possesses at least one vertex whose deletion leaves the graph 1-twin-free; for any $r \geq 1$, any r -twin-free tree with at least $2r + 2$ vertices always possesses at least one vertex whose deletion leaves the graph r -twin-free; on the other hand, for any $r \geq 3$, there exist r -twin-free graphs such that the deletion of any vertex makes the graph not r -twin-free. The case $r = 2$ remains open.

Of what interest this study is, can be illustrated by the watching of a museum: we place ourselves in the case $r = 1$ and assume that we have to protect a museum, or any other type of premises, using smoke detectors. The museum can be viewed as a graph, where the vertices represent the rooms, and the edges, the doors or corridors between rooms. The detectors are located in some of the rooms and give the alarm whenever there is smoke in their room or in one of the adjacent rooms. If there is smoke in one room and if the detectors are located in rooms corresponding to a 1-identifying

code, then, only by knowing which detectors gave the alarm, we can identify the room where someone is smoking.

Of course we want to use as few detectors as possible. Now, what are the consequences, beneficial or not, of closing or opening one room or one door? This is exactly the object of our investigation, in the more general case when r can take values other than 1.

In the conclusion of [22], it is already observed, somewhat paradoxically, that a cycle with one vertex less can require more codewords/detectors. We shall exhibit examples of large variations for the minimum size of an identifying code.

A related issue is that of *t-edge-robust* identifying codes, which remain identifying when at most t edges are added or deleted, in any possible way; see, e.g., [15]–[17], [19] or [20].

Let us mention that in the sequel, we shall consider two cases,

- (i) both graphs G and G^* are connected,
- (ii) the graph with one edge less or one vertex less may be disconnected, and observe one significant difference in our results for vertex addition/deletion, whereas our constructions can always be made such that (i) holds when edge addition/deletion is concerned.

Before we proceed, we still need some additional definitions and notation, and we also give three lemmata which, although very easy, will prove useful in the sequel, even implicitly.

When we delete the edge $e \in E$ in a graph $G = (V, E)$, we denote the resulting subgraph by $G_e = G \setminus e = (V, E_e)$. For a vertex $v \in V$, we denote by G_v or $G \setminus v$ the graph with vertex set V' and edge set E' , where

$$V' = V \setminus \{v\}, E' = \{xy \in E : x \in V', y \in V'\}.$$

If $G = (V, E)$ is a graph and \mathcal{S} is a subset of V , we say that two vertices $x \in V$ and $y \in V$ are (G, \mathcal{S}, r) -twins if

$$I_{G, \mathcal{S}, r}(x) = I_{G, \mathcal{S}, r}(y).$$

In other words, x and y are not r -separated by \mathcal{S} in G . By definition, if \mathcal{C} is r -identifying in G , then no (G, \mathcal{C}, r) -twins exist.

Lemma 1 [*(G, S, r)-twin transitivity*] *In a graph $G = (V, E)$, if x, y, z are three distinct vertices, if \mathcal{S} is a subset of V , if x and y are (G, \mathcal{S}, r) -twins and if y and z are (G, \mathcal{S}, r) -twins, then x and z are (G, \mathcal{S}, r) -twins. \triangle*

Lemma 2 *If \mathcal{C} is an r -identifying code in a graph $G = (V, E)$, then so is any set \mathcal{S} such that*

$$\mathcal{C} \subseteq \mathcal{S} \subseteq V.$$

\triangle

Lemma 3 *If a graph $G = (V, E)$ is 1-twin-free and contains a vertex v which is linked to all the other vertices, then there is an optimal 1-identifying code \mathcal{C} not containing v .*

Proof. Assume that an optimal 1-identifying code \mathcal{C} contains v . Since v cannot 1-separate any pair of vertices in G , its only purpose as a codeword is to 1-cover some vertices not 1-covered by any other codeword; because these vertices are 1-separated by \mathcal{C} , only one of them, which we denote by x , can be such that $I_{G, \mathcal{C}, 1}(x) = \{v\}$. Then $\mathcal{C} \setminus \{v\} \cup \{x\}$ is also optimal and 1-identifying. \triangle

Part I: Addition and deletion of one vertex

We present our main results in the following way. In Section 3 we consider the case $r = 1$: we study how *large* $\gamma_1(G_x) - \gamma_1(G)$ and $\gamma_1(G_x)/\gamma_1(G)$ can be (Proposition 4), then Theorem 5 states exactly how *small* the difference can be (namely, -1).

In Section 4, we study how *large* the difference $\gamma_r(G_x) - \gamma_r(G)$ can be, in the following three cases: (i) $r \geq 2$, r is even and the graphs are connected (Proposition 9); (ii) $r \geq 3$, r is odd and the graphs are connected (Proposition 11); (iii) $r \geq 2$ and the graph G_x is disconnected (Proposition 13). Then we consider how *large* the ratio $\gamma_r(G_x)/\gamma_r(G)$ can be (Proposition 15), and it so happens that the graphs we use are connected.

Finally, we study how *small* $\gamma_r(G_x) - \gamma_r(G)$ can be for any $r \geq 2$ (Proposition 17) and how *small* $\gamma_r(G_x)/\gamma_r(G)$ can be for any $r \geq 2$ (Proposition 18), and again it so happens that the graphs we use are connected.

In these sections, the number n represents the order of either G or G_x , or an approximation. A general conclusion recapitulates our results in a Table.

3 The case $r = 1$

Note that we obtain the following result with connected graphs: we found no better with disconnected graphs.

Proposition 4 *Let $k \geq 1$ be an arbitrary integer. There exist two (connected) 1-twin-free graphs G and G_x , where G has $2k + \lceil \log_2(k+1) \rceil + 2$ vertices, such that $\gamma_1(G) \leq \lceil \log_2(k+1) \rceil + 2$ and $\gamma_1(G_x) \geq k$.*

Proof. We put the cart before the horse and, before defining G , we describe G_x (see Figure 4 with $r = 1$): we begin by choosing k vertices x_1, \dots, x_k , none of them adjacent with each other, and then build a graph G_x with a "small" 1-identifying code in the following way: we take $s = \lceil \log_2(k+1) \rceil + 1$ auxiliary vertices a_1, \dots, a_s . We first connect each x_i to a_1 ; then we connect each x_i to the vertices of a unique nonempty subset A_i of the set $A = \{a_2, \dots, a_s\}$. The sets A_i can indeed be chosen in this way, because there are $2^{s-1} - 1$ nonempty subsets of A , and $s-1 = \lceil \log_2(k+1) \rceil$. Without loss of generality, we can choose the sets A_i in such a way that the graph constructed so far is connected.

Clearly the auxiliary vertices form a 1-identifying code in this graph: the 1-identifying set of each auxiliary vertex is a singleton consisting of the vertex itself; and for all the vertices x_i , the 1-identifying set contains a_1 and at least one more vertex, and no two of these sets are the same by the construction.

As the next step, we take another set of k vertices, y_1, \dots, y_k , none of them adjacent with each other, and each y_i connected to exactly the same auxiliary vertices a_j as x_i . In this new graph G_x , which is connected, every 1-identifying code must contain at least one of the vertices x_i and y_i for each i : otherwise we cannot 1-separate between x_i and its "copy" y_i . But certainly if for each i we take at least one of x_i and y_i into the code and take all the auxiliary vertices a_j into the code, then the code is 1-identifying, and G_x is 1-twin-free. All in all, for this graph G_x , the smallest 1-identifying code has size at least k .

However, if we add one more vertex x , and connect it to each x_i (but not to any y_i nor any a_j), then in the resulting graph G the set consisting of x and all the auxiliary vertices a_j is a 1-identifying code.

Therefore, $\gamma_1(G) \leq \lceil \log_2(k+1) \rceil + 2$ and $\gamma_1(G_x) \geq k$. \triangle

Remark. The difference $\gamma_1(G_x) - \gamma_1(G)$ and ratio $\gamma_1(G_x)/\gamma_1(G)$ can be made arbitrarily large:

$$\gamma_1(G_x) - \gamma_1(G) \geq k - \lceil \log_2(k+1) \rceil - 2, \quad (1)$$

$$\frac{\gamma_1(G_x)}{\gamma_1(G)} \geq \frac{k}{\lceil \log_2(k+1) \rceil + 2}. \quad (2)$$

In terms of $n = 2k + \lceil \log_2(k+1) \rceil$, which is the approximate order of G and G_x , we can approximate these two lower bounds by $\frac{n}{2} - \frac{3}{2} \log_2 n$ and $\frac{n}{2 \log_2 n}$, respectively.

An open question is whether these difference or ratio can be made substantially larger.

Theorem 5 *Let $G = (V, E)$ be any 1-twin-free graph with at least three vertices. For any vertex $x \in V$ such that G_x is 1-twin-free, we have:*

$$\gamma_1(G_x) \geq \gamma_1(G) - 1. \quad (3)$$

Proof. Cf. [13, Prop. 3]. For completeness, we still give a proof. Let $x \in V$ be such that G_x is 1-twin-free. Let \mathcal{C}_x be a minimum 1-identifying code in G_x : $|\mathcal{C}_x| = \gamma_1(G_x)$. There are two cases: either (a) x is not 1-covered (in G) by any codeword of \mathcal{C}_x , or (b) x is 1-covered (in G) by at least one codeword of \mathcal{C}_x .

(a) In this case, let $\mathcal{C} = \mathcal{C}_x \cup \{x\}$. Then \mathcal{C} is clearly 1-identifying in G (in particular, thanks to Lemma 2); therefore, $\gamma_1(G) \leq \gamma_1(G_x) + 1$.

(b) x is 1-covered by $y \in \mathcal{C}_x$. If \mathcal{C}_x is 1-identifying in G , then $\gamma_1(G) \leq \gamma_1(G_x)$, and we are done. So we assume that \mathcal{C}_x is not 1-identifying in G . This means that either (i) at least one vertex in G is not 1-covered by \mathcal{C}_x , or (ii) at least two vertices in G are not 1-separated by \mathcal{C}_x .

(i) Since \mathcal{C}_x 1-covers any vertex in G_x and x is linked to $y \in \mathcal{C}_x$, this case is impossible.

(ii) Let $u, v \in V$ be two distinct vertices which are not 1-separated by \mathcal{C}_x . One of them is necessarily x , and without loss of generality, we assume that $x = u$.

Now, v is unique by Lemma 1: \mathcal{C}_x is not 1-identifying in G only because one pair of vertices, x and v , is not 1-separated by \mathcal{C}_x .

Since G is 1-twin-free, there is a vertex z which 1-covers exactly one of the vertices v and x . We set $\mathcal{C} = \mathcal{C}_x \cup \{z\}$, and we obtain a 1-identifying code in G , so $\gamma_1(G) \leq \gamma_1(G_x) + 1$. \triangle

Corollary 6 *If $\gamma_1(G_x) \leq a$ and $\gamma_1(G) \geq a + 1$, then $\gamma_1(G_x) = a$ and $\gamma_1(G) = a + 1$.* \triangle

Note that we made no assumption on the connectivity of G or G_x . Examples where $\gamma_1(G_x) = \gamma_1(G) - 1$, or $\gamma_1(G_x) = \gamma_1(G)$, are numerous and easy to find.

Conclusion 7 *Provided that the graphs considered are 1-twin-free, we can see, using Proposition 4 and Theorem 5, that $\gamma_1(G_x) - \gamma_1(G)$ cannot be smaller than -1 , but examples exist where it can be as large as, approximately, $\frac{n}{2} - \frac{3}{2} \log_2 n$, and where the ratio $\frac{\gamma_1(G_x)}{\gamma_1(G)}$ can be as large as, approximately, $\frac{n}{2 \log_2 n}$. This can even be obtained with connected examples.*

4 The case $r \geq 2$

Things are different for $r \geq 2$, since we can exhibit pairs of graphs (G, G_x) proving that $\gamma_r(G_x) - \gamma_r(G)$ and $\gamma_r(G_x)/\gamma_r(G)$ can be arbitrarily large or small.

