# Extremal values for the maximum degree in a twin-free graph

## *Valeurs extrémales du degré maximum dans un graphe sans jumeaux*

Irène Charon
Olivier Hudry
Antoine Lobstein

2008D007

Mai 2008

Département Informatique et Réseaux
Groupe MIC2 : Mathématiques de l'Information, des Communications et du Calcul

# Valeurs extrémales du degré maximum dans un graphe sans jumeaux

# Extremal Values for the Maximum Degree in a Twin-Free Graph

**Irène Charon**
Institut TELECOM - TELECOM ParisTech & CNRS - LTCI UMR 5141
46, rue Barrault, 75634 Paris Cedex 13 - France
irene.charon@telecom-paristech.fr

**Olivier Hudry**
Institut TELECOM - TELECOM ParisTech & CNRS - LTCI UMR 5141
46, rue Barrault, 75634 Paris Cedex 13 - France
olivier.hudry@telecom-paristech.fr

**Antoine Lobstein**
CNRS - LTCI UMR 5141 & Institut TELECOM - TELECOM ParisTech
46, rue Barrault, 75634 Paris Cedex 13 - France
antoine.lobstein@telecom-paristech.fr

May 16, 2008

**Résumé.**
On considère un graphe connexe et non orienté $G = (V, E)$ et un entier $r \geq 1$; pour tout sommet $v \in V$, on désigne par $B_r(v)$ la boule de rayon $r$ centrée sur $v$, i.e., l'ensemble de tous les sommets reliés à $v$ par un chemin d'au plus $r$ arêtes. Si pour tous les sommets $v \in V$, les ensembles $B_r(v)$ sont différents, alors on dit que $G$ est sans $r$-jumeaux.

Dans les graphes sans $r$-jumeaux, nous prolongeons l'étude des valeurs extrémales pouvant être atteintes par certains paramètres classiques en théorie des graphes, et nous étudions ici le degré maximum.

**Mots Clés:**
Théorie des graphes, Codes identifiants, Jumeaux, Degré maximum

**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 $r$ centred 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.

In $r$-twin-free graphs, we prolong the study of the extremal values that can be reached by some classical parameters in graph theory, and investigate here the maximum degree.

**Key Words:**
Graph Theory, Identifying Codes, Twins, Maximum Degree

# 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 $r$ centred 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*.

The trivial graph with one vertex is twin-free, the trivial connected graph with two vertices is not, and generally we consider graphs with at least three vertices.

Twin-free graphs are of interest because they are strongly connected with *identifying codes* [6], 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 [6]. 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.

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 $v_1, v_2, \ldots, v_q$ 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 $1 \leq i < q$.

**Remark 2**. It is easy to observe that a connected $r$-twin-free graph has one vertex or at least $2r+1$ vertices. Actually, it can be shown that a connected $r$-twin-free graph with at least two vertices has $P_{2r+1}$ as a subgraph [3], and even as an induced subgraph [1],[2].

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 $\mathcal{C}_q$. The following graph will be used in the sequel: we shall call it the *n-star*, or 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 Illustration

The motivations for identifying codes come, for instance, from fault diagnosis in multiprocessor systems [6]. 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 $r$-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 ones of the codewords gave the alarm.

Identifying codes were introduced in [6], 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, planar graphs, or the hypercube), dealing with complexity issues, or using heuristics such as the noising methods for the construction of small codes. For a bibliography, see [9].

Therefore, it is quite natural to study some of the parameters of twin-free graphs, since these graphs, and only these graphs, admit identifying codes.

## 1.3 Scope of the paper

We intend to investigate the extremal values that some parameters, classical in graph theory, can reach in connected twin-free graphs. More precisely, for a parameter $p$ such as the number of edges, the maximum degree, the diameter, $\ldots$, 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 $G$:

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

where $\mathcal{G}_r = \{G : G$ connected, $r$-twin-free with at least $2r+1$ vertices$\}$;

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

where $\mathcal{G}_{r,n} = \{G : G$ connected, $r$-twin-free with $n \geq 2r+1$ vertices$\}$.

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

In this paper, we are interested in the maximum degree, $\Delta_{\max}$, and we shall study the functions $f_r(\Delta_{\max})$, $f_{r,n}(\Delta_{\max})$ and $F_{r,n}(\Delta_{\max})$.

In [4], the same study was run for the following four parameters: number of vertices, minimum size of an $r$-identifying code, $r$-domination number, maximum size of a clique. The number of edges was investigated in [8, Sec. 4.1.2], and the minimum degree in [7],[5].

## 2 The maximum degree, $\Delta_{\max}$

In any connected graph with $n$ vertices, the maximum degree is comprised between 1 (if $n = 2$) or 2 (paths $P_n$ and cycles $\mathcal{C}_n$, $n \geq 3$), and $n-1$ (clique, star, ...).

It is straightforward to obtain the exact values for $f_r(\Delta_{\max})$ and $f_{r,n}(\Delta_{\max})$.

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

We distinguish between two cases in the study of $F_{r,n}(\Delta_{\max})$, $r = 1$ and $r > 1$.

**Theorem 2** *For all $n \geq 3$, we have: $F_{1,n}(\Delta_{\max}) = n-1$.*

**Proof.** The $n$-star, defined in the Introduction, is a connected 1-twin-free graph with maximum degree equal to $n-1$, which is of course the upper bound. △

For $r \geq 2$, we first give an upper bound.

