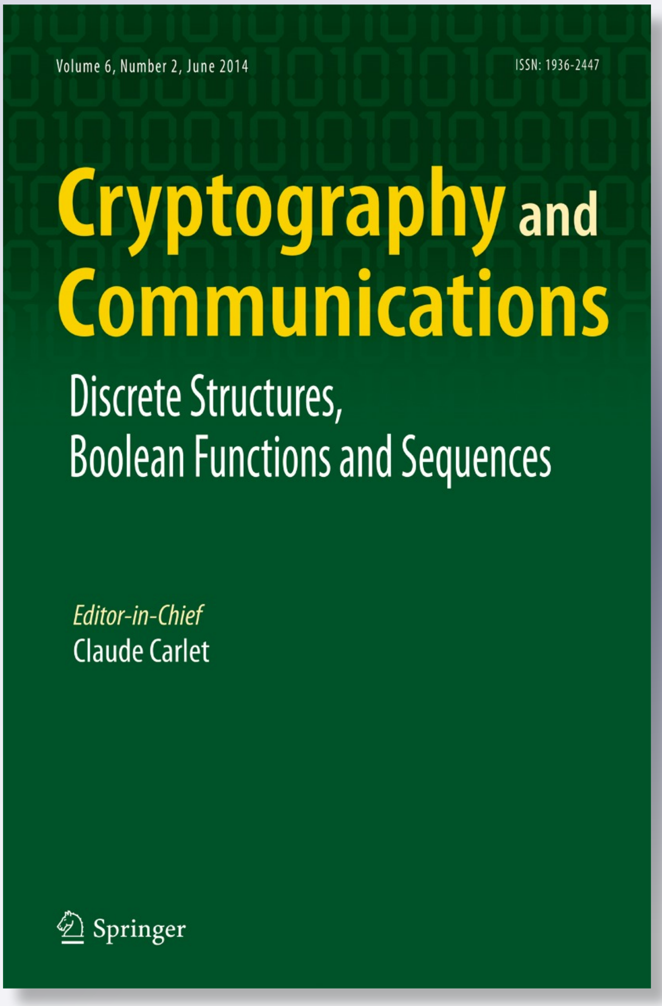
Minimum sizes of identifying codes in graphs differing by one edge

Irène Charon, Iiro Honkala, Olivier Hudry & Antoine Lobstein

Cryptography and Communications
Discrete Structures, Boolean Functions and Sequences

ISSN 1936-2447
Volume 6
Number 2

Cryptogr. Commun. (2014) 6:157-170
DOI 10.1007/s12095-013-0094-x



Springer

Cryptogr. Commun. (2014) 6:157–170
DOI 10.1007/s12095-013-0094-x

Minimum sizes of identifying codes in graphs differing by one edge

Irène Charon · Iiro Honkala · Olivier Hudry · Antoine Lobstein

Received: 14 May 2013 / Accepted: 28 October 2013 / Published online: 29 November 2013

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 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.

Keywords Graph theory · Twin-free graphs · Identifiable graphs · Identifying codes

1 Introduction

We introduce basic definitions and notation for graphs, for which we refer to, e.g., [1] and [11], 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 denoted by xy or yx.

I. Charon · O. Hudry · A. Lobstein
Institut Télécom - Télécom ParisTech & CNRS - LTCI UMR 5141, 46, rue Barrault, 75634 Paris Cedex 13, France
e-mail: charon@telecom-paristech.fr

O. Hudry
e-mail: hudry@telecom-paristech.fr

A. Lobstein
e-mail: lobstein@telecom-paristech.fr

I. Honkala (✉)
Department of Mathematics and Statistics, University of Turku, 20014 Turku, Finland
e-mail: honkala@utu.fi

We denote by P_n the path with n vertices and by C_n the cycle with n vertices; the length of P_n is $n - 1$.

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, 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 paper and in [5], 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 [7]: 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. As we shall see, examples of large variations may exist for the minimum size of an identifying code, which means that there are configurations for which opening or closing a door can save a significant number of detectors.

In the conclusion of [22], it is already observed, somewhat paradoxically, that there exist cycles C_{n-1} which require more codewords/detectors than C_n. See also [13] for the concept of criticality with respect to vertex or edge deletion and addition.

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, 20].

