

# Maximum difference about the size of optimal identifying codes in graphs differing by one vertex

Mikko Peltó

► **To cite this version:**

Mikko Peltó. Maximum difference about the size of optimal identifying codes in graphs differing by one vertex. *Discrete Mathematics and Theoretical Computer Science, DMTCS*, 2015, Vol. 17 no. 1 (in progress) (1), pp.339–356. <hal-01196851>

**HAL Id: hal-01196851**

**<https://hal.inria.fr/hal-01196851>**

Submitted on 10 Sep 2015

**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.

# Maximum difference about the size of optimal identifying codes in graphs differing by one vertex

Mikko Pelto\*

Department of Mathematics and Statistics, University of Turku, Finland

received 25<sup>th</sup> June 2013, revised 9<sup>th</sup> Feb. 2015, accepted 12<sup>th</sup> Feb. 2015.

Let  $G = (V, E)$  be a simple undirected graph. We call any subset  $C \subseteq V$  an identifying code if the sets  $I(v) = \{c \in C \mid \{v, c\} \in E \text{ or } v = c\}$  are distinct and non-empty for all vertices  $v \in V$ . A graph is called twin-free if there is an identifying code in the graph. The identifying code with minimum size in a twin-free graph  $G$  is called the optimal identifying code and the size of such a code is denoted by  $\gamma(G)$ . Let  $G_S$  denote the induced subgraph of  $G$  where the vertex set  $S \subset V$  is deleted. We provide a tight upper bound for  $\gamma(G_S) - \gamma(G)$  when both graphs are twin-free and  $|V|$  is large enough with respect to  $|S|$ . Moreover, we prove tight upper bound when  $G$  is a bipartite graph and  $|S| = 1$ .

**Keywords:** identifying codes, graph theory, twin-free graphs

## 1 Introduction

Let  $G = (V, E)$  be an undirected graph where  $V$  is the vertex set and  $E$  the edge set. The *order* of the graph  $G$  is the number of vertices  $|V|$  and it is denoted by  $n$  or  $n_G$ . The *neighbourhood* of a vertex  $v$  is

$$N(v, G) = \{v\} \cup \{u \in V \mid \{u, v\} \in E\}.$$

If the graph is clear from the context, we usually denote  $N(v)$  instead of  $N(v, G)$ . Two vertices  $u$  and  $v$  are called *twins* if their neighbourhoods are the same, i.e.  $N(u) = N(v)$ . A graph is called *twin-free* if there are no twins, i.e. all sets  $N(v)$  are distinct.

We call any  $C \subseteq V$  a *code*. The vertices of  $C$  are called *codewords*. A code is called an *identifying code* if the sets

$$I(v, C) = \{c \in C \mid \{v, c\} \in E \text{ or } c = v\}$$

are distinct and non-empty for all vertices  $v \in V$ . Usually we denote these sets by  $I(v)$  instead of  $I(v, C)$  if the code  $C$  is clear from the context. Furthermore, if  $I(v, C) \neq I(u, C)$  and  $w \in (I(v, C) \setminus I(u, C)) \cup (I(u, C) \setminus I(v, C))$  we say that  $C$  or  $w$  *separates*  $v$  from  $u$ . If  $w \in I(v, C)$ , we say that  $v$  is *dominated* by  $w$  or  $C$ . Otherwise stated,  $C$  is an identifying code if all vertices are dominated and separated from each other by  $C$ .

\*Email: mikko.pelto@utu.fi. Research supported by Finnish Academy of Science and Letters.

Identifying codes were introduced in the late 1990s by Karpovsky, Chakrabarty and Levitin [9]. The purpose of such codes is to create a safeguard analysis of a facility [11, 12] or a fault diagnosis of multiprocessor systems [9]. More papers on identifying codes and related topics can be found in the web bibliography [10]. It has been shown that there is an identifying code in a given graph if and only if the graph is twin-free [3, Remark 1]. Therefore, in this paper we restrict to twin-free graphs. In this case, the code which consists of all the vertices of a graph is always an identifying code. Moreover, if we add a vertex to an identifying code, the extended code is also an identifying code.

An interesting question among identifying codes is what the minimum size of any identifying code in a given graph is. A minimum sized identifying code is called *optimal*. The size of an optimal identifying code of  $G$  is denoted by  $\gamma(G)$ .

This paper is about how much the size of an optimal identifying code can increase if some vertices are deleted from the graph. In other words, we study the difference of  $\gamma(G_S) - \gamma(G)$  where  $G = (V, E)$  is any graph and

$$G_S = (V \setminus S, \{\{u, v\} \in E \mid u \notin S \text{ and } v \notin S\}),$$

i.e. the induced subgraph of  $G$  where the subset  $S$  of vertices is deleted. If  $S = \{x\}$  contains only one vertex we typically refer to  $G_x$  instead of  $G_S$ . Moreover, we denote the order of graphs  $G_x$  and  $G_S$  by  $n_x$  and  $n_S$  respectively. Foucaud et al [6] proved that  $\gamma(G_S) - \gamma(G) \geq -|S|$ . Moreover, many such graphs as  $G$  and  $G_S$  which satisfy the equation  $\gamma(G_S) - \gamma(G) = -|S|$  are known. On the other hand, Charon et al [2] showed that there are twin-free graphs  $G$  and  $G_x$  such as  $\gamma(G_x) - \gamma(G) \approx \frac{n}{2} - \frac{3}{2} \log_2 n$ . This in particular means that  $\gamma(G)$  is not a monotone parameter with respect to the subgraph inclusion order.

In this paper, we prove the following inequalities of the difference between the sizes of optimal identifying codes in  $G$  and  $G_S$ :

$$\begin{aligned} \gamma(G_x) - \gamma(G) &\leq \begin{cases} \lfloor \frac{n}{2} \rfloor - 2 & \text{if } n \in \{2, 4, 5, 6, 8\} \\ \lfloor \frac{n}{2} \rfloor - 1 & \text{otherwise} \end{cases} \\ \gamma(G_S) - \gamma(G) &\leq n - 2|S| - \left\lfloor \frac{n - |S|}{2^{|S|}} \right\rfloor \quad \text{if } n \geq 2^{|S|-1}. \end{aligned}$$

Besides, we show that the first inequality is tight for all  $n$  and the second inequality is tight when  $n$  is large enough with respect to  $S$ . We also show that there are connected graphs  $G$  and  $G_x$  which attain this maximal value if  $n$  is odd and greater than 3. Instead, if  $n$  is even, then there are no such connected graphs except if  $n = 6$  or  $n = 8$ . Finally, we improve the previous upper bound for the case when  $G$  is bipartite. We prove the bound

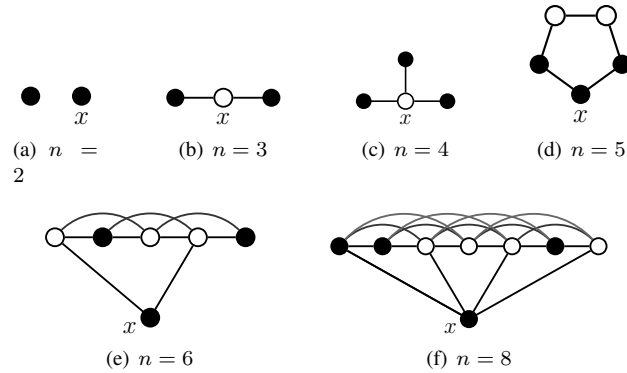
$$\gamma(G_x) - \gamma(G) \leq \left\lfloor \frac{n - \log_2(n - \log_2 n)}{2} \right\rfloor - 1$$

for such graphs and show that it is tight for all  $n \geq 3$ .

## 2 Constructions

**Theorem 1.** *For every  $n \geq 2$  there exists a twin-free graph  $G = (V, E)$  and a vertex  $x \in V$  such that  $G_x$  is twin-free and*

$$\gamma(G_x) - \gamma(G) \geq \begin{cases} \lfloor \frac{n}{2} \rfloor - 2 & \text{if } n \in \{2, 4, 5, 6, 8\} \\ \lfloor \frac{n}{2} \rfloor - 1 & \text{otherwise.} \end{cases}$$



**Fig. 1:** A graph  $G$  where  $\gamma(G_x) - \gamma(G)$  is maximal for small orders  $n$ . Black dots are codewords of an optimal identifying code in  $G$ .