**Theorem 3** *For all $r \geq 2$ and $n \geq 2r+1$, we have: $F_{r,n}(\Delta_{\max}) \leq k$, where $k$ is the largest integer such that*

$$k + (r-2)\lceil \log_3(k+1) \rceil + \lceil \log_2(k+1) \rceil \leq n-1. \quad (2)$$

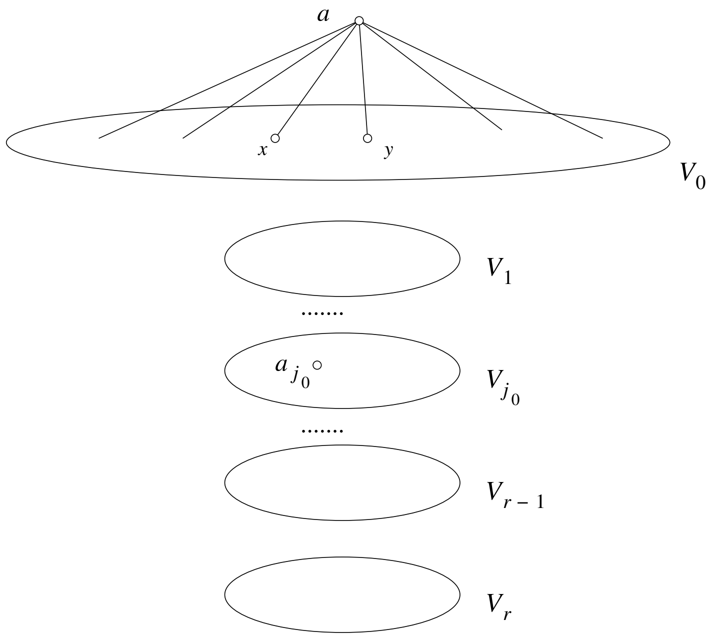


Figure 1: A partial representation of the vertex $a$ and the sets $V_i$.

**Proof**. Let $G = (V, E)$ be any connected $r$-twin-free graph with $n$ vertices, let $a$ be any vertex in $V$, the degree of which we denote by $\deg(a)$ or $\delta$. Let $V_0(a) = V_0$ be the set of vertices adjacent to $a$: $|V_0| = \delta$. For $i = 1, 2, \ldots, r, \ldots$, let $V_i(a) = V_i = \{x \in V \setminus \{a\} : d(x, V_0) = i\}$, 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 $V \setminus \{a\}$, and we have the following property:
• (P) any vertex in $V_i$ is at distance $i$, $i+1$ or $i+2$ from any vertex in $V_0$, and is at distance exactly $i+1$ from $a$.

We also observe that edges in $G$ can exist only between $a$ and $V_0$, 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 $V_i$). See Figure 1, where the sets $V_i$ are represented up to $i = r$, in a partial view of the graph.

For each vertex $x \in V_0 \cup \{a\}$, we define the couples $C_i(x)$, $1 \leq i \leq r-1$, as follows:

$$C_i(x) = \big(B_i(x) \cap V_i, B_{i+1}(x) \cap V_i\big).$$

Obviously, $C_i(a) = \big(\emptyset, V_i\big)$.

First, we show that the couples $C_i(x)$, $i \in \{1, \ldots, r-1\}$, are all distinct when $x$ runs through $V_0$. Assume on the contrary that $x$ and $y$ exist in $V_0$, such that, for some $j_0$ between 1 and $r-1$,

$$B_{j_0}(x) \cap V_{j_0} = B_{j_0}(y) \cap V_{j_0}, \tag{3}$$

and

$$B_{j_0+1}(x) \cap V_{j_0} = B_{j_0+1}(y) \cap V_{j_0}. \tag{4}$$

Since there are no $r$-twins in $G$, there is a vertex $z \in V$ which, say, $r$-covers $x$, not $y$. Obviously, $z \neq a$ and $z \notin V_i$, $i \leq r-2$, otherwise $z$ would $r$-cover $y$; and $z \notin V_i$, $i \geq r+1$, otherwise $z$ would not $r$-cover $x$. So $z \in V_{r-1} \cup V_r$.

Assume first that $z \in V_r$. Then the distance between $x$ and $z$ is exactly $r$, and there is a path $x, a_1, a_2, \ldots, a_{r-1}, z$, where $a_i \in V_i$ is at distance $i$ from $x$ and at distance $r-i$ from $z$. In particular $a_{j_0}$ belongs to $B_{j_0}(x) \cap V_{j_0}$, which by (3) implies that $a_{j_0} \in B_{j_0}(y)$, i.e., $d(a_{j_0}, y) \leq j_0$ and finally $d(z,y) \leq (r-j_0)+j_0 = r$, a contradiction.

Next, assume that $z \in V_{r-1}$. Then the distance between $x$ and $z$ is $r-1$ or $r$. Consider a shortest path between $x$ and $z$; it goes through at least one vertex in $V_{j_0}$, which we call $a_{j_0}$. Now the distance between $a_{j_0}$ and $x$ is $j_0$ or $j_0+1$: if it were $j_0+2$, then the distance from $x$ to $z$ would be at least $(j_0+2)+(r-1-j_0) = r+1 > r$. If $d(a_{j_0}, x) = j_0$, i.e., $a_{j_0} \in B_{j_0}(x)$, then (3) shows that $d(a_{j_0}, y) = j_0$. If $d(a_{j_0}, x) = j_0+1$, then by (4), we have $d(a_{j_0}, y) = j_0+1$. In both cases, $d(a_{j_0}, x) = d(a_{j_0}, y)$, which implies that $d(y,z) \leq r$, again a contradiction.

Therefore we have proved that for any $i \in \{1, \ldots, r-1\}$, the couples $C_i(x)$, $x \in V_0$, are all distinct. Next, we show that, for any $x \in V_0$ and any $i \in \{1, \ldots, r-1\}$, we have $C_i(x) \neq C_i(a)$. Assume on the contrary that there is a vertex $x \in V_0$ such that $C_{j_0}(x) = C_{j_0}(a) = \big(\emptyset, V_{j_0}\big)$ for some $j_0$ between 1 and $r-1$.

Consider a vertex $z \in V_r$. We have $d(z,a) = r+1$ and $d(z,x) \geq r$. But the assumption that $B_{j_0}(x) \cap V_{j_0} = \emptyset$ shows that $d(z,x) \geq r+1$, and so $z$ $r$-covers neither $x$ nor $a$.

