Extremal Values for Identification, Domination and Maximum Cliques in Twin-Free Graphs

Irène Charon
GET - Télécom Paris & CNRS - LTCI UMR 5141
46, rue Barrault, 75634 Paris Cedex 13 - France
irene.charon@enst.fr

Olivier Hudry
GET - Télécom Paris & CNRS - LTCI UMR 5141
46, rue Barrault, 75634 Paris Cedex 13 - France
olivier.hudry@enst.fr

Antoine Lobstein
CNRS - LTCI UMR 5141 & GET - Télécom Paris
46, rue Barrault, 75634 Paris Cedex 13 - France
antoine.lobstein@enst.fr

Abstract

Consider a connected undirected graph $G = (V, E)$ and an integer $r \geq 1$; for any vertex $v \in V$, let $B_r(v)$ denote the ball of radius r centered at v, i.e., the set of all vertices linked to v by a path of at most r edges. If for all vertices $v \in V$, the sets $B_r(v)$ are different, then we say that G is r-twin-free.

Studies have been made, e.g., on the number of edges or the minimum degree in one-twin-free graphs. We extend these investigations and in particular we determine the exact size of the largest clique in a connected r-twin-free graph.

Key Words: Graph Theory, Identifying Codes, Twins, Clique, Domination

1 Introduction

1.1 Definitions and notation

Given a connected, undirected, finite graph $G = (V, E)$ and an integer $r \geq 1$, we define $B_r(v)$, the *ball* of radius r centered at $v \in V$, by

$$B_r(v) = \{x \in V : d(x, v) \leq r\},$$

where $d(x, v)$ denotes the number of edges in any shortest path between v and x.

Whenever $d(x, v) \leq r$, we say that x and v *r-cover* each other (or simply *cover* if there is no ambiguity). A set $X \subseteq V$ covers a set $Y \subseteq V$ if every vertex in Y is covered by at least one vertex in X.

Two vertices $v_1, v_2 \in V$ such that $B_r(v_1) = B_r(v_2)$ are called *r-twins* or *twins*. If G has no r-twins, that is, if

$$\forall\, v_1, v_2 \in V \text{ with } v_1 \neq v_2,\ B_r(v_1) \neq B_r(v_2), \tag{1}$$

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

A graph with one vertex is trivially twin-free, and generally we consider graphs with at least two vertices.

Twin-free graphs are of interest because they are strongly connected with *identifying codes* [18], which we now define.

A *code* C is a nonempty set of vertices, and its elements are called *codewords*. For each vertex $v \in V$, we denote by

$$K_{C,r}(v) = C \cap B_r(v)$$

the set of codewords which r-cover v. Two vertices v_1 and v_2 with $K_{C,r}(v_1) \neq K_{C,r}(v_2)$ are said to be *r-separated*, or *separated*, by code C.

A code C is called *r-identifying*, or *identifying*, if the sets $K_{C,r}(v), v \in V$, are all nonempty and distinct [18]. In other words, all vertices must be covered and pairwise separated by C.

Remark 1. For given $G = (V, E)$ and integer r, the graph G admits at least one r-identifying code if and only if it is r-twin-free. Indeed, if for all $v_1, v_2 \in V$, $B_r(v_1)$ and $B_r(v_2)$ are different, then $C = V$ is r-identifying. Conversely, if for some $v_1, v_2 \in V$, $B_r(v_1) = B_r(v_2)$, then for any code $C \subseteq V$, we have $K_{C,r}(v_1) = K_{C,r}(v_2)$. This is why r-twin-free graphs are also called *r-identifiable*. For instance, there is no r-identifying code in a complete graph (or clique) with at least two vertices.

If G is r-twin-free, we denote by $c_r(G)$ the size of the smallest r-identifying code in G.

We define the *minimum distance* of a code C as follows:

$$d_{\min}(C) = \min\{d(z_i, z_j) : z_i \in C, z_j \in C, z_i \neq z_j\}.$$

Remark 2. If C is r-identifying, then $d_{\min}(C) \leq r + 1$ (otherwise, a codeword c and a vertex adjacent to c could not be r-separated by C).

Remark 3. If G is not connected, we simply consider each of its connected components, and apply the above definitions.

In the following, n will denote the number of vertices of G. For any integer $q > 0$, P_q will denote the path on q vertices, and the length of P_q will be equal to $q - 1$, its number of edges. Moreover, if v_1, v_2, ..., v_q denote the vertices in P_q, we shall assume that these vertices are numbered in such a way that the edges in P_q are $\{v_i, v_{i+1}\}$ for $1 \leq i < q$. The cycle of length q (with q vertices and q edges), consisting of P_q to which we add the edge $\{v_q, v_1\}$, will be denoted by C_q. The following graph will be used several times in the sequel: we shall call it the *star*, and it consists of n vertices $0, 1, \ldots, n-1$, and $n - 1$ edges $\{0, i\}$, $1 \leq i \leq n - 1$.

1.2 Motivations

The motivations for identifying codes come, for instance, from fault diagnosis in multiprocessor systems. Such a system can be modeled as a graph where vertices are processors and edges are links between processors. Assume that at most one of the processors is malfunctioning and we wish to test the system and locate the faulty processor. For this purpose, some processors (constituting the code) will be selected and assigned the task of testing their neighbourhoods (i.e., the vertices at distance at most r). Whenever a selected processor (i.e., a codeword) detects a fault, it sends an alarm signal, saying that one element in its neighbourhood is malfunctioning, and we require that we can uniquely tell the location of the malfunctioning processor based only on the information which codewords gave the alarm.

Identifying codes were introduced in [18], and they constitute now a topic of their own, studied in a large number of various papers, investigating particular graphs or families of graphs (such as certain infinite regular grids, trees, chains, cycles, or the k-cube), dealing with complexity issues, or using heuristics such as the noising methods for the construction of small codes. See, e.g., [1], [2], [3], [5], [6], [16], [22], and references therein, or [23].

Therefore, it is quite natural to study some of the parameters of twin-free graphs, since these graphs admit identifying codes. This is what we shall do in the present paper as well as in the forthcoming paper [9].

1.3 Plan of the study

We intend to investigate the extremal values that some parameters, often studied in graph theory, can reach in connected twin-free graphs. More precisely, for a parameter p (such as the number of vertices, the number of edges, the minimum degree, ...), we fix r and search for the smallest value, $f_r(p)$, that this parameter can reach in G, or we fix r and n and search for the smallest and largest values, $f_{r,n}(p)$ and $F_{r,n}(p)$, respectively, that this parameter can reach in G:

$$f_r(p) = \min\{p : G \in \mathcal{G}_r\},$$

$$\text{where } \mathcal{G}_r = \{G : G \text{ connected, } r\text{-twin-free with at least two vertices}\};$$

$$f_{r,n}(p) = \min\{p : G \in \mathcal{G}_{r,n}\} \quad \text{and} \quad F_{r,n}(p) = \max\{p : G \in \mathcal{G}_{r,n}\},$$

$$\text{where } \mathcal{G}_{r,n} = \{G : G \text{ connected, } r\text{-twin-free with } n \geq 2 \text{ vertices}\}.$$

The function $F_r(p) = \max\{p : G \in \mathcal{G}_r\}$ would present much less interest, since, for the parameters that we shall deal with, F_r is not bounded by above.

In this paper, we are interested in the following four parameters:
- number of vertices, n,
- minimum cardinality, c_r, of an r-identifying code,
- r-domination number, γ_r,
- maximum size of a clique, ω,

and study, for each of them, the functions f_r, $f_{r,n}$ and $F_{r,n}$: each of the Sections 2–5 deals with one parameter; at the beginning of each section, for comparison, the extremal values of the parameter are given for connected graphs — when relevant. Section 6 recapitulates what we obtained in three tables ($r = 1$, $r = 2$, $r \geq 3$). One of the main results is the exact value of $F_{r,n}(\omega)$.

In [9], we intend to run the same study for the following five parameters: number of edges ε, minimum degree $\Delta_{\min}$, maximum degree $\Delta_{\max}$, diameter δ, maximum size of a stable α.

We shall see that some of these parameters are connected in some way. For instance, we shall derive results on the number of edges using knowledge on the maximum clique size and the maximum degree.

So far, only the parameters n (cf. Sec. 2), c_r (cf. Sec. 3), ε [20, Sec. 4.1.2] and, to a lesser extent, $\Delta_{\min}$ [19] [12] and γ_r [17], had been investigated, to our knowledge, within the framework of twin-free graphs.

Note that, for parameters not depending on r, the monotonicity of the functions $f_{r,n}$ and $F_{r,n}$, with respect to r, can be proved using the following lemma, in which the condition $n \geq 2r+3$ is not restrictive, since we shall see

(Prop. 3) that any connected r-twin-free graph has at least $2r+1$ vertices, unless $n=1$.