**Proof:** For  $n \leq 6$  or  $n = 8$ , an example of a graph for which the inequality is tight is given in Figure 1.

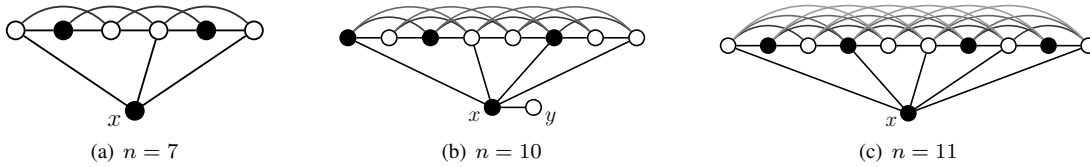
Assume that  $n = 7$  or  $n \geq 9$ . First, we define  $t = \lfloor \frac{n-1}{2} \rfloor$ . We can now construct the following family of graphs. Let  $G$  be the graph with the vertex set

$$V = \begin{cases} \{v_1, v_2, \dots, v_{n-1} = v_{2t}; x\} & \text{if } n \text{ is odd} \\ \{v_1, v_2, \dots, v_{n-2} = v_{2t}; x, y\} & \text{if } n \text{ is even,} \end{cases}$$

and the edge set

$$E = \{\{v_i, v_j\} : |i - j| < t\} \cup \{\{v_i, x\} : i \geq 4 \text{ and even}\} \cup \{\{v_1, x\}\}.$$

Moreover, there is an edge  $\{x, y\} \in E$  if  $n$  is even. Some examples of these constructions are shown in Figure 2.



**Fig. 2:** Three graphs  $G$  such that  $\gamma(G_x) - \gamma(G)$  is maximal. The black dots are an identifying code of  $G$ .

Now, we show that

$$C = \begin{cases} \{x; v_2, v_4, \dots, v_{t-1}; v_{t+2}, v_{t+4}, \dots, v_{2t-1}\} & \text{if } t \text{ is odd} \\ \{x; v_1, v_3, \dots, v_{t-1}; v_{t+2}, v_{t+4}, \dots, v_{2t-2}\} & \text{if } t \text{ is even} \end{cases}$$

is an identifying code of  $G$ . The black dots in Figures 2(a), 2(b) and 2(c) illustrate the code  $C$  when  $n = 7$ ,  $n = 10$  and  $n = 11$  respectively.

We first show that  $C$  separates  $x$  from all the other vertices. First, if  $n = 7$ ,  $x$  is the only vertex whose neighbourhood does not contain any codeword other than  $x$ . If  $t$  is even, so  $x$  is the only vertex whose neighbourhood contains  $v_1$  but not  $v_3$ . Otherwise  $t$  is odd and at least 5. Then  $v_2 \notin I(x)$  and  $v_{t+2} \notin I(x)$ . Now,  $v_1$  and  $v_2$  are the only vertices in addition to  $x$  whose neighbourhoods do not contain  $v_{t+2}$ . However, the neighbourhoods of both  $v_1$  and  $v_2$  contain  $v_2$ , so  $x$  is also separated from these vertices by  $C$ .

Next, we prove that  $v_i$  and  $v_j$  are separated by  $C$  for all  $i$  and  $j$  when  $i \neq j$ . Without loss of generality, we assume  $i < j$ . If  $j - i = 1$  then exactly one of  $I(v_i)$  and  $I(v_j)$  contains  $x$  except if  $i = 2$  and  $j = 3$ . However, since  $v_{t+2} \in I(v_3) \setminus I(v_2)$ , then  $v_2$  and  $v_3$  are also separated. Now assume that  $j - i \geq 2$ . This claim essentially follows since the only couples of consecutive non-codewords are  $v_t$  and  $v_{t+1}$  and (for  $n$  even)  $v_{2t-1}$  and  $v_{2t}$ . Strictly speaking, if  $1 \leq i \leq t - 2$  and then  $i + 2 \leq j \leq 2t$ , then  $v_{i+t} \in N(v_j) \setminus N(v_i)$  and  $v_{i+t+1} \in N(v_j) \setminus N(v_i)$  and one of  $v_{i+t}$  and  $v_{i+t+1}$  is a codeword and separates  $v_i$  from  $v_j$ . Analogously, if  $t + 2 \leq j \leq 2t$  and then  $1 \leq i \leq j - 2$ , then either  $v_{j-t-1}$  or  $v_{j-t}$  separates  $v_i$  from  $v_j$ . Otherwise,  $t - 1 \leq i \leq j \leq t + 1$  and since  $j - i \geq 2$ , then  $i = t - 1$  and  $j = t + 1$ . Now,  $v_1 \in N(v_{t-1}) \setminus N(v_{t+1})$  and  $v_{2t-1} \in N(v_{t+1}) \setminus N(v_{t-1})$  and one of these two vertices is a codeword which separates  $v_{t+1}$  and  $v_{t-1}$ .

If  $n$  is even and  $n \geq 10$ , we must also separate  $y$  from the other vertices. Now,  $I(y) = \{x\}$  and for all  $v \in V \setminus \{y\}$ , the set  $I(v)$  contains either  $v_{t-1}$  or  $v_{t+2}$  (including when  $v = x$  since  $t = \lfloor \frac{n-1}{2} \rfloor \geq 4$ ). Therefore,  $y$  is separated from all other vertices by  $C$ . We can also see that the set  $I(v)$  is non-empty for every  $v \in V$ .

Thus,  $C$  is an identifying code of size  $t = \lfloor \frac{n-1}{2} \rfloor$ , and so  $\gamma(G) = \lfloor \frac{n-1}{2} \rfloor$ .

Notice that  $G_x$  is twin-free since it is the disjoint union of a vertex and the power of a path. Besides, it has been shown in [6] that such power of a path is a tight example for Lemma 7 and hence, any minimum identifying code of  $G_x$  has size  $n - 2$ .

Hence

$$\gamma(G_x) - \gamma(G) = n - 2 - \left\lfloor \frac{n-1}{2} \right\rfloor = \left\lfloor \frac{n}{2} \right\rfloor - 1.$$

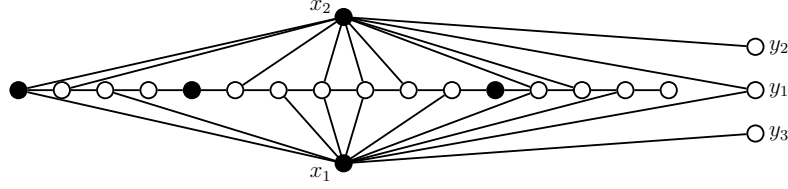
□

**Theorem 2.** For every  $s \geq 1$  and  $n$  such that  $n \geq 4 \cdot 2^s + s$ , there exists a twin-free graph  $G = (V, E)$  and a set  $S = \{x_1, \dots, x_s\} \subseteq V$  such that  $G_S$  is twin-free and

$$\gamma(G_S) - \gamma(G) \geq n - 2s - \left\lfloor \frac{n-s}{2^s} \right\rfloor.$$

**Proof:** Let  $n = s + 2^s \cdot k + l$  such that  $0 \leq l < 2^s$ . From this follows that  $k = \lfloor \frac{n-s}{2^s} \rfloor$ . Moreover, we define  $t = 2^s \cdot k$ . First, we define the graph  $G_S = (V_S, E_S)$  that consists on vertices  $P = \{v_1, v_2, \dots, v_t\}$  which form a power of a path such that there is an edge between  $v_i$  and  $v_j$  if  $|i - j| < \frac{t}{2}$ . Moreover, there are  $l$  isolated vertices  $Y = \{y_1, y_2, \dots, y_l\}$ . Now we can show in the same way as in the proof of Theorem 1 that  $\gamma(G_S) = n_S - 1$ .

The graph  $G$  is constructed from  $G_S$  by adding the set of vertices  $S = \{x_1, x_2, \dots, x_s\}$ . Every vertex of  $Y$  will have a unique set of vertices of  $S$  in its neighbourhood. Similarly, we can partition vertices of  $P$  into  $k$  sets which all contain  $2^s$  consecutive vertices  $(v_{i \cdot 2^s + 1}, \dots, v_{(i+1)2^s})$ . Every vertex of such a subset has a unique set of vertices of  $S$  in its neighbourhood. Thus, all the vertices in a given subset are separated from each other by  $S$ .



**Fig. 3:** A such graph  $G$  where  $\gamma(G_S) - \gamma(G)$  is maximal when  $|S| = 2$  and  $n = 21$ . The black dots are an identifying code of  $G$ . Notice that all edges between the vertices in the path are not drawn in the figure.