Consider a vertex $z \in V_{r-1}$. By property (P), we have $d(z,a) = r$ and $d(z,x) \leq r+1$. A path of length $r$ from $z$ to $a$ goes through a vertex $a_{j_0} \in V_{j_0}$. Obviously, $d(z, a_{j_0}) = r-1-j_0$. By assumption, $a_{j_0} \in V_{j_0}$ implies that $a_{j_0} \in B_{j_0+1}(x) \cap V_{j_0}$, so $d(a_{j_0}, x) \leq j_0+1$ and finally $d(z,x) \leq (r-1-j_0)+(j_0+1) = r$, i.e., $z$ $r$-covers both $x$ and $a$. Since this is also true for vertices $z \in V_i$ with $i < r-1$, we see that $a$ and $x$ are $r$-twins, a contradiction.

All in all, we have just proved that the couples $C_i(x)$, $i \in \{1, \ldots, r-1\}$, are all distinct when $x$ runs through $V_0 \cup \{a\}$.

When, inside $V_i$, the set $B_i(x) \cap V_i$ is fixed with size $s$, there are $2^{|V_i|-s}$ possibilities for the choice of $B_{i+1}(x) \cap V_i$, so there are at most

$$\sum_{0 \leq s \leq |V_i|} \binom{|V_i|}{s} 2^{|V_i|-s} = (2+1)^{|V_i|} = 3^{|V_i|}$$

different couples $C_i(x)$, for $x \in V_0 \cup \{a\}$ and $i = 1, \ldots, r-1$. This establishes that

$$|V_0 \cup \{a\}| = \deg(a) + 1 = \delta + 1 \leq 3^{|V_i|}, \tag{5}$$

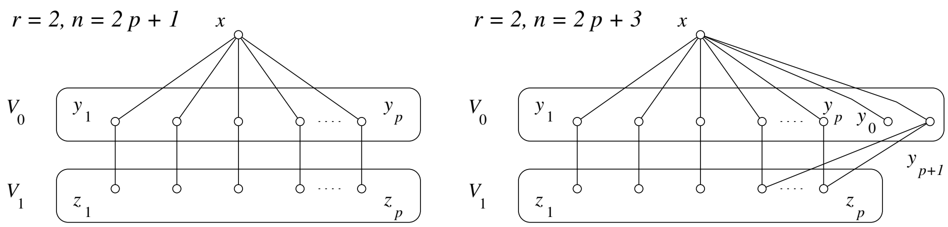


Figure 2: Constructions for Theorem 4, when $r = 2$, $n = 2p+1$ or $n = 2p+3$.

for every $i$ between 1 and $r-1$.

We have seen earlier that if $z$ $r$-covers $x \in V_0$ and not $y \in V_0$, then $z \in V_r \cup V_{r-1}$. The same is true for a vertex $r$-covering $x \in V_0$ and not $a$, or a vertex $r$-covering $a$ and not $x \in V_0$. This implies that

$$|V_0 \cup \{a\}| = \delta + 1 \leq 2^{|V_{r-1} \cup V_r|}. \tag{6}$$

Therefore, using (5) and (6), we obtain

$$1+\delta+(r-2)\lceil \log_3(\delta+1)\rceil+\lceil \log_2(\delta+1)\rceil \leq |V_0 \cup \{a\}|+\sum_{i=1}^{r-2}|V_i|+|V_{r-1}\cup V_r| \leq n,$$

which shows that any degree in the graph satisfies inequality (2). △

We now give constructions where there is one vertex with large degree. We distinguish between the cases $r = 2$ and $r \geq 3$, because in the case $r = 2$, we are able to determine the exact value for $F_{r,n}(\Delta_{\max})$.

**Theorem 4** *Let $n \geq 5$ and $p \geq 2$ be two integers. If $2p+1 \leq n \leq 2^p + p$, then there exists a connected 2-twin-free graph with $n$ vertices and maximum degree $n-p-1$.*

**Proof.** Giving overlapping intervals for $n$ makes the proof easier. We shall see in the next theorem that this construction is optimum.

We start from a basic construction for $n = 2p+1$, where $G = (V, E)$ with

$$V = \{x\} \cup \{y_i, z_i : 1 \leq i \leq p\} \text{ and } E = \{\{x, y_i\}, \{y_i, z_i\} : 1 \leq i \leq p\},$$

see the left part of Figure 2. We set $V_0 = \{y_i : 1 \leq i \leq p\}$ and $V_1 = \{z_i : 1 \leq i \leq p\}$. This construction is easily seen to give a 2-twin-free graph, and the degree of $x$ is $p = n-p-1$.

If $n = 2p+2$, then we can add a vertex $y_0$ in $V_0$, adjacent only to $x$. The resulting graph is still twin-free, and the degree of $x$ becomes equal to $p+1 = n-p-1$.

Similarly, when increasing $n$ up to $n = 2^p + p$, we can successively add vertices $y_{p+1}, y_{p+2}, \ldots$, in $V_0$, provided that each new vertex $y_j$, $j > p$, is linked to $x$ and to a subset of $V_1$ which is different from $V_1$ and different from $B_1(y_i) \cap V_1$ for any $y_i$ added so far in $V_0$ $(1 \le i < j)$. See the right part of Figure 2 when $n = 2p + 3$. The degree of $x$ is equal to $|V_0| = n - p - 1$, and we now show that the graph is 2-twin-free.

Two vertices $z_i, z_j \in V_1$, $1 \le i < j \le p$, cannot be twins, since $y_i$ is within distance two from $z_i$ and not from $z_j$.

Two vertices $z_i \in V_1, y_j \in V_0$, $1 \le i \le p$, $0 \le j \le n - p - 1$, cannot be twins, since any vertex $y_k$, $k \neq i, 0 \le k \le p$, is within distance two from $y_j$ and not from $z_i$.

Two vertices $y_i, y_j \in V_0$, $0 \le i < j \le n - p - 1$, cannot be twins, since $B_2(y_i) \cap V_1 = B_1(y_i) \cap V_1$, $B_2(y_j) \cap V_1 = B_1(y_j) \cap V_1$, and, by construction, these two sets are different.