We first give a result with $\gamma_r(G_x) - \gamma_r(G)$ arbitrarily large. We start with connected graphs, and have two subcases, r even and r odd. In both cases, we shall use the following result on cycles of even length.

Theorem 8 [3] *For all $r \geq 1$ and for all even n , $n \geq 2r + 4$, we have:*

$$\gamma_r(C_n) = \frac{n}{2}.$$

\triangle

• (i) Case of a connected graph G_x and $r \geq 2$, r even

Proposition 9 *There exist two (connected) r -twin-free graphs G and G_x , with $n + 1$ and n vertices respectively, such that*

$$\gamma_r(G_x) - \gamma_r(G) \geq \frac{n}{4} - (r + 1), \quad (4)$$

$$\frac{\gamma_r(G_x)}{\gamma_r(G)} \geq \frac{2n}{n + 4r + 4}. \quad (5)$$

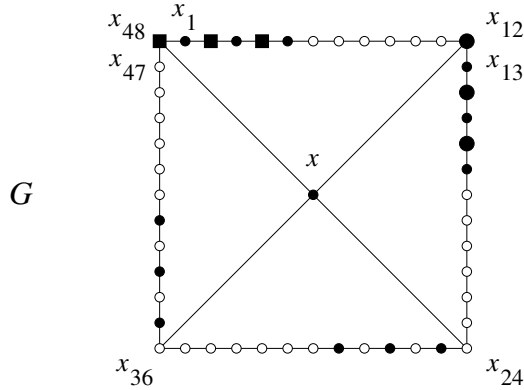


Figure 1: Graph G in Proposition 9, for $r = 6$ and $k = 4$. Squares and circles, white or black, small or large, are vertices. The 19 black vertices constitute a 6-identifying code in G .

Remark preceding the proof. The lower bound (4) is equivalent to $n/4$ when n increases with respect to r . An open question is whether this can be improved. The lower bound (5) is equivalent to 2, but will be strongly improved in Proposition 15.

Proof of Proposition 9. Let $r \geq 2$ be an even integer, and n be an (even) integer such that $n = k \cdot 2r$, $k \geq 2$; let $G_x = C_n = x_1x_2 \dots x_n$ be the cycle of length n and G be the graph obtained from G_x by adding the vertex x and linking it to the k vertices $x_{j \cdot 2r}$, $1 \leq j \leq k$. See Figure 1, which illustrates the case $r = 6, k = 4, n = 48$ and G has 49 vertices.

We know by Theorem 8 that $\gamma_r(G_x) = \frac{n}{2}$, and we claim that

$$\gamma_r(G) \leq 1 + (k + 2) \frac{n}{4k} = \frac{n}{4} + r + 1,$$

from which (4) and (5) follow. Proving this claim, by exhibiting an r -identifying code for G , is tedious and of no special interest; therefore, we content ourselves with showing how it works in the case $r = 6, n = 48$, hoping that this will help the reader to gain an insight into the general case. We consider a first set

$$\mathcal{S} = \{x, x_1, x_3, x_5, x_{13}, x_{15}, x_{17}, x_{25}, x_{27}, x_{29}, x_{37}, x_{39}, x_{41}\},$$

see the small black circles in Figure 1. It is now quite straightforward to observe that the pairs $\{x_{48}, x_1\}$, $\{x_2, x_3\}$ and $\{x_4, x_5\}$ are pairs of $(G, \mathcal{S}, 6)$ -twins, as well as $\{x_{12}, x_{13}\}$, $\{x_{14}, x_{15}\}$, $\{x_{16}, x_{17}\}$, $\{x_{24}, x_{25}\}$, $\{x_{26}, x_{27}\}$, $\{x_{28}, x_{29}\}$, $\{x_{36}, x_{37}\}$, $\{x_{38}, x_{39}\}$ and $\{x_{40}, x_{41}\}$, for reasons of symmetry, and that they are the only ones.

Let us consider the first three pairs, $\{x_{48}, x_1\}$, $\{x_2, x_3\}$, $\{x_4, x_5\}$. Using edges going through x , they can be 6-separated, for instance, by the vertices

x_{16} , x_{14} and x_{12} (see the large black circles), and these three vertices also 6-separate the other pairs of $(G, \mathcal{S}, 6)$ -twins, except for $\{x_{12}, x_{13}\}$, $\{x_{14}, x_{15}\}$, $\{x_{16}, x_{17}\}$. These three pairs can however be 6-separated by three more codewords, for instance x_4 , x_2 and x_{48} , see the black squares in Figure 1. Now the code

$$\mathcal{C} = \mathcal{S} \cup \{x_{12}, x_{14}, x_{16}, x_{48}, x_2, x_4\}$$

is 6-identifying in G and has $1 + (4 \times 3) + (2 \times 3) = 19$ codewords.

In the general case,

$$\mathcal{S} = \{x\} \cup \{x_{1+j \cdot 2r}, x_{3+j \cdot 2r}, \dots, x_{r-1+j \cdot 2r} : 0 \leq j \leq k-1\},$$

there are $k \times \frac{r}{2}$ pairs of (G, \mathcal{S}, r) -twins, and \mathcal{C} can be chosen, for instance, as

$$\mathcal{C} = \mathcal{S} \cup \{x_n, x_2, \dots, x_{r-2}\} \cup \{x_{2r}, x_{2r+2}, \dots, x_{2r+(r-2)}\},$$

which shows that the cardinality of \mathcal{C} is

$$1 + (k \times \frac{r}{2}) + (2 \times \frac{r}{2}) = 1 + (k+2) \frac{n}{4k},$$

and so $\gamma_r(G) \leq 1 + (k+2) \frac{n}{4k}$. \triangle

Conclusion 10 *When r is even, Proposition 9 gives pairs of connected graphs proving that $\gamma_r(G_x) - \gamma_r(G)$ can be, asymptotically, as large as approximately $\frac{n}{4}$.*

• (ii) **Case of a connected graph G_x and $r \geq 3$, r odd**

Proposition 11 *There exist two (connected) r -twin-free graphs G and G_x , with $n+1$ and n vertices respectively, such that*

$$\gamma_r(G_x) - \gamma_r(G) \geq \frac{n(3r-1)}{12r} - r, \quad (6)$$

$$\frac{\gamma_r(G_x)}{\gamma_r(G)} \geq \frac{6nr}{n(3r+1) + 12r^2}. \quad (7)$$

Remark preceding the proof. An open question is whether the first lower bound, which is equivalent to $\frac{n(3r-1)}{12r}$ when r is fixed and n goes to infinity, can be improved. The second lower bound, equivalent to $\frac{6r}{3r+1}$, will be improved in Proposition 15.

Proof of Proposition 11. Let $r \geq 3$ be an odd integer, and n be an (even) integer such that $n = k \cdot 2r$, where $k \geq 3$ is a multiple of 3; let $G_x = C_n = x_1 x_2 \dots x_n$ be the cycle of length n and G be the graph obtained from G_x by adding the vertex x and linking it to the k vertices $x_{j \cdot 2r}$, $1 \leq j \leq k$. See Figure 2, which illustrates the case $r = 5, k = 6, n = 60$ and G has 61 vertices.

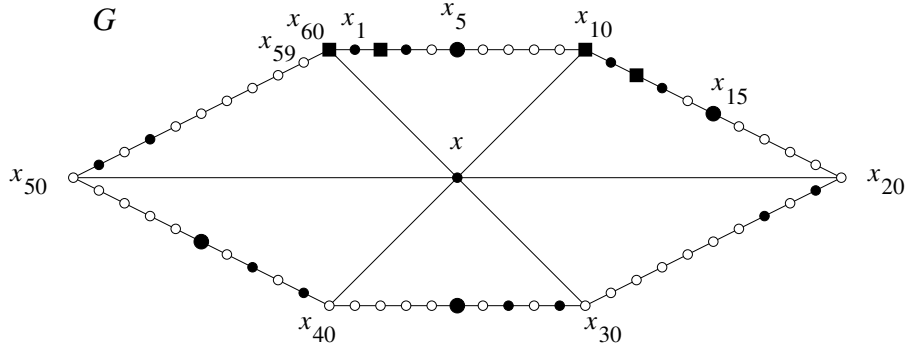


Figure 2: Graph G in Proposition 11, for $r = 5$ and $k = 6$. Squares and circles, white or black, small or large, are vertices. The 21 black vertices constitute a 5-identifying code in G .

We know by Theorem 8 that $\gamma_r(G_x) = \frac{n}{2}$, and we claim that

$$\gamma_r(G) \leq \frac{n}{4} + \frac{n}{12r} + r,$$

from which (6) and (7) follow. Again, proving this claim is of no interest here, and we just show how it works in the case $r = 5$, $n = 60$. We consider a first set

$$\mathcal{S} = \{x, x_1, x_3, x_{11}, x_{13}, x_{21}, x_{23}, x_{31}, x_{33}, x_{41}, x_{43}, x_{51}, x_{53}\},$$

see the small black circles in Figure 2. It is straightforward to see that only the following sets of $(G, \mathcal{S}, 5)$ -twins exist:

- (i) $\{x, x_{10}, x_{20}, x_{30}, x_{40}, x_{50}, x_{60}\}$,
- (ii) $\{x_{59}, x_1, x_2\}$ together with the five symmetrical sets $\{x_9, x_{11}, x_{12}\}, \dots$,
- (iii) $\{x_3, x_4\}$ together with the five symmetrical sets $\{x_{13}, x_{14}\}, \dots$

The first two cases are annoying and will be “expensive” because they present symmetries with respect to x . Define the set \mathcal{T} as follows:

$$\mathcal{T} = \mathcal{S} \cup \{x_5, x_{15}, x_{35}, x_{45}\},$$

see the large black circles in Figure 2. Now in Case (i), all the vertices are 5-separated by the vertices in $\mathcal{T} \setminus \mathcal{S}$, and so are x_{59} on the one hand and x_1, x_2 on the other hand, as well as their symmetrical counterparts from Case (ii). The remaining pairs of $(G, \mathcal{T}, 5)$ -twins are $\{x_1, x_2\}$, $\{x_3, x_4\}$ and the 10 pairs obtained by symmetry. As in the proof of Proposition 9, these handle very economically: the vertex x_{60} 5-separates the 5 pairs $\{x_{13}, x_{14}\}, \dots, \{x_{53}, x_{54}\}$, and so does x_2 for $\{x_{11}, x_{12}\}, \dots, \{x_{51}, x_{52}\}$; finally, $\{x_1, x_2\}$ and $\{x_3, x_4\}$ can be 5-separated, for instance, by x_{10} and x_{12} , see the black squares in Figure 2:

$$\mathcal{C} = \mathcal{T} \cup \{x_{60}, x_2, x_{10}, x_{12}\}$$

is a 5-identifying code in G and has $1 + (6 \times 2) + (4 \times 1) + (2 \times 2) = 21$ codewords. In the general case,

$$\mathcal{S} = \{x\} \cup \{x_{1+j \cdot 2r}, x_{3+j \cdot 2r}, \dots, x_{r-2+j \cdot 2r} : 0 \leq j \leq k-1\}$$

contains $1 + (k \times \frac{r-1}{2})$ vertices; then

$$\mathcal{T} = \mathcal{S} \cup \{x_{r+j \cdot 2r} : 0 \leq j \leq k-1, j \text{ not congruent to } 2 \text{ modulo } 3\}$$

contains $|\mathcal{S}| + \frac{2k}{3}$ elements, and finally we take

$$\mathcal{C} = \mathcal{T} \cup \{x_n, x_2, \dots, x_{r-3}\} \cup \{x_{2r}, x_{2r+2}, \dots, x_{2r+(r-3)}\},$$

which shows that

$$\gamma_r(G) \leq 1 + (k \times \frac{r-1}{2}) + \frac{2k}{3} + (2 \times \frac{r-1}{2}) = \frac{n}{4} + \frac{n}{12r} + r.$$

△

Conclusion 12 *When $r \geq 3$ and r is odd, Proposition 11 gives pairs of connected graphs proving that $\gamma_r(G_x) - \gamma_r(G)$ can be, asymptotically, as large as approximately $\frac{n(3r-1)}{12r}$.*

If we do not require to consider a connected graph G_x , then we can obtain a larger difference or ratio than in (4)-(7), we need consider only one case, whatever the parity of r is, and moreover the construction is easy to understand; see next paragraph.