The graph  $G = (V, E)$  can be defined as follows:

$$V = \{v_1, v_2, \dots, v_t; x_1, x_2, \dots, x_s; y_1, y_2, \dots, y_l\}$$

and

$$E = \begin{cases} E_v \cup E_y \cup E_o & \text{if } k \text{ is odd} \\ E_v \cup E_y \cup E_e & \text{if } k \text{ is even,} \end{cases}$$

where

$$E_v = \left\{ \{v_i, v_j\} \mid |i - j| < 2^{s-1}k = 2^{s-1} \cdot \frac{n-s-l}{2^s} \right\},$$

$$E_y = \{ \{x_i, y_j\} \mid j \equiv b \pmod{2^i} \text{ for some } b \in \{1, 2, \dots, 2^{i-1}\} \},$$

$$E_e = \left\{ \{x_i, v_j\} \mid i = 1, 2, \dots, s; j \neq \frac{t}{2} + 1 - 2^s \text{ and } j \equiv b \pmod{2^i} \text{ for some } b = 1, 2, \dots, 2^{i-1} \right\} \\ \cup \left\{ \{x_i, v_{\frac{t}{2}}\} \mid i = 1, 2, \dots, s \right\}$$

and

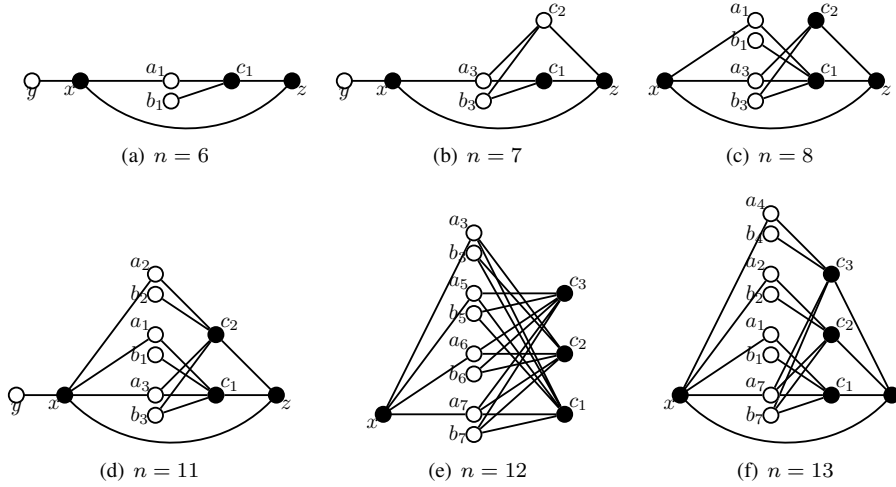
$$E_o = \left\{ \{x_i, v_j\} \mid i = 1, 2, \dots, s-1; j \neq \frac{t}{2} + 1 - 2^s \text{ and } j \equiv b \pmod{2^i} \text{ for some } b = 1, 2, \dots, 2^{i-1} \right\} \\ \cup \left\{ \{x_s, v_j\} \mid j \neq \frac{t}{2} + 1 - 2^s \text{ and } j \equiv b + 1 \pmod{2^i} \text{ for some } b = 1, 2, \dots, 2^{s-1} \right\} \\ \cup \left\{ \{x_i, v_{\frac{t}{2} - 2^{s-1}}\} \mid i = 1, 2, \dots, s \right\}.$$

We can now see that

$$C = S \cup \left\{ v_i \mid i = \frac{t}{2} + 1 - 2^s \cdot j, j = 1, 2, \dots, \left\lfloor \frac{k}{2} \right\rfloor \right\} \cup \left\{ v_i \mid i = \frac{t}{2} + 2^s \cdot j, j = 1, 2, \dots, \left\lfloor \frac{k-1}{2} \right\rfloor \right\}$$

is an optimal identifying code in  $G$  and it contains  $\left\lfloor \frac{n-s}{2^s} \right\rfloor + s - 1$  codewords. An example of such a graph and an optimal code is given in Figure 3, when  $|S| = 2$  and  $n = 21$ . On the other hand,  $V_S \setminus \{v_1\}$  is an optimal identifying code in  $G_S$ .

The details of the proof can be proved in a similar way as in the proof of Theorem 1 and the details of this proof are ignored.  $\square$



**Fig. 4:** Bipartite graphs  $G$  such that  $\gamma(G_x) - \gamma(G)$  is maximal within these bipartite graphs. The maximality of Constructions 4(a), 4(d) and 4(f) follows from Theorem 4. The maximality of Constructions 4(c) and 4(b) and 4(e) follows from Remarks 5 and 6. The black dots are an identifying code of  $G$ . The vertices  $a_i, c_j, x$  and  $y$  are an optimal identifying code of  $G_x$ .

**Remark 3.** Theorem 2 is also valid if  $3 \cdot 2^s + s \leq n < 4 \cdot 2^s$ . The proof of Theorem 2 is valid if the set  $Y$  is chosen so that all these vertices have at least two neighbours which belong to the set  $S$ .

**Theorem 4.** Let  $k$  be the smallest integer such that  $n \leq 2^k + k$ . There is a bipartite graph  $G = (V, E)$  and a vertex  $x \in V$  such that

$$\gamma(G_x) - \gamma(G) \geq \left\lfloor \frac{n - k}{2} \right\rfloor - 1$$

if  $n = 5, 6, 10, 11$  or  $n \geq 13$ .

**Proof:** Let  $t = \lfloor \frac{n-k-1}{2} \rfloor$ . Then  $2^{k-2} \leq t \leq 2^{k-1} - 1$ . We define  $G = (V, E)$  as follows:

$$V = \{a_i, b_i \mid i = 1, 2, \dots, t - 1; 2^{k-1} - 1\} \cup \{c_j \mid j = 1, 2, \dots, k - 1\} \cup \{x, z\}$$

If  $n - k$  is even, then a vertex  $y$  is added to  $V$ . If  $t = 2^{k-2}$  and  $n \geq 13$ , then  $a_3$  and  $b_3$  are replaced by  $a_{2^{k-2}}$  and  $b_{2^{k-2}}$ . Thus,  $a_{2^i} \in V$  and  $b_{2^i} \in V$  belong to  $V$  for all  $i = 0, 1, \dots, k - 2$ . The edge set is

$$E = \{\{a_i, c_j\}, \{b_i, c_j\} \mid i \equiv -1, -2, \dots, -2^{j-1} \pmod{2^j}\} \cup \{\{a_i, x\} \mid \forall i\} \cup \{\{z, c_j\} \mid \forall j\} \cup \{z, x\} \cup \{y, x\}$$

The edge  $\{y, x\}$  naturally belongs to  $E$  only if  $y \in V$ . The graph  $G$  is bipartite since exactly one of the ends of each edge is  $x$  or  $c_i$  for some  $i$ . The graph  $G$  is given in Figures 4(a), 4(d) and 4(f) when  $n$  is 6, 11 and 13, respectively.

Now, we show that

$$C = \{x, z\} \cup \{c_i \mid \forall i\}$$

is an identifying code of  $G$ , and we see that  $\gamma(G) \leq |C| = k + 1$ . First,  $z$  is the only vertex which is dominated by  $x, z$  and  $c_1$ . Second,  $x$  is the only vertex which is dominated by  $x$  and  $z$  but not  $c_1$ . Moreover,  $y$  is the only vertex in addition to  $x$  which is not dominated by  $c_i$  for any  $i$ . Then,  $x, y$  and  $z$  are separated from all other vertices. Now,  $c_i$  is the only vertex which is dominated by  $c_i$  and  $z$ . Therefore, every  $c_i$  is separated from all other vertices. Furthermore,  $x$  dominates  $a_i$  for all  $i$ , but not any  $b_j$ , i.e.  $a_i$  is separated from  $b_j$  for all  $i$  and  $j$ . Finally, every  $a_i$  is dominated by a unique set of vertices  $c_j$ . Thus,  $a_i$  is separated from  $a_j$  if  $i \neq j$ . In the same way we show that  $b_i$  and  $b_j$  are separated. Thus, we have shown that  $C$  is an identifying code.

Let

$$C_x = \{y, z\} \cup \{a_i \mid \forall i\} \cup \{c_i \mid \forall i\}$$

be a code of  $G_x$ . Again,  $y$  naturally belongs to  $C_x$  only if  $y \in V$ . We prove that there is not an identifying code of  $G_x$  with smaller cardinality than  $|C_x|$ . First, if  $y \in V$ , it is isolated and must be a codeword of every identifying code of  $G_x$ . Second,  $a_i$  and  $b_i$  are the only vertices which are able to separate  $a_i$  and  $b_i$ . Without loss of generality, we can assume that  $a_i$  belongs to the identifying code for all  $i$ . The vertices  $z$  and  $b_{2^{k-1}-1}$  must still be separated and  $z$  and  $b_{2^{k-1}-1}$  are the only vertices which can do it since every  $c_i$  belongs to the neighbourhood of both  $z$  and  $b_{2^{k-1}-1}$ . Again, we can assume without loss of generality that  $z$  belongs to  $C_x$ . Finally, the vertices  $b_i$  are not dominated by  $z$  and  $a_j$ . Moreover, the neighbourhood of the vertex  $b_{2^i}$  contains only  $c_i$  in addition to  $b_{2^i}$ . Then, either  $c_i$  or  $b_{2^i}$  must belong to any identifying code. Now we have seen that the cardinality of every identifying code of  $G_x$  is at least

$$|C_x| = \begin{cases} t + k + 1 = \lfloor \frac{n-k-1}{2} \rfloor + k + 1 = \frac{n+k}{2} & \text{if } y \in V, \text{ i.e. } n - k \text{ is even,} \\ t + k = \lfloor \frac{n-k-1}{2} \rfloor + k = \frac{n+k-1}{2} & \text{if } y \notin V, \text{ i.e. } n - k \text{ is odd} \end{cases}$$

since  $C_x$  contains  $t$  vertices of  $a_i, k - 1$  vertices of  $c_j$ , and  $z$  and possibly  $y$ . Hence,

$$\gamma(G_x) - \gamma(G) \geq |C_x| - |C| = \left\lfloor \frac{n+k}{2} \right\rfloor - (k+1) = \left\lfloor \frac{n-k}{2} \right\rfloor - 1.$$

□

**Remark 5.** Theorem 4 and the construction of the proof of this theorem are also valid if  $n = 8$  or  $n = 9$ . The reasons why  $|C|$  and  $|C_x|$  are optimal identifying codes of  $G$  and  $G_x$  respectively, are valid despite the fact that there is not a vertex  $b_2$  in  $G$ . The vertex  $b_2$  only has  $b_2$  and  $c_2$  in its neighbourhood.