Lastly, $B_2(x)$ is the set of all vertices, and since we forbade $B_1(y) \cap V_1 = V_1$ for any $y \in V_0$, $x$ is the only vertex in this case.

So the graph is 2-twin-free, and we see that we can have up to $2^p - 1$ vertices in $V_0$, leading to all the values of $n$ comprised between $2p + 1$ and $2^p + p$. △

We now show that this constructive lower bound is the exact value.

**Theorem 5** *For all $p \ge 2$ and $n \ge 5$, if $2^{p-1} + p - 1 < n \le 2^p + p$, then we have: $F_{2,n}(\Delta_{\max}) = n - p - 1$.*

**Proof.** For alleviation of notation, in this proof we denote $F_{2,n}(\Delta_{\max})$ simply by $F_2$. By Theorem 3, for all $n \ge 5$, we have:

$$F_2 + \lceil \log_2(F_2 + 1) \rceil \le n - 1. \tag{7}$$

We assume that $n$ satisfies the conditions of Theorem 5, and that $F_2 \ge n - p$. The former implies that

$$n - p + 1 > 2^{p-1} = (2^{p-1} + p - 1) - p + 1, \tag{8}$$

the latter that

$$F_2 + \lceil \log_2(F_2 + 1) \rceil \ge (n - p) + \lceil \log_2(n - p + 1) \rceil. \tag{9}$$

By (8), $\lceil \log_2(n-p+1) \rceil \ge p$, and, plugging this inequality into (9), we obtain $F_2 + \lceil \log_2(F_2 + 1) \rceil \ge n$, contradicting (7) and proving that $F_2 \le n - p - 1$. Combining this result with Theorem 4, we obtain our claim, provided that $n \ge 2p + 1$. Since $n > 2^{p-1} + p - 1$, this is always true, unless $p = 2$; in this case however, the fact that $n \ge 5$ implies that $n \ge 2p + 1$. △

We now show various constructions for $r \ge 3$, giving lower bounds which will be compared to the upper bound given by Theorem 3. Note that the cases $r = 3$ and $r > 3$ are somewhat different.

The intervals given in the following theorem are mutually exclusive and give, for each value of $r \geq 3$, all the values of $n$ starting from $2r+1$.

**Theorem 6** *(a) We consider the case $r = 3$.*

*(a1) For $n = 7, 8$, we have: $F_{3,n}(\Delta_{\max}) \geq n - 5$;*
*(a2) for $n = 9, 10, 11$, we have: $F_{3,n}(\Delta_{\max}) \geq n - 6$;*
*(a3) for $n = 12, 13$, we have: $F_{3,n}(\Delta_{\max}) \geq n - 7$;*
*(a4) for $14 \leq n \leq 18$, we have: $F_{3,n}(\Delta_{\max}) \geq n - 8$;*
*(a5) for $19 \leq n \leq 25$, we have: $F_{3,n}(\Delta_{\max}) \geq n - 9$;*
*(a6) for $26 \leq n \leq 33$, we have: $F_{3,n}(\Delta_{\max}) \geq n - 10$.*

*For all $p \geq 4$,*

*(a7) if $3^{p-1}+2p-1 \leq n \leq 4\times 3^{p-2}+2p+2$, then $F_{3,n}(\Delta_{\max}) \geq n-3p+1$;*
*(a8) if $4\times 3^{p-2}+2p+3 \leq n \leq 2\times 3^{p-1}+2p+1$, then $F_{3,n}(\Delta_{\max}) \geq n-3p$;*
*(a9) if $2 \times 3^{p-1} + 2p + 2 \leq n \leq 3^p + 2p$, then $F_{3,n}(\Delta_{\max}) \geq n - 3p - 1$.*

*(b) For all $p \geq 2$, we define $\lambda(p) = 1$ if $p = 2$, $0$ otherwise, and $\mu(p) = 1$ if $p = 3$, $0$ otherwise. For all $r \geq 4$, $n \geq 2r + 1$, and $p \geq 2$,*

*(b1) if $n = 2r + 1$, then $F_{r,n}(\Delta_{\max}) = 2$;*
*(b2) if $n = 2r + 2$, then $F_{r,n}(\Delta_{\max}) = 3$;*
*(b3) if $p = 3$ and $pr + 8 \leq n \leq pr + 9$, or if $p \geq 4$ and $3^{p-1} + pr - 1 \leq n \leq 4 \times 3^{p-2} + p(r-2) + 2$, then $F_{r,n}(\Delta_{\max}) \geq n - rp + 1$;*
*(b4) if $4 \times 3^{p-2} + p(r-2) + 3 + \mu(p) \leq n \leq 2 \times 3^{p-1} + p(r-2) + 3$, then $F_{r,n}(\Delta_{\max}) \geq n - rp$;*
*(b5) if $2\times 3^{p-1}+p(r-2)+4 \leq n \leq 3^p+p(r-2)+\lambda(p)$, then $F_{r,n}(\Delta_{\max}) \geq n - rp - 1$;*
*(b6) if $3^p + p(r-2) + 1 + \lambda(p) \leq n \leq 3^p + pr + 1$, then $F_{r,n}(\Delta_{\max}) \geq n - rp - 2$;*
*(b7) if $3^p + pr + 2 \leq n \leq 3^p + (p+1)r - 2$, then $F_{r,n}(\Delta_{\max}) \geq 3^p - 1$.*

**Sketch of proof.** If $n = 2r + 1$, then it is not difficult to see that the only connected $r$-twin-free graph is the path $P_{2r+1}$, which has maximum degree two. Thus claim (b1) is true.

If $n = 2r + 2$, then $k = 4$ does not satisfy inequality (2) in Theorem 3. On the other hand, a degree equal to three is possible, for instance in the following $r$-twin-free graph: $G = (V, E)$, where $V = \{x_i : 1 \leq i \leq 2r + 2\}$ and $E = \{\{x_i, x_{i+1}\} : 1 \leq i \leq 2r\} \cup \{\{x_{r+1}, x_{2r+2}\}\}$. Thus claim (b2) is true.