• **(iii) Case of a disconnected graph G_x and $r \geq 2$, r even or odd**

Proposition 13 *There exist two graphs G and G_x , with $p(2r+1)+1$ and $n = p(2r+1)$ vertices respectively, such that*

$$\gamma_r(G_x) - \gamma_r(G) \geq \frac{n(2r-2)}{2r+1} - 2r, \quad (8)$$

$$\frac{\gamma_r(G_x)}{\gamma_r(G)} \geq \frac{nr}{n+4r^2+2r}. \quad (9)$$

Remark preceding the proof. Can the first lower bound, equivalent to $\frac{n(2r-2)}{2r+1}$, be improved? The second bound, equivalent to r , is still improved in Proposition 15.

Proof of Proposition 13. Let $r \geq 2$ and $p \geq 3$ be integers; the graph G_x consists of p copies of the path P_{2r+1} , and G is obtained by adding the vertex x and linking it to all the middle vertices of the path copies, see Figure 3. We claim that: (a) $\gamma_r(G_x) = 2pr$ and (b) $\gamma_r(G) \leq 2p + 2r$, from which (8) and (9) follow.

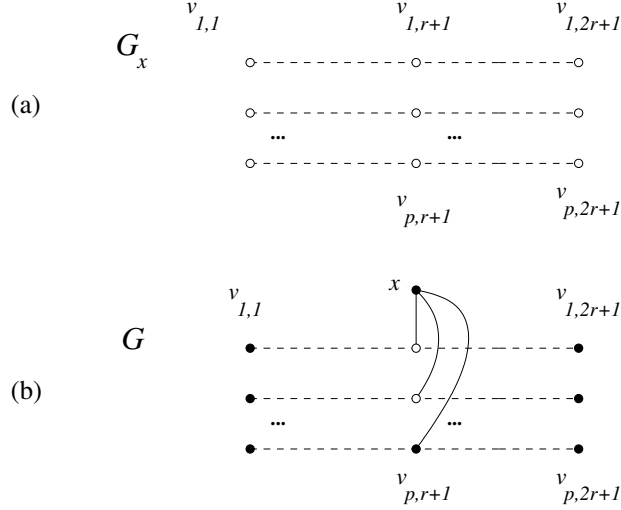


Figure 3: The graphs G_x and G in Proposition 13.

Proof of (a). The result comes from the obvious fact that $\gamma_r(P_{2r+1}) = 2r$.
 Proof of (b). It is not difficult to check that

$$\mathcal{C} = \{x\} \cup \{v_{i,1}, v_{i,2r+1} : 1 \leq i \leq p-1\} \cup \{v_{p,j} : 1 \leq j \leq 2r+1\}$$

(see the black circles in Figure 3) is indeed r -identifying in G . Note however that, for simplicity, we chose to give the bound $2p+2r$, when actually, with a little more care, $2p+2r-3$ can be reached, which would improve only slightly on (8) and (9). \triangle

Conclusion 14 Proposition 13 gives pairs of graphs (G, G_x) , where G_x is not connected, proving that $\gamma_r(G_x) - \gamma_r(G)$ can be, asymptotically, as large as approximately $\frac{n(2r-2)}{2r+1}$.

Finally, we give a construction (obtained with connected graphs) with a ratio $\gamma_r(G_x)/\gamma_r(G)$ arbitrarily large, but where the difference $\gamma_r(G_x) - \gamma_r(G)$ is not as large as in (4) and (6).

Proposition 15 Let $k \geq 2$ be an arbitrary integer. There exist two (connected) r -twin-free graphs G and G_x , where G has $2rk + r \lceil \log_2(k+1) \rceil + r + 1$ vertices, such that

$$\frac{\gamma_r(G_x)}{\gamma_r(G)} \geq \frac{k}{r \lceil \log_2(k+1) \rceil + r + 1}. \quad (10)$$

Proof. The construction is a straightforward generalization to any $r \geq 2$ of the one used in the proof of Proposition 4, see Figure 4; the basic idea is similar, but the implementation becomes somewhat more involved.

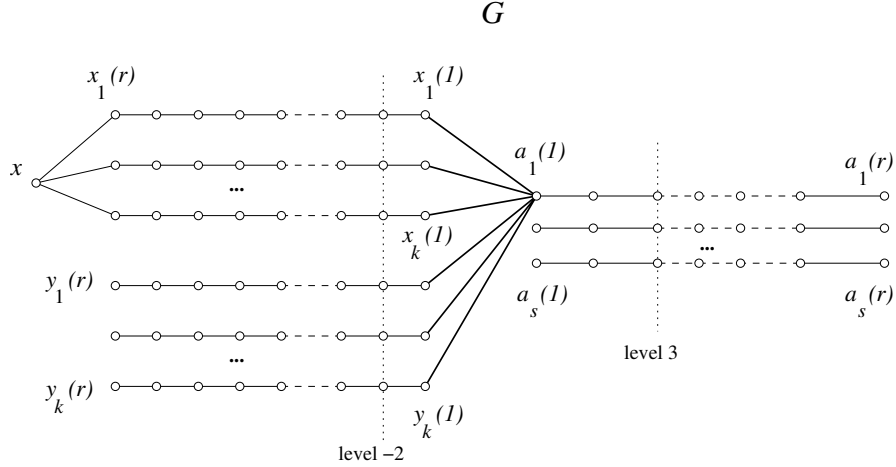


Figure 4: A partial representation of the graph G in Proposition 15: more edges exist between the vertices $x_i(1)$ and $y_i(1)$ on the one hand, and the vertices $a_j(1)$ on the other hand. The case $r = 1$ can be used to illustrate Proposition 4.

We consider, for each i between 1 and k , the paths $x_i(1)x_i(2)\dots x_i(r)$, and $y_i(1)y_i(2)\dots y_i(r)$. We need also some auxiliary vertices. Denoting again $s = \lceil \log_2(k+1) \rceil + 1$, for each $j = 1, 2, \dots, s$, we consider the path $a_j(1)a_j(2)\dots a_j(r)$; we denote the set of these sr auxiliary vertices by A . We say that the vertices $x_i(-h)$, $y_i(-h)$ and $a_j(h)$ are on the h -th level (cf. Figure 4).

We now imitate the proof of Proposition 4, and for each $i \in \{1, \dots, k\}$ choose a unique nonempty subset A_i of the set $\{a_2(1), \dots, a_s(1)\}$ and connect $x_i(1)$ and $y_i(1)$ by an edge to the vertices $a_j(1)$ for which $j \in \{1\} \cup A_i$.

In the resulting graph G_x , we first take all the vertices in A as codewords. Then we observe that for an arbitrary, unknown vertex v ,

- $B_r(v)$ contains at least two vertices $a_j(r)$ if v is on the level -1;
- $B_r(v)$ does not contain any vertices $a_j(r)$ if v is on the h -th level for some $h \leq -2$; and
- $B_r(v)$ contains exactly one $a_j(r)$ if $v \in A$.

From the last case we see that we can uniquely tell whether or not $v \in A$ simply by looking which vertices of A are in $B_r(v)$. We can in fact do even more: if j is the only index for which $a_j(r)$ is in $B_r(v)$, then v is one of the vertices $a_j(h)$ for some $h = 1, 2, \dots, r$. We know that $a_j(1)$ is connected to at least one $x_i(1)$ (as we chose s to be as small as possible) and $x_i(1)$ is connected to at least one $a_{j'}(1)$ with $j' \neq j$. Then exactly $r - h$ of the vertices $a_{j'}(1), \dots, a_{j'}(r)$ are in $B_r(v)$, and this uniquely identifies v .

Assume now that we already know that $v \notin A$. Let h be the highest level for which some $a_j(h)$ belongs to $B_r(v)$. Then v must be one of the vertices $x_i(r+1-h)$ or $y_i(r+1-h)$, and moreover, we can uniquely tell i by looking at the indices j for which $a_j(h)$ belong to $B_r(v)$, because by the construction $\{j : a_j(h) \in B_r(v)\} = \{1\} \cup A_i$ (as we can only reach these vertices from v by going from v to $x_i(1)$ or $y_i(1)$ and from it directly to those $a_j(1)$ to which $x_i(1)$ or $y_i(1)$ was connected to by an edge).

In conclusion, by only looking at which auxiliary vertices are in $B_r(v)$ we can "almost" identify v : we find indices i and m such that v is either $x_i(m)$ or $y_i(m)$. This implies that the graph is clearly r -twin-free. Indeed, if all the vertices are in the code, then the only remaining task, i.e., separating each $x_i(m)$ from $y_i(m)$, becomes easy: if $x_i(r)$ is in $B_r(v)$ then $v = x_i(m)$; if not then $v = y_i(m)$.

Moreover, every r -identifying code must contain at least one element of the set $\{x_i(1), x_i(2), \dots, x_i(r), y_i(1), y_i(2), \dots, y_i(r)\}$: otherwise we cannot r -separate $x_i(1)$ and $y_i(1)$. Consequently, any r -identifying code in this graph has size at least k .

We now add one more vertex x , and connect it by an edge to each $x_i(r)$. We claim that the vertex x together with all the vertices in A form an r -identifying code. By the construction, the set $B_r(v)$, $v \neq x$, contains exactly the same vertices of A as before adding the vertex x (and the set $B_r(x)$ contains none), so the only thing to check is that $x_i(m)$ and $y_i(m)$ can now be r -separated: but this is indeed done by x . \triangle

Remark. In terms of $n = 2rk + r\lceil \log_2(k+1) \rceil$, the lower bound (10) can be approximated by $\frac{n}{2r^2 \log_2 n}$ and is open to improvements.

Conclusion 16 *Proposition 15 gives pairs of (connected) graphs proving that $\gamma_r(G_x)/\gamma_r(G)$ can be, asymptotically, as large as approximately $\frac{n}{2r^2 \log_2 n}$.*

Then we turn to examples where $\gamma_r(G) - \gamma_r(G_x)$ is arbitrarily large. Note that we obtain this result with connected graphs.

Proposition 17 *There exist two (connected) r -twin-free graphs G_x and G , with $n = pr + 1$ and $pr + 2$ vertices respectively, such that*

$$\gamma_r(G_x) = p + 2r - 3 = \frac{n + 2r^2 - 3r - 1}{r} \quad \text{and} \quad \gamma_r(G) = r(p - 1) + 1 = n - r,$$

where p is any integer greater than or equal to 3.