Lemma 1 *Let $r \geq 1$ and $n \geq 2(r+1)+1$ be two integers. If G is a connected $(r+1)$-twin-free graph with n vertices, then G is also r-twin-free.*

Proof. Assume that $G=(V,E)$ is $(r+1)$-twin-free and not r-twin-free. Then there exist two vertices $x, y \in V$ such that $B_r(x) = B_r(y)$, and one vertex $z \in V$ which $(r+1)$-covers x and not y (or the other way round). Necessarily, $d(z,x) = r+1$; consider a vertex $t \in V$ at distance one from z and r from x. Then t is also within distance r from y, therefore $d(y,z) \leq r+1$, a contradiction. △

Corollary 2 *Let $r \geq 1$ and $n \geq 2r+3$ be two integers, and let $p = \omega$, ε, $\Delta_{\min}$, $\Delta_{\max}$, δ, or α. Then $F_{r,n}(p) \geq F_{r+1,n}(p)$ and $f_{r,n}(p) \leq f_{r+1,n}(p)$.*

Proof. Since, by Lemma 1, the family $\mathcal{G}_{r,n}$ of connected r-twin-free graphs with n vertices includes $\mathcal{G}_{r+1,n}$, we find a better optimum when searching for it in $\mathcal{G}_{r,n}$ than in $\mathcal{G}_{r+1,n}$. △

Corollary 2 holds also for any other parameter which does not depend on r. On the other hand, we cannot use it with parameters such as c_r or γ_r, since, when studying, e.g., $f_{r,n}$ with respect to domination, we chose to restrict ourselves to the case when the *same* r parametrizes both f and γ, i.e., we consider only $f_{r,n}(\gamma_r)$.

2 The number of vertices, n

This parameter has already been studied, and we have the following result.

Proposition 3 *[7],[20],[4] Let $r \geq 1$ and G be any connected r-twin-free graph with at least two vertices. Then we have: $n \geq 2r+1$. Moreover, P_{2r+1} is the only connected r-twin-free graph with $2r+1$ vertices.* △

Corollary 4 *For all $r \geq 1$, we have: $f_r(n) = 2r+1$.* △

Observe that obviously, $f_{r,n}(n) = F_{r,n}(n) = n$ for all $r \geq 1$ and $n \geq 2r+1$.

3 The minimum size, c_r, of an r-identifying code

This parameter has already been studied, and we recapitulate here the results obtained so far. First, we give an upper bound on $F_{r,n}(c_r)$.

Proposition 5 *[8],[13] Let $r \geq 1$ and $n \geq 2r + 1$ be two integers. Let G be any connected r-twin-free graph with n vertices. Then we have: $c_r(G) \leq n - 1$.* △

When $r = 1$, this bound is sharp, as we now show.

Proposition 6 *[8] For all $n \geq 3$, there exists a connected one-twin-free graph G with n vertices such that $c_1(G) = n - 1$.* △

One example is the star, defined in the Introduction.

Corollary 7 *For all $n \geq 3$, we have: $F_{1,n}(c_1) = n - 1$.* △

For $r > 1$, the following proposition shows that the upper bound can also be attained, if n is large enough with respect to r.

Proposition 8 *[8] Let $r \geq 2$. If $n \geq 3r^2$ is even, or if $n \geq 3r^2 + 1$ is odd, then there exists a connected r-twin-free graph G with n vertices such that $c_r(G) = n - 1$.* △

For all $r \geq 1$, it is quite straightforward to observe that we also have:

$$c_r(P_{2r+1}) = 2r = n - 1. \tag{2}$$

Corollary 9 *If $n = 2r + 1$, or if $n \geq 3r^2$ is even, or if $n \geq 3r^2 + 1$ is odd, then we have: $F_{r,n}(c_r) = n - 1$.* △

Next, we have an easy lower bound on $f_{r,n}(c_r)$.

Proposition 10 *[18] Let $r \geq 1$ and $n \geq 2r + 1$ be two integers. Let G be any connected r-twin-free graph with n vertices. Then we have: $c_r(G) \geq \lceil \log_2(n + 1) \rceil$.* △

Next proposition shows that this bound can be reached with conditions on n.

Proposition 11 *[8] Let $r \geq 1$. If $n \geq 2^{2r}$, then there exists a connected r-twin-free graph G with n vertices such that $c_r(G) = \lceil \log_2(n + 1) \rceil$.* △

A similar result was established independently in [11]. The case $r = 1$ was previously solved in [18]. When $r = 1$, Proposition 11 is true for all $n \geq 3$; for $r = 2$, it holds if and only if $n \geq 6$.

Corollary 12 *If $r = 1$ and $n \geq 3$, or $r = 2$ and $n \geq 6$, or $r \geq 3$ and $n \geq 2^{2r}$, then we have: $f_{r,n}(c_r) = \lceil \log_2(n + 1) \rceil$.* △

We feel that significant improvements on the bound $n \geq 2^{2r}$ would involve complex arguments. We now turn to $f_r(c_r)$, for which no previous results are known, and first give an easy lower bound.

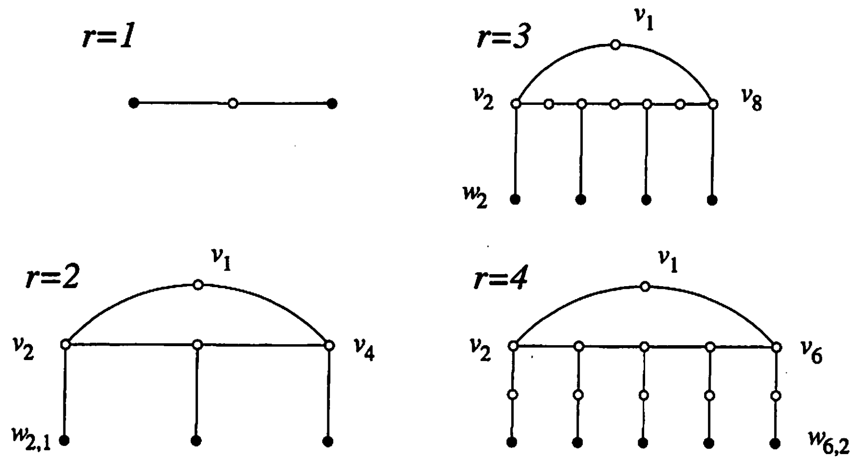

Figure 1: Different twin-free graphs constructed in the proof of Theorem 14. Codewords are in black.

Theorem 13 *We have:* $f_1(c_1) \geq 2$ *and, for all* $r \geq 2$*, we have:* $f_r(c_r) \geq \lceil \log_2(2r+4) \rceil$.

Proof. Let G be any connected r-twin-free graph with n vertices. By Proposition 10, we have $c_r(G) \geq \lceil \log_2(n+1) \rceil$, and by Proposition 3, we have $n \geq 2r+1$. Hence, for all $r \geq 1$, $f_r(c_r) \geq \lceil \log_2(2r+2) \rceil$, which gives the result for $r = 1$.

Now if $n = 2r+1$, then, by Proposition 3, $G = P_{2r+1}$, and by (2), $c_r(G) = 2r$. Since $2r > \lceil \log_2(2r+2) \rceil$ when $r > 1$, we have: $f_r(c_r) \geq \lceil \log_2(2r+3) \rceil$. The final result simply comes from the fact that $2r+3$ is odd. △

In the following theorem, a construction gives $r+1$ as an upper bound.

Theorem 14 *For all* $r \geq 1$*, we have:* $f_r(c_r) \leq r+1$.

Sketch of proof. Besides the particular case $r = 1$, there are two different constructions, depending on the parity of r.

(a) $r = 1$; we set $G = P_3$, and the two ends of the path constitute a one-identifying code.

(b) r odd, $r \geq 3$; we construct the graph $G = (V, E)$, where

$$V = \{v_i : 1 \leq i \leq 2r+2\} \cup \{w_j : j \text{ even}, 2 \leq j \leq 2r+2\},$$

$$E = \{\{v_i, v_{i+1}\} : 1 \leq i \leq 2r+2\} \cup \{\{v_j, w_j\} : j \text{ even}, 2 \leq j \leq 2r+2\},$$

with $v_{2r+3} = v_1$. The $r+1$ vertices w_j will constitute the code C (see Figure 1). In other words, we consider the cycle C_{2r+2}, and we link every second vertex to a codeword.

(c) $r = 2t, t \geq 1$; we construct the graph $G = (V, E)$, where

$$V = \{v_i : 1 \leq i \leq r+2\} \cup \{w_{i,j} : 2 \leq i \leq r+2, 1 \leq j \leq t\},$$

$$E = \{\{v_i, v_{i+1}\} : 1 \leq i \leq r+2\} \cup$$

$$\{\{v_i, w_{i,1}\}, \{w_{i,j}, w_{i,j+1}\} : 2 \leq i \leq r+2, 1 \leq j \leq t-1\},$$

with $v_{r+3} = v_1$. The $r+1$ vertices $w_{i,t}$ will constitute the code C (see Figure 1). In other words, we consider the cycle C_{r+2}, and we link every vertex but one to a path with t vertices, the end of which is a codeword.