Now we give a basic construction, for $n$ comprised between $pr + 3$ and $3^p+pr+1$, in which there is a vertex with degree $n-pr-2$. This construction is easy to understand and will be the starting point for many variations. We start with $n = pr + 3$, and build the graph $G = (V, E)$, where

$$V = \{x, y_1, w\} \cup \{a_{i,j} : 1 \leq i \leq p, 1 \leq j \leq r\},$$

$$E = \{\{x, y_1\}\} \cup \{\{y_1, a_{i,1}\}, \{w, a_{i,r}\} : 1 \leq i \leq p\} \cup \{\{a_{i,j}, a_{i,j+1}\} :$$

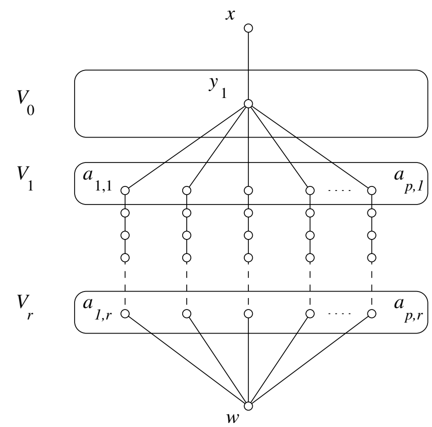


Figure 3: A first construction in the proof of Theorem 6.

$$1 \leq i \leq p, 1 \leq j \leq r-1\},$$

see Figure 3. It is rather straightforward to check that this graph is $r$-twin-free, because it mainly consists of chordless cycles of length $2r+2$. We set $I_0 = \{1, 2, \ldots, p\}$, $V_0 = \{y_1\}$, and $V_j = \{a_{i,j} : i \in I_0\}$, for $j = 1, 2, \ldots, r$.

In a way similar to the proof of Theorem 4, we can progressively add vertices in $V_0$, all adjacent to $x$, and keep the graph twin-free while increasing the degree of $x$. First, we add a vertex $y_0$ which is linked to $x$ only. Then we add vertices $y_{J,J}$, where the sets $J$ are distinct subsets of $I_0$ (different from $I_0$ and from the empty set), and where $y_{J,J}$ is linked to $x$ and to the vertices $a_{j,1} \in V_1$ for $j \in J$. With this notation, we see that $y_1 = y_{I_0,I_0}$ and $y_0 = y_{\emptyset,\emptyset}$.

Once we have all possible vertices $y_{J,J}$ in $V_0$, we can add even more vertices to $V_0$: we denote these vertices by $y_{I,J}$, where $J \subsetneq I \subseteq I_0$, $(I, J) \neq (I_0, \emptyset)$, and the couple $(I, J)$ has not been used already; any new vertex $y_{I,J}$ is linked to $x$, to $y_{I,I}$, and to the vertices $a_{j,1}$, for $j \in J$. See Figure 4 for an illustration.

We claim that all the successive graphs obtained from the graph given in Figure 3 by adding, one by one, vertices $y_{I,I}$, then, one by one, vertices $y_{I,J}$, as described above, are $r$-twin-free, have a number of vertices $n$ comprised between $pr+3$ and $3^p+pr+1$, and that the degree of $x$ is equal to $n-pr-2$.

First, a vertex in $V_0 \cup \{x\}$ and a vertex in $V_1 \cup \ldots \cup V_r \cup \{w\}$ cannot be twins, because the latter is $r$-covered by $w$ and the former is not. It is equally easy to see that $w$ cannot be the twin of any vertex. And $x$ has no twins either: the only candidate would be a vertex in $V_0$ which is linked to $x$ and to the only vertex in $V_0$ which is linked to all the vertices in $V_1$, namely, $y_{I_0,I_0}$; but precisely, we mentioned that we forbid the couple $(I, J) = (I_0, \emptyset)$. So all we have to check is that there are no twins inside $V_0$, or inside $V_1 \cup \ldots \cup V_r$.

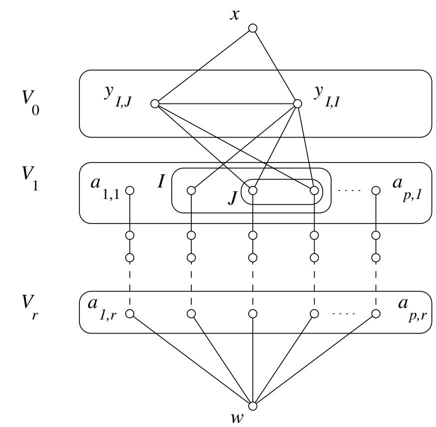


Figure 4: How the vertex $y_{I,J}$ is added to $V_0$, after $y_{I,I}$ has been. Here, the sets $I$ and $J$ in $V_1$ stand for $\{a_{i,1} : i \in I\}$ and $\{a_{i,1} : i \in J\}$, respectively.

It is quite straightforward to notice that for any vertex $y_{I,J} \in V_0$, where we allow $I = J$, we have:

$$B_r(y_{I,J}) = \{x\} \cup V_0 \cup \ldots \cup V_{r-2} \cup \big(V_{r-1} \cap \{a_{i,r-1} : i \in I\}\big) \cup \big(V_r \cap \{a_{i,r} : i \in J\}\big),$$

i.e.,

$$B_r(y_{I,J}) = \{x\} \cup V_0 \cup \ldots \cup V_{r-2} \cup \{a_{i,r-1} : i \in I\} \cup \{a_{i,r} : i \in J\}. \quad (10)$$

By the successive choices for the couples $(I, I)$ or $(I, J)$, these sets are all different, and therefore there are no twins inside $V_0$.

Two vertices $a_{i,j} \in V_j$ and $a_{i',j'} \in V_{j'}$ $(i \neq i')$ are not twins because $a_{i',r-j+1}$ $r$-covers $a_{i',j'}$ and not $a_{i,j}$; and two vertices $a_{i,j} \in V_j$ and $a_{i,j'} \in V_{j'}$ $(j \neq j')$ are not twins because if $i' \neq i$, then $a_{i',r-j+1}$ $r$-covers $a_{i,j'}$ and not $a_{i,j}$. This is easily understood if one sees any such two vertices as belonging to a chordless cycle of length $2r + 2$ (containing $w$ and a vertex in $V_0$, for instance $y_1 = y_{I_0,I_0}$). We can now conclude that our graph is twin-free.