We can assume as in the proof of Theorem 4 that  $a_1, a_3$  and  $z$  are codewords of  $C_x$ . Now,  $b_1$  and  $b_3$  are not dominated. These two vertices can not both be dominated and separated by only one codeword. Therefore, the set  $I(b_1) \cup I(b_3) \subseteq N(b_1) \cup N(b_3) = \{b_1, b_3, c_1, c_2\}$  must contain at least two vertices. Thus,  $C_x$  is optimal.

**Remark 6.** Theorem 4 is valid also if  $n \in \{3, 4, 7, 12\}$  but either  $G$  is not defined or  $C_x$  is not an optimal identifying code of  $G_x$ . The claim follows from the constructions in Figures 1(b) and 1(c) in the cases  $n = 3$  and  $n = 4$  respectively. If  $n = 7$ , the same construction as the proof of Theorem 4 is a valid example, but the optimal codes of  $G$  and  $G_x$  are  $\{z, c_1, x\}$  and  $\{z, c_1, a, y\}$  respectively. The case  $n = 12$  follows from the construction where  $V = \{x, c_1, c_2, c_3\} \cup \{a_i, b_i : i = 3, 5, 6, 7\}$  and the edges define in a similar way as in the proof of Theorem 4. Now, it is possible to show that  $C = \{x, c_1, c_2, c_3\}$  and  $C_x = \{a_3, a_5, a_6, a_7, c_1, c_2, c_3\}$  are optimal identifying codes of  $G$  and  $G_x$  respectively.



### 3 Upper bounds for $\gamma(G_x) - \gamma(G)$

We first give some basic results which are needed in this section.

**Lemma 7** ([7, 1]). *Let  $G = (V, E)$  be a twin-free graph with  $n$  vertices, and at least one edge, then*

$$\gamma(G) \leq n - 1.$$

**Lemma 8** ([9]). *Let  $G = (V, E)$  be a twin-free graph with  $n$  vertices. Then,*

$$\gamma(G) \geq \lceil \log_2(n + 1) \rceil.$$

**Lemma 9** ([9]). *Let  $G = (V, E)$  be a twin-free graph of order  $n$  and  $E \neq \emptyset$ . If (exactly) one end of each edge of  $E$  is  $x \in V$ , then*

$$\gamma(G) = n - 1.$$

We next show the upper bound for  $\gamma(G_S) - \gamma(G)$  in general case.

**Theorem 10.** *Let  $G$  be a twin-free graph on  $n$  vertices and let  $S \subseteq V$  such that satisfying  $n \geq 2^{|S|-1}$ . If  $G_S$  is twin-free, then*

$$\gamma(G_S) - \gamma(G) \leq n - 2|S| - \left\lfloor \frac{n - |S|}{2^{|S|}} \right\rfloor. \quad (1)$$

**Proof:** First, if  $2^{|S|-1} \leq n \leq 2^{|S|} - 1$ , then the claim follows since  $|S| = \lceil \log_2(n + 1) \rceil$  by Lemma 8 and since an optimal identifying code of  $G_S$  contains at most  $n_S = n - |S|$  codewords.

Assume then that  $n \geq 2^{|S|}$ . Let  $C$  be an optimal identifying code of  $G$ . Now, we construct an identifying code  $C_S$  of  $G_S$ . Let  $C_S = C \setminus S$  in the beginning. Next, we partition the vertices of  $G_S$  into equivalence classes  $T_R$  such that the vertex  $v$  belongs to  $T_R$  if  $I(v, C \setminus S) = R$ . We can also assume that all vertices of  $S$  belong to  $C$ . Otherwise we can reduce it to a smaller graph. Moreover, the right hand side of Equation (1) increases when  $|S|$  increases and  $n \geq 2^{|S|}$ .

First, we can observe that every class  $T_R$  can contain at most  $2^{|S|}$  vertices. However, there can be at most  $2^{|S|} - 1$  vertices in  $T_\emptyset$ . Now we can separate vertices of  $T_R$  by adding at most  $|T_R| - 1$  codewords to  $C_S$  if  $R \neq \emptyset$ . Instead we may need to add  $|T_\emptyset|$  vertices to  $C_S$  so that all vertices of  $T_\emptyset$  are dominated (and separated from each other).

There have to be at least  $\frac{|G_S|}{2^{|S|}}$  non-empty equivalence classes if  $T_\emptyset = \emptyset$ . Otherwise, there are at least  $\frac{|G_S| - (2^{|S|} - 1)}{2^{|S|}}$  equivalence classes in addition to  $T_\emptyset$ . In all cases, the number of the other than  $T_\emptyset$  equivalence classes is at least

$$\min \left\{ \left\lceil \frac{|G_S|}{2^{|S|}} \right\rceil, \left\lceil \frac{|G_S| - (2^{|S|} - 1)}{2^{|S|}} \right\rceil \right\} = \left\lceil \frac{|G_S| - (2^{|S|} - 1)}{2^{|S|}} \right\rceil = \left\lfloor \frac{|G_S|}{2^{|S|}} \right\rfloor.$$

From this follows that

$$\begin{aligned}
 |C_S| &\leq |C \setminus S| + \left( |T_\emptyset| + \sum_{R \neq \emptyset, T_R \neq \emptyset} (|T_R| - 1) \right) \\
 &= (|C| - |S|) + \sum_{T_R \neq \emptyset} |T_R| - \sum_{R \neq \emptyset, T_R \neq \emptyset} 1 \leq |C| - |S| + |G_S| - \left\lfloor \frac{|G_S|}{2^{|S|}} \right\rfloor \\
 &= |C| - |S| + (n - |S|) - \left\lfloor \frac{n - |S|}{2^{|S|}} \right\rfloor = |C| + n - 2|S| - \left\lfloor \frac{n - |S|}{2^{|S|}} \right\rfloor.
 \end{aligned}$$

□

**Theorem 11.** *Let  $G$  be a twin-free graph on  $n \in \{2, 4, 5, 6, 8\}$  vertices. For every  $x \in V$ , if  $G_x$  is twin-free, then*

$$\gamma(G_x) - \gamma(G) \leq \left\lfloor \frac{n}{2} \right\rfloor - 2.$$

**Proof:** The case  $n = 2$  is straightforward to check.

Assume now that  $n \geq 4$ . If an optimal identifying code of  $G_x$  contains  $n_x$  codewords, there are no edges in  $G_x$ . Then one end of all edges in  $G$  is  $x$  and  $\gamma(G) \geq n - 1 = n_x = \gamma(G_x)$  by Lemma 9, i.e.  $\gamma(G_x) - \gamma(G) \leq 0 \leq \left\lfloor \frac{n}{2} \right\rfloor - 2$ . Otherwise,  $\gamma(G_x) \leq n_x - 1$  and now the claim follows again by Lemma 8. □

**Corollary 12.** *Let  $G = (V, E)$  be a twin-free graph with  $n$  vertices. Then,*

$$\gamma(G_x) - \gamma(G) \leq \begin{cases} \left\lfloor \frac{n}{2} \right\rfloor - 2 & \text{if } n \in \{2, 4, 5, 6, 8\} \\ \left\lfloor \frac{n}{2} \right\rfloor - 1 & \text{in other cases} \end{cases}$$

and the inequality is tight.

**Proof:** The claim follows by Theorems 1, 10 and 11. □

**Corollary 13.**

$$\frac{\gamma(G_x)}{\gamma(G)} \leq \frac{\lceil \log_2(n+1) \rceil + \left\lfloor \frac{n}{2} \right\rfloor - 1}{\lceil \log_2(n+1) \rceil} \approx \frac{n}{2 \log_2 n}.$$

**Proof:** The claim follows by Theorem 10 and Lemma 8. □

The inequality of Corollary 13 is almost tight. Indeed, Charon et al [2] proved that there are graphs  $G$  and  $G_x$  such that  $\frac{\gamma(G_x)}{\gamma(G)} \approx \frac{n}{2 \log_2 n}$  (Conclusion 6 [2]). There is nevertheless a small gap between the construction of Conclusion 6 [2] and Corollary 13. In fact, the construction of the proof in Theorem 4 is very similar to the construction of Conclusion 6 [2] and it also satisfies the approximate equation  $\frac{\gamma(G_x)}{\gamma(G)} \approx \frac{n}{2 \log_2 n}$ .

Next, we show that if  $n$  is even and large enough and  $\gamma(G_x) - \gamma(G)$  reaches the upper bound  $\gamma(G_x) - \gamma(G) = \left\lfloor \frac{n}{2} \right\rfloor - 1$ , then  $G_x$  must be disconnected.