Proof. Let $r \geq 2$ and $p \geq 3$ be integers; before defining G , we describe G_x in the following informal way, illustrated in Figure 5(a): G_x consists of p copies of the path P_r , and in each copy the last vertex is linked to v . This graph has $n = pr + 1$ vertices. Next, we construct the graph G consisting of G_x to which we add one vertex x , linked to each first vertex of all the

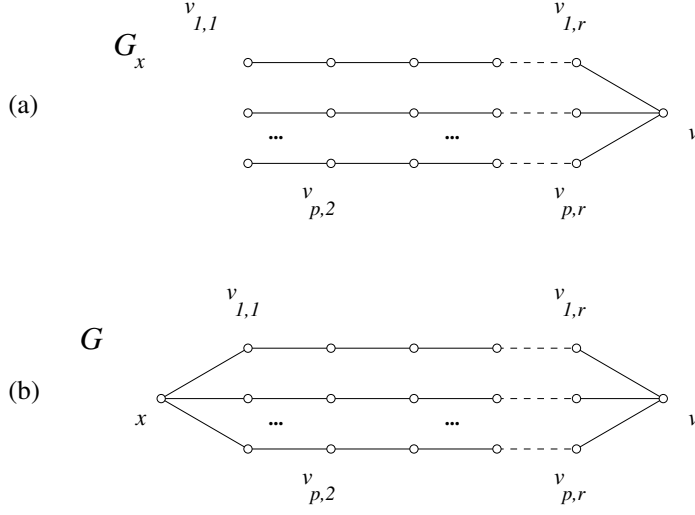


Figure 5: The graphs G_x and G in Proposition 17.

copies of P_r . See Figure 5(b). We claim that: (a) $\gamma_r(G_x) = p + 2r - 3$, and (b) $\gamma_r(G) = r(p - 1) + 1$, from which (13) and (14) follow.

Proof of (a). The code

$$\mathcal{C} = \{v_{1,i} : 2 \leq i \leq r\} \cup \{v_{2,i} : 1 \leq i \leq r\} \cup \{v_{j,1} : 3 \leq j \leq p\},$$

i.e., the code consisting of all the vertices of the first two copies of P_r , except $v_{1,1}$, and the first vertex of each of the following copies, is r -identifying in G_x ; this it is straightforward to check. So $\gamma_r(G_x) \leq (r - 1) + r + (p - 2) = p + 2r - 3$. We now prove that $\gamma_r(G_x) \geq p + 2r - 3$. The following two observations will be useful. For $1 \leq i \leq p$ and $2 \leq k \leq r$, we have:

$$B_{G_x,r}(v_{i,r-k+1}) \Delta B_{G_x,r}(v_{i,r-k+2}) = \{v_{j,k} : 1 \leq j \leq p, j \neq i\}, \quad (11)$$

where Δ stands for the symmetric difference, and for $1 \leq i < j \leq p$:

$$B_{G_x,r}(v_{i,r}) \Delta B_{G_x,r}(v_{j,r}) = \{v_{i,1}, v_{j,1}\}. \quad (12)$$

The consequences are immediate. First, in order to have the vertices $v_{i,r}$, $1 \leq i \leq p$, pairwise r -separated in G_x , we see by (12) that we need at least $p - 1$ codewords among the p vertices $v_{i,1}$; second, for k fixed between 2 and r , we see, using (11), that we need at least two codewords among the p vertices $v_{i,k}$. So $\gamma_r(G_x) \geq (p - 1) + 2(r - 1) = p + 2r - 3$, and Claim (a) is proved.

Proof of (b). Note that in G , for i and j such that $1 \leq i < j \leq p$, the set of vertices

$$\{x\} \cup \{v_{i,k} : 1 \leq k \leq r\} \cup \{v\} \cup \{v_{j,k} : 1 \leq k \leq r\}$$

forms the cycle C_{2r+2} , which is r -twin-free and is denoted by $C(i, j)$. On such a cycle, we say that the vertex z is the *opposite* of the vertex y if z is the (only) vertex at distance $r + 1$ from y .

We claim that, for k fixed between 1 and r , among the p vertices $v_{i,k}$, at least $p - 1$ of them belong to any r -identifying code \mathcal{C} in G . Indeed, assume on the contrary that two vertices, say $v_{1,k}$ and $v_{2,k}$, are not in \mathcal{C} ; then their opposite vertices in $C(1, 2)$, $v_{2,r-k+1}$ and $v_{1,r-k+1}$ respectively, cannot be r -separated by \mathcal{C} .

Finally, the fact that $B_{G,r}(v) \Delta B_{G,r}(x) = \{v, x\}$ shows that v or x belong to \mathcal{C} , and finally $\gamma_r(G) \geq (p - 1)r + 1$. On the other hand,

$$\{v\} \cup \{v_{i,k} : 2 \leq i \leq p, 1 \leq k \leq r\}$$

is an r -identifying code in G , with size $(p - 1)r + 1$, thus Claim (b) is proved. Observe that this code contains all the vertices in G , except the $r + 1$ vertices x and $v_{1,k}$, $1 \leq k \leq r$. \triangle

Note that we could have contented ourselves with the inequalities $\gamma_r(G_x) \leq p + 2r - 3$ and $\gamma_r(G) \geq r(p - 1) + 1$, so as to obtain $\gamma_r(G) - \gamma_r(G_x) \geq p(r - 1) - 3r + 4$ and $\frac{\gamma_r(G)}{\gamma_r(G_x)} \geq \frac{r(p-1)+1}{p+2r-3}$.

Remark. The difference

$$\gamma_r(G) - \gamma_r(G_x) = p(r - 1) - 3r + 4$$

can be made arbitrarily large; in terms of n , the number of vertices of G_x , we can see that we have:

$$\gamma_r(G) - \gamma_r(G_x) = \frac{(n - 3r)(r - 1) + 1}{r}, \quad (13)$$

which is equivalent to $\frac{n(r-1)}{r}$ when r is fixed and n goes to infinity. As far as the ratio given by Proposition 17 is concerned, we have:

$$\frac{\gamma_r(G)}{\gamma_r(G_x)} = \frac{r(n - r)}{n + 2r^2 - 3r - 1}, \quad (14)$$

which is equivalent to r when we increase n . This can be improved, with a ratio which becomes arbitrarily large; again, it so happens that the graphs are connected:

Proposition 18 *Let $k \geq 2$ be an arbitrary integer.*

There exist two (connected) 2-twin-free graphs G and G_x , where G has $3k + 2\lceil \log_2(k + 2) \rceil + 4$ vertices, such that

$$\frac{\gamma_2(G)}{\gamma_2(G_x)} \geq \frac{k}{2\lceil \log_2(k + 2) \rceil + 3}. \quad (15)$$

Let $r \geq 3$. There exist two (connected) r -twin-free graphs G and G_x , where G has $(r+1)k + r \lceil \log_2(k+2) \rceil + 2r + 1$ vertices, such that

$$\frac{\gamma_r(G)}{\gamma_r(G_x)} \geq \frac{k}{r \lceil \log_2(k+2) \rceil + r + 3}. \quad (16)$$

Proof. We first deal with the general case $r \geq 3$. We construct the graph G for a given $k \geq 2$ in the following way, see Figure 6: G consists of the paths $x_i(0)x(1)x(2) \dots x(r-2)x(r-1)x$ and $y_i(0)y_i(1) \dots y_i(r-1)$, for $i = 1, \dots, k$, of the path $a_1(1) \dots a_1(r+1)$, of the paths $a_j(1) \dots a_j(r)$ for $j = 2, \dots, s$, where $s = 1 + \lceil \log_2(k+2) \rceil$, plus the edge $xa_1(1)$ and the following edges, joining exclusively the vertices $x_i(0)$ and $y_i(0)$ on the one hand, and the vertices $a_j(1)$ on the other hand: for each i we choose a unique nonempty proper subset A_i of the set $A = \{2, 3, \dots, s\}$, and connect every $x_i(0)$ and every $y_i(0)$ to every vertex $a_j(1)$ for which $j \in A_i$. Moreover, we connect every $x_i(0)$ and every $y_i(0)$ to $a_1(1)$. The sets A_i can indeed be chosen in this way, because there are $2^{s-1} - 2$ proper nonempty subsets of A , and $s - 1 = \lceil \log_2(k+2) \rceil$. Without loss of generality, we can choose the sets A_i in such a way that each $a_j(1)$ has degree at least two, and so the graph constructed is connected, as will be G_x .

We say that the vertices $x(-h)$, $x_i(-h)$, $y_i(-h)$ and $a_j(h)$ are on the h -th level, cf. Figure 6 (and x is not given any level). Let

$$\mathcal{A} = \{a_j(h) : 1 \leq j \leq s, 1 \leq h \leq r\} \cup \{a_1(r+1)\}.$$

Let us first consider G_x , and let $\mathcal{C} = \mathcal{A} \cup \{x_1(0), x(r-1)\}$. We show that \mathcal{C} is r -identifying, so that $\gamma_r(G_x) \leq sr + 3$. The argument is very similar to the first part of the proof of Proposition 15: let v be an arbitrary, unknown vertex in G_x .

If v belongs to \mathcal{A} , then v is r -covered by exactly one codeword $a_j(r)$, whereas every vertex of level 0 is r -covered by at least two codewords of level r , and no vertex with negative level is r -covered by any codeword of level r ; if $v \in \mathcal{A}$ is r -covered by $a_j(r)$, we know moreover that $v = a_j(h)$ for some h between 1 and $r+1$. If $h < r$, then h is given by the highest level ℓ of any codeword $a_{j'}(\ell)$ r -covering $a_j(h)$, with $j' \neq j$ (such a j' exists because $a_j(1)$ is connected to at least one $x_i(0)$, which in turn is connected to at least one $a_{j'}(1)$). If $j \neq 1$ and $h = r$, then h is given by the fact that no codeword $a_{j'}(\ell)$ ($j' \neq j$) r -covers $a_j(h)$. And if $j = 1$ and $h \in \{r, r+1\}$, then the codeword $x_1(0)$ tells whether $h = r$ or $h = r+1$. This means that we can determine first that $v \in \mathcal{A}$, then on which path and at which level it is located.

If $v \notin \mathcal{A}$, then its level can be determined by the highest level, say ℓ , of the codewords in \mathcal{A} which r -cover it. Then the codeword $x(r-1)$ tells if v is of type x or y ; and finally, if $v = x_i(0)$ or $v = y_i(h)$ for some h between 0 and $r-1$, then we can uniquely tell i by looking at the indices j for which

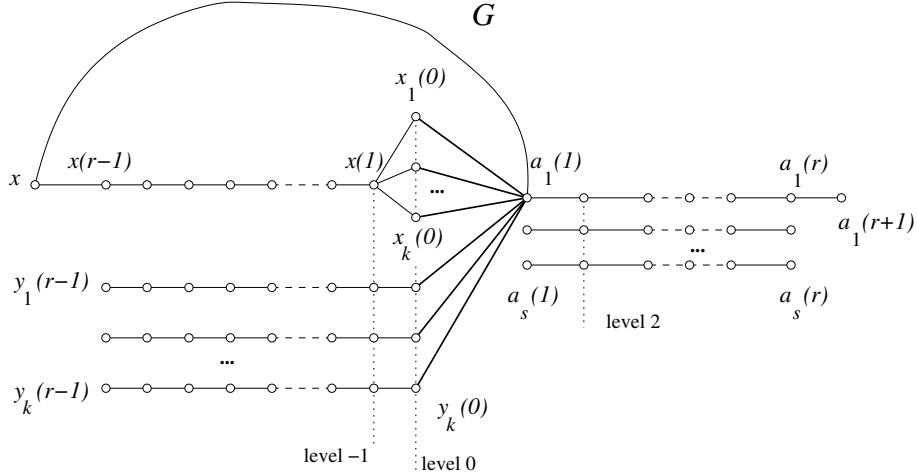


Figure 6: A partial representation of the graph G in Proposition 18, in the general case $r \geq 3$: more edges exist between the vertices $x_i(0)$ and $y_i(0)$ on the one hand, and the vertices $a_j(1)$ on the other hand.

$a_j(\ell) \in B_r(v)$, because by the construction $\{j : a_j(\ell) \in B_r(v)\} = \{1\} \cup A_i$. This ends the study of G_x .

