Identifying codes and searching with balls in graphs

Younjin Kim[*]   Mohit Kumbhat[†]   Zoltán Lóránt Nagy[‡]

Balázs Patkós[§]   Alexey Pokrovskiy [¶]   Máté Vizer[‖]

June 2, 2014

Abstract

Given a graph G and a positive integer R we address the following combinatorial search theoretic problem: What is the minimum number of queries of the form "does an unknown vertex $v \in V(G)$ belong to the ball of radius r around u?" with $u \in V(G)$ and $r \leq R$ that is needed to determine v. We consider both the adaptive case when the jth query might depend on the answers to the previous queries and the non-adaptive case when all queries must be made at once. We obtain bounds on the minimum number of queries for hypercubes, the Erdős-Rényi random graphs and graphs of bounded maximum degree .

Keywords: identifying codes, combinatorial search, hypercube, Erdős-Rényi random graph, bounded degree graphs

1   Introduction

Combinatorial search theory is concerned with problems of the following type "given a finite set S and an unknown element $x \in S$, find x as quickly as possible". In order to find x, we are allowed to ask questions of the type "is x contained in a subset $B \subseteq S$?" for various

$B \subseteq S$. There are many real-world search problems which are of this type. For example given a set of blood samples, we might want to identify one which is infected. Or given a building, we might want to determine which room a person is hiding in.

Notice that to guarantee finding the unknown element x in a set S we always need to ask at least $\log_2 |S|$ questions. Indeed, to be able to distinguish between two elements $x, x' \in S$, we must at some point query a set $B \subseteq S$ such that $x \in B$ and $x' \notin B$. However if we ask less than $\log_2 |S|$ questions, then by the Pigeonhole Principle, there would be two elements x and $x' \in S$ which receive the same sequence of yes/no answers.

The lower bound $\log_2 |S|$ on the number of questions can be tight. In fact it is always possible to find x using $\lceil \log_2 |S| \rceil$ questions. Indeed first we divide S is half and ask which half x is in. Then we divide the half containing x in half again, and find out which one x is in. Proceeding this way, we reduce the subset of S which x may be in by half at every step, and hence we find x after at most $\lceil \log_2 |S| \rceil$ queries. This is the so called *halving search*.

Notice that in the above search the set we query at each step depends on the answers we got in the previous steps. We say that a search is *adaptive* if a question is allowed to depend on answers to previous questions. In practice we often would like to make all our queries at the same time (for example if each individual query takes a long time, it would save time to make queries simultaneously). We say that a search is *non-adaptive* if the queries do not depend on answers obtained from previous ones. Just like in the adaptive case, there exist non-adaptive searches using the optimal $\log_2 |S|$ number of queries. For example if S is the set of $0 - 1$ sequences of length n, then we can find $x = (x_1, \ldots, x_n)$ using the $n = \log_2 |S|$ queries "is x_i equal to zero or one?"

Note that the answers to a set Q of query sets surely determine the unknown element x if and only if for any pair of elements $s_1, s_2 \in S$ there exists a query $B \in Q$ with $|B \cap \{s_1, s_2\}| = 1$. In this case we say that B *separates* s_1 and s_2, and Q is said to be a *separating system* or said to possess the *separating property*.

In this paper we study a search problem when the set S is a metric space, and we are only allowed to query sets B which are balls in the metric space. We call this the *ball search*. There are natural search problems of this type. For example if we want to track someone moving through a building, we might set up sensors at various points in the building which tell us if the person is within a certain range. This problem is inherently non-adaptive since the person is moving (and it would usually be inpractical to change the location of the sensors as the person moves). Another example of a ball search is to locate some object in the night sky using a telescope. This problem is inherently adaptive since once we capture the object with the telescope, we would want to zoom in in order to get a better idea of its location.