**Proposition 14.** *There does not exist graphs  $G$  and  $G_x$  such that  $G_x$  is connected and  $\gamma(G_x) - \gamma(G) = \lfloor \frac{n}{2} \rfloor - 1$  if  $n$  is even and at least 10.*

**Proof:** Assume the contrary that  $G$  and  $G_x$  would be such graphs. Let  $C$  be an optimal identifying code of  $G$ . The vertices of  $G_x$  can be divided into equivalence classes  $T_R$  as the proof of Theorem 10. In particular, every equivalence class contains at most 2 vertices. Therefore, two vertices are called a *pair* if the vertices belong to the same equivalence class. Since  $n$  is even, then  $n_x$  is odd. From this follows that there are at most  $\frac{n_x-1}{2}$  pairs and at least one equivalence class which contains only one vertex. Now, if  $\gamma(G_x) - \gamma(G) = \lfloor \frac{n}{2} \rfloor - 1$ , then every vertex must belong to a pair except one vertex  $y$  which satisfies the condition  $I(y, C) = \{x\}$ . Moreover, any two pairs are not separated by the same vertex of  $G_x$  and any vertex which dominates  $y$  does not separate any of these pairs.

Since  $G_x$  is connected,  $y$  has at least one neighbour in  $G_x$ . Assume that  $v_1$  and  $y$  are connected by an edge. Let  $u_1$  be the pair of  $v_1$  that is  $C \setminus \{x\}$  does not separate  $v_1$  and  $u_1$ . Now, if  $u_1$  and  $y$  were not neighbours then  $y$  would separate  $v_1$  and  $u_1$  in addition to dominating  $y$ . Thus, all the vertices  $y$ ,  $v_1$  and  $u_1$  can dominate  $y$ . Therefore any of these three vertices can not separate any of these  $\frac{n_x-1}{2}$  pairs.

Now, we show the claim by induction. Assume that  $y$ ,  $v_1, \dots, v_k$  and  $u_1, \dots, u_k$  can separate only the pairs  $\{v_i, u_i\}$ , where  $i = 1, 2, \dots, k-1$ . Let  $\{v_{k+1}, u_{k+1}\}$  be the pair such that at least one of these vertices separate the pair  $\{v_k, u_k\}$ . Without loss of generality, we can assume that there is an edge between  $v_k$  and  $v_{k+1}$  but not between  $u_k$  and  $v_{k+1}$ . Now, if there is not an edge between  $v_k$  and  $u_{k+1}$ , then  $v_k$  separates the pair  $\{v_{k+1}, u_{k+1}\}$  against the assumption of induction. Similarly, if there is an edge between  $u_k$  and  $u_{k+1}$ , then  $u_k$  separates  $\{v_{k+1}, u_{k+1}\}$ . Again, this is against the assumption of induction. Therefore both  $u_{k+1}$  and  $v_{k+1}$  can separate the pair  $\{v_k, u_k\}$ . Thus, none of these two vertices can separate any other pair. Hence, any of the vertices  $y$ ,  $v_1, \dots, v_{k+1}$  and  $u_1, \dots, u_{k+1}$  can not separate any other pairs except  $\{v_i, u_i\}$ , when  $i = 1, 2, \dots, k$ . In other words,  $y$  and  $k+1$  pairs can separate only  $k$  pairs and this holds for all  $k$ . This is nevertheless impossible since the graph is finite. Therefore it is impossible that  $G_x$  is connected if  $n \geq 10$ ,  $n$  is even and  $\gamma(G_x) - \gamma(G) = \lfloor \frac{n}{2} \rfloor - 1$ .  $\square$

Next we prove an upper bound for  $\gamma(G_x) - \gamma(G)$  among the bipartite graphs.

**Theorem 15.** *For bipartite twin-free graphs  $G$  and  $G_x$ , the maximum difference between  $\gamma(G_x)$  and  $\gamma(G)$  is*

$$\gamma(G_x) - \gamma(G) \leq \frac{n-3}{2} - \frac{1}{3} \left\lfloor \log_2 \left( n - 3 - \lfloor \log_2 (n - \lfloor \log_2 (4n) \rfloor) \rfloor \right) \right\rfloor - \frac{1}{6} \left\lfloor \log_2 \left( n - \lfloor \log_2 (4n) \rfloor \right) \right\rfloor$$

if  $n \geq 7$ .

**Proof:** Let  $C$  be an optimal identifying code in  $G$  and  $C \setminus \{x\}$  be a code in  $G_x$ . Let  $A$  and  $B$  be a partition of all vertices  $V$  such that if there is an edge between vertices  $u$  and  $v$ , then one of these two vertices belongs to  $A$  and the other belongs to  $B$ . Assume also that  $x$  belongs to  $B$ . Moreover, if there is a vertex  $v$  such that  $I(v, C) = \{x\}$  and  $v \neq x$ , the vertex is denoted by  $y$ . Observe that  $y$  is not dominated by  $C \setminus \{x\}$ .

### Step 1. The partition of pairs

If  $C \setminus \{x\}$  does not separate  $u$  and  $v$ , then  $u$  and  $v$  is said to be a *pair*. A pair of  $u$  and  $v$  is called *AA-pair* if both  $u$  and  $v$  belong to  $A$ . Also, the set  $AA$  denotes the set of vertices which belong to some *AA-pair*. Furthermore,  $|AA|$  denotes the number of vertices of *AA-pairs*. Thus, the number of *AA-pairs*

is  $\frac{1}{2}|AA|$ . Similarly, a pair is called *AB-pair* if exactly one of the vertices of the pair belongs to  $A$  and the other one belongs to  $B$ . Notice that there can not be a pair between two vertices of the set  $B$  since  $x$  can not belong to the neighbourhood of any vertex of  $B$  except the neighbourhood of  $x$ .

We still divide *AB*-pairs into two disjoint sets  $AB^*$ - and  $AB'$ -pairs. Let  $a \in A$  and  $b \in B$  be an *AB*-pair.

- If there is an *AA*-pair of  $a'$  and  $a''$  such that both  $a'$  and  $a''$  are the neighbours of  $b$ , then  $a$  and  $b$  are an  $AB^*$ -pair.
- Otherwise, the pair of  $a$  and  $b$  is an  $AB'$ -pair.

We furthermore partition  $AB'$ -pairs into four disjoint sets  $AB'_{\alpha_1}$ ,  $AB'_{\alpha_2}$ ,  $AB'_{\alpha_3}$  and  $AB'_{\beta}$  or two sets  $AB'_{\alpha} = AB'_{\alpha_1} \cup AB'_{\alpha_2} \cup AB'_{\alpha_3}$  and  $AB'_{\beta}$ . We first define an induced subgraph  $G' = (AB', E')$  of  $G$ , where  $AB'$  is the set of vertices in  $AB'$ -pairs and  $E' = \{\{u, v\} \in E \mid u \in AB', v \in AB'\}$ . Let  $a \in AB'$  and  $b \in AB'$  be an  $AB'$ -pair and  $a'$  and  $b'$  be another  $AB'$ -pair. Notice that  $a$  and  $b$  belong to the same connected component in  $G'$  since there must be an edge between  $a$  and  $b$  or else the neighbourhoods of  $a$  and  $b$  would be totally distinct in  $G$ .

- If the pair of  $a$  and  $b$  belongs to the connected component whose cardinality is at least six vertices in  $G'$ , the pair belongs to  $AB'_{\alpha_3}$ .
- If  $a, b, a'$  and  $b'$  form a connected component of four vertices and if at least one of the pairs has no neighbours in  $G$  except  $x, a, a', b$  and  $b'$ , then all the vertices  $a, b, a'$  and  $b'$  belong to  $AB'_{\alpha_2}$ .
- If there is an  $AB'$ -pair of  $a \in A$  and  $b \in B$  such that the vertices do not belong to the set  $AB'_{\alpha_2} \cup AB'_{\alpha_3}$  and  $b$  has a neighbour which belongs to the set  $AA$ , then  $a$  and  $b$  belong to  $AB'_{\alpha_1}$ .
- Otherwise, an  $AB'$ -pair is an  $AB'_{\beta}$ -pair.

Notice that at least one of the vertices of  $AB'_{\beta}$ -pair must have such a non-codeword neighbour that does not belong to  $AA \cup AB'$ . Otherwise, the pair would be an  $AB'_{\alpha}$ -pair or  $G_x$  would not be twin-free.

Let us have a look some basic properties of *AA* and *AB*-pairs. First, no vertex of *AA* belongs to the code  $C$  since there are no edges between two vertices of *AA*. Secondly, if there is an *AB*-pair of  $a \in A$  and  $b \in B$ , then  $a$  and  $b$  are the only vertices which can belong to the neighbourhoods of both  $a$  and  $b$ . Then  $a, b$  and  $x$  are the only possible codewords in their neighbourhoods. On the other hand,  $a$  or  $b$  belongs to  $C$ , since  $b$  has to be dominated by  $C$ . Especially, if  $a$  and  $b$  belong to  $AB'_{\alpha}$ , then  $a$  or  $b$  is able to separate another pair, and from this follows that this vertex can not be a codeword. It also means that either  $a$  or  $b$  is a leaf of  $G'$  and a codeword of  $C$ .