In Cases (b) and (c), we leave it to the reader to check that C is indeed r-identifying. △

Unfortunately, the gap remains important between lower and upper bounds. For values of r up to 4, we can prove that the exact value for $f_r(c_r)$ is the upper bound.

Theorem 15 *For $r \in \{1, 2, 3, 4\}$, we have: $f_r(c_r) = r+1$.*

Proof. If $r = 1, 2$ or 3, then the lower and upper bounds given by Theorems 13 and 14 coincide.

If $r = 4$, all we have to show is that $f_4(c_4) \neq 4$. Assume on the contrary that G is a connected graph with n vertices admitting a 4-identifying code C of size four. By Propositions 10 and 3, and the fact that, by (2), $c_4(P_9) = 8$, we know that $10 \leq n \leq 15$. By Remark 2, we have: $1 \leq d_{\min}(C) \leq 5$.

(i) $d_{\min}(C) = 1$; there are at least two codewords, z_1 and z_2, with $d(z_1, z_2) = 1$. Then there is a codeword z_3 4-covering exactly one of z_1 and z_2, say z_2. The only possibility is given in Figure 2(i). Then, for $i \in \{1, 2, 3\}$, we have: $B_4(v_i) \cap \{z_1, z_2, z_3\} = \{z_1, z_2, z_3\}$, and we see that with only one additional codeword we are unable to separate the three vertices v_i.

(ii) $d_{\min}(C) = 2$; there are at least two codewords, z_1 and z_2, with $d(z_1, z_2) = 2$. Then there is a codeword z_3 4-covering exactly one of z_1 and z_2, say z_2. The only two possibilities are given in Figure 2(ii). In both cases, similarly to the previous case, there is no way to separate z_2, v_2 and v_3 with only one more codeword.

(iii) $d_{\min}(C) = 3$; there are at least two codewords, z_1 and z_2, with $d(z_1, z_2) = 3$. Then there is a codeword z_3 4-covering exactly one of z_1 and z_2, say z_2. The only three possibilities are given in Figure 2(iii). In the first and third cases, it will be impossible to separate z_2, v_2 and v_3 with only one more codeword. In the middle case, the fourth codeword, z_4, must be used to separate z_2 and v_3: for instance z_4 covers z_2 and not v_3,

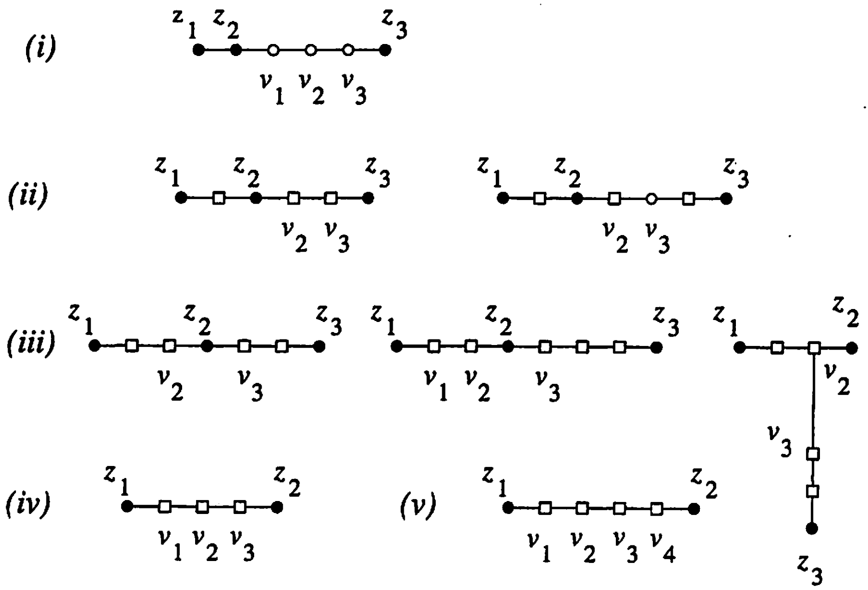

Figure 2: The five cases in Theorem 15. Not all edges are drawn. Black vertices are codewords, white circles are codewords or not, squares are noncodewords.

so that: $K_{C,4}(z_2) = \{z_1, z_2, z_3, z_4\}$ and $K_{C,4}(v_3) = \{z_1, z_2, z_3\}$. Now, z_1, v_1 and v_2 are covered by z_1, z_2, and possibly by z_3, z_4. Therefore the possible sets $K_{C,4}(.)$ for these three vertices are among $\{z_1, z_2\}$, $\{z_1, z_2, z_3\}$, $\{z_1, z_2, z_4\}$, $\{z_1, z_2, z_3, z_4\}$. But $\{z_1, z_2, z_3, z_4\}$ and $\{z_1, z_2, z_3\}$ have already been "used", therefore it will be impossible to separate z_1, v_1 and v_2.

(iv) $d_{\min}(C) = 4$; there are at least two codewords, z_1 and z_2, with $d(z_1, z_2) = 4$, see Figure 2(iv). Then z_1, z_2, v_1, v_2 and v_3 are covered by z_1, z_2, and two more codewords are not sufficient to separate these five vertices.

(v) $d_{\min}(C) = 5$; there are at least two codewords, z_1 and z_2, with $d(z_1, z_2) = 5$, see Figure 2(v). We see that v_1, v_2, v_3 and v_4 are covered by z_1, z_2. Therefore, the sets $K_{C,4}(.)$ will be, for these four vertices: $\{z_1, z_2\}$, $\{z_1, z_2, z_3\}$, $\{z_1, z_2, z_4\}$, $\{z_1, z_2, z_3, z_4\}$. So all the sets containing z_1 and z_2 have been "used", which implies that there is no vertex adjacent to v_2 or v_3 other than those already in Figure 2(v) — otherwise, this new vertex, which would also be 4-covered by z_1 and z_2, could not be separated from v_1, v_2, v_3, v_4. Now a new codeword, say z_3, must 4-cover exactly one of v_2 and v_3, say v_2: $d(z_3, v_2) \leq 4$ and $d(z_3, v_3) > 4$. Since any shortest path linking z_3 and v_2 must go through v_1, we have $d(z_3, z_1) = d(z_3, v_2) \leq 4$, which contradicts $d_{\min}(C) = 5$. △

Open Problem. Find the exact value of $f_r(c_r)$ for $r \geq 5$.

4 The r-domination number, γ_r

The r-domination number in a graph $G = (V, E)$ is the smallest size of a code $C \subseteq V$ which r-covers all the vertices of the graph. Such a code is called an *r-covering*, or a *dominating set* in the case $r = 1$ and more generally a *distance-r dominating set.* We denote by $\gamma_r(G)$ the size of the smallest r-covering in G. Note that dominating sets have been intensively studied in graph theory, see for instance [14], and that r-coverings in the k-cube are well known in coding and information theory, see [10].

Here, the discrimination "connected twin-free graphs" *vs* "connected graphs" is null, since we shall see that in both cases the bounds on γ_r are the same.

We start with a result on one-domination.

Proposition 16 *[21] Let G be a connected graph with at least two vertices. Then we have:* $1 \leq \gamma_1(G) \leq \lfloor \frac{n}{2} \rfloor$. *Moreover, these bounds are sharp.* △

This is immediately transposable to our problem.

Corollary 17 *Let G be a connected one-twin-free graph with at least three vertices. Then we have:* $1 \leq \gamma_1(G) \leq \lfloor \frac{n}{2} \rfloor$. *Moreover, these bounds are sharp.*

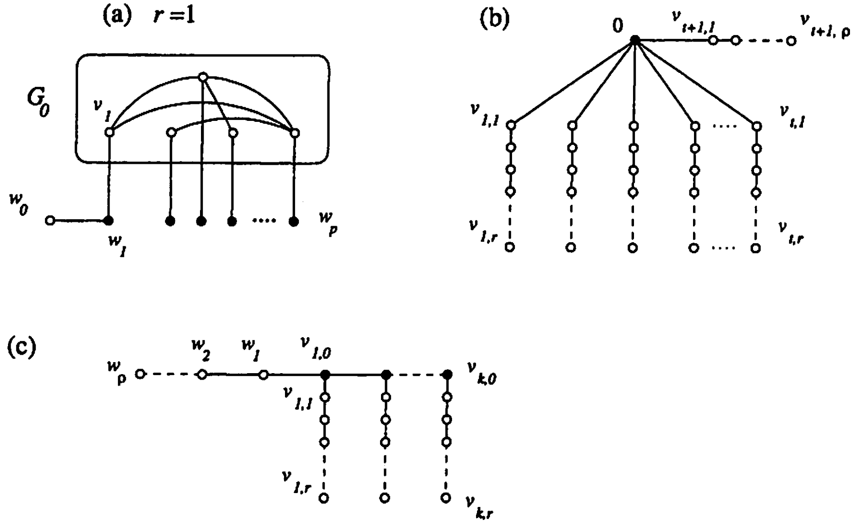

Figure 3: Different twin-free graphs constructed in Section 4. Codewords are in black.