Ball searches have been studied before, under the name of *identifying codes on graphs*. A ball $B_r(v)$ of radius r at vertex v in a graph G is the set of vertices in G that have distance at most r from v. An $(r, \leq l)$-identifying code of the graph G is a set C of vertices such that for every pair of distinct subsets $X, Y \subset V(G)$ with $|X|, |Y| \leq l$ the sets $B_r(X) \cap C$ and $B_r(Y) \cap C$ are different and non-empty. Notice that given a graph with an $(r, \leq 1)$-identifying code C, we can always find an unknown vertex using the nonadaptive search querying the balls $\{B_r(c) : c \in C\}$. Identifying codes were introduced by Karpovsky, Chakrabarty, and Levitin [14] and the problem of determining the minimum size $i_r^{(l)}(G)$ of an identifying code in G has since attracted the attention of many researchers (for a full bibliography see [19]).

An adaptive version of identifying codes was introduced by Moncel in his doctoral thesis [20] and in [2]. The minimum number of queries needed in the adaptive case is denoted by $a_r^{(l)}(G)$.

We will study a slight variant of identifying codes in which the restriction that the balls in the search must cover *all* the vertices of G is omitted. The reason we study this variant is that it seems to have links to a natural combinatorial object (namely the Fano Plane. See Section 2 for details). We denote by $M(G, r)$ the minimum number of queries needed to nonadaptively find an unknown vertex $v \in G$ using balls of radius r. Similarly we denote by $A(G, r)$ the minimum number of queries needed to adaptively find an unknown vertex $v \in G$ using balls of radius r. Notice that we always have $i_r^{(1)}(G) \leq M(G, r) \leq i_r^{(1)}(G) + 1$ and $a_r^{(1)}(G) \leq A(G, r) \leq a_r^{(1)}(G) + 1$. Therefore, since we will mostly be interested in asymptotic estimates, it usually won't matter whether we study the quantity $M(G, r)$ or $i_r^{(1)}(G)$.

Notice that in the above ball searches we specified that we could only query balls of radius *exactly* r. With this restriction it is not always possible to find an unknown vertex v—for example if r is bigger than the diameter of G (the maximum distance between two vertices in G), then querying balls of radius r gives no information about the location of v. At the 5th Emléktábla Workshop, Gyula Katona asked what happens [16] if balls of radius *at most* r should be allowed as queries. We'll denote by $M(G, \leq r)$ and $A(G, \leq r)$ the minimum number of queries to find an unknown vertex $v \in G$ using balls of radius r in the nonadaptive and adaptive searches respectively. We'll denote by $M(G)$ and $A(G)$ the minimum number of queries to find an unknown vertex $v \in G$ using balls (of any radius) in the nonadaptive and adaptive searches respectively. Notice that since balls of radius zero are just single vertices in G, we always have $M(G, \leq r), A(G, \leq r) \leq |G|$. Also, we trivially have $M(G, \leq r) \leq M(G, r)$ and $A(G, \leq r) \leq A(G, r)$. The quantities $M(G, \leq r)$ and $A(G, \leq r)$ are monotonic in the sense that $M(G) \leq M(G, \leq r + 1) \leq M(G, \leq r)$ and $A(G) \leq A(G, \leq r + 1) \leq A(G, \leq r)$ hold for all r.

The first metric spaces in which we will study the ball search are hypercubes. Let Q_n be the n dimensional hypercube—the set of all $0 - 1$ vectors of length n. We define a graph on Q_n by placing an edge between two vectors whenever they differ in exactly one entry. In this graph the distance $d(\underline{u}, \underline{v})$ between two vertices is exactly the *Hamming distance* between $\underline{u}$ and $\underline{v}$ i.e. the number of entries on which $\underline{u}$ and $\underline{v}$ differ. If $\underline{c}$ is an element of Q_n and r is a non-negative integer, then $B(\underline{c}, r)$ denotes the ball of center $\underline{c}$ and radius r, that is, $B(\underline{c}, r) := \{\underline{v} : \underline{v} \in Q_n, d(\underline{c}, \underline{v}) \leq r\}$. The most recent upper and lower bounds on $i_r^{(l)}(Q_n)$ were proved in [4, 8, 6, 7, 12]. Adaptive identification in Q_n was studied by Junnila [13] who obtained lower and upper bounds on $a_1^{(1)}(Q_n)$. To be able to state our results let $K(n, r)$ denote the minimum number of balls of radius r that cover all vertices of Q_n. The centers of such balls are said to form *covering codes* in Q_n [5].