We now consider the graph G , and prove that it is r -twin-free. Comparing with the previous graph G_x , it is still true that every vertex in \mathcal{A} is r -covered by exactly one vertex $a_j(r)$, whereas every vertex of level 0 is r -covered by at least two vertices of level r , and no vertex with negative level is r -covered by any vertex of level r – and note that x is r -covered by exactly one $a_j(r)$, namely $a_1(r)$; it is still true that no two vertices inside \mathcal{A} are r -twins, that one vertex in \mathcal{A} and one vertex of type y or x (except maybe x itself) are not r -twins, and that no two vertices of type y are r -twins; also, thanks to the vertices $y_i(r-1)$, no vertex of type y can be r -twin with a vertex of type x ; but we have to see what happens with the vertices of type x between themselves, and with the vertex x and one vertex in \mathcal{A} .

Now x is not r -twin with any $a_j(h)$, $j > 1$, thanks to $a_j(r)$, and not either with any $a_1(h)$, thanks to $a_1(r+1)$ – note in particular that $a_1(r+1)$ is the only vertex r -separating x and $a_1(2)$. Assume finally that v is of type x , $v \neq x$. If $v = x_i(0)$ for some i , the set of indices j for which $a_j(r) \in B_r(v)$ equals $\{1\} \cup A_i$, has size at least two, and identifies v . So assume that v is not on level 0, and denote by $h \in \{1, 2, \dots, r-1\}$ the largest level for which at least one $a_j(h)$ belongs to $B_r(v)$. If the **only** shortest path between v and $a_1(1)$ goes via x , then $\{j : a_j(h) \in B_r(v)\} = \{1\}$; if there is a shortest path between v and $a_1(1)$ that goes via one (and hence all) $x_i(0)$, then $\{j : a_j(h) \in B_r(v)\} = \{1, 2, \dots, s\}$: in both cases, h uniquely identifies v .

Ultimately, what is the smallest size of an r -identifying code in G ? For

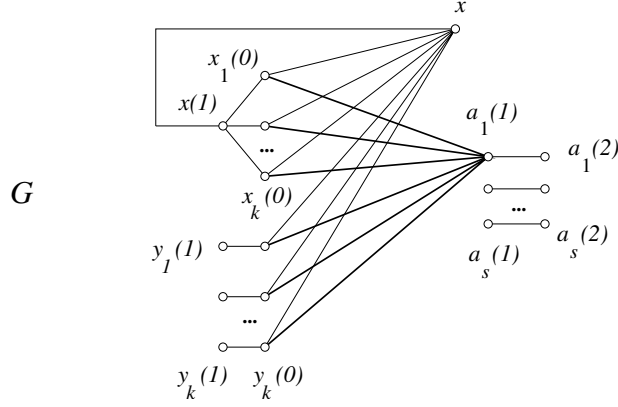


Figure 7: A partial representation of the graph G in Proposition 18, in the particular case $r = 2$: more edges exist between the vertices $x_i(0)$ and $y_i(0)$ on the one hand, and the vertices $a_j(1)$ on the other hand.

a given i between 1 and k , it is easy to see that we have:

$$B_r(y_i(0)) = B_r(x_i(0)) \cup \{y_i(r-1)\}, \quad (17)$$

where the right-hand side is a disjoint union; this shows that any r -identifying code in G contains at least k elements, and ends the case $r \geq 3$. Note that if we had considered this construction for $r = 2$, then (17) would not be true, since $x(1)$ would be in $B_2(x_i(0)) \setminus B_2(y_i(0))$.

When $r = 2$, the previous construction does not work, as we have just seen, but the following does: the x -paths are again $x_i(0)x(1)x$, and the y -paths are $y_i(0)y_i(1)$ as before; the vertex $a_1(3)$ is removed, and, keeping all the edges between the vertices of level 1 in \mathcal{A} and the vertices of level 0 as before, we add all the edges between x and the vertices of level 0; see Figure 7.

It is then rather straightforward, using the same kind of argument as in the general case, to check that $\mathcal{C} = \{a_j(h) : 1 \leq j \leq s, 1 \leq h \leq 2\} \cup \{x(1)\}$ is 2-identifying in G_x , that G is 2-twin-free, and that any 2-identifying code in G needs at least k codewords. \triangle

Remark. In terms of $n = (r+1)k + r \lceil \log_2(k+2) \rceil$, the approximate order of G and G_x , we can approximate the lower bounds in (15) and (16) by $\frac{n}{r(r+1) \log_2 n}$. Again, can the bounds given in (13), (15) and (16) be significantly improved?

Conclusion 19 *When $r \geq 2$, Propositions 17 and 18 provide pairs of graphs proving that $\gamma_r(G_x) - \gamma_r(G)$ can be, asymptotically, as small as approximately $-\frac{n(r-1)}{r}$, and $\frac{\gamma_r(H_x)}{\gamma_r(H)}$ can be, asymptotically, as small as approximately $\frac{r(r+1) \log_2 n}{n}$, and both can even be obtained with connected examples.*

5 General conclusion

Table 1 recapitulates the results obtained in the previous sections, using in particular the Conclusions 7, 10, 12, 14, 16 and 19; these are stated for n large with respect to r , where n is the approximate order of G or of G_x ; when using $\gtrsim X$ (respectively, $\lesssim X$), we mean that we have a lower bound (respectively, an upper bound), for the difference or ratio, which is approximately X . We only consider the difference $\gamma_r(G_x) - \gamma_r(G)$ and the ratio $\frac{\gamma_r(G_x)}{\gamma_r(G)}$.

r	r	comment	$\gamma_r(G_x) - \gamma_r(G)$	$\frac{\gamma_r(G_x)}{\gamma_r(G)}$	reference
$r = 1$		impossible to have (connected) graphs	< -1 $\gtrsim \frac{n}{2} - \frac{3}{2} \log_2 n$	$\gtrsim \frac{n}{2 \log_2 n}$	Concl. 7
	even	connected graphs	$\gtrsim \frac{n}{4}$		Concl. 10
≥ 2	odd	connected graphs	$\gtrsim \frac{n(3r-1)}{12r}$		Concl. 12
		graphs	$\gtrsim \frac{n(2r-2)}{2r+1}$		Concl. 14
	any	(connected) graphs		$\gtrsim \frac{n}{2r^2 \log_2 n}$	Concl. 16
≥ 2	any	(connected) graphs	$\lesssim -\frac{n(r-1)}{r}$	$\lesssim \frac{r(r+1) \log_2 n}{n}$	Concl. 19

Table 1: The difference $\gamma_r(G_x) - \gamma_r(G)$ and ratio $\frac{\gamma_r(G_x)}{\gamma_r(G)}$, as functions of n and r .

Part II: Addition and deletion of one edge

This part is organized as follows. Section 6 is devoted to the case $r = 1$; here, the difference $\gamma_1(G_e) - \gamma_1(G)$ must lie between -2 and $+2$. Then in the beginning of Section 7, we study how *small* $\gamma_r(G_e) - \gamma_r(G)$ and $\gamma_r(G_e)/\gamma_r(G)$ can be for any $r \geq 2$, and it so happens that the graphs we use are connected (Corollary 27 for $r \geq 5$ and Proposition 28 for $r \in \{2, 3, 4\}$); finally, we study how *large* these difference and ratio can be, for $r \geq 3$ in Corollary 29 and for $r = 2$ in Proposition 30 (in both cases, the graphs can be made connected).

A conclusion recapitulates our results in a Table.

6 The case $r = 1$

The difference $\gamma_1(G_e) - \gamma_1(G)$ can vary only inside the set $\{-2, -1, 0, 1, 2\}$ (Theorem 24), and these five values can be reached (Examples 21, 23 and 25).

We first study how small $\gamma_1(G_e) - \gamma_1(G)$ can be. Putting the cart before the horse, in the next theorem we first define G_e , and only then, G .

Theorem 20 *Let $G_e = (V, E_e)$ be a 1-twin-free graph with at least four vertices, let x and y be two distinct vertices in V such that $e = xy \notin E_e$, and let $G = (V, E)$ with $E = E_e \cup \{xy\}$. Assume that G is also 1-twin-free.*

If \mathcal{C}_e is a 1-identifying code in G_e , then there exists a 1-identifying code \mathcal{C} in G with

$$|\mathcal{C}| \leq |\mathcal{C}_e| + 2.$$

As a consequence, we have:

$$\gamma_1(G_e) - \gamma_1(G) \geq -2. \tag{18}$$

Proof. Since we add an edge when going from G_e to G , all vertices remain 1-covered, in G , by at least one codeword in \mathcal{C}_e .

Since we only add the edge xy , only the balls of x and y are modified in G . As a consequence, only the following pairs are possible $(G, \mathcal{C}_e, 1)$ -twins:

- x and y ,
- x and u with $u \neq x, u \neq y$,
- y and v with $v \neq x, v \neq y$.

Moreover, x and u' , with $u' \neq u, u' \neq x, u' \neq y$, cannot be $(G, \mathcal{C}_e, 1)$ -twins since this would imply, by Lemma 1, that u and u' are $(G, \mathcal{C}_e, 1)$ -twins, hence $(G_e, \mathcal{C}_e, 1)$ -twins, which would contradict the fact that \mathcal{C}_e is 1-identifying in G_e . The same is true for y and v' , with $v' \neq v, v' \neq x, v' \neq y$. So at most three pairs of $(G, \mathcal{C}_e, 1)$ -twins can appear.

Similarly, if these three pairs of $(G, \mathcal{C}_e, 1)$ -twins all do appear, then u and v are $(G, \mathcal{C}_e, 1)$ -twins, which leads to the same contradiction, *unless* $u = v$. In this case, because G is 1-twin-free, we can pick an additional

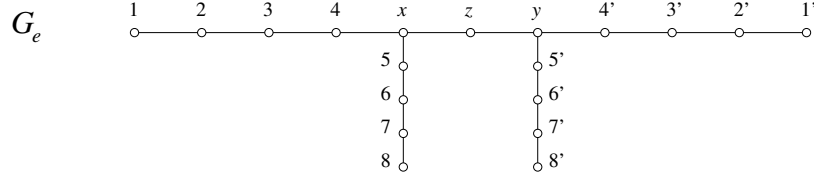


Figure 8: Graph G_e in Example 21.

codeword c_1 1-separating x and u by, say, 1-covering x and not u . If c_1 1-covers y , then c_1 also 1-separates y and u ; if c_1 does not 1-cover y , then c_1 also 1-separates y and x . In both cases, we are left with one pair of vertices not yet 1-separated by a codeword, which we can do with a second additional codeword c_2 . Now $\mathcal{C} = \mathcal{C}_e \cup \{c_1, c_2\}$ is 1-identifying in G , and it has $|\mathcal{C}_e| + 2$ elements.

When at most two pairs of $(G, \mathcal{C}_e, 1)$ -twins appear, then obviously with at most two more codewords added to \mathcal{C}_e we can 1-separate them. \triangle

Note that we made no assumption on the connectivity of G_e . The following example shows that graphs G_e and G with $\gamma_1(G) = \gamma_1(G_e) + 2$ do exist; we do not know if this is the smallest possible example.

Example 21 Let $G_e = (V, E_e)$ be the graph represented in Figure 8, and G the graph obtained by adding the edge $e = xy$. We claim that: (a) $\gamma_1(G_e) \leq 10$ and (b) $\gamma_1(G) \geq 12$, which by (18) implies that $\gamma_1(G) = 12 = \gamma_1(G_e) + 2$.

Proof of (a). It is quite straightforward to check that $\mathcal{C}_e = \{1, 3, x, 6, 8, 8', 6', y, 3', 1'\}$ is 1-identifying in G_e . Hence $\gamma_1(G_e) \leq 10$.

Proof of (b). Let \mathcal{C} be a 1-identifying code in G . Because 1 and 2 must be 1-separated by \mathcal{C} , we have $3 \in \mathcal{C}$; and because 1 must be 1-covered by at least one codeword, we have $1 \in \mathcal{C}$ or $2 \in \mathcal{C}$. Similarly, \mathcal{C} contains $6, 6', 3'$ and at least one element in each of the 2-sets $\{7, 8\}$, $\{8', 7'\}$ and $\{2', 1'\}$, which amounts to eight codewords.