In this paper, we mainly focus on the addition or deletion of one edge, whereas in [5] we study the consequences of adding or removing one vertex. For our purpose, when we construct two graphs G and G^* differing by one edge, only the following two cases should be considered:

(i) both graphs G and G^* are connected,
(ii) the graph with one edge less may be disconnected;

however, we shall always be able to give constructions such that (i) holds.

Before we proceed, we still need some additional definitions and notation, and we also give three lemmas 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 *[$(G, \mathcal{S}, r)$-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.*

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*

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

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. □

This article is organized as follows. Section 2 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 3, 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 11 for $r \geq 5$ and Proposition 12 for $r \in \{2, 3, 4\}$); finally, we study how *large* these difference and ratio can be, for $r \geq 3$ in Corollary 13 and for $r = 2$ in Proposition 14 (in both cases, the graphs can be made connected).

A conclusion recapitulates our results in a Table.

2 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 8), and these five values can be reached (Examples 5, 7 and 9).

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 4 *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:

$$\gamma_1(G_e) - \gamma_1(G) \geq -2. \tag{1}$$

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 $u \notin \{x, y\}$,
- y and v with $v \notin \{x, y\}$.

(Note that this argument would not work for $r > 1$.) Moreover, x and u', with $u' \notin \{u, x, y\}$, 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 $v' \notin \{v, x, y\}$. 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 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. □

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.

Example 5 Let $G_e = (V, E_e)$ be the graph represented in Fig. 1, 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 (1) 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, one obtains 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). □

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

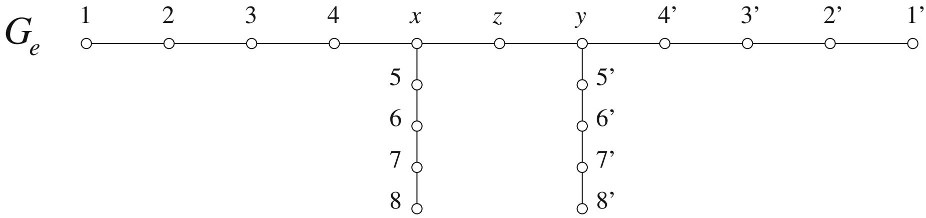


Fig. 1 Graph G_e in Example 5

Theorem 6 *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:

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

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) all vertices remain 1-separated by $\mathcal{C}$ in G_e, but 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.

Since our assumption is that all vertices are 1-separated by $\mathcal{C}$ in G_e, the code $\mathcal{C} \cup \{x\}$ is 1-identifying in G_e.

2) $(G_e, \mathcal{C}, 1)$-twins appear (with no assumption on whether x or y are still 1-covered by $\mathcal{C}$ in G_e); because only the edge xy is deleted when going from G to G_e, and similarly to the proof of Theorem 4, only the following pairs can be $(G_e, \mathcal{C}, 1)$-twins:

- x and y,
- x and u with $u \notin \{x, y\}$,
- y and v with $v \notin \{x, y\}$.

(Again, note that this would not be true for $r > 1$.) 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

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

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.

□

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.

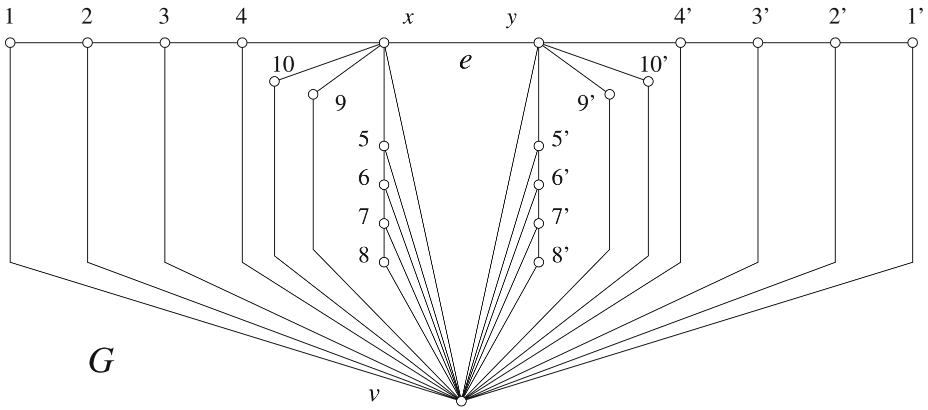


Fig. 2 Graph G in Example 7

Example 7 Let $G = (V, E)$ be the graph represented in Fig. 2, 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 (2) 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 5, 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 5, 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 $x \notin \mathcal{C}_e$, there are, because of 9 and 10, at least $3 + 3 + 2 = 8$ codewords, and we are done. We now assume that $x \in \mathcal{C}_e$, 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. □

By Theorems 4 and 6, we have the following result.