Theorem 1.1. *The functions $A(Q_n)$, $A(Q_n, \leq r)$ and $A(Q_n, r)$ satisfy the following inequalities:*

(i) $n \leq A(Q_n) \leq n - 1 + \lceil \log(n + 1) \rceil$.

(ii) $K(n, r) - 1 \leq A(Q_n, \leq r) \leq K(n, r) + n - 1 + \lceil \log(r + 1) \rceil$. *There is a constant C such that if $r \leq n/2 - C\sqrt{n}\log n$, then $A(Q_n, \leq r) = (1 + o(1))K(n, r)$.*

(iii) $K(n,r) - 1 \leq A(Q_n, r) \leq K(n,r) + \sum_{i=0}^{r} \binom{n}{r}$. *There is a constant C' such that if $r \leq C'n$, then $A(Q_n, r) = (1 + o(1))K(n,r)$.*

In Section 2, we will also study $M(Q_n)$ for small values of n.

Then we turn our attention to the Erdős-Rényi random graph model $G(n,p)$. The asymptotics of $i_1^{(1)}(G(n,p))$ was determined by Frieze, Martin, Moncel, Ruszinkó, and Smyth [10] provided $p, 1 - p \geq 4 \log \log n / \log n$ holds. Using results of Bollobás [3] and Katona [15] we obtain upper and lower bounds on $M(G(n,p), r)$ that differ by an $O(\ln n)$-factor for any r that is smaller than the diameter of $G(n,p)$ whenever $p = \Omega(n^{-1+\varepsilon})$. As the formulation of our theorem is somewhat technical, we defer its statement until Section 3.

Finally, we consider graphs of bounded maximum degree. Results concerning the parameter $i_1^{(1)}(G)$ for such graphs, for random d-regular graphs and for some other graph classes defined by their degree sequences were obtained by Foucaud and Perarnau [9]. Let us denote by $\mathcal{G}_{\Delta,n}$ the set of finite graphs on n vertices with maximum degree at most Δ.

Theorem 1.2. *For all connected graphs $G \in \mathcal{G}_{\Delta,n}$ we have $A(G) = \Theta_\Delta(\log n)$.*

Notice that the connectivity assumption cannot be ommited in the above theorem. Indeed, otherwise G could be a union of disjoint cliques on $\Delta + 1$ vertices. A search in this graph must ask at least one query in each component, and hence use at least $|G|/(\Delta+1) = O_{Delta}(n)$ queries. Similarly, the bound on the maximal degree of G cannot be removed since an optimal search on a star with n vertices needs $n - 1$ queries. Finally, note that Theorem 1.2 cannot be strengthened to give a logarithmic bound on the length of a non-adaptive search on bounded degree graphs. Indeed it is easy to show that $M(P_n) = \lceil \frac{n-1}{2} \rceil$, and so even for graphs with $\Delta = 2$, the nonadaptive search may need linearly many queries.

Let us observe the connection of non-adaptive ball search to another well studied parameter of graphs. A set D of vertices of G is called a *resolving set* of G if for any vertices $x, y \in V(G)$ there exists $z \in D$ such that $d(x,z) \neq d(y,z)$. The minimum size of a resolving set is called the *metric dimension of G* (introduced in [11, 21]) and is denoted by $\beta(G)$. Clearly, if a set of ball queries finds an unknown vertex v, then the centers of the balls form a resolving set in G and thus we have $\beta(G) \leq M(G)$. Also, if D is a resolving set in G, then the query set $\{B_G(d, i) : d \in D, 1 \leq i \leq diam(G)\}$ finds the unknown vertex of G, where $diam(G)$ denotes the diameter of G. Consequently, we have $M(G) \leq diam(G)\beta(G)$.

Throughout the paper "log" stands for the logarithm of base 2, and "ln" denotes the natural logarithm.