How many vertices can we have? In $V_0$, we put vertices $y_{I,J}$, where $J \subseteq I \subseteq I_0$, with $(I, J) \neq (I_0, \emptyset)$. Therefore $V_0$ can contain up to

$$\Big(\sum_{k=0}^{p} \binom{p}{k} 2^k\Big) - 1 = (2 + 1)^p - 1 = 3^p - 1$$

vertices, and $n$ can be as large as $3^p + pr + 1$. In all cases, the degree of $x$ is $n - pr - 2$.

This ends the description of our basic construction, and proves the penultimate claim, (b6), in Theorem 6, since $pr + 3 \leq 3^p + p(r - 2) + 1$ for all $p \geq 2$.

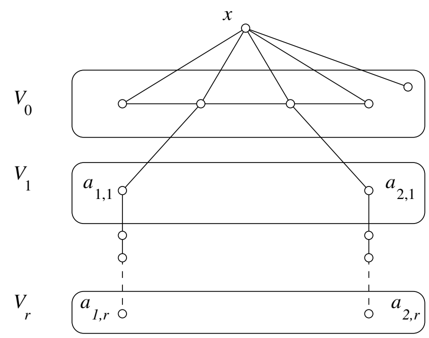


Figure 5: The particular case $n = 2r + 6$.

We denote by $G^{(n)} = (V^{(n)}, E^{(n)})$ the graph obtained, with $n$ ranging between $pr + 3$ and $3^p + pr + 1$.

There is a first trivial modification to $G^{(n)}$, which works best when $n = 3^p + pr + 1$: just add a "tail" to $w$, that is, a path $z_1, z_2, \ldots$, with $z_1$ linked to $w$. The degree of $x$ remains the same, equal to $3^p - 1$. This proves the last claim, (b7), in Theorem 6, since the tail can be as long as necessary.

Slightly more difficult is to try to remove the vertex $w$. If we remove $w$ from $G^{(3^p+pr+1)}$, then there are twins: the sets $B_r(a_{i,1})$, $i \in I_0$, consist of all vertices except $\{a_{j,r} : j \in I_0 \setminus \{i\}\}$, and if we compare to (10), which still holds after the removal of $w$, we see that $B_r(a_{i,1}) = B_r(y_{I_0,\{i\}})$. Therefore, in the basic construction of $G^{(n)}$, we are led to forbid the vertices $y_{I_0,\{i\}}$, $i \in I_0$. Similarly, the sets $B_r(a_{i,2})$ consist of all vertices except $\{a_{j,r}, a_{j,r-1} : j \in I_0 \setminus \{i\}\}$ when $r \geq 4$, whereas, when $r = 3$, in addition the vertex $y_{\emptyset,\emptyset}$ does not belong to any set $B_r(a_{i,2})$. We see that in the former case, $B_r(a_{i,2}) = B_r(y_{\{i\},\{i\}})$, and we are moreover led to banish the vertices $y_{\{i\},\{i\}}$. In conclusion, we modify $G^{(3^p+pr+1)}$ by removing the vertex $w$, but forbidding the $p$ vertices $y_{I_0,\{i\}}$ if $r = 3$, or the $2p$ vertices $y_{I_0,\{i\}}, y_{\{i\},\{i\}}$ if $r \geq 4$; we leave it to the reader to check that the new graph is indeed twin-free. If $r = 3$ (respectively, $r \geq 4$), this graph has $3^p + 2p$ (respectively, $3^p + (r-2)p$) vertices, the degree of $x$ is $3^p - p - 1 = n - 3p - 1$ (respectively, $3^p - 2p - 1 = n - rp - 1$). Obviously, more vertices $y_{I,J}$ can be removed, down to $pr + 2$. This proves claim (a9), since $3p + 2 \leq 2 \times 3^{p-1} + 2p + 2$ for all $p \geq 4$, and claim (b5) (except for $\lambda(p) = 1$, i.e., $p = 2$ and $n = 2r + 6$), since $pr + 2 \leq 2 \times 3^{p-1} + p(r-2) + 4$ for all $p \geq 2$. When $n = 2r + 6$, we use the special construction given by Figure 5, where $x$ has degree $5 = n - 2r - 1$; we leave the checking to the reader.

Still more difficult are the proofs of the remaining claims, of which we only give a sketch.

One idea is to remove the vertex $w$ and merge two vertices, $a_{1,r}$ and $a_{2,r}$, into one, see Figure 6(a). Then of course, vertices have to be removed in $V_0$,

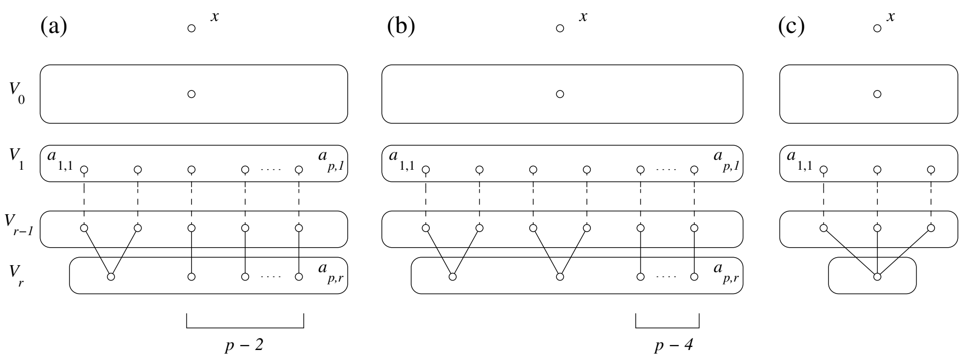


Figure 6: Different ways of contracting vertices in the proof of Theorem 6.

or, equivalently, couples $(I, J)$ must be forbidden. This is where we drop the delicate details, only to mention that the set $V_0$ can have up to $2\times 3^{p-1}-p+1$ vertices if $r = 3$ and up to $2 \times 3^{p-1} - 2p + 3$ vertices if $r \geq 4$. In the former case, the graph has between $3p + 1$ and $2 \times 3^{p-1} + 2p + 1$ vertices, in the latter case, it has between $rp + 1$ and $2 \times 3^{p-1} + (r - 2)p + 3$ vertices, and in both cases $x$ has degree $n - rp$, which establishes claims (a8) and (b4).