With simple arguments, we obtain the following fact:

- there are at least three codewords in $\{1, 2, 3, 4, x\}$.

The same is true for $\{x, 5, 6, 7, 8\}$, $\{y, 5', 6', 7', 8'\}$ and $\{y, 4', 3', 2', 1'\}$. So, if neither x nor y belongs to \mathcal{C} , there are at least $3 \times 4 = 12$ codewords, and we are done.

If, on the other hand, both x and y belong to \mathcal{C} , then we have already ten codewords, and still x , y and z are not 1-separated by any codeword; this will require two additional codewords, and again, $|\mathcal{C}| \geq 12$.

If we assume finally, without loss of generality, that $x \in \mathcal{C}$ and $y \notin \mathcal{C}$, then we have already chosen $(3 \times 2) + 5 = 11$ codewords: three in each of the

sets $\{4', 3', 2', 1'\}$ and $\{5', 6', 7', 8'\}$, one in each of the sets $\{1, 2\}$ and $\{7, 8\}$, plus 3, 6 and x ; still, x and z are not 1-separated by any codeword, so again we need at least twelve codewords, which proves Claim (b). \triangle

Next, we establish how large $\gamma_1(G_e) - \gamma_1(G)$ can be.

Theorem 22 *Let $G = (V, E)$ be a 1-twin-free graph with at least four vertices, let x and y be two vertices in V such that $e = xy \in E$, and let $G_e = (V, E_e)$ with $E_e = E \setminus \{xy\}$. Assume that G_e is also 1-twin-free.*

If \mathcal{C} is a 1-identifying code in G , then there exists a 1-identifying code \mathcal{C}_e in G_e with

$$|\mathcal{C}_e| \leq |\mathcal{C}| + 2.$$

As a consequence, we have:

$$\gamma_1(G_e) - \gamma_1(G) \leq 2. \tag{19}$$

Proof. We assume that \mathcal{C} is not 1-identifying in G_e anymore, otherwise we are done. There can be two reasons why \mathcal{C} is not 1-identifying:

1) at least one of the two vertices x and y , say x , is not 1-covered by any codeword anymore:

$$B_{G_e,1}(x) \cap \mathcal{C} = \emptyset = (B_{G,1}(x) \setminus \{y\}) \cap \mathcal{C},$$

which implies that $B_{G,1}(x) \cap \mathcal{C} = \{y\}$, $y \in \mathcal{C}$ and $x \notin \mathcal{C}$; we see that in this case y is still 1-covered by a codeword, namely itself.

If meanwhile all vertices remain 1-separated by \mathcal{C} in G_e , then $\mathcal{C} \cup \{x\}$ is 1-identifying in G_e . But this first reason can go along with the second reason:

2) $(G_e, \mathcal{C}, 1)$ -twins appear;

because only the edge xy is deleted when going from G to G_e , and similarly to the proof of Theorem 20, only the following pairs can be $(G_e, \mathcal{C}, 1)$ -twins:

- x and y ,
- x and u with $u \neq x, u \neq y$,
- y and v with $v \neq x, v \neq y$.

If x and y are $(G_e, \mathcal{C}, 1)$ -twins, this means that

$$B_{G_e,1}(x) \cap \mathcal{C} = B_{G_e,1}(y) \cap \mathcal{C},$$

which implies that $x \notin \mathcal{C}$, $y \notin \mathcal{C}$, and so

$$B_{G,1}(x) \cap \mathcal{C} = B_{G,1}(y) \cap \mathcal{C},$$

contradicting the fact that \mathcal{C} is 1-identifying in G .

Assume next that x and u are $(G_e, \mathcal{C}, 1)$ -twins. Then

$$B_{G_e,1}(x) \cap \mathcal{C} = B_{G_e,1}(u) \cap \mathcal{C} = B_{G,1}(u) \cap \mathcal{C} \neq B_{G,1}(x) \cap \mathcal{C},$$

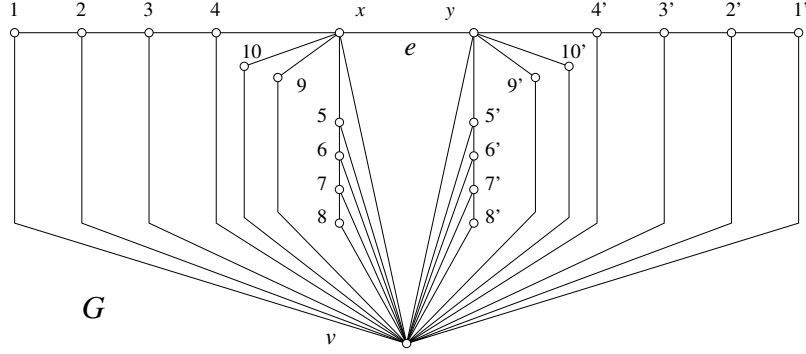


Figure 9: Graph G in Example 23.

and so

$$y \in \mathcal{C} \text{ and } B_{G_e,1}(x) \cap \mathcal{C} = (B_{G,1}(x) \cap \mathcal{C}) \setminus \{y\}.$$

If x and u are the only $(G_e, \mathcal{C}, 1)$ -twins, then with two more codewords we can both 1-cover x if necessary and 1-separate x and u in G_e . The same argument would work if y and v were the only $(G_e, \mathcal{C}, 1)$ -twins. So we assume that x and u , and y and v are $(G_e, \mathcal{C}, 1)$ -twins. This implies that both x and y are codewords, each 1-covered by itself. All there is left to do is to 1-separate two pairs of $(G_e, \mathcal{C}, 1)$ -twins in G_e , which can be done using two more codewords. \triangle

Note that we made no assumption on the connectivity of G and G_e . The following example shows that (connected) graphs G and G_e with $\gamma_1(G_e) = \gamma_1(G) + 2$ exist.

Example 23 Let $G = (V, E)$ be the graph represented in Figure 9, and G_e the graph obtained by deleting the edge xy . We claim that: (a) $\gamma_1(G) \leq 12$ and (b) $\gamma_1(G_e) \geq 14$, which by (19) will imply that $\gamma_1(G_e) = 14 = \gamma_1(G) + 2$.

Proof of (a). It is quite straightforward to check that $\mathcal{C} = \{1, 3, x, 6, 8, 9, 9', 8', 6', y, 3', 1'\}$ is 1-identifying in G . Hence $\gamma_1(G) \leq 12$.

Proof of (b). Let \mathcal{C}_e be an optimal 1-identifying code in G_e , not containing v : thanks to Lemma 3, we know that this is possible. We are going to show that the left part of the graph G_e , consisting of the vertices 1 to 10 and x , requires at least seven codewords.

As in Example 21, we have $3 \in \mathcal{C}_e$, $6 \in \mathcal{C}_e$, and, because $v \notin \mathcal{C}_e$, \mathcal{C}_e also contains at least one element in each of the 2-sets $\{1, 2\}$ and $\{7, 8\}$, which amounts to four codewords.

As in Example 21, we also have that:

- there are at least three codewords in $\{1, 2, 3, 4, x\}$,

and three codewords in $\{x, 5, 6, 7, 8\}$. So, if $x \notin C_e$, there are, because of 9 and 10, at least $3 + 3 + 2 = 8$ codewords, and we are done. We now assume that $x \in C_e$, so that we have already taken five codewords. One more codeword is not sufficient to 1-separate both 9 and 10, 9 and x , and 10 and x . This proves Claim (b), by symmetry. \triangle

By Theorems 20 and 22, we have the following result.

Theorem 24 *Let G_1 and G_2 be two 1-twin-free graphs, with same vertex set and differing by one edge. Then*

$$\gamma_1(G_1) - 2 \leq \gamma_1(G_2) \leq \gamma_1(G_1) + 2.$$

As a consequence, if for instance $\gamma_1(G_1) \leq a$ and $\gamma_1(G_2) \geq a + 2$, then $\gamma_1(G_1) = a$ and $\gamma_1(G_2) = a + 2$. \triangle

We conclude the case $r = 1$ by mentioning that pairs of graphs G and G_e such that $\gamma_1(G_e) - \gamma_1(G) = 0$ or $\gamma_1(G_e) - \gamma_1(G) = \pm 1$ exist.

Example 25 *We give simple examples with (a) $\gamma_1(G_e) - \gamma_1(G) = -1$ and (b) $\gamma_1(G_e) - \gamma_1(G) = 1$, omitting the easy case when the difference is 0.*

(a) *Let $G_e = P_9 = x_1x_2 \dots x_9$, and add the edge $\{x_3, x_5\}$ in order to obtain G . It is known ([3, Th. 3]) that $\gamma_1(P_9) = 5$, and it is easy to see that $\gamma_1(G) = 6$, so $\gamma_1(G_e) - \gamma_1(G) = -1$.*

(b) *Let G_e be the graph consisting of P_1 and P_4 , and G be the graph obtained by adding an edge between one extremity of P_4 and the vertex of P_1 , so that $G = P_5$. We have $\gamma_1(P_1) = 1$, $\gamma_1(P_4) = 3$, and $\gamma_1(P_5) = 3$, which shows that $\gamma_1(G_e) - \gamma_1(G) = 1$.* \triangle

7 The case $r \geq 2$

We now give our central result, Theorem 26. It describes graphs for which we delete edges and/or vertices, because we think that it is interesting to have such a "mixed construction", cf. Introduction and the forthcoming paper [6]. It also presents the remarkable feature that, starting from the graph G and performing two consecutive deletions, we first decrease the function γ_r , then increase it. The consequences of this result for edge deletion are detailed in Corollaries 27 and 29, and are extended in Propositions 28 and 30. For simplicity, we give constructions where two of the graphs, namely

$$(G \setminus e) \setminus f \text{ and } (G \setminus u) \setminus f,$$

are disconnected, but the remarks after the proofs of Corollary 29 and of Proposition 30 show an easy way to have connected graphs, with a slightly different result, when needed. Since we estimate the value of γ_r for all these graphs, this means that all are r -twin-free, a fact not stated explicitly in the theorem.

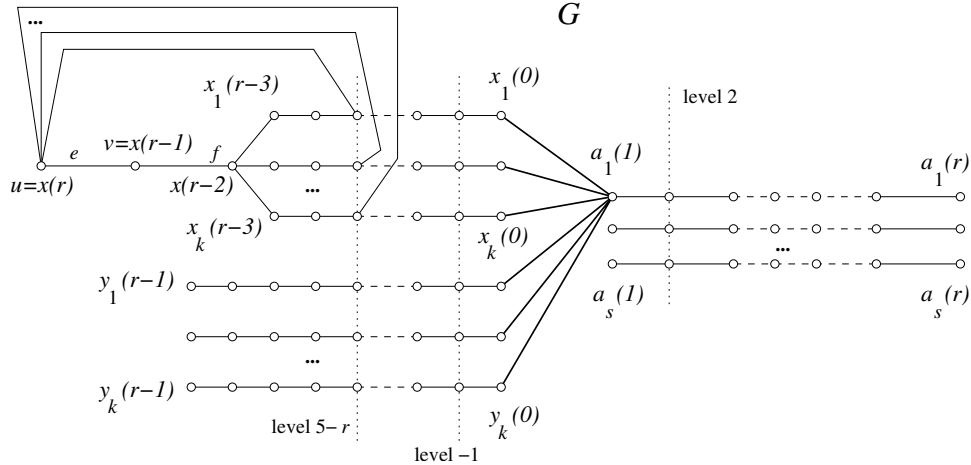


Figure 10: A partial representation of the graph G in Theorem 26: more edges exist between the vertices $x_i(0)$ and $y_i(0)$ on the one hand, and the vertices $a_j(1)$ on the other hand.