2 The hypercube

In this section we consider the adaptive search on Q_n. We find bounds on $A(Q_n)$ and $A(Q_n, \leq r)$. We also find a code in Q_7 of balls of radius 3 that is in some sense better than the best previously known one. We start by proving Theorem 1.1.

Throughout this section we will denote the number of vertices in a ball of radius r in Q_n by $V(n,r) = \sum_{i=0}^{r} \binom{n}{r}$. Recall that that the following bound on $V(v,r)$ follows from the Chernoff Bound (see [1]) $|V(n,r)| \leq 2^n \exp(-(n - 2r)/\sqrt{n})$.

Proof of Theorem 1.1:   (i) For the lower bound notice that the sequence of answers must be different for each vertex. Therefore there must be at least $\log(|Q_n|) = n$ questions.

For the upper bound, let $\underline{u}^*$ be the unknown element of Q_n. We give an adaptive algorithm that finds $\underline{u}^*$ and uses at most $n - 1 + \lceil \log(n+1) \rceil$ queries.

The first part of our algorithm uses $\lceil \log(n+1) \rceil$ queries to determine the distance $d(\underline{0}, \underline{u}^*)$. Notice that asking a ball $B(\underline{0}, r)$ is equivalent to asking whether we have $d(\underline{0}, \underline{u}^*) \le r$ or $d(\underline{0}, \underline{u}^*) > r$. Note that $d(\underline{0}, \underline{u}^*) > r$ can be anything in the set $\{0, 1, \ldots n\}$. In the first query we ask the ball $B(\underline{0}, \lfloor n/2 \rfloor)$. Depending on whether we get a "yes" or "no" answer, we obtain that $d(\underline{0}, \underline{u}^*)$ is either in $\{0, \ldots \lfloor n/2 \rfloor\}$ or in $\{\lfloor n/2 \rfloor + 1, \ldots, n\}$. Continuing in this fashion, always asking $B(\underline{0}, r)$ where r is the median of the possible (not already excluded) distance values, we see that $\lceil \log(n+1) \rceil$ determine $d(\underline{0}, \underline{u}^*)$. Let us denote this distance by d.

To finish the algorithm we show that for every $1 \le i \le n$ one further query is enough to determine the value of $\underline{u}^*$'s ith coordinate. Indeed, $\underline{u}^* \in B(\underline{e}_i, d-1)$ holds if and only if the ith coordinate of $\underline{u}^*$ is 1. We ask the balls $B(\underline{e}_i, d-1)$ for $i = 1, \ldots, n-1$ in order to determine $\underline{u}_1^*, \ldots, \underline{u}_{n-1}^*$. Since $\underline{u}_n^* = d - \sum_{i=1}^{n-1} \underline{u}_i^*$, this is enough to determine $\underline{u}^*$.

(ii) For the lower bound, observe that if after asking balls $B_1, B_2, \ldots, B_t$ we are able to determine the unknown vertex, then at most one vertex is not covered by the union of the balls.

For the upper bound, consider the following algorithm. First, let us ask a set of balls $\mathcal{B}$ of radius r that cover all vertices and with $|\mathcal{B}| = K(n, r)$. After finding a ball $B \in \mathcal{B}$ containing the marked vertex, we repeat the algorithm of part (i) with the modification that we know that the marked vertex is of Hamming distance at most r from the center of B. Therefore, to determine this distance exactly, we only need $\lceil \log(r+1) \rceil$ queries.

We now show that there is a constant C such that if $r \le n/2 - C\sqrt{n} \log n$, then $A(Q_n, \le r) = (1 + o(1))K(n, r)$. The exact value of $K(n, r)$ is known only for a short range of n and r and even determining the asymptotics of $K(n, r)$ is one of the major open problems of coding theory. For an introduction to covering codes, see the book [5]. Clearly $K(n, r) \ge 2^n / V(n, r)$ holds. It is known that for any fixed r, we have $K(n, r) = \Theta(2^n / V(n, r))$ (the best known constant is due to Krivelevich, Sudakov and Vu [18]) and that for any function $r = r(n)$ we have $K(n, r) = O(n2^n / V(n, r))$.