Proof. The lower bound is trivial, and the upper bound derives from Proposition 16. Graphs used to reach the bounds in Proposition 16 are one-twin-free, so they will work for Corollary 17 too. We describe here possible constructions.

The star, defined in the Introduction, is a connected one-twin-free graph for which $C = \{0\}$ is a one-covering. More generally, any connected twin-free graph with one vertex of degree $n-1$ will do.

If n is even, $n = 2p$, $p \geq 2$, we consider any connected graph with p vertices, $G_0 = (V_0, E_0)$, where $V_0 = \{v_i : 1 \leq i \leq p\}$, and construct the graph $G = (V, E)$, where $V = V_0 \cup \{w_i : 1 \leq i \leq p\}$ and $E = E_0 \cup \{\{v_i, w_i\} : 1 \leq i \leq p\}$. If $n = 2p+1$, to G we add one vertex w_0 which we link to w_1, see Figure 3(a). In both cases, we obtain a graph which is connected and one-twin-free, and $C = \{w_i : 1 \leq i \leq p\}$ is a one-covering of size $\lfloor \frac{n}{2} \rfloor$, which is easily seen to be the minimum in G. If $n = 3$, then $G = P_3$ and $\gamma_1(G) = 1$. △

And this is immediately generalizable to any r.

Proposition 18 *[15] Let G be a connected graph with at least two vertices. Then $\gamma_r(G) \leq \lfloor \frac{n}{r+1} \rfloor$. Moreover, this bound is sharp.* △

Corollary 19 *Let G be a connected r-twin-free graph with at least $2r+1$ vertices. Then we have: $1 \leq \gamma_r(G) \leq \lfloor \frac{n}{r+1} \rfloor$. Moreover, these bounds are sharp.*

Proof. The trivial lower bound can be attained by the following generalization of the star, see Figure 3(b); we build the tree $G = (V, E)$ as follows: first, we divide $n-1$ by r, so $n = tr + \rho + 1$, with $t \geq 2$ and $0 \leq \rho < r$. Then we set $V = \{0\} \cup \{v_{i,j} : 1 \leq i \leq t, 1 \leq j \leq r\} \cup \{v_{t+1,j} : 1 \leq j \leq \rho\}$, and $E = \{\{0, v_{i,1}\}, \{v_{i,j}, v_{i,j+1}\} : 1 \leq i \leq t, 1 \leq j \leq r-1\} \cup \{\{0, v_{t+1,1}\}\} \cup \{\{v_{t+1,j}, v_{t+1,j+1}\} : 1 \leq j \leq \rho - 1\}$. In other words, starting from 0, we have t "branches" with r vertices, and one with ρ vertices. This graph has n vertices, $n-1$ edges, is connected and r-twin-free (because there are at least two "complete branches"), and $C = \{0\}$ is an r-covering in G.

The upper bound derives from Proposition 18. A possible construction is the following, see Figure 3(c). Divide n by $r+1$: $n = (r+1)k + \rho$, with $k \geq 1$ and $0 \leq \rho < r+1$; note that if $k = 1$, then $\rho = r$ and $n = 2r+1$, because $n \geq 2r+1$. Set $G = (V, E)$, with $V = \{v_{i,j} : 1 \leq i \leq k, 0 \leq j \leq r\} \cup \{w_i : 1 \leq i \leq \rho\}$ and $E = \{\{v_{i,0}, v_{i+1,0}\} : 1 \leq i \leq k-1\} \cup \{\{v_{i,j}, v_{i,j+1}\} : 1 \leq i \leq k,\ 0 \leq j \leq r-1\} \cup \{\{v_{1,0}, w_1\}, \{w_i, w_{i+1}\} : 1 \leq i \leq \rho - 1\}$. Then G has n vertices, is connected and r-twin-free, and $C = \{v_{i,0} : 1 \leq i \leq k\}$ is an r-covering of size $k = \lfloor \frac{n}{r+1} \rfloor$, which is easily seen to be the minimum in G. △

Therefore, we have exact values for $f_r(\gamma_r)$, $f_{r,n}(\gamma_r)$ and $F_{r,n}(\gamma_r)$.

Corollary 20 *For all $r \geq 1$, we have: $f_r(\gamma_r) = 1$. For all $r \geq 1$ and $n \geq 2r+1$, we have: $f_{r,n}(\gamma_r) = 1$.* △

Corollary 21 *For all $r \geq 1$, $n \geq 2r+1$, we have: $F_{r,n}(\gamma_r) = \lfloor \frac{n}{r+1} \rfloor$.* △

5 The maximum size of a clique, ω

In any connected graph with n vertices, the maximum size of a clique trivially lies between 2 (e.g., trees) and n. In this section we obtain, for all $r \geq 1$ and $n \geq 2r+1$, the exact values for $f_r(\omega)$, $f_{r,n}(\omega)$ and, much more significantly, for $F_{r,n}(\omega)$. In particular, we shall see that, when r is fixed and n grows, $F_{r,n}(\omega)$ is close to $n - r\log_2 n$, which means that there exist twin-free graphs with rather large cliques; conversely, if we wish to have a system which is a clique, then with few additional processors, we can make it twin-free.

We give a first easy result.

Theorem 22 *For all $r \geq 1$, we have: $f_r(\omega) = 2$. For all $r \geq 1$ and $n \geq 2r+1$, we have: $f_{r,n}(\omega) = 2$.*

Proof. Any maximum clique has size at least two because we consider only connected graphs. The largest cliques have size two in the paths P_n, $n \geq 2r+1$. $\triangle$

The core of this section however is the study of $F_{r,n}(\omega)$. The case $r = 1$ is not too difficult; its underlying idea can already be found in [18, Sec. IV].

Theorem 23 *For all $n \geq 3$, we have: $F_{1,n}(\omega) = k$, where k is the largest integer such that $k + \lceil \log_2 k \rceil \leq n$.*

Proof. Let $G = (V, E)$ be a connected one-twin-free graph with n vertices, $n \geq 3$. Let k be the size of a clique in G, with vertex set V_1, and let $V_2 = V \setminus V_1$. The vertices in V_1 must have k different neighbourhoods, and this can be done by the vertices in V_2 only, so $2^{n-k} \geq k$, which implies the upper bound on $F_{1,n}(\omega)$. Note that this argument works for any r.

We show that this bound can be reached for any $n \geq 3$, with the following construction: let k be the largest integer such that $k + \lceil \log_2 k \rceil \leq n$, and let $G_1 = (V_1, E_1)$ be a clique with k vertices. To V_1 we add a set V_2 of $n-k$ vertices, and create new edges as follows: we link each of the k vertices of V_1 to a different subset of vertices in V_2. This is possible because $k \leq 2^{n-k}$, and this may include the empty set for a particular vertex in V_1. Moreover, we can easily make G connected, for instance by choosing one vertex in V_1 which is linked to all vertices in V_2. The graph G is also one-twin-free: it is easy to check that two vertices in V_1, two vertices in V_2, or one vertex in V_1 and one vertex in V_2, cannot be one-twins. $\triangle$

How does k vary with n? If $n = 3$, then $k = 2$. Let $\ell = \lfloor \log_2 n \rfloor \geq 2$.

If $2^\ell \leq n \leq 2^\ell + \ell$, then $k = n - \ell$, because $2^\ell - \ell \leq n - \ell \leq 2^\ell$ and $\lceil \log_2(n-\ell) \rceil = \ell$, so $n - \ell + \lceil \log_2(n-\ell) \rceil = n$.

If $2^\ell + \ell + 1 \leq n < 2^{\ell+1}$, then we prove that $k = n - \ell - 1$. Because $2^\ell \leq n - \ell - 1 < 2^{\ell+1} - \ell - 1$, we have $\lceil \log_2(n-\ell-1) \rceil \in \{\ell, \ell+1\}$, and $n - \ell - 1 + \lceil \log_2(n-\ell-1) \rceil \in \{n-1, n\}$. If this sum is equal to n, we are done, so assume that $n - \ell - 1 + \lceil \log_2(n-\ell-1) \rceil = n-1$. This implies that $\lceil \log_2(n-\ell-1) \rceil = \ell$, and $2^\ell = n - \ell - 1$. Therefore $\lceil \log_2(n-\ell) \rceil = \ell + 1$, and $n - \ell + \lceil \log_2(n-\ell) \rceil = n+1$, so $n - \ell$ is too large.

As a consequence, we can reformulate Theorem 23.