### Step 2. The set $B^*$

We first prove that the cardinality of  $B^* = (B \cap C \setminus \{x\} \setminus AB')$  is  $|B^*| \geq \lfloor \log_2 |AA| \rfloor$  where  $|AA|$  is the number of vertices in *AA*-pairs. We see that all codeword neighbours of  $a \in AA$  belong to the set  $B^* \cup \{x\}$  since  $a$  and its pair  $a'$  must have the same codeword neighbours apart from  $x$ , and if  $b \in AB$  is a codeword neighbour of both  $a$  and  $a'$ , then  $b$  belongs to  $AB^*$ . Moreover, every vertex in *AA* has to be dominated by at least one codeword of  $B^*$  since vertices of *AA* have to be non-codewords and only one vertex of each pair is dominated by  $x$ . Furthermore, codewords in  $B^*$  are the only codewords which can

separate  $AA$ -pairs from other  $AA$ -pairs. Since we can form at most  $2^{|B^*|} - 1$  non-empty subsets from  $|B^*|$  codewords, then the number of  $AA$ -pairs is  $\frac{1}{2}|AA| \leq 2^{|B^*|} - 1$ , i.e.

$$|B^*| \geq \left\lceil \log_2 \left( \frac{1}{2}|AA| + 1 \right) \right\rceil = 1 + \left\lceil \log_2 \left( \frac{1}{2}|AA| \right) \right\rceil = \lfloor \log_2 |AA| \rfloor. \quad (2)$$

### Step 3. The construction of $C_x$

Next, we construct an identifying code  $C_x$  of  $G_x$  and show that  $|C_x| - |C| \leq \frac{1}{2}|AA| + \frac{1}{3}|AB'_\alpha| + |S| + |y| - 1$  where  $|S| \leq \frac{1}{2}|AB'_\beta|$  and  $|y| \in \{0, 1\}$  is the number of non-codewords which are dominated only by  $x$ .

Let  $C_x = C \setminus \{x\}$  or  $C_x = C \cup \{y\} \setminus \{x\}$  in the beginning, depending on whether there exists a vertex  $y$  in  $G$ . Then, we add and delete vertices to/from  $C_x$  in the following way:

**Step 3.1. Separation of  $AB'_{\alpha 3}$ -pairs:** Firstly, we study an  $AB'_{\alpha 3}$ -pair of  $a \in A$  and  $b \in B$ . Let  $T$  be a connected component of  $G'$  that contains  $a$  and  $b$ . (Since  $ab$  is an  $AB'_{\alpha 3}$ -pair, the size of the connected component is at least 6.) Let  $T'$  be an induced subgraph of  $T$  which contains non-codewords of  $C$ , i.e.  $T'$  contains one vertex of each pair of  $T$  and the vertices of  $T'$  are exactly all non-leaves of  $T$ . In particular,  $T'$  is connected and its order is at least three. Let  $T_D$  be a minimum total dominating set of  $T'$ , i.e. such a set that every vertex of  $T'$  has a neighbour which belongs to  $T_D$ . We now note that  $C_x \cup T_D$  separates  $a$  and  $b$  and also all the other pairs of  $T$ . Since either  $a$  or  $b$  belongs to  $T'$ , this vertex must have a neighbour which belongs to  $T_D$ . Moreover, it is known that every connected graph  $T'$  of order at least three has a total dominating set with at most  $\frac{2}{3} \cdot |T'|$  codewords [5, 8]. Moreover,  $T$  contains  $|T'|$   $AB'_{\alpha 3}$ -pairs. Therefore all  $AB'_{\alpha 3}$ -pairs can be separated adding  $\frac{1}{3}|AB'_{\alpha 3}|$  codewords to  $C_x$ . (Notice that there are  $\frac{1}{2}|AB'_{\alpha 3}|$  different  $AB'_{\alpha 3}$ -pairs.)

**Step 3.2. Separation of  $AB'_{\alpha 2}$ -pairs:** Let  $a, b, a'$  and  $b'$  be the vertices of a connected component of  $G'$  such that  $a$  and  $b$  is an  $AB'_{\alpha 2}$ -pair and  $a'$  and  $b'$  is another  $AB'_{\alpha 2}$ -pair. As we formerly mentioned in the end of Step 1 there has to be exactly two leaves of the connected component and the leaves are codewords of  $C$ , but the other vertices are not. Therefore, the component has to be a path (either  $a - b - a' - b'$  or  $a' - b' - a - b$ ). Without loss of generality, we can assume that the path is  $a - b - a' - b'$  and  $a$  and  $b$  have no neighbours apart from  $x, a, b$  and  $a'$  in  $G$ . Now,  $a$  and  $b'$  belong to  $C$ . We show that  $C \cup \{b, a'\} \setminus \{a, x\}$  separates any vertices  $u \in V \setminus \{x\}$  and  $v \in V \setminus \{x, u\}$  if  $C \setminus \{x\}$  also separates them or  $u$  or  $v$  belongs to the set  $\{a, b, a', b'\}$ . First,  $a$  and  $b$  are the only vertices of  $G_x$  such that a vertex has been deleted from their identifying set. Then it is enough to show that  $a$  and  $b$  are separated from all other vertices of  $G_x$  and  $a'$  is separated from  $b'$ . This is valid, since  $a, b$  and  $a'$  are the only vertices which are dominated by  $b$  and the identifying sets  $I(a) = \{b\}$ ,  $I(b) = \{b, a'\}$  and  $I(a') \supseteq \{b, a', b'\}$  are different. All in all, we need to add two vertices to  $C_x$  and we can delete a vertex from  $C_x$ . Therefore, all  $AB'_{\alpha 2}$ -pairs can be separated by increasing the number of codewords of  $C_x$  by  $\frac{1}{4}|AB'_{\alpha 2}|$ .

**Step 3.3. Separation of  $AB'_\beta$ -pairs:** If  $a$  and  $b$  belong to an  $AB'_\beta$ -pair, then  $a$  or  $b$  must have a non-codeword neighbour  $v$  such that  $v \notin AA \cup AB'$ . We add  $v$  to  $C_x$ . Notice that  $a$  and  $b$  may have many such neighbours  $v$  but we add only one of such vertices to  $C_x$ . Moreover, it is possible that we add the same vertex  $v$  because of more than one  $AB'_\beta$ -pair. All in all, we need to add at most  $\frac{1}{2}|AB'_\beta|$  codewords to  $C_x$ . The set of the vertices which add to  $C_x$  by  $AB'_\beta$ -pairs is called the set  $S$ . Thus,  $|S| \leq \frac{1}{2}|AB'_\beta|$  and we also know that  $S \cap (AA \cup AB' \cup B^*) = \emptyset$  since  $B^* \subseteq C$  and  $S \cap C = \emptyset$ .

**Step 3.4. Separation of  $AA$ -pairs and  $AB'_{\alpha 1}$ -pairs:** Finally, we add the following non-codewords of  $AA$  to the code  $C_x$ . First, every  $AB'_{\alpha 1}$ -pair has at least one neighbour which belongs to  $AA$ . We add one

of these neighbours to  $C_x$  for every  $AB'_{\alpha_1}$ -pair. Second, if none of the vertices of an  $AA$ -pair is not yet added to  $C_x$ , then we add one of these two vertices to  $C_x$ . Now we have added at least one of the vertices of each  $AA$ -pair to  $C_x$ . However, if there is an  $AA$ -pair of  $a$  and  $a'$  such that both  $a$  and  $a'$  belong to  $C_x$  now, this means that there have to be two  $AB'_{\alpha_1}$ -pairs such that  $a$  has been added by another  $AB'_{\alpha_1}$ -pair and  $a'$  by the other  $AB'_{\alpha_1}$ -pair. Therefore,  $a$  and  $a'$  are able to separate three pairs: an  $AA$ -pair and two  $AB'_{\alpha_1}$ -pairs. In particular, we add at most  $\frac{1}{2}|AA| + \frac{1}{4}|AB'_{\alpha_1}|$  vertices of the set  $AA$  to  $C_x$  in the final state.

**Step 3.5. The Conclusion of Step 3 and the separation of  $AB^*$ -pairs:** Now, we have seen that  $C_x$  separates  $AA$ -pairs and  $AB'$ -pairs. Moreover, it separates  $AB^*$ -pairs since at least one of the vertices of every  $AA$ -pair is a codeword of  $C_x$ . Therefore,  $C_x$  is an identifying code of  $G_x$ . Moreover,

$$\begin{aligned} |C_x| - |C| &\leq \frac{1}{2}|AA| + \frac{1}{4}|AB'_{\alpha_1}| + \frac{1}{4}|AB'_{\alpha_2}| + \frac{1}{3}|AB'_{\alpha_3}| + |S| + |y| - 1 \\ &\leq \frac{1}{2}|AA| + \frac{1}{3}|AB'_{\alpha}| + |S| + |y| - 1, \end{aligned} \quad (3)$$

where the term  $-1$  follows since  $x$  belongs to  $C$  but not  $C_x$ .