Another idea is, after removing $w$, to merge two vertices twice, namely, $a_{1,r}$ and $a_{2,r}$ on the one hand, and $a_{3,r}$ and $a_{4,r}$ on the other (so necessarily, $p \geq 4$), see Figure 6(b). This leads to graphs with a number of vertices ranging between $3p$ and $4 \times 3^{p-2} + 2p + 2$ if $r = 3$, or between $pr$ and $4\times 3^{p-2}+p(r-2)+2$ if $r \geq 4$. In both cases, $x$ has degree $n-rp+1$, which treats claim (a7) and the second part of claim (b3).

When $p = 3$ and $r \geq 4$, it is possible to merge the three vertices $a_{1,r}$, $a_{2,r}$, and $a_{3,r}$ into one, see Figure 6(c), to obtain the first part of claim (b3).

Finally, claims (a1)–(a6) are shown using the same kinds of constructions. There is no point describing these constructions here. One of the difficulties, as for the previous constructions, is to find the most efficient balance between the number of vertices in $V_r$ and in $V_0$. △

Tables 7 and 8 give the results obtained by Theorems 3 and 6, for $r = 3, 4, 5$ and 10 and $2r + 1 \leq n \leq 66$, as well as for some larger values of $n$.

Note that we have the *exact* value of $F_{r,n}(\Delta_{\max})$ for *infinitely many* values of $n$ and $r$; for instance, if $k = 3^p - 1$, $p \geq 1$, then (2) in Theorem 3 is satisfied by $k$ and not by $k + 1$ for $r \geq 2$ and $n$ between

$$A = 3^p + pr - (2p - \lceil p \log_2 3\rceil) \text{ and } B = 3^p + pr + r - 2 - (2p - \lceil p \log_2 3\rceil),$$

and $F_{r,n}(\Delta_{\max}) \leq 3^p - 1$ for these values. On the other hand, using Theorem 6(b7), and 6(b6) with $n = 3^p + pr + 1$, shows that $F_{r,n}(\Delta_{\max}) \geq 3^p - 1$ for $r \geq 4$ and $n$ between

$$C = 3^p + pr + 1 \quad \text{and} \quad D = 3^p + pr + r - 2.$$

| $n$ | $r=3$ | $r=4$ | $r=5$ | $r=10$ | $n$ | $r=3$ | $r=4$ | $r=5$ | $r=10$ |
|---|---|---|---|---|---|---|---|---|---|
| 7 | **2** | | | | 37 | 26–27 | 23–25 | 20–22 | **8** |
| 8 | **3** | | | | 38 | 27–28 | 24–26 | 21–23 | **9** |
| 9 | **3** | **2** | | | 39 | 28–29 | 25–26 | 22–24 | **10** |
| 10 | **4** | **3** | | | 40 | 29–30 | **26** | 23–25 | 10–11 |
| 11 | **5** | **3** | **2** | | 41 | 30–31 | 26–27 | 24–26 | 11–12 |
| 12 | 5–6 | **4** | **3** | | 42 | **31** | 27–28 | 25–26 | 12–13 |
| 13 | 6–7 | **5** | **3** | | 43 | **32** | 28–29 | **26** | 13–14 |
| 14 | 6–7 | 5–6 | **4** | | 44 | **33** | 29–30 | **26** | 14–15 |
| 15 | 7–8 | 5–7 | **5** | | 45 | **34** | 30–31 | 26–27 | **15** |
| 16 | **8** | 6–7 | 5–6 | | 46 | **35** | **31** | 27–28 | 15–16 |
| 17 | **9** | 7–8 | 5–7 | | 47 | 35–36 | 31–32 | 28–29 | 16–17 |
| 18 | **10** | **8** | 6–7 | | 48 | 36–37 | 32–33 | 29–30 | 17–18 |
| 19 | 10–11 | **8** | 7–8 | | 49 | 37–38 | 33–34 | 30–31 | 18–19 |
| 20 | 11–12 | **9** | **8** | | 50 | 38–39 | 34–35 | **31** | 19–20 |
| 21 | 12–13 | **10** | **8** | **2** | 51 | 39–40 | 35–36 | 31–32 | 20–21 |
| 22 | 13–14 | 10–11 | **8** | **3** | 52 | 40–41 | 36–37 | 32–33 | 20–22 |
| 23 | 14–15 | 11–12 | **9** | **3** | 53 | 41–42 | 37–38 | 33–34 | 21–23 |
| 24 | **15** | 12–13 | **10** | **4** | 54 | 42–43 | 38–39 | 34–35 | 22–24 |
| 25 | **16** | 13–14 | 10–11 | **5** | 55 | 43–44 | 39–40 | 35–36 | 23–25 |
| 26 | 16–17 | 14–15 | 11–12 | 5–6 | 56 | 44–45 | 40–41 | 36–37 | 24–26 |
| 27 | 17–18 | **15** | 12–13 | 5–7 | 57 | 45–46 | 41–42 | 37–38 | 25–26 |
| 28 | 18–19 | 15–16 | 13–14 | 6–7 | 58 | 46–47 | 42–43 | 38–39 | **26** |
| 29 | 19–20 | 16–17 | 14–15 | 7–8 | 59 | 47–48 | 43–44 | 39–40 | **26** |
| 30 | 20–21 | 17–18 | **15** | **8** | 60 | 48–49 | 44–45 | 40–41 | **26** |
| 31 | 21–22 | 18–19 | 15–16 | **8** | 61 | 49–50 | 45–46 | 41–42 | **26** |
| 32 | 22–23 | 19–20 | 16–17 | **8** | 62 | 50–51 | 46–47 | 42–43 | **26** |
| 33 | 23–24 | 20–21 | 17–18 | **8** | 63 | 51–52 | 47–48 | 43–44 | **26** |
| 34 | 23–25 | 20–22 | 18–19 | **8** | 64 | 51–53 | 48–49 | 44–45 | **26** |
| 35 | 24–26 | 21–23 | 19–20 | **8** | 65 | 52–54 | 49–50 | 45–46 | 26–27 |
| 36 | 25–26 | 22–24 | 20–21 | **8** | 66 | 53–55 | 49–51 | 46–47 | 27–28 |