Theorem 26 *Let $k \geq 2$ be arbitrary and $r \geq 5$. There exists a graph G with $(2r - 2)k + r \lceil \log_2(k + 2) \rceil + r + 3$ vertices with the following properties:*

- (i) $\gamma_r(G) \geq k$.
- (ii) *There is an edge e of G such that $\gamma_r(G \setminus e) \leq 1 + r + r \lceil \log_2(k + 2) \rceil$.*
- (iii) *There is a vertex u of G such that $\gamma_r(G \setminus u) \leq 1 + r + r \lceil \log_2(k + 2) \rceil$.*
- (iv) *There is a vertex v of $G \setminus e$ such that $\gamma_r((G \setminus e) \setminus v) \geq k$.*
- (v) *There is an edge f of $G \setminus e$ such that $\gamma_r((G \setminus e) \setminus f) \geq k$.*
- (vi) *There is an edge f of $G \setminus u$ such that $\gamma_r((G \setminus u) \setminus f) \geq k$.*
- (vii) *There is a vertex v of $G \setminus u$ such that $\gamma_r((G \setminus u) \setminus v) \geq k$.*

Proof. We first construct the graph G for the given $k \geq 2$ and $r \geq 5$, see Figure 10. Denote $s = 1 + \lceil \log_2(k + 2) \rceil$.

For each $j = 1, 2, \dots, s$ we form the paths $a_j(1)a_j(2) \dots a_j(r)$ (i.e., $a_j(h)$ and $a_j(h + 1)$ are connected by an edge for all $h = 1, 2, \dots, r - 1$). Each vertex $a_j(h)$ is said to be **on level h** (cf. Figure 10).

For each $i = 1, 2, \dots, k$ we form the paths $y_i(0)y_i(1) \dots y_i(r - 1)$. Each vertex $y_i(h)$ is said to be **on level $-h$** .

Also, we form the paths $x_i(0)x_i(1) \dots x_i(r - 3)x_i(r - 2)x_i(r - 1)x_i(r)$, where now the same three vertices $x_i(r - 2)$, $x_i(r - 1)$ and $x_i(r)$ appear on all these paths. Again, each vertex $x_i(h)$ is said to be **on level $-h$** .

Now for each i we choose a unique nonempty proper subset A_i of the set $A = \{2, 3, \dots, s\}$, and connect every $x_i(0)$ and every $y_i(0)$ to every vertex $a_j(1)$ for which $j \in A_i$. Moreover, we connect every $x_i(0)$ and every $y_i(0)$ to $a_1(1)$. The sets A_i can indeed be chosen in this way, because there are $2^{s-1} - 2$ proper nonempty subsets of A , and $s - 1 = \lceil \log_2(k + 2) \rceil$. Without

loss of generality, we can choose the sets A_i in such a way that each $a_j(1)$ has degree at least two, and so already the graph constructed so far is connected.

The construction of G is now almost complete. As the final step, we connect the vertex $x(r)$ by an edge to every $x_i(r-5)$ (which is fine as we have assumed that $r \geq 5$).

In the statement of the theorem $u = x(r)$, $v = x(r-1)$, e is the edge connecting these two, and finally f is the edge connecting $x(r-1)$ and $x(r-2)$.

The first step of the proof consists of working out that if we take

$$\mathcal{C} = \{a_j(h) : j = 1, 2, \dots, s, h = 1, 2, \dots, r\},$$

that is, the sr vertices of type a , then \mathcal{C} is not r -identifying, but it does a lot, for all the graphs in the theorem: as we shall see, the only thing we need to worry about is to make sure that for each i , $x_i(h)$ and $y_i(h)$ can be r -separated for all $h = 0, 1, \dots, r-3$ and that $x(r-2)$, $x(r-1)$ and $x(r)$ can be identified.

In what follows, for a vertex w we always denote

$$I(w) = B_r(w) \cap \mathcal{C}$$

for this particular choice of \mathcal{C} , whatever the graph is. To begin with, we observe that

- $I(w)$ contains exactly one vertex from the r -th level, if $w = a_j(h)$ for some j and h (and then of course this one vertex is $a_j(r)$);
- $I(w)$ contains at least two vertices from the r -th level, if $w = x_i(0)$ or $y_i(0)$ for some i (and one of them is $a_1(r)$);
- $I(w)$ does not contain any vertices from the r -th level, if w is any other vertex.

All the vertices $a_j(h)$ can now be identified. A vertex w is one of the vertices $a_j(h)$ if and only if $I(w)$ contains exactly one vertex on level r , and this unique vertex already tells us j . Moreover, for any $j' \neq j$, in $I(w)$ the vertex $a_{j'}(h')$ with the largest level is $a_{j'}(r-h)$ if $h < r$ and there are no vertices $a_{j'}(h')$ at all in $I(w)$ if $h = r$. Either way, we can determine h .

In all the graphs mentioned in the statement of the theorem the following facts are clearly valid:

- Fact 1: If $i \neq i'$, then the distance between $y_i(h)$ and $y_{i'}(0)$ is $h+2$ (as we can always go via $a_1(1)$) and the distance between $y_i(h)$ and $x_{i'}(0)$ is $h+2$; the latter holds also for $i = i'$.
- Fact 2: If $w \notin \{x(r-2), x(r-1), x(r)\}$ is on level $h \leq 0$, then the highest level containing at least one vertex in $I(w)$ is $h+r$, and moreover, if $w = x_i(h)$ or $y_i(h)$, then the set $\{j \geq 2 : a_j(h+r) \in I(w)\}$,

which we call the **signature** of w , equals A_i , and since A_i is unique for each i , this tells us i .

By Fact 2, the only two remaining things are that we always have to be able to decide whether w belongs to the x -path or the corresponding y -path, and we have to make sure that the three vertices $x(r-2)$, $x(r-1)$ and $x(r)$ (when they exist in the graph) are identified.

Let us first consider the graph G itself. If $w = x(r)$, then the highest level h for which at least one $a_j(h)$ is in $I(w)$ is $h = 4$ (as we can take a shortcut and jump directly from $x(r)$ to **every** $x_i(r-5)$) and $\{j \geq 2 : a_j(4) \in I(w)\} = A$. In the same way, if $w = x(r-1)$, then the highest level points in $I(w)$ are on level 3 and $\{j \geq 2 : a_j(3) \in I(w)\} = A$; and if $w = x(r-2)$, then the highest level points in $I(w)$ are on level 2 and again $\{j \geq 2 : a_j(2) \in I(w)\} = A$. As all the signatures referred to in Fact 2 were proper subsets of A , the vertices $x(r)$, $x(r-1)$ and $x(r-2)$ are identified by \mathcal{C} .

The vertex $y_i(r-1)$ is within distance r from all the vertices $y_i(h)$ and by Fact 1, its distance to all the x -vertices is larger than r . Therefore G is r -twin-free as the addition of all the vertices $y_i(r-1)$ to \mathcal{C} would yield an r -identifying code.

Exactly the same argument shows that in fact all the graphs mentioned in the theorem are r -twin-free (but notice that the highest level points in $I(x(r-1))$ move two levels down, from 3 to 1, if e (or $x(r)$) has been removed, and that $x(r-1)$ is an isolated vertex if also f has been removed).

Let us now prove (i). Look at the vertices $x_i(0)$ and $y_i(0)$ for any fixed i . By Fact 1, no $y_{i'}(h)$ with $i' \neq i$ can r -separate them; neither can any $a_j(h)$. By the construction, every x -vertex is within distance $r-2$ from at least one $x_{i'}(0)$. As $x_{i'}(0)$ is connected by an edge to $a_1(1)$, which in turn is connected by an edge to every vertex on level 0, we see that no x -vertex can r -separate $x_i(0)$ and $y_i(0)$. Therefore at least one $y_i(h)$ has to do the job, and therefore any r -identifying code must contain at least k codewords.

Exactly the same argument gives us (iv)-(vii).

To prove (iii), it suffices to observe that in this graph $x(r-1)$ is within distance $r-1$ from all the x -vertices, but at distance greater than r from all the y -vertices, so the vertex $x(r-1)$ together with the codewords in \mathcal{C} form an r -identifying code.

It remains to prove (ii). Now the vertex $x(r-1)$ is within distance r from all the x -vertices including $x(r)$, and at distance greater than r from all the y -vertices, and again the codewords in \mathcal{C} together with $x(r-1)$ will do. \triangle

The next corollary studies how small $\gamma_r(G_e) - \gamma_r(G)$ and $\frac{\gamma_r(G_e)}{\gamma_r(G)}$ can be for a graph G , when $r \geq 5$.

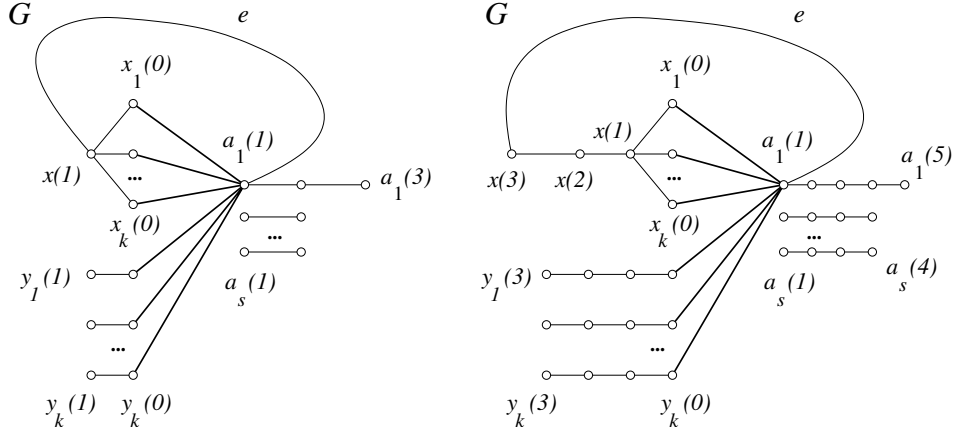


Figure 11: A partial representation of the graph G in Proposition 28, for $r = 2$, and for $r = 4$.

Corollary 27 *Let $k \geq 2$ be arbitrary and $r \geq 5$. There exist two (connected) r -twin-free graphs G and G_e with $(2r - 2)k + r \lceil \log_2(k + 2) \rceil + r + 3$ vertices, such that*

$$\gamma_r(G) - \gamma_r(G_e) \geq k - r \lceil \log_2(k + 2) \rceil - r - 1, \quad (20)$$

$$\frac{\gamma_r(G)}{\gamma_r(G_e)} \geq \frac{k}{r \lceil \log_2(k + 2) \rceil + r + 1}. \quad (21)$$

Proof. Use (i) and (ii) in the previous theorem. \triangle

The following proposition gives a very similar result for $r \in \{2, 3, 4\}$ (and also for $r \geq 5$).

Proposition 28 *Let $k \geq 2$ be arbitrary and $r \geq 2$. There exist two (connected) r -twin-free graphs G and G_e with $(r + 1)k + r \lceil \log_2(k + 2) \rceil + 2r$ vertices, such that*

$$\gamma_r(G) - \gamma_r(G_e) \geq k - r \lceil \log_2(k + 2) \rceil - r - 3, \quad (22)$$

$$\frac{\gamma_r(G)}{\gamma_r(G_e)} \geq \frac{k}{r \lceil \log_2(k + 2) \rceil + r + 3}. \quad (23)$$

Proof. We slightly modify the construction of G in Theorem 26, so that the x -paths are $x_i(0)x(1) \dots x(r - 1)$, the vertices $x(r - 1)$ and $a_1(1)$ are connected by the edge e , and there is one additional vertex $a_1(r + 1)$ which is only connected to $a_1(r)$, see Figure 11.