By the above reasoning we obtained that to prove $A(Q_n, \le r) = (1 + o(1))K(n, r)$, it is sufficient to show that $n = o(2^n / V(n, r))$ holds. This holds since we have $2^n / V(n, r) \ge \exp((n - 2r)/\sqrt{n})$ and $r \le n/2 - C\sqrt{n} \log n$.

(iii) The lower bound follows from part (ii). For the upper bound, proceed as we did in part (ii) to use $K(n, r)$ queries to find one ball of radius r containing the marked vertex. After that we can eliminate one possible vertex by each further query thus finding $\underline{u}^*$ with $K(n, r) + V(n, r)$ queries.

The second part follows from the fact that if $r = r(n) = c_n n$ with $c_n \to 0$, then $\log V(n, r)/n \to 0$.

$\square$

We finish this section with considering the special case $n = 7$. There are many results on $i_r^{(1)}(Q_n)$ for small values of r and n. In [4], Charon, Cohen, Hudry, and Lobstein proved that $i_3^{(1)}(Q_7) = 8$. This is true, but the lower bound is due to the fact that the definition of identifying code includes the condition that balls around the codewords should cover all vertices of the graph. On the other hand, no 7 of their 8 codewords possess the separation property. This can be checked easily as if they did, then for any two of them $B(\underline{c_1}, 3) \cap B(\underline{c_2}, 3), B(\underline{c_1}, 3) \setminus B(\underline{c_2}, 3), B(\underline{c_2}, 3) \setminus B(\underline{c_1}, 3), V(Q_7) \setminus (B(\underline{c_1}, 3) \cup B(\underline{c_2}, 3))$ should have size 32 and consequently those 7 codewords should have pairwise Hamming distance 4. In the remainder of this section we show a construction of seven codewords with the separation property.

Lemma 2.1. *There exist vertices $\underline{v_1}, \underline{v_2}, ..., \underline{v_7} \in V(Q_7)$ such that the balls $\{B(\underline{v_i}, 3), i = 1, 2, \ldots, 7\}$ separate $V(Q_7)$.*

Proof. Consider the incidence matrix of the Fano plane $PG(2, 2)$ consisting of 7 rows and columns corresponding to its line and point set, see Figure 1. Let v_i be the ith row corresponding to the ith line in the finite geometry. We show that the following query set provides a separating ball system, implying the lemma:

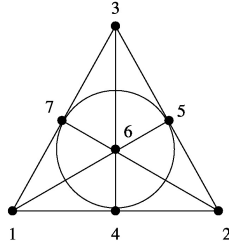$$\{B(\underline{v_i}, 3) : i = 1 \ldots 7\}.$$

Figure 1: Fano plane of order 2

Let $\underline{u^*}$ denote the marked element in Q_7 and $S(\underline{u^*})$ denote the subset of the set of balls $\{B(\underline{v_i}, 3) : i = 1 \ldots 7\}$ for which $\underline{u^*} \in B(\underline{v_i}, 3)$. We show that any subset S is squarely determined by the choice of $\underline{u^*}$. To this end, we study the connection between the support of $\underline{u^*}$ and the cardinality of $S(\underline{u^*})$.

Observation 2.2. .

(0) $|supp(\underline{u^*})| = 0 \rightarrow |S(\underline{u^*})| = 7$.

(1) $|supp(\underline{u^*})| = 1 \rightarrow |S(\underline{u^*})| = 3$, and $S(\underline{u^*})$ consists of the balls, whose centers' support contains $supp(\underline{u^*})$.

(2) $|supp(\underline{u^*})| = 2 \rightarrow |S(\underline{u^*})| = 5$, and $S(\underline{u^*})$ consists of the balls, whose centers' support intersects $supp(\underline{u^*})$.

(3) $|supp(\underline{u^*})| = 3$ and $supp(\underline{u^*})$ corresponds to a line in the Fano plane, i.e. $\underline{u^*} = v_j$ $\rightarrow |S(\underline{u^*})| = 1$ and $S(\underline{u^*}) = \{B(\underline{v_j}, 3)\}$.