Figure 7: Some lower and upper bounds on the maximum degree.

| $n$ | $r=3$ | $r=4$ | $r=5$ | $r=10$ |
|---|---|---|---|---|
| 1 000 | 979–982 | 972–975 | 965–968 | 931–933 |
| 10 000 | 9 973–9 976 | 9 964–9 967 | 9 955–9 958 | 9 910–9 913 |
| 20 000 | 19 971–19 974 | 19 961–19 964 | 19 951–19 954 | 19 901–19 904 |
| 100 000 | 99 967–99 971 | 99 956–99 960 | 99 945–99 949 | 99 890–99 894 |

Figure 8: Some lower and upper bounds on the maximum degree.

Now we have: $A \leq C, B \leq D$, and, for $r \geq 3 + 2p - \lceil p \log_2 3 \rceil$: $C \leq B$. Therefore, for $r \geq 3 + 2p - \lceil p \log_2 3 \rceil$ and $n$ between $C$ and $B$, the value of $F_{r,n}(\Delta_{\max})$ is $3^p - 1$.

For instance we see in Table 7 that $F_{10,n}(\Delta_{\max}) = 26$ for $n$ between 58 and 64. Another example, not from the tables, is $p = 10, r = 100$, $60050 \leq n \leq 60143$, and $F_{100,n}(\Delta_{\max}) = 59048$.

We conclude this section by grossly approximating the lower and upper bounds given by Theorems 6 and 3, then their difference, for $r \geq 3$ fixed and $n$ sufficiently large.

Inequalities in Theorem 6, where $p$ is approximately equal to $\log_3 n$, show that $F_{r,n}(\Delta_{\max})$ is approximately greater than

$$n - r \log_3 n. \tag{11}$$

On the other hand, we can estimate the greatest integer $k$ satisfying (2) in Theorem 3 by approximating it with the greatest $k'$ such that

$$k' + (r-2) \log_3 k' + \log_2 k' \leq n,$$

which, if we set $s = 2 - \log_2 3$ (the value of which is around 0.4), reads

$$k' + (r-s) \log_3 k' \leq n.$$

Now let $k_1 = n - (r-s) \log_3 n$ and $k_2 = n - (r-s) \log_3 n + (r-s)$. We see that

$$k_1 + (r-s) \log_3 k_1 \leq k_1 + (r-s) \log_3 n = n,$$

which shows that $k' \geq k_1$. And, if $r$ is fixed and $n$ grows (and even with weaker constraints actually), then $\log_3 k_2 \geq \log_3 n - 1$; therefore,

$$k_2 + (r-s) \log_3 k_2 \geq k_2 + (r-s) \log_3 n - (r-s) = n,$$

so $k' \leq k_2$. This shows that $k'$, hence $k$, behaves approximately like

$$n - (r-s) \log_3 n, \tag{12}$$

which, comparing to the approximate lower bound $n - r \log_3 n$, shows that the difference between lower and upper bounds can be roughly estimated by $0.4 \times \log_3 n$ (independent of $r$). This could already be empirically observed for the large values of $n$ given in Table 8.

# 3 Conclusion

We have a fairly good estimation of $F_{r,n}(\Delta_{\max})$. The table below recapitulates some of the results obtained, for $r = 1$, $r = 2$, and $r \geq 3$.

| $r$ | $f_r(\Delta_{\max})$ | $f_{r,n}(\Delta_{\max})$ | $F_{r,n}(\Delta_{\max})$ |
|---|---|---|---|
| 1 | 2 [Th. 1] | 2 [Th. 1] | $n-1$ [Th. 2] |
| 2 | 2 [Th. 1] | 2 [Th. 1] | $n-p-1$ for $2^{p-1}+p-1<n\leq 2^p+p$ [Th. 5] |
| $\geq 3$ | 2 [Th. 1] | 2 [Th. 1] | $\gtrsim n-r\log_3 n$ [Th. 6 with (11)]<br>$\lesssim n-(r-0.4)\log_3 n$ [Th. 3 with (12)] |

# References


[1] D. Auger, Problèmes d'identification métrique dans les graphes, Rapport Interne ENST-2007D013, Paris, France, 44 pp., 2007.

[2] D. Auger, Induced paths in twin-free graphs, submitted.

[3] I. Charon, I. Honkala, O. Hudry and A. Lobstein, Structural properties of twin-free graphs, Electron. J. Combin. 14(1) (2007), R16.

[4] I. Charon, O. Hudry and A. Lobstein, Extremal values for identification, domination and maximum cliques in twin-free graphs, Ars Combin., to appear.

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

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

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

[8] J. Moncel, Codes identifiants dans les graphes, Thèse de Doctorat de l'Université de Grenoble, France, 165 pp., 2005.

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

Dépôt légal : 2008 – 2ème trimestre
Imprimé à l'Ecole Nationale Supérieure des Télécommunications – Paris
ISSN 0751-1345 ENST D (Paris) (France 1983-9999)

**TELECOM ParisTech**

Institut TELECOM - membre de ParisTech

46, rue Barrault - 75634 Paris Cedex 13 - Tél. + 33 (0)1 45 81 77 77 - www.telecom-paristech.frfr

**Département INFRES**