Theorem 24 *For all $n \geq 3$, we have:*

$$F_{1,n}(\omega) = \begin{cases} n - \lfloor \log_2 n \rfloor & \text{if } 2^\ell \leq n \leq 2^\ell + \ell,\ \ell \geq 1, \\ n - \lfloor \log_2 n \rfloor - 1 & \text{if } 2^\ell + \ell + 1 \leq n < 2^{\ell+1},\ \ell \geq 2. \end{cases}$$

$\triangle$

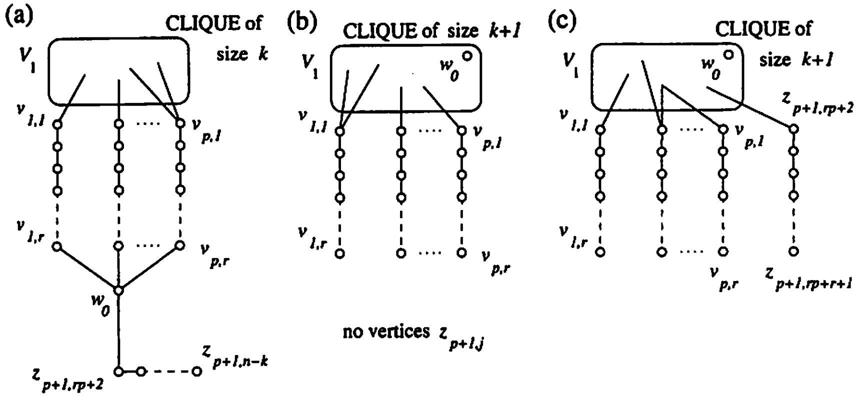

Figure 4: Partial representations of the graphs constructed for Theorem 25.

We now turn to the general case for r, and give constructions where there is a large clique. By "large", we mean that it will meet the upper bound (see Theorem 27). Let $r \geq 2$ and $n \geq 2r+1$, let k be the largest integer such that $k + r\lceil \log_2 k \rceil \leq n-1$, and let $p = \lceil \log_2 k \rceil$.

If $2^{p-1} < k < 2^p$, then, since $(k+1) + r\lceil \log_2(k+1) \rceil \geq n$ and $p = \lceil \log_2(k+1) \rceil$, we have:

$$rp + 1 = n - k. \tag{3}$$

If $k = 2^p$, then, since $(k+1) + r\lceil \log_2(k+1) \rceil = k + 1 + r(p+1) \geq n$, we obtain:

$$rp + 1 \leq n - k \leq rp + 1 + r. \tag{4}$$

Now let $G_1 = (V_1, E_1)$ be a clique with k vertices. To V_1 we add a set V_2 of $n-k$ vertices, where

$$V_2 = \{w_0\} \cup \{v_{i,j} : 1 \leq i \leq p, 1 \leq j \leq r\} \cup \{z_{p+1,j} : rp + 2 \leq j \leq n - k\},$$

see Figure 4(a). Note that by (3) and (4), there are between 0 and r vertices $z_{p+1,j}$ in V_2. We construct the edges $\{v_{i,j}, v_{i,j+1}\}$ for $1 \leq i \leq p$, $1 \leq j \leq r-1$, and $\{z_{p+1,j}, z_{p+1,j+1}\}$ for $rp + 2 \leq j \leq n-k-1$, to which we add the edges $\{w_0, z_{p+1,rp+2}\}$ and $\{w_0, v_{i,r}\}$, $1 \leq i \leq p$.

Then we create new edges as follows: we link each of the k vertices in V_1 to a different subset of vertices taken among the vertices $v_{i,1}$, $1 \leq i \leq p$. This is possible because $k \leq 2^p$, and this may include the empty set for a particular vertex in V_1. Obviously, this can be done in such a way that G is connected.

The graph G is r-twin-free: one can see that two vertices in V_1 are not r-twins, since by construction the vertices $v_{i,r}$, $1 \leq i \leq p$, all contribute differently to the balls of radius r centered at the vertices in the clique; one vertex in V_1 and one vertex in V_2 cannot be r-twins, because w_0 r-covers all the vertices in V_2, and no vertex in V_1. Finally, we have to consider two vertices in V_2. If $p = 1$, then $k = 2$ and our construction yields the path P_n, which is r-twin-free for $n \geq 2r + 1$. If $p > 1$, the vertices w_0, $v_{i,j}$ and $v_{i',j'}$ $(i \neq i')$ belong to chordless cycles of length $2r+2$ or more, which suffices to show that they are not twins. Lastly it is easy to check that the vertices $z_{p+1,m}$ are twins neither with themselves, nor with w_0, nor with any vertex $v_{i,j}$.

In some cases, we can improve on this construction by one. First, we place ourselves in the case when $2^{p-1} < k < 2^p$. We have seen that (3) holds: $rp + 1 = n - k$, which means that there are no vertices $z_{p+1,j}$. Now we assume that $k \leq 2^p - p - 1$, and we modify the previous graph as follows: we remove the edges going to w_0, and we move w_0 into the clique, which has now size $k+1$, see Figure 4(b). We keep the edges $\{v_{i,j}, v_{i,j+1}\}$, $1 \leq i \leq p$, $1 \leq j \leq r - 1$, and we link each of the $k + 1$ vertices in the clique to a different subset of vertices taken among the vertices $v_{i,1}$, $1 \leq i \leq p$, in a way slightly different from previously: subsets containing *exactly one element* $v_{i,1}$ are *forbidden.* This is possible because $k+1 \leq 2^p - p$, and this may include the empty set for a particular vertex in the clique. We can also require that each $v_{i,1}$ is linked to at least one vertex in the clique (so that G is connected). We leave it to the reader to check that this graph is r-twin-free.

Next, we assume that $k = 2^p$ with $p \geq 2$ and that $n - k = rp + 1 + r$ — see (4). Then, since $n - (k+1) = (p+1)r$, we can, starting from the graph in Figure 4(a), remove the edges going to w_0, move w_0 into the clique, and directly link $z_{p+1,rp+2}$ to vertices belonging to the clique, so that the chain of vertices $z_{p+1,rp+2}, \ldots, z_{p+1,rp+1+r}$ assumes the same role as any of the chains $v_{i,1}, \ldots, v_{i,r}$; again, because $k + 1 = 2^p + 1 \leq 2^{p+1} - (p + 1)$ when $p \geq 2$, we can avoid the singletons, cf. above paragraph. See Figure 4(c). This new graph, which contains a clique of size $k + 1$, is r-twin-free (and can be made connected). We leave the details to the reader.

Finally, using the two particular constructions given in Figure 5, which are easy to check, we obtain the following lower bounds on $F_{r,n}(\omega)$.

Theorem 25 *For all $r \geq 2$ and $n \geq 2r + 1$, we have:*

$$F_{r,n}(\omega) \geq k,$$

where k is the largest integer such that $k + r\lceil \log_2 k \rceil \leq n - 1$.

If the same k satisfies either $k \leq 2^{\lceil \log_2 k \rceil} - \lceil \log_2 k \rceil - 1$, or $4 \leq k = 2^{\lceil \log_2 k \rceil}$ and $n = k + 1 + r(\lceil \log_2 k \rceil + 1)$, or $k = 2$ and $n = 2r + 3$, or $k = 5$

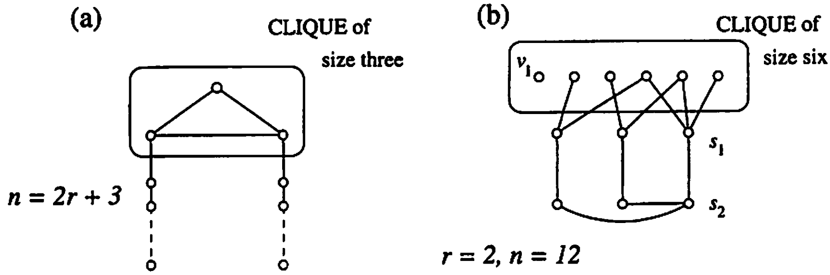

Figure 5: Two particular constructions of twin-free graphs.

and $n = 12$ *with* $r = 2$*, then we have:*

$$F_{r,n}(\omega) \geq k + 1.$$

△

Next, we give upper bounds on $F_{r,n}(\omega)$.

Theorem 26 *Let* $r \geq 2$ *and* $n \geq 2r + 1$ *be two integers, let* $G = (V, E)$ *be a connected* r*-twin-free graph with* n *vertices, and let* G_0 *be a clique of* G *with size* k_0*. Then we have:*

$$k_0 + r\lceil \log_2 k_0 \rceil \leq n.$$

If k_0 *satisfies* $2^{\lceil \log_2 k_0 \rceil} - \lceil \log_2 k_0 \rceil < k_0 \ (\leq 2^{\lceil \log_2 k_0 \rceil})$*, then we have:*

$$k_0 + r\lceil \log_2 k_0 \rceil \leq n - 1,$$

unless either $r = 2$, $\lceil \log_2 k_0 \rceil = 3$, $k_0 = 6$ *and* $n = 12$, or $\lceil \log_2 k_0 \rceil = 2$, $k_0 = 3$ *and* $n = 2r + 3$.