(3') $|supp(\underline{u}^*)| = 3$ and $supp(\underline{u}^*)$ does not correspond to any line in the Fano plane, $\rightarrow$ $|S(\underline{u}^*)| = 3$ and $S(\underline{u}^*)$ consists of the balls, whose centers' support intersection with $supp(\underline{u}^*)$ has cardinality 2. (Hence, the intersection of the supports of all these 3 centers is the empty set.)

Claim 2.3. $S(\underline{u}^*) \dot{\bigcup} S(\overline{\underline{u}^*}) = \{B(\underline{v_i}, 3) : i = 1 \ldots 7\}$.

Proof. $d(\overline{\underline{u}}, \underline{v}) + d(\underline{u}, \underline{v}) = 7$ for all $\underline{u}, \underline{v} \in Q_7$, thus exactly one of the summands is less then or equal to 3. $\qquad \square$

Corollary 2.4. $|supp(\underline{u}^*)| > 3 \rightarrow |S(\underline{u})|$ *is even.*

To complete the proof of Lemma 2.1, it is enough to verify that no set $S \subseteq \{B(\underline{v_i}, 3) : i = 1 \ldots 7\}$, $|S|$ odd, is assigned to different elements of Q_7. Indeed, the other case $S \subseteq \{B(\underline{v_i}, 3) : i = 1 \ldots 7\}$, $|S|$ even will follow from this in view of Claim 2.3, since $S(\underline{u}^*) = S(\underline{u}')$ is equivalent to $S(\overline{\underline{u}^*}) = S(\overline{\underline{u}'})$, where $|S(\overline{\underline{u}^*})| = 7 - |S(\underline{u}^*)|$ is odd if $|S(\underline{u}^*)|$ is even.
But this verification is straightforward according to the case analysis of Observation 2.2, as it is easy to see how the sets $S \subseteq \{B(\underline{v_i}, 3) : i = 1 \ldots 7\}$ are determined accordingly in the cases $|S| = 1, 3, 5$ and 7.

$\qquad \square$

3 The Erdős-Rényi random graph

In this section we consider the Erdős-Rényi random graph $G(n, p)$. By $G(n, p)$ we mean the probability space of all labeled graphs on n vertices, where every edge appears randomly and independently with probability $p = p(n)$. We say that $G(n, p)$ possesses a property $\mathcal{P}$ *asymptotically almost surely*, or a.a.s. for brevity, if the probability that $G(n, p)$ satisfies $\mathcal{P}$ tends to 1 as n tends to infinity. We will find bounds for the quantities $M(G(n, p), r)$, $M(G(n, p), \leq r)$, and $M(G(n, p))$ which hold asymptotically almost surely. All our searches will use balls of equal radius, so the bounds we find for $M(G(n, p), \leq r)$ will hold for $M(G(n, p), r)$ as well.

Before we start to investigate the parameter $M(G(n, p), \leq r)$ generally, let us consider the special case when $r = 1$. In this case the asymptotics of $M(G(n, p), \leq r)$ follow immediately from a result of Frieze, Martin, Moncel, Ruszinkó, and Smyth [10]. As we mentioned in the introduction, for any graph G we always have $M(G, 1) \leq i_1^{(1)}(G)$. The parameter $i_1^{(1)}(G(n, p))$ was studied by Frieze, Martin, Moncel, Ruszinkó, and Smyth who proved the following theorem.

Theorem 3.1 (Frieze, Martin, Moncel, Ruszinkó, Smyth [10])**.** *Let $q = p^2 + (1 - p)^2$ and assume $p, 1 - p \geq \frac{\log \log n}{\log n}$. Then asymptotically almost surely $i_1^{(1)}(G(n, p)) = (1 + o(1))\frac{2 \log n}{\log(1/q)}$ holds.*