Theorem 8 *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 $\gamma_1(G_2) = a + 2$.

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

Example 9

(a) Let $G_e = P_9 = x_1x_2 \dots x_9$, and add the edge x_3x_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 plus an additional vertex which is linked to the five vertices 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 consists of P_5 plus one vertex linked to all the vertices of P_5. It is easy to see that $\gamma_1(G_e) = 4$ and $\gamma_1(G) = 3$, which shows that $\gamma_1(G_e) - \gamma_1(G) = 1$.

(c) $\gamma_1(C_4) = \gamma_1(P_4) = 3$.

3 The case $r \geq 2$

We now give our central result, Theorem 10. 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 [5], where we focus on additon/deletion of one vertex. 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 11 and 13, and are extended in Propositions 12 and 14. 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 13 and of Proposition 14 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.

Theorem 10 *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 Fig. 3. 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. Fig. 3).

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

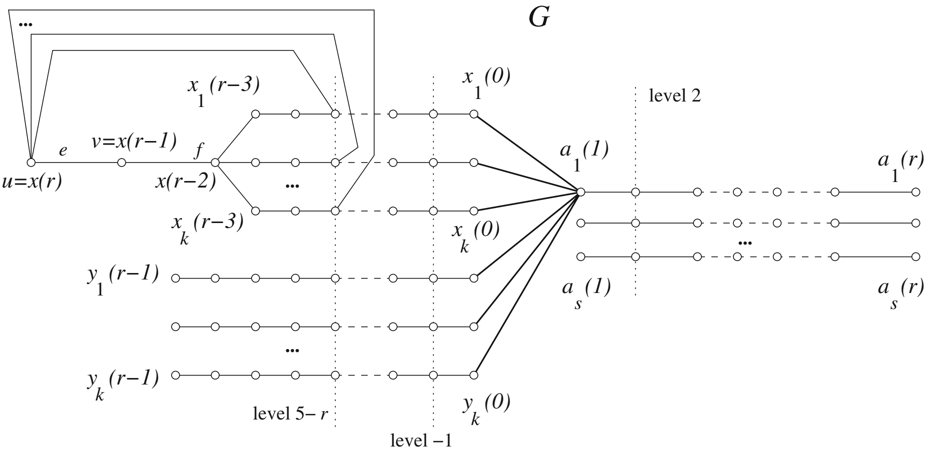


Fig. 3 A partial representation of the graph G in Theorem 10: more edges exist between the vertices $x_i(0)$ and $y_i(0)$ on the one hand, and the vertices $a_j(1)$ on the other hand

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, all the vertices $a_j(h)$, 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 the lower bound on $\gamma_r(G)$ given in (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 the lower bounds (iv)–(vii).

To prove the upper bound (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. □

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 $r \geq 5$.

Corollary 11 *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*

$$\gamma_r(G) - \gamma_r(G_e) \geq k - r\lceil \log_2(k+2) \rceil - r - 1, \tag{3}$$

$$\frac{\gamma_r(G)}{\gamma_r(G_e)} \geq \frac{k}{r\lceil \log_2(k+2) \rceil + r + 1}. \tag{4}$$

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

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

Proposition 12 *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*

$$\gamma_r(G) - \gamma_r(G_e) \geq k - r\lceil \log_2(k+2) \rceil - r - 3, \tag{5}$$

$$\frac{\gamma_r(G)}{\gamma_r(G_e)} \geq \frac{k}{r\lceil \log_2(k+2) \rceil + r + 3}. \tag{6}$$

Proof We slightly modify the construction of G in Theorem 10, so that the x-paths are $x_i(0)x(1)\ldots x(r-1)$, 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 Fig. 4.