**Step 4. An upper bound for  $|C_x| - |C|$**

We recall that the sets  $AA$ ,  $AB'_{\alpha}$ ,  $AB'_{\beta}$ ,  $S$  and  $B^*$  are disjoint and none of these contain  $x$  or  $y$ . If we denote the other vertices of  $G$  by  $S' = V \setminus (AA \cup AB' \cup S \cup B^* \cup \{x, y\})$ , then we know that

$$\begin{aligned} n - 1 &= n_x = |AA| + |AB'_{\alpha}| + |AB'_{\beta}| + |S| + |B^*| + |y| + |S'| \\ &\geq |AA| + |AB'_{\alpha}| + 3|S| + |B^*| + |y| + |S'|. \end{aligned} \quad (4)$$

Now, the next inequality follows from Equations (3) and (4):

$$\begin{aligned} |C_x| - |C| &\leq \frac{1}{2}|AA| + \frac{1}{3}|AB'_{\alpha}| + |S| + |y| - 1 \\ &= \frac{1}{3} \left( |AA| + |AB'_{\alpha}| + 3|S| + |y| + |B^*| + |S'| \right) + \frac{1}{6}|AA| + \frac{2}{3}|y| - \frac{1}{3}|B^*| - \frac{1}{3}|S'| - 1 \\ &\leq \frac{1}{3}(n - 1) + \frac{1}{6}|AA| - \frac{1}{3}|B^*| + \frac{2}{3}|y| - \frac{1}{3}|S'| - 1 \\ &= \frac{1}{3} \left( n - 4 + \frac{|AA|}{2} - |B^*| + 2|y| - |S'| \right). \end{aligned}$$

Next, we denote  $S'' = AB' \cup S$ . Then, the first equality of Equation 4 can be written as follows:

$$n - 1 = |AA| + |B^*| + |S'| + |S''| + |y| \quad (5)$$

**Step 4.1. Case  $|AA| \geq 2$  and  $|S'| + |S''| + |y| \leq n - \lfloor \log_2 n \rfloor - 2$ :**

We now make two assumptions. We first assume that  $|AA| \geq 2$ , i.e. there is at least one  $AA$ -pair. Secondly, we assume that  $|S'| + |S''| + |y| \leq n - \lfloor \log_2 n \rfloor - 2$ . Steps 4.2 and 4.3 handle the cases when these assumptions are not valid.

Next, we show that

$$|AA| \leq n - 1 - |S'| - |S''| - |y| - \lfloor \log_2(n - |S'| - |S''| - |y| - \lfloor \log_2 n \rfloor) \rfloor.$$

Let us assume to the contrary, that

$$|AA| \geq n - |S'| - |S''| - |y| - \lfloor \log_2 (n - |S'| - |S''| - |y| - \lfloor \log_2 n \rfloor) \rfloor.$$

Now, we can use Equation (2) and write

$$|AA| + |B^*| + |S'| + |S''| + |y| \geq |AA| + \lfloor \log_2 |AA| \rfloor + |S'| + |S''| + |y| \geq n$$

which is a contradiction according to Equation (5). We still notice that the right hand side of the inequality

$$\frac{|AA|}{2} - |B^*| \leq \frac{|AA|}{2} - \lfloor \log_2 |AA| \rfloor$$

is increasing when  $|AA|$  increases since  $|AA|$  has to be even and  $|AA| \geq 2$ . Now,

$$\begin{aligned} |C_x| - |C| &\leq \frac{1}{3} \left( n - 4 + \frac{|AA|}{2} - |B^*| + 2|y| - |S'| \right) \\ &\leq \frac{1}{3} \left( n - 4 + 2|y| - |S'| + \frac{|AA|}{2} - \lfloor \log_2 |AA| \rfloor \right) \\ &\leq \frac{1}{3} \left( n - 4 + 2|y| - |S'| \right. \\ &\quad \left. + \frac{1}{2} \left( n - 1 - |S'| - |S''| - |y| - \lfloor \log_2 (n - |S'| - |S''| - |y| - \lfloor \log_2 n \rfloor) \rfloor \right) \right. \\ &\quad \left. - \left\lfloor \log_2 \left( n - 1 - |S'| - |S''| - |y| - \lfloor \log_2 (n - |S'| - |S''| - |y| - \lfloor \log_2 n \rfloor) \rfloor \right) \right\rfloor \right) \\ &= \frac{n - 3 - |S'| - \frac{1}{3}|S''| + |y|}{2} - \frac{1}{6} \left\lfloor \log_2 \left( n - |S'| - |S''| - |y| - \lfloor \log_2 n \rfloor \right) \right\rfloor \\ &\quad - \frac{1}{3} \left\lfloor \log_2 \left( n - 1 - |S'| - |S''| - |y| - \lfloor \log_2 (n - |S'| - |S''| - |y| - \lfloor \log_2 n \rfloor) \rfloor \right) \right\rfloor \\ &:= f(n, |S'|, |S''|, |y|), \end{aligned}$$

**Step 4.1.1. The analysis of  $f(n, |S'|, |S''|, |y|)$ :** Now, we show that

$$\gamma(G_x) - \gamma(G) \leq \max(f(n, 0, 0, 0), f(n, 1, 0, 1), f(n, 0, 3, 1) - 1, f(n, 0, 4, 1)).$$

First,  $|S'|$ ,  $|S''|$  and  $|y|$  are non-negative integers. If  $y$  is constant, then  $f$  is decreasing when  $|S'|$  or  $|S''|$  increases. Therefore,  $f(n, 0, 0, 0)$  is an upper bound for  $|C_x| - |C|$  if  $|y| = 0$ . Otherwise,  $|y| = 1$ . Again, if  $|S'| \geq 1$ , then  $f(n, 1, 0, 1)$  is an upper bound for  $|C_x| - |C|$ . Assume then that  $|S'| = 0$  and  $|y| = 1$ . Now  $x$  must have a codeword neighbour  $c \in C$  in  $G$ , since  $I(y, C) = \{x\}$ . Moreover, since no codeword belongs to  $S \cup AA$ , then  $c$  has to belong to  $AB' = S'' \setminus S$ . Furthermore, at least either  $c$  or  $x$  must have another neighbour  $c' \in C$  or else  $I(x, C) = \{x, c\} = I(c, C)$ . If  $c'$  is a neighbour of  $x$ , then there are at least two  $AB'$ -pairs and then  $|S''| \geq |AB'| \geq 4$ . If there are at least 4 vertices in  $S'$ , then  $|C_x| - |C| \leq f(n, 0, 4, 1)$ . Otherwise, there are at most 3 vertices in  $S'$ , and  $c'$  is a neighbour of  $c$ . Then,  $c \in A$  and  $c' \in B$  must be the only  $AB'$ -pair. Now  $c$  or  $c'$  has to be a neighbour  $v$  in addition to  $x$ ,  $c$  and  $c'$ , since  $G_x$  is twin-free. On the other hand,  $v$  does not belong to any pair. Indeed,  $c$  or  $c'$  would separate this

pair since  $c$  and  $c'$  do not belong to  $AB^*$ . Then,  $v$  has to belong to  $S \subseteq S''$  (or  $S'$  which is a contradiction according to our assumption) and  $v$  is the only neighbour for  $c$  and  $c'$  in addition to  $x$ ,  $c$  and  $c'$ . Now,  $C_x$  contains all the vertices  $v$ ,  $c$  and  $c'$  following from Step 3.3. However,  $C_x \setminus \{c\}$  or  $C_x \setminus \{c'\}$  is also an identifying code depending on which are the neighbours of  $v$ . Thus,  $\gamma(G_x) - \gamma(G) \leq f(n, 0, 3, 1) - 1$  in this case.

Now,  $f(n, 1, 0, 1) \geq f(n, 0, 0, 0) \geq f(n, 0, 3, 1) - 1$  for all  $n \geq 7$ . Moreover,  $f(n, 0, 4, 1) \leq f(n, 1, 0, 1) + \frac{1}{3}$  for all  $n \geq 10$ . However,  $f(n, 1, 0, 1)$  is a multiple of  $\frac{1}{2}$  except if  $n = 2^t + t + 2$ , where  $t$  is an integer. Furthermore,  $f(n, 1, 0, 1) = f(n, 0, 4, 1)$ , if  $n = 2^t + t + 2$ . Therefore,

$$\gamma(G_x) - \gamma(G) \leq \lfloor f(n, 1, 0, 1) \rfloor = \left\lfloor f(n, 1, 0, 1) + \frac{1}{3} \right\rfloor \geq \lfloor f(n, 0, 4, 1) \rfloor$$

for all  $n \geq 10$ , since  $\gamma(G_x) - \gamma(G)$  is an integer. If  $n \leq 9$ , then the case  $f(n, 0, 4, 1)$  is impossible, since  $|S'| + |S''| + |y|$  would be  $5 > n - \lfloor \log_2 n \rfloor - 2$ . All in all,  $\gamma(G_x) - \gamma(G) \leq \lfloor f(n, 1, 0, 1) \rfloor$  if  $|S'| + |S''| + |y| \leq n - \lfloor \log_2 n \rfloor - 2$  and  $AA \neq \emptyset$ .