Proof. Let $G = (V, E)$ be any connected r-twin-free graph with n vertices, containing a clique G_0 of size k_0, whose set of vertices is denoted by V_0, and let $p_0 = \lceil \log_2 k_0 \rceil$. For $i = 1, 2, \ldots, r, \ldots$, let $V_i = \{x \in V : d(x, V_0) = i\}$, where as usual $d(x, V_0)$ is the smallest distance between x and the vertices in V_0. Obviously, the sets V_i, $i = 0, 1, \ldots$, partition V. Because G_0 is a clique, any vertex in V_i is at distance i or $i + 1$ from the vertices in V_0. We observe that edges in G can exist only inside the sets V_i or between V_i and V_{i+1}, for $i = 0, 1, \ldots$ (there is no jump between non-consecutive sets V_i).

Assume that for some i between 1 and r, we have $|V_i| < p_0$. Then there are two vertices in V_0, x and y, such that $B_i(x) \cap V_i = B_i(y) \cap V_i$. We claim that x and y are r-twins.

Assume on the contrary that there is a vertex w which, say, r-covers x and not y: $d(w,x) \le r$ and $d(w,y) > r$. Then necessarily $w \in V_r$ and $d(w,x) = r$, $d(w,y) = r+1$. Consequently, $B_i(x) \cap V_i \neq \emptyset$ and any vertex $z \in B_i(x) \cap V_i$ is at distance exactly i from x and exactly $r-i$ from w. Since z is also in $B_i(y) \cap V_i$, this implies that $d(w,y) \le r$, a contradiction which shows that x and y are twins.

So $|V_i| \ge p_0$ and

$$n \ge k_0 + rp_0. \tag{5}$$

Next, we assume that $2^{p_0} - p_0 < k_0 \ (\le 2^{p_0})$, and show that we can improve on (5) by one.

We can assume that for $i = 1, 2, \ldots, r$, $|V_i| = p_0$ and $V_j = \emptyset$ for $j > r$ (otherwise, $n > k_0 + rp_0$ and we are done), so that $n = k_0 + rp_0$. We distinguish between two cases, each of them ultimately yielding a contradiction.

Case 1: $\forall x \in V_1, \forall y \in V_1, \exists z \in V_0$ such that $\{z,x\} \in E$ and $\{z,y\} \in E$.

Let x_0, x_1, x_{r-1} be any vertices in V_0, V_1 and V_{r-1}, respectively. We know that $d(x_{r-1}, x_0) \le r$, and, since x_{r-1} is at distance $r-1$ from V_0, there is a vertex $t \in V_1$ such that $d(x_{r-1}, t) = r-2$. By assumption, there is a vertex $z \in V_0$ such that $\{z, x_1\} \in E$ and $\{z, t\} \in E$, so $d(x_{r-1}, x_1) \le r$ (see Figure 6). This shows that for all $x \in V_0 \cup V_1$, we have: $V_{r-1} \subseteq B_r(x)$, and, *a fortiori*, $V \setminus V_r \subseteq B_r(x)$. This implies that the sets $B_r(x) \cap V_r$, $x \in V_0 \cup V_1$, must be pairwise distinct. This in turn implies that $2^{p_0} \ge |V_0| + |V_1| = k_0 + p_0$, which contradicts our assumption $2^{p_0} - p_0 < k_0$.

Case 2: $\exists x \in V_1, \exists y \in V_1$, such that $\forall z \in V_0$, $\{z,x\} \in E \Rightarrow \{z,y\} \notin E$.

We know that the sets $B_1(z) \cap V_1$, $z \in V_0$, must be pairwise distinct. Moreover, they do not contain both vertices x and y. Therefore, since there are 2^{p_0-2} subsets of V_1 which contain simultaneously x and y, we have only $2^{p_0} - 2^{p_0-2}$ subsets available, and we obtain: $k_0 \le 2^{p_0} - 2^{p_0-2}$. This contradicts $2^{p_0} - p_0 < k_0$, *unless either* $p_0 = 3$, $k_0 = 6$ and $n = 3r+6$, *or* $p_0 = 2$, $k_0 = 3$ and $n = 2r+3$.

Finally, among the above cases which do not give a contradiction, when $p_0 = 3$, $n = 3r+6$, $r \ge 3$, we show that actually $k_0 = 6$ is impossible, so $k_0 \le 5$ and in this case also, $k_0 + rp_0 \le n-1$. Strangely enough, the proof of this particular case is not that easy; see Appendix. △

The following theorem shows that the upper and lower bounds given in the previous theorems actually coincide.

Theorem 27 *For all $r \ge 2$ and $n \ge 2r+1$, let k be the largest integer such that $k + r\lceil \log_2 k \rceil \le n-1$. If k satisfies either $k \le 2^{\lceil \log_2 k \rceil} - \lceil \log_2 k \rceil - 1$, or $4 \le k = 2^{\lceil \log_2 k \rceil}$ and $n = k + 1 + r(\lceil \log_2 k \rceil + 1)$, or $k = 2$ and $n = 2r+3$, or $k = 5$ and $n = 12$ with $r = 2$, then we have:*

$$F_{r,n}(\omega) = k+1.$$

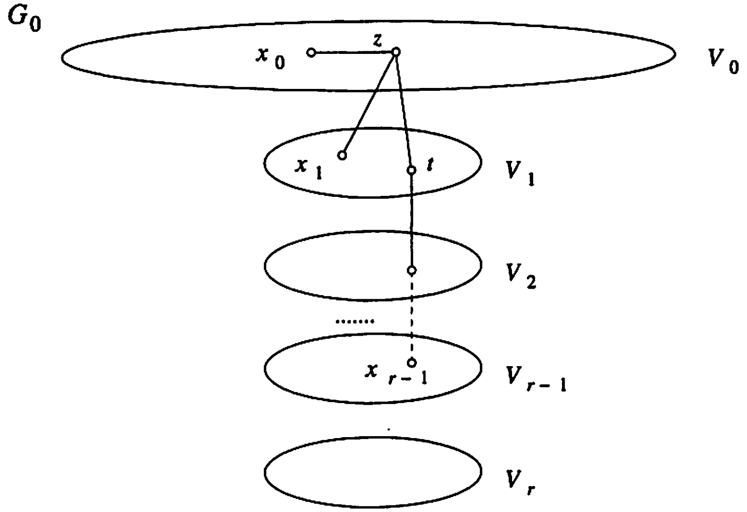

Figure 6: A partial representation of the clique G_0 and the sets V_i.

Otherwise, we have:

$$F_{r,n}(\omega) = k.$$

Proof. Let $p = \lceil \log_2 k \rceil$, $K = k + 1$ if $k \leq 2^p - p - 1$, or $4 \leq k = 2^p$ and $n = k + 1 + r(p+1)$, or $k = 2$ and $n = 2r + 3$, or $k = 5$ and $n = 12$ with $r = 2$, and $K = k$ otherwise; this means that, by Theorem 25, we have: $F_{r,n}(\omega) \geq K$.

Let k_0 be the size of the largest clique in a connected r-twin-free graph with n vertices, and $p_0 = \lceil \log_2 k_0 \rceil$. We have seen that k_0 satisfies $k_0 + rp_0 \leq n$, and even $k_0 + rp_0 \leq n - 1$ if k_0 is such that $2^{p_0} - p_0 < k_0 \leq 2^{p_0}$, excluding the case (a) $r = 2$, $p_0 = 3$, $k_0 = 6$ and $n = 12$, and the case (b) $p_0 = 2$, $k_0 = 3$ and $n = 2r + 3$.

We now prove that $K \geq k_0$, which will establish the theorem.

If $k_0 + rp_0 \leq n - 1$, then the very definitions of k and K show that $k_0 \leq k \leq K$. So from now on, we assume that $k_0 + rp_0 = n$, and the definition of k now shows that $k = k_0 - 1$. Because $k_0 + rp_0 = n$, either we have $2^{p_0} - p_0 \geq k_0$, or we are in cases (a) or (b). Since these two cases have precisely been treated in Theorem 25 (see Figure 5), we can assume from now on that $2^{p_0} - p_0 \geq k_0$, so $2^{p_0} - p_0 - 1 \geq k$.

If $p = p_0$, then $k \leq 2^p - p - 1$, and we are in the case when $K = k + 1$, i.e., $K = k_0$.