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

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

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 $x_i(0)$). □

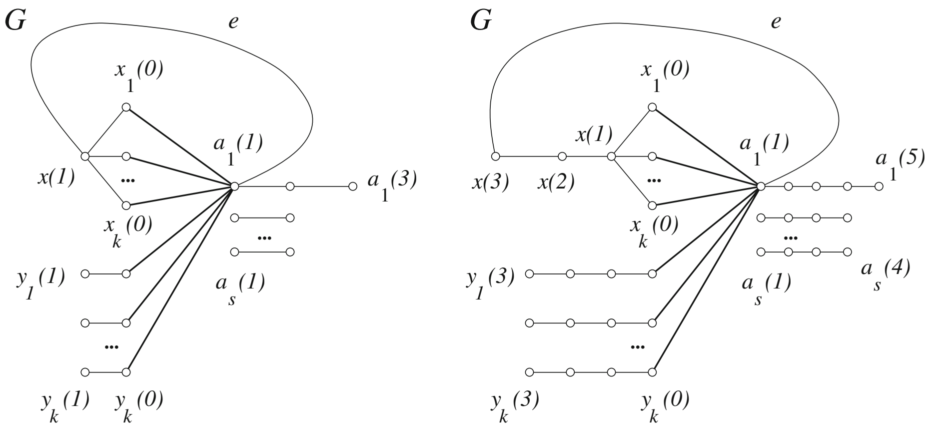


Fig. 4 A partial representation of the graph G in Proposition 12, for $r = 2$, and for $r = 4$

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 13 *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*

$$\gamma_r(H_f) - \gamma_r(H) \geq k - r\lceil\log_2(k+2)\rceil - r - 1, \tag{7}$$

$$\frac{\gamma_r(H_f)}{\gamma_r(H)} \geq \frac{k}{r\lceil\log_2(k+2)\rceil + r + 1}. \tag{8}$$

Proof Consider Theorem 10 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 (7) and (8). □

Remark 1 If we want connected graphs, we can slightly modify the construction for Corollary 13, 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 13.

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

Proposition 14 *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*

$$\gamma_2(G_e) - \gamma_2(G) \geq k - 2\lceil\log_2(k+2)\rceil - 5, \tag{9}$$

$$\frac{\gamma_2(G_e)}{\gamma_2(G)} \geq \frac{k}{2\lceil\log_2(k+2)\rceil + 5}. \tag{10}$$

Proof We bring only a very small modification to the graph described in the proof of Proposition 12: see Fig. 5, where, compared to the left part of Fig. 4, 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

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

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. □

Remark 2 Similarly to Corollary 13, 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.

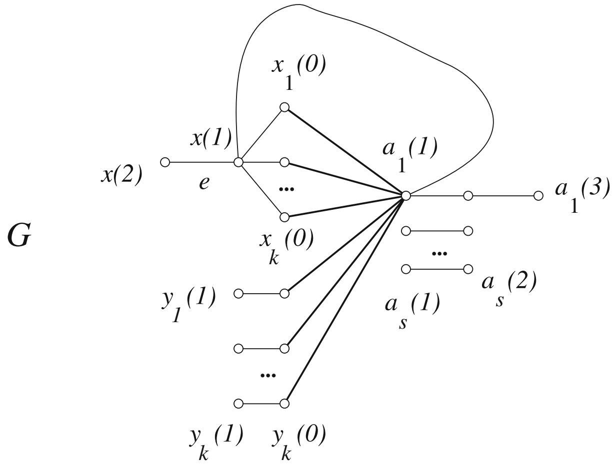


Fig. 5 A partial representation of the graph G in Proposition 14

4 Conclusion

Table 1 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.

In the previous two corollaries and two propositions, 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) takes the values $(2, -2, 1, 3)$ in Corollary 11, $(1, 1, 2, 0)$ in Proposition 12, $(2, -2, 1, 2)$ in Corollary 13, and $(1, 1, 2, 1)$ in Proposition 14, where $r = 2$; 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

Table 1 $\gamma_r(G_e) - \gamma_r(G)$ and $\frac{\gamma_r(G_e)}{\gamma_r(G)}$, as a function of r and k

r	$\gamma_r(G_e) - \gamma_r(G)$	
$= 1$	must be inside $\{-2, -1, 0, 1, 2\}$	(1), (2), Th. 8
$= 1$	graphs with $= -2, = -1, = 0, = 1, = 2$	Ex. 5, 7, 9
r	$\gamma_r(G_e) - \gamma_r(G)$	$\gamma_r(G_e)/\gamma_r(G)$
$= 2, 3, 4$	$\leq -(k - r\lceil \log_2(k+2)\rceil - r - 3)$ (5)	$\leq \frac{r\lceil \log_2(k+2)\rceil + r + 3}{k}$ (6)
≥ 5	$\leq -(k - r\lceil \log_2(k+2)\rceil - r - 1)$ (3)	$\leq \frac{r\lceil \log_2(k+2)\rceil + r + 1}{k}$ (4)
$= 2$	$\geq k - 2\lceil \log_2(k+2)\rceil - 5$ (9)	$\geq \frac{k}{2\lceil \log_2(k+2)\rceil + 5}$ (10)
≥ 3	$\geq k - r\lceil \log_2(k+2)\rceil - r - 1$ (7)	$\geq \frac{k}{r\lceil \log_2(k+2)\rceil + r + 1}$ (8)