**Step 4.2. Case  $|AA| \geq 4$  and  $|S'| + |S''| + |y| \geq n - \lfloor \log_2 n \rfloor - 1$ :**

If  $|S'| + |S''| + |y| \geq n - \lfloor \log_2 n \rfloor - 1$ , then  $|AA| \leq \lfloor \log_2 n \rfloor - |B^*|$  according to Equation (5). If  $|AA| \geq 4$ , then  $|B^*| \geq \lfloor \log_2 |AA| \rfloor \geq 2$  and then  $|C_x| - |C| \leq \frac{n-4}{3} + \frac{\lfloor \log_2 n \rfloor - 2}{6} \leq f(n, 1, 0, 1)$ , when  $n \geq 7$ .

**Step 4.3. Case  $|AA| \leq 2$ :** If  $|AA| \leq 2$ , then  $\frac{|AA|}{2} - |B^*| \leq 0$  and  $|C_x| - |C| \leq \frac{n-2}{3} < f(n, 1, 0, 1)$  when  $n \geq 11$ . In fact,  $|C_x| - |C|$  is an integer, then  $|C_x| - |C| \leq \lfloor \frac{n-2}{3} \rfloor \leq \lfloor f(n, 1, 0, 1) \rfloor$ , when  $n = 7$  or  $n \geq 9$ .

Finally, we have proved the claim which is equivalent with  $\gamma(G_x) - \gamma(G) \leq \lfloor f(n, 1, 0, 1) \rfloor$  except if  $n = 8$  and  $|AA| \leq 2$ . The case  $n = 8$  follows from Lemma 16.  $\square$

**Lemma 16.** *If  $G$  is a bipartite graph of order 8, then*

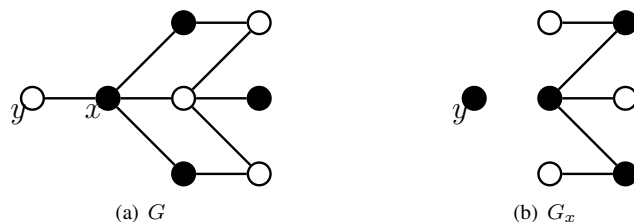
$$\gamma(G_x) - \gamma(G) \leq 1.$$

**Proof:** We use the same markings as the proof of Theorem 15. Assume the contrary that the claim is invalid, i.e. there is a graph  $G$  of order 8 such that  $\gamma(G_x) - \gamma(G) \geq 2$ . This is possible only if  $|C| = 4$  and  $|C_x| = \gamma(G_x) = 6$  according to Lemma 8 and the proof of Theorem 11. Now,  $G$  has to contain at least three pairs or two pairs and the vertex  $y$ . These pairs and the vertex  $y$  should provide three codewords to  $C_x$ .

Assume first that there is an  $AA$ -pair of  $a'$  and  $a''$ . Without loss of generality, we can assume that  $\{a', x\} \in E$  and  $\{a'', x\} \notin E$ . Now,  $a'$  and  $a''$  must have a common codeword neighbour  $b^* \in B^*$ . This  $AA$ -pair can provide at most one codeword to  $C_x$  in addition to  $C \setminus \{x\}$ . Moreover, at least one codeword must be provided to  $C_x$  by another pair, i.e. by  $AB' \cup S$ . However, vertices of  $AB' \cup S$  can provide at most  $\frac{1}{3}|AB' \cup S|$  codewords to  $C_x$ . Thus, there must be an  $AB'$ -pair and a vertex  $y$ . Moreover, the  $AB'$ -pair has to be an  $AB'_\beta$ -pair since a single  $AB'_\alpha$ -pair does not provide any codeword to  $C_x$ . Thus, the eightest vertex belongs to  $S$  and it is denoted by  $s$ . Now,  $\{s, y, a', a''\} \cap C = \emptyset$ . Moreover,  $b^*$  or  $a''$  can not have  $a$  or  $b$  in its neighbourhood or else  $a$  and  $b$  would not be an  $AB'_\beta$ -pair. Therefore,  $I(a'') = \{b^*\} = I(b^*)$  which is a contradiction.

Assume second that there is not an  $AA$ -pair. Now, at least two codewords must be provided to  $C_x$  by  $AB'$ -pairs and this is possible only if there are three  $AB'_{\alpha 3}$ -pairs or two  $AB'_\beta$ -pairs and 2 vertices of  $|S|$ .





**Fig. 5:** The bipartite graph of order 8 with three  $AB'_{\alpha_3}$ -pairs. It is possible that there are more edges between  $y$  and the other vertices. The black dots represent an identifying code.

The third codeword must be provided by  $y$ . First, assume that  $s \in S$  separates an  $AB'_\beta$ -pair of  $a$  and  $b$  and  $s'$  separates another  $AB'_\beta$ -pair of  $a'$  and  $b'$ . Moreover, there can not be edges  $\{s', a\}$ ,  $\{s', b\}$ ,  $\{s, a'\}$  and  $\{s, b'\}$  since  $s$  or  $s'$  can not separate both of the pairs. Thus,  $I(a) \cup I(b) \cup I(s) \subseteq \{a, b, x\}$ . Moreover, since none of  $I(a)$ ,  $I(b)$  and  $I(s)$  can be empty or  $\{x\} = I(y)$ , both  $a$  and  $b$  in addition to  $x$  have to be codewords. Similarly,  $a'$  and  $b'$  are codewords and  $|C| = 5$ , which is a contradiction.

Finally, let us assume that  $G$  contains three  $AB'_{\alpha_3}$ -pairs and  $x$  and  $y$ . Now,  $x$  must have at least two codeword neighbours, since  $I(x) \neq I(y) = \{x\}$  and  $|I(v)| \leq 2$  for all  $v \in AB'_{\alpha_3}$ . Then,  $G$  is uniquely defined apart from the edges between  $y$  and vertices of  $AB'$ . See Figure 5. Now,  $\gamma(G_x) \leq 4$  regardless of the edges from  $y$ .  $\square$

**Corollary 17.** Let  $G = (V, E)$  be a twin-free bipartite graph with  $n$  vertices. Then,

$$\gamma(G_x) - \gamma(G) \leq \left\lfloor \frac{n - \log_2(n - \log_2 n)}{2} \right\rfloor - 1$$

and the inequality is tight for all  $n = 3, 4, 5, \dots$

**Proof:** The claim follows by Theorems 1, 4, 10, 11 and 15 and Remarks 5 and 6. Notice that the values of the equations in Theorems 4 and 15 and this corollary are same for all integers  $n \geq 3$  when the equations are defined.  $\square$

## 4 Acknowledgements

I would like to thank anonymous referees for numerous linguistic and mathematical comments. Also I am grateful to referees for suggestions to extend this paper to bipartite graphs and to cases where  $|S|$  is larger than 1.

## References

- [1] Bertrand, N.: Codes identifiants et codes localisateurs-dominateurs sur certains graphes, Mémoire de stage de maîtrise, ENST, Paris, France, 28 pages (2001).
- [2] Charon, I.; Honkala, I.; Hudry, O.; Lobstein A.: Minimum sizes of identifying codes in graphs differing by one vertex, *Cryptography and Communications*, Vol. 5, 119–136 (2013).
- [3] Charon, I.; Honkala, I.; Hudry, O.; Lobstein A.: Structural Properties of Twin-Free Graphs, *Electronic Journal of Combinatorics*, Vol. 14(1), R16 (2007).
- [4] Charon, I.; Hudry, O.; Lobstein A.: Extremal cardinalities for identifying and locating-dominating codes in graphs, *Discrete Mathematics*, Vol. 307, 356–366 (2007).
- [5] Cockayne, E. J.; Dawes, R. M.; Hedetniemi, S. T.: Total domination in graphs, *Networks*, Vol. 10, 211–219 (1980).
- [6] Foucaud, F.; Guerrini, E.; Kovse, M.; Naserasr, R.; Parreau, A.; Valicov, P.: Extremal graphs for the identifying code problem, *European Journal of Combinatorics*, Vol. 32, 628–638 (2011).
- [7] Gravier, S.; Moncel, J.: On graphs having a  $V \setminus \{x\}$  set as an identifying code, *Discrete Mathematics*, Vol. 307, 432–434 (2007).
- [8] Haynes, T. W.; Hedetniemi, S. T., Slater, P. J.: *Fundamentals of Domination in graphs*, Marcel Dekker, Inc., New York (1998).
- [9] Karpovsky, M. G.; Chakrabarty, K.; Levitin, L. B.: On a new class of codes for identifying vertices in graphs, *IEEE Transactions on Information Theory*, Vol. IT-44, 599–611 (1998).
- [10] Lobstein, A.: Watching systems, identifying, locating-dominating and discriminating codes in graphs, *Bibliography*.  
<http://www.infres.enst.fr/~lobstein/bibLOCDOMetID.html>
- [11] Slater, P. J.: Domination and location in acyclic graphs, *Networks*, Vol. 17, 55–64 (1987).
- [12] Slater, P. J.: Dominating and reference sets in a graph, *Journal of Mathematical and Physical Sciences*, Vol. 22, 445–455 (1988).