If $p \neq p_0$, then $p = p_0 - 1$, $k_0 = 2^{p_0 - 1} + 1$ and $k = 2^p$. If $k \leq 2$, then

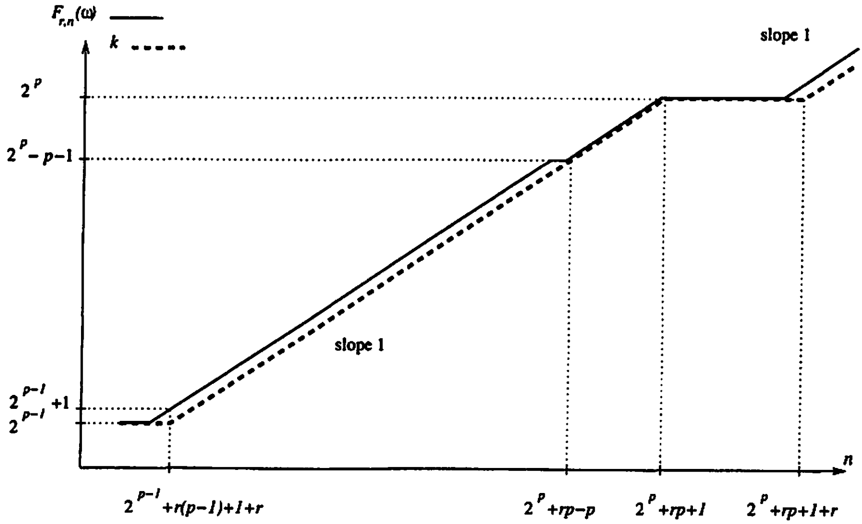

Figure 7: General behaviour of k and $F_{r,n}(\omega)$ with respect to n.

k_0 does not satisfy $2^{p_0} - p_0 \geq k_0$; so $k \geq 4$. Since $n = k_0 + rp_0$, we have $n = k + 1 + rp + r$, and again we are in a case when $K = k + 1 = k_0$.

In both cases ($p = p_0$ and $p \neq p_0$), $K \geq k_0$. △

The very general behaviour of k and $F_{r,n}(\omega)$ with n is given by Figure 7.

We conclude this section by showing that $F_{r,n}(\omega)$ is close to $n - r\log_2 n$, when r is fixed and n grows, and that $F_{r,n}(\omega)$ is bounded by above by a constant when r is a fraction of n.

For simplicity, we study not exactly the k defined in Theorems 25 and 27, but the largest k such that $k + r\lceil \log_2 k \rceil \leq n$, since this new k is also close to $F_{r,n}(\omega)$.

Proposition 28 *Let $r \geq 2$ and $n \geq 2r + 1$ be two integers, and let k be the largest integer such that $k + r\lceil \log_2 k \rceil \leq n$. If*

$$r < \frac{n - 2^{\lfloor \log_2 n \rfloor - 1}}{\lfloor \log_2 n \rfloor}, \tag{6}$$

or equivalently $2^{\lfloor \log_2 n \rfloor - 1} < n - r\lfloor \log_2 n \rfloor$, then

$$n - r\lfloor \log_2 n \rfloor - r \leq k \leq n - r\lfloor \log_2 n \rfloor, \tag{7}$$

and these bounds are sharp.

Proof. Let $\ell = \lfloor \log_2 n \rfloor$. By assumption, $2^{\ell-1} < n - r\ell \leq 2^{\ell+1}$.

Because $\lceil \log_2(n - r\ell - r) \rceil \leq \ell + 1$, we have $(n - r\ell - r) + r\lceil \log_2(n - r\ell - r) \rceil \leq n$, so $k \geq n - r\ell - r$, with possible equality.

Because $\lceil \log_2(n - r\ell) \rceil \geq \ell$, we have $(n - r\ell) + r\lceil \log_2(n - r\ell) \rceil \geq n$, so $k \leq n - r\ell$, with possible equality. △

So we see in particular that, if r is fixed, then there exists an n_0 such that all $n \geq n_0$ satisfy inequalities (7), which means that when n goes to infinity, $F_{r,n}(\omega)$ behaves like $n - r \log_2 n$.

If now r is a fraction of n, say $r = \rho n$, where ρ is a constant satisfying $0 < \rho < 1/2$, then the very definition of k gives the rough estimate

$$\lceil \log_2 k \rceil \leq \frac{n-k}{\rho n} \leq \frac{1}{\rho},$$

showing that $k \leq$ constant. The extremal case is when $r = (n-1)/2$ and $F_{r,n}(\omega) = 2$ (path P_n).

6 Recapitulatory

The tables below recapitulate the results obtained in the previous four sections, without the specific conditions on n or r. They are given for $r = 1$, $r = 2$, and $r \geq 3$. We can see that, apart from the conditions on n and r, the only uncertainty left is for $f_r(c_r)$, $r \geq 5$.

$r = 1$	$f_1(.)$	$f_{1,n}(.)$	$F_{1,n}(.)$
n	3 (Cor. 4)	n	n
c_1	2 (Th. 15)	$\lceil \log_2(n+1) \rceil$ (Cor. 12)	$n - 1$ (Cor. 7)
γ_1	1 (Cor. 20)	1 (Cor. 20)	$\lfloor n/2 \rfloor$ (Cor. 21)
ω	2 (Th. 22)	2 (Th. 22)	$= n - \lfloor \log_2 n \rfloor - 1$ or $= n - \lfloor \log_2 n \rfloor$ (Th. 24)

$r = 2$	$f_2(.)$	$f_{2,n}(.)$	$F_{2,n}(.)$
n	5 (Cor. 4)	n	n
c_2	3 (Th. 15)	$\lceil \log_2(n+1) \rceil$ (Cor. 12)	$n - 1$ (Cor. 9)
γ_2	1 (Cor. 20)	1 (Cor. 20)	$\lfloor n/3 \rfloor$ (Cor. 21)
ω	2 (Th. 22)	2 (Th. 22)	$= k$ or $= k + 1$, k max s.t. $k + 2\lceil \log_2 k \rceil \leq n - 1$ (Th. 27)

$r \geq 3$	$f_r(.)$	$f_{r,n}(.)$	$F_{r,n}(.)$
n	$2r+1$ (Cor. 4)	n	n
c_r	$\geq \lceil \log_2(2r+4) \rceil$ (Th. 13) $\leq r+1$ (Th. 14)	$\lceil \log_2(n+1) \rceil$ (Cor. 12)	$n-1$ (Cor. 9)
γ_r	1 (Cor. 20)	1 (Cor. 20)	$\lfloor n/(r+1) \rfloor$ (Cor. 21)
ω	2 (Th. 22)	2 (Th. 22)	$= k$ or $= k+1$, k max such that $k + r\lceil \log_2 k \rceil \leq n-1$ (Th. 27)

7 Appendix

Continuation of the proof of Theorem 26, the case $p_0 = 3$, $n = 3r+6$, $r \geq 3$.

Assume that a clique $G_0 = (V_0, E_0)$ of size six exists. We denote by $v_1, \ldots, v_6$ the vertices in V_0, and by a_i, b_i, s_i the three vertices in V_i, for i between 1 and r. Going through Cases 1 and 2 in the proof of Theorem 26, we can see that the absence of contradiction appeared only in Case 2, so there are two vertices in V_1, say a_1 and b_1, such that for all i between one and six, the edges $\{a_1, v_i\}$ and $\{b_1, v_i\}$ do not exist simultaneously. It is then straightforward to see that, up to permutations, there is a unique way of linking V_0 and V_1, given in Figure 8(a). We will denote the vertices in V_0 by the set of vertices in V_1 to which they are linked; so, for instance, $v_1 = \emptyset$, or $v_4 = \{a_1, s_1\}$.

Now there is a path of length $r-1$ between V_r and a_1 (otherwise, $\emptyset$ and $\{a_1\}$ would be twins). The same is true for b_1 and s_1. Without loss of generality, one path of length $r-1$ between s_1 and V_r is $s_1, s_2, \ldots, s_r$, see Figure 8(b).

(i) Assume that there is a second vertex in V_r, say b_r, which is at distance $r-1$ from s_1. Then the three vertices $\{s_1\}$, $\{b_1, s_1\}$ and $\{a_1, s_1\}$ are within distance r from all vertices in V, except maybe a_r. Therefore, among these three vertices, at least two of them are r-twins. This shows that in V_r, s_r is the only vertex at distance $r-1$ from s_1. In particular:

$$d(a_r, s_1) \geq r. \tag{8}$$

We can also see that all vertices in V, except maybe a_r and b_r, are within distance r from s_1.