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. Berge, C.: Graphes, Gauthier-Villars: Paris, 1983. English translation: Graphs. North-Holland Publishing Co., Amsterdam (1985)
2. Bertrand, N.: Codes Identifiants et Codes Localisateurs-Dominateurs sur Certains Graphes, Mémoire de Stage de Maîtrise, 28 pp. ENST, Paris (2001)
3. Bertrand, N., Charon, I., Hudry, O., Lobstein, A.: Identifying and locating-dominating codes on chains and cycles. Eur. J. Comb. **25**, 969–987 (2004)
4. Charon, I., Honkala, I., Hudry, O., Lobstein, A.: Structural properties of twin-free graphs. Electron. J. Comb. **14**(1), R16 (2007)
5. Charon, I., Honkala, I., Hudry, O., Lobstein, A.: Minimum sizes of identifying codes in graphs differing by one vertex. Cryptogr. Commun.–Discret. Struct. Boolean Funct. Seq. **5**, 119–136 (2013)
6. Charon, I., Hudry, O., Lobstein, A.: Extremal cardinalities for identifying and locating-dominating codes in graphs, Technical Report Télécom Paris-2003D006, 18 pp. Paris (2003)
7. Charon, I., Hudry, O., Lobstein, A.: On the structure of identifiable graphs. Electron. Notes Discret. Math. **22**, 491–495 (2005)
8. Charon, I., Hudry, O., Lobstein, A.: Possible cardinalities for identifying codes in graphs. Australas. J. Comb. **32**, 177–195 (2005)
9. Charon, I., Hudry, O., Lobstein, A.: Extremal cardinalities for identifying and locating-dominating codes in graphs. Discret. Math. **307**, 356–366 (2007)
10. Chen, C., Lu, C., Miao, Z.: Identifying codes and locating-dominating sets on paths and cycles. Discret. Appl. Math. **159**, 1540–1547 (2011)
11. Diestel, R.: Graph Theory. Springer, Berlin (2005)
12. Foucaud, F., Guerrini, E., Kovše, M., Naserasr, R., Parreau, A., Valicov, P.: Extremal graphs for the identifying code problem. Eur. J. Comb. **32**, 628–638 (2011)
13. Frick, M., Fricke, G.H., Mynhardt, C.M., Skaggs, R.D.: Critical graphs with respect to vertex identification. Util. Math. **76**, 213–227 (2008)
14. Gravier, S., Moncel, J.: On graphs having a $V \setminus \{x\}$ set as an identifying code. Discret. Math. **307**, 432–434 (2007)
15. Honkala, I.: An optimal edge-robust identifying code in the triangular lattice. Ann. Comb. **8**, 303–323 (2004)
16. Honkala, I.: On 2-edge-robust r-identifying codes in the king grid. Australas. J. Comb. **36**, 151–165 (2006)
17. Honkala, I., Laihonen, T.: On identifying codes that are robust against edge changes. Inf. Comput. **205**, 1078–1095 (2007)
18. Karpovsky, M.G., Chakrabarty, K., Levitin, L.B.: On a new class of codes for identifying vertices in graphs. IEEE Trans. Inf. Theory **IT-44**, 599–611 (1998)
19. Laihonen, T.: Optimal t-edge-robust r-identifying codes in the king lattice. Graphs Comb. **22**, 487–496 (2006)
20. Laihonen, T.: On edge-robust $(1, \leq \ell)$-identifying codes in binary Hamming spaces. Int. J. Pure Appl. Math. **36**, 87–102 (2007)
21. Lobstein, A.: A bibliography on watching systems, identifying, locating-dominating and discriminating codes in graphs. http://perso.telecom-paristech.fr/~lobstein/debutBIBidetlocdom.pdf
22. Roberts, D.L., Roberts, F.S.: Locating sensors in paths and cycles: the case of 2-identifying codes. Eur. J. Comb. **29**, 72–82 (2008)