The same argument as in the proof of the theorem shows that in G , which is r -twin-free (in particular because $a_1(r + 1)$ can r -separate $x(r - 1)$ and $a_1(2)$), any r -identifying code has at least k elements, in order to have

each pair of vertices $x_i(0), y_i(0)$ r -separated by the code. On the other hand, it is straightforward to check that

$$\mathcal{C} = \{a_j(h) : j = 1, 2, \dots, s, h = 1, 2, \dots, r\} \cup \{a_1(r+1), x(r-1), y_1(0)\}$$

is r -identifying in G_e : in particular, $x(r-1)$ r -separates $x_i(0)$ and $y_i(0)$ for every i , and $y_1(0)$ r -separates $a_1(r)$ and $a_1(r+1)$ (this job could have been done by any $y_i(0)$ or $x_i(0)$). \triangle

The next corollary studies how large $\gamma_r(G_e) - \gamma_r(G)$ and $\frac{\gamma_r(G_e)}{\gamma_r(G)}$ can be for a graph G , when $r \geq 3$.

Corollary 29 *Let $k \geq 2$ be arbitrary and $r \geq 3$. There exist two r -twin-free graphs H and H_f with $(2r-2)k + r\lceil\log_2(k+2)\rceil + r + 2$ vertices, such that*

$$\gamma_r(H_f) - \gamma_r(H) \geq k - r\lceil\log_2(k+2)\rceil - r - 1, \quad (24)$$

$$\frac{\gamma_r(H_f)}{\gamma_r(H)} \geq \frac{k}{r\lceil\log_2(k+2)\rceil + r + 1}. \quad (25)$$

Proof. Consider Theorem 26 and let $H = G_u = G_{x(r)}$. The condition $r \geq 5$ can be relaxed and changed into $r \geq 3$ because the vertex $u = x(r)$, connected to every $x_i(r-5)$, does not exist here. We can then mimic the proof of the theorem for the cases (iii) and (vi), see that it works also for $r = 3$ and $r = 4$, and retrieve (24) and (25). \triangle

Remark. If we want connected graphs, we can slightly modify the construction for Corollary 29, e.g., by introducing $2r+1$ new vertices to form a path from $x(r-1)$ to $a_1(r)$ and by taking them all as codewords. Then, an arbitrary vertex is one of the new vertices if and only if it contains the middle one of the new vertices in its r -identifying set. The slightly different resulting numbers of vertices and of codewords do not fundamentally alter the meaning of Corollary 29.

The following proposition gives a very similar result for $r = 2$.

Proposition 30 *Let $k \geq 2$ be arbitrary. There exist two 2-twin-free graphs G and G_e with $3k + 2\lceil\log_2(k+2)\rceil + 5$ vertices, such that*

$$\gamma_2(G_e) - \gamma_2(G) \geq k - 2\lceil\log_2(k+2)\rceil - 5, \quad (26)$$

$$\frac{\gamma_2(G_e)}{\gamma_2(G)} \geq \frac{k}{2\lceil\log_2(k+2)\rceil + 5}. \quad (27)$$

Proof. We bring only a very small modification to the graph described in the proof of Proposition 28: see Figure 12, where, compared to the left part of Figure 11, we have added the vertex $x(2)$ and connected it to $x(1)$, renaming this edge $x(1)x(2)$ by e . Now G is 2-twin-free, and

$$\mathcal{C} = \{a_j(h) : j = 1, 2, \dots, s, h = 1, 2\} \cup \{a_1(3), x(2), y_1(0)\}$$

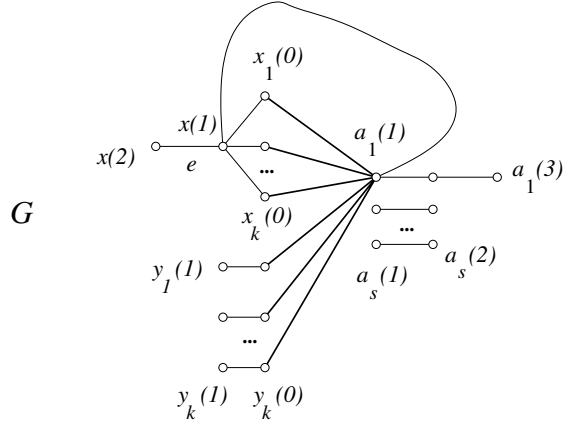


Figure 12: A partial representation of the graph G in Proposition 30.

is 2-identifying in G . In particular, $x(2)$ 2-separates $x_i(0)$ and $y_i(0)$ for every i , and $y_1(0)$ 2-separates $a_1(2)$ and $a_1(3)$. But in G_e , which is 2-twin-free, by the standard argument we need at least k codewords. \triangle

Remark. Similarly to Corollary 29, the addition of a chain of $2r + 1 = 5$ vertices linking $x(2)$ to $a_1(3)$ would give a slightly different result, this time with G_e connected.

8 Conclusion

Table 2 recapitulates the results obtained in the previous sections, for $\gamma_r(G_e) - \gamma_r(G)$ and, when appropriate, $\gamma_r(G_e)/\gamma_r(G)$. In the lower part of the table, the inequalities mean that there exist pairs of graphs G, G_e such that these inequalities hold.

Whether these inequalities can be substantially improved is left as an open problem.

r	$\gamma_r(G_e) - \gamma_r(G)$		$\gamma_r(G_e)/\gamma_r(G)$
$= 1$	must be inside $\{-2, -1, 0, 1, 2\}$	(18), (19), Th. 24	
$= 1$	graphs with $= -2, = -1, = 0, = 1, = 2$	Ex. 21, 23, 25	
r	$\gamma_r(G_e) - \gamma_r(G)$		$\gamma_r(G_e)/\gamma_r(G)$
$= 2, 3, 4$	$\leq -(k - r \lceil \log_2(k+2) \rceil - r - 3)$ (22)		$\leq \frac{r \lceil \log_2(k+2) \rceil + r + 3}{k}$ (23)
≥ 5	$\leq -(k - r \lceil \log_2(k+2) \rceil - r - 1)$ (20)		$\leq \frac{r \lceil \log_2(k+2) \rceil + r + 1}{k}$ (21)
$= 2$	$\geq k - 2 \lceil \log_2(k+2) \rceil - 5$ (26)		$\geq \frac{k}{2 \lceil \log_2(k+2) \rceil + 5}$ (27)
≥ 3	$\geq k - r \lceil \log_2(k+2) \rceil - r - 1$ (24)		$\geq \frac{k}{r \lceil \log_2(k+2) \rceil + r + 1}$ (25)

Table 2: $\gamma_r(G_e) - \gamma_r(G)$ and $\frac{\gamma_r(G_e)}{\gamma_r(G)}$, as a function of r and k .

The integer $k \geq 2$ can be taken arbitrarily, and is linked to n , the order of G and G_e , by the relation

$$n = (c_1 r + c_2)k + r \lceil \log_2(k + 2) \rceil + (c_3 r + c_4),$$

where the quadruple (c_1, c_2, c_3, c_4) can take the values $(2, -2, 1, 3)$, $(1, 1, 2, 0)$, $(2, -2, 1, 2)$, $(1, 1, 2, 1)$; this means, roughly speaking, that k is a fraction, depending on r , of n ; therefore, given $r \geq 2$, what we have shown is that there is an infinite sequence of graphs G and two positive constants α and β such that $\gamma_r(G) \geq \alpha n$, but, after deleting a suitable edge e , we have $\gamma_r(G_e) \leq \beta \log_2 n$ (or the other way round).

References

- [1] C. BERGE: *Graphes*, Gauthier-Villars: Paris, 1983. English translation: *Graphs*, North-Holland Publishing Co.: Amsterdam, 1985.
- [2] N. BERTRAND: Codes identifiants et codes localisateurs-dominateurs sur certains graphes, Mémoire de stage de maîtrise, ENST, Paris, France, 28 pages, June 2001.
- [3] N. BERTRAND, I. CHARON, O. HUDRY and A. LOBSTEIN: Identifying and locating-dominating codes on chains and cycles, *European Journal of Combinatorics*, Vol. 25, pp. 969–987, 2004.
- [4] I. CHARON, I. HONKALA, O. HUDRY and A. LOBSTEIN: Structural properties of twin-free graphs, *Electronic Journal of Combinatorics*, Vol. 14(1), R16, 2007.
- [5] I. CHARON, I. HONKALA, O. HUDRY and A. LOBSTEIN: Minimum sizes of identifying codes in graphs differing by one edge, submitted.
- [6] I. CHARON, I. HONKALA, O. HUDRY and A. LOBSTEIN: Minimum sizes of identifying codes in graphs differing by one vertex, submitted.
- [7] I. CHARON, O. HUDRY and A. LOBSTEIN: Extremal cardinalities for identifying and locating-dominating codes in graphs, Technical Report Télécom Paris-2003D006, Paris, France, 18 pages, August 2003.
- [8] I. CHARON, O. HUDRY and A. LOBSTEIN: On the structure of identifiable graphs, *Electronic Notes in Discrete Mathematics*, Vol. 22, pp. 491–495, 2005.
- [9] I. CHARON, O. HUDRY and A. LOBSTEIN: Possible cardinalities for identifying codes in graphs, *Australasian Journal of Combinatorics*, Vol. 32, pp. 177–195, 2005.

- [10] I. CHARON, O. HUDRY and A. LOBSTEIN: Extremal cardinalities for identifying and locating-dominating codes in graphs, *Discrete Mathematics*, Vol. 307, pp. 356–366, 2007.
- [11] C. CHEN, C. LU and Z. MIAO: Identifying codes and locating-dominating sets on paths and cycles, *Discrete Applied Mathematics*, Vol. 159, pp. 1540–1547, 2011.
- [12] R. DIESTEL: *Graph Theory*, Springer-Verlag: Berlin, 2005.
- [13] F. FOUCAUD, E. GUERRINI, M. KOVŠE, R. NASERASR, A. PARRAU and P. VALICOV: Extremal graphs for the identifying code problem, *European Journal of Combinatorics*, Vol. 32, pp. 628–638, 2011.
- [14] S. GRAVIER and J. MONCEL: On graphs having a $V \setminus \{x\}$ set as an identifying code, *Discrete Mathematics*, Vol. 307, pp. 432–434, 2007.
- [15] I. HONKALA: An optimal edge-robust identifying code in the triangular lattice, *Annals of Combinatorics*, Vol. 8, pp. 303–323, 2004.
- [16] I. HONKALA: On 2-edge-robust r -identifying codes in the king grid, *Australasian Journal of Combinatorics*, Vol. 36, pp. 151–165, 2006.
- [17] I. HONKALA and T. LAIHONEN: On identifying codes that are robust against edge changes, *Information and Computation*, Vol. 205, pp. 1078–1095, 2007.
- [18] M. G. KARPOVSKY, K. CHAKRABARTY and L. B. LEVITIN: On a new class of codes for identifying vertices in graphs, *IEEE Transactions on Information Theory*, Vol. IT-44, pp. 599–611, 1998.
- [19] T. LAIHONEN: Optimal t -edge-robust r -identifying codes in the king lattice, *Graphs and Combinatorics*, Vol. 22, pp. 487–496, 2006.
- [20] T. LAIHONEN: On edge-robust $(1, \leq \ell)$ -identifying codes in binary Hamming spaces, *International Journal of Pure and Applied Mathematics*, Vol. 36, pp. 87–102, 2007.
- [21] A. LOBSTEIN: A bibliography on watching systems, identifying, locating-dominating and discriminating codes in graphs, <http://perso.telecom-paristech.fr/~lobstein/debutBIBidetlocdom.pdf>
- [22] D. L. ROBERTS and F. S. ROBERTS: Locating sensors in paths and cycles: the case of 2-identifying codes, *European Journal of Combinatorics*, Vol. 29, pp. 72–82, 2008.