(ii) Next, we assume that $d(b_1, s_r) \leq r-1$. Then the five vertices s_1, $\{s_1\}$, $\{s_1, b_1\}$, $\{s_1, a_1\}$ and $\{b_1\}$ are within distance r from all vertices in V, except maybe a_r and b_r, and some of them are necessarily r-twins (two vertices are not sufficient for distinguishing between five vertices). So: $d(b_1, s_r) \geq r$, and similarly,

$$d(a_1, s_r) \geq r. \tag{9}$$

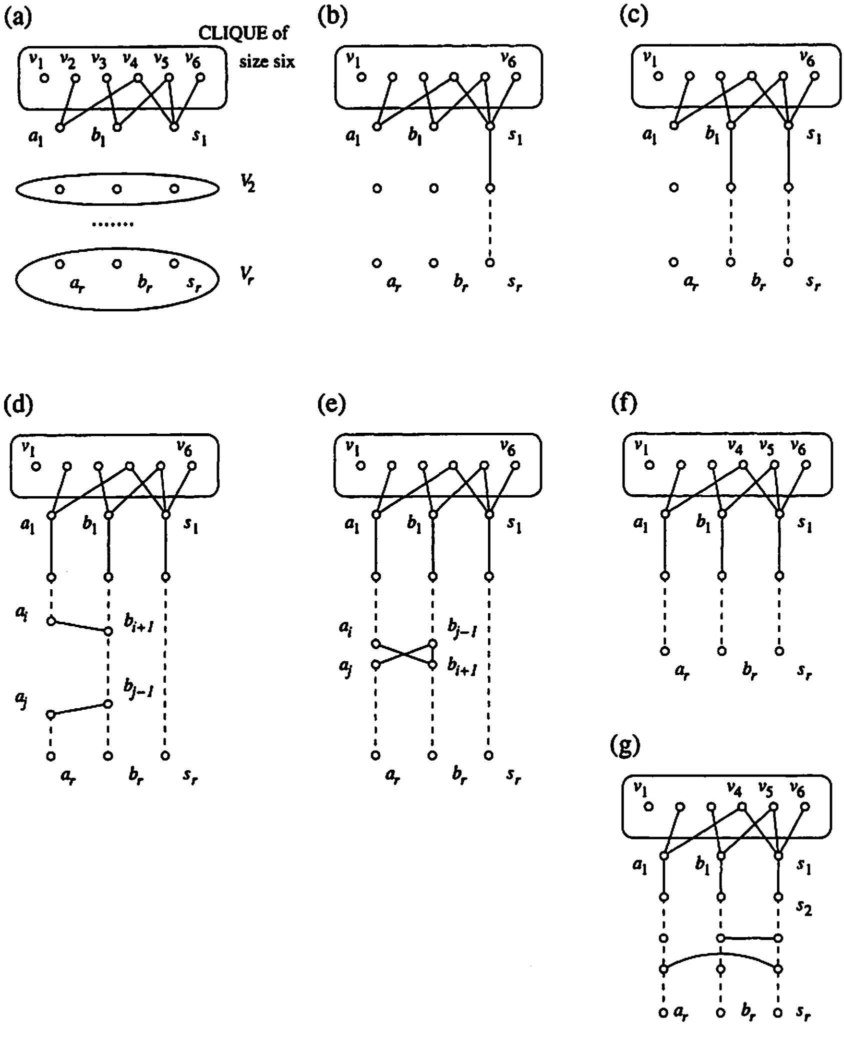

Figure 8: The case $r > 2$, $p_0 = 3$, $n = 3r + 6$: there is no clique of size six.

Now, by (i) and (ii), any path of length $r-1$ going from V_r to b_1 cannot go through vertices s_i. Therefore, without loss of generality, we can assume that one such path is $b_1, b_2, \ldots, b_r$, and no edge $\{s_i, b_{i+1}\}$ or $\{b_i, s_{i+1}\}$ exists in G, for i between 1 and $r-1$. See Figure 8(c).

We now study the possible connexions between vertices a_i and b_{i+1} or b_{i-1}. Assume that the path $a_1, a_2, \ldots, a_r$ does not exist, and let i be the smallest integer such that the edge $\{a_i, a_{i+1}\}$ does not exist, and j the largest such that $\{a_{j-1}, a_j\}$ does not exist.

Since there is a vertex in V_r which is at distance $r-1$ from a_1 and since inequality (9) holds, the edge $\{a_i, b_{i+1}\}$ exists; also, since there is a vertex in V_1 which is at distance $r-1$ from a_r and since inequality (8) holds, the edge $\{a_j, b_{j-1}\}$ exists, see Figure 8(d) and (e).

Now a_1 and b_1 are within distance r from all vertices except maybe those in V_r, and the same is true for s_1 and the six vertices in V_0. But then it is impossible to distinguish between nine vertices with only the three vertices in V_r.

This shows that there is the path $a_1, a_2, \ldots, a_r$, and furthermore there are no edges $\{a_i, b_{i+1}\}$ or $\{a_{i+1}, b_i\}$: Figure 8(f) gives the only possibility for paths of length $r-1$ between the vertices a_1, b_1 and s_1 on the one hand, and V_r on the other hand; additional edges in this figure can be only horizontal, i.e., of type $\{a_i, b_i\}$, $\{a_i, s_i\}$ or $\{s_i, b_i\}$.

So far, in Figure 8(f), we have $B_r(\{s_1\}) = B_r(s_1) = V \setminus \{a_r, b_r\}$, $B_r(\{a_1, s_1\}) = V \setminus \{b_r\}$, and $B_r(\{b_1, s_1\}) = V \setminus \{a_r\}$. Therefore, we must have two edges $\{s_i, b_i\}$ and $\{s_j, a_j\}$, $1 \le i, j \le r$, so that $B_r(s_1) = V$; see Figure 8(g). This however implies that $B_r(s_2) = V$, i.e., s_1 and s_2 are r-twins (this works *only for* $r \ge 3$, since $v_1 \notin B_r(s_2)$ when $r = 2$, cf. Figure 5(b) which gives a 2-twin-free graph). $\triangle$

References

[1] N. Bertrand, I. Charon, O. Hudry, A. Lobstein: Identifying and locating-dominating codes on chains and cycles. Europ. J. Combin. 25/7, 969–987 (2004)

[2] N. Bertrand, I. Charon, O. Hudry, A. Lobstein: 1-identifying codes on trees. Austral. J. Combin. 31, 21–35 (2005)

[3] U. Blass, I. Honkala, S. Litsyn: Bounds on identifying codes. Discrete Math. 241, 119–128 (2001)

[4] I. Charon, I. Honkala, O. Hudry, A. Lobstein: Structural properties of twin-free graphs. Electron. J. Combin., submitted

[5] I. Charon, O. Hudry, A. Lobstein: Identifying codes with small radius in some infinite regular graphs. Electron. J. Combin. 9(1), R11 (2002)

[6] I. Charon, O. Hudry, A. Lobstein: Minimizing the size of an identifying or locating-dominating code in a graph is NP-hard. Theoret. Comput. Sci. 290(3), 2109–2120 (2003)

[7] I. Charon, O. Hudry, A. Lobstein: On the structure of identifiable graphs: results, conjectures, and open problems. Proc. 29th Australasian Conf. in Combinatorial Mathematics and Combinatorial Computing, Taupo, New Zealand, pp. 37–38 (2004)

[8] I. Charon, O. Hudry, A. Lobstein: Extremal cardinalities for identifying and locating-dominating codes in graphs. Discrete Math., to appear

[9] I. Charon, O. Hudry, A. Lobstein: Extremal values of some classical parameters in twin-free graphs, in preparation

[10] G. Cohen, I. Honkala, S. Litsyn, A. Lobstein: Covering Codes. Amsterdam: Elsevier 1997

[11] M. Daniel: Codes identifiants. Mémoire de DEA, Université de Grenoble (France), June 2003

[12] S. Gravier, J. Moncel: Constructions of codes identifying sets of vertices. Electron. J. Combin. 12(1), R13 (2005)

[13] S. Gravier, J. Moncel: On graphs having a $V \setminus \{x\}$ set as an identifying code. Discrete Math., to appear

[14] T.W. Haynes, S.T. Hedetniemi, P.J. Slater: Fundamentals of Domination in Graphs. New York: Marcel Dekker 1998

[15] M.A. Henning, O.R. Oellermann, H.C. Swart: Bounds on distance domination parameters. J. Combin. Inform. System Sci. 16, 11–18 (1991)

[16] I. Honkala, T. Laihonen, S. Ranto: On strongly identifying codes. Discrete Math. 254, 191–205 (2002)

[17] I. Honkala, A. Lobstein: On identifying codes in binary Hamming spaces. J. Combin. Theory Ser. A 99, 232–243 (2002)

[18] M.G. Karpovsky, K. Chakrabarty, L.B. Levitin: On a new class of codes for identifying vertices in graphs. IEEE Trans. Inform. Theory 44, 599–611 (1998)

[19] T. Laihonen, S. Ranto: Codes identifying sets of vertices. Lecture Notes in Computer Science 2227, 82–91 (2001)

[20] J. Moncel: Codes identifiants dans les graphes. Thèse de Doctorat de l'Université de Grenoble, France, June 2005

[21] O. Ore: Theory of Graphs (Amer. Math. Soc. Colloq. Publ. 38) Providence: Amer. Math. Soc. 1962

[22] S. Ranto, I. Honkala, T. Laihonen: Two families of optimal identifying codes in binary Hamming spaces. IEEE Trans. Inform. Theory 48, 1200–1203 (2002)

[23] http://www.infres.enst.fr/~lobstein/bibLOCDOMetID.html