It is quite natural to think that apart from some exceptional graphs, querying balls of radius 0 (i.e. asking whether the center is the unknown vertex or not) is not efficient and therefore $i_1^{(1)}(G) = M(G, \leq 1)$ or at least $i_1^{(1)}(G) = (1 - o(1))M(G, \leq 1)$ holds. It is indeed the case for $G(n, p)$ and the proof of Lemma 2 in [10] suffices. Frieze, Martin, Moncel, Ruszinkó, and Smyth prove an upper bound on the probability that there exists an

identifying code C of size $\frac{(2-\varepsilon)\log n}{\log(1/q)}$ by considering the probability that there exists a pair of vertices in $V(G(n,p)) \setminus C$ that is not separated by the closed neighborhoods (i.e. 1-balls) of vertices of C. Since the probability that a set B_1 of 1-balls and a set B_0 of 0-balls together form a separating system is not larger than the probability that the closed neighborhoods of the centers C of the balls in $B_1 \cup B_0$ separate all pairs of vertices in $V(G(n,p)) \setminus C$, we obtain the following corollary of Theorem 3.1.

Corollary 3.2. *Let $q = p^2 + (1-p)^2$ and assume $p, 1-p \geq \frac{\log\log n}{\log n}$. Then almost surely $M(G(n,p), 1) = (1 + o(1))\frac{2\log n}{\log(1/q)}$ holds.*

Let us now consider the problem for general r. Note that all our constructions use balls of equal radius thus all our bounds will apply to $i_r^{(1)}(G(n,p))$ as well. To obtain our bounds on $M(G(n,p), \leq r)$ and $M(G(n,p))$ we will need two tools. The first tool is the following theorem of Katona.

Theorem 3.3 (Katona [15]). *Let X be an M-element set and $\mathcal{A} \subseteq 2^X$ be a separating system of subsets of X such that for all $A \in \mathcal{A}$ we have $|A| \leq m$ for some integer $m < M/2$. Then the following inequality holds*

$$|\mathcal{A}| \geq \frac{M}{m}\frac{\log M}{\log \frac{eM}{m}}.$$

Secondly, we will apply the Chernoff inequalities (see [1]).
Chernoff bound: *If X is a binomially distributed random variable with parameters n and p and $\lambda = np$, then for any $t \geq 0$ we have*

$$\mathbb{P}(X \geq \mathbb{E}X + t) \leq \exp\left(-\frac{t^2}{2(\lambda + t/3)}\right)$$

and

$$\mathbb{P}(X \leq \mathbb{E}X - t) \leq \exp\left(-\frac{t^2}{2\lambda}\right).$$

Our strategy will be quite simple. Whenever we are allowed to use balls of radius at most r, we will indeed use balls of radius r and place the centers of the balls randomly. Then we have to analyze how many randomly placed balls will guarantee the separation of all pairs of vertices. Note that in any graph G two vertices $x, y \in V(G)$ are separated by a ball of radius r with center $z \in V(G)$ if and only if $z \in B_G(x,r) \triangle B_G(y,r)$, where $X \triangle Y = X \setminus Y \cup Y \setminus X$ for any two sets X and Y. Therefore to obtain our results we will need bounds on $|B(x,r)|$, the sizes of all balls of radius r in $G(n,p)$ and also on the sizes of pairwise intersections of balls of radius r in $G(n,p)$.

These required bounds are strongly connected to the results of Bollobás. In [3], he determined the diameter $d = d(n)$ of $G(n,p)$ (it was also discovered by Klee and Larmann [17] when $d(n)$ is constant). Clearly, in any graph, it is only worth trying to search with balls of radius at most $d - 1$. First we state his results on the diameter of $G(n,p)$, then we summarize his proof approach and also cite Lemma 5 of his paper as we will need to use that as well.

Theorem 3.4 (Bollobás, [3]). *(i) Suppose $p^2n - 2\log n \to \infty$ and $n^2(1-p) \to \infty$. Then $G(n,p)$ has diameter 2 asymptotically almost surely.*

(ii) Suppose the functions $d = d(n) \geq 3$ and $0 < p = p(n) < 1$ satisfy $(\log n)/d - 3\log\log n \to \infty$, $p^d n^{d-1} - 2\log n \to \infty$ and $p^{d-1} n^{d-2} - 2\log n \to -\infty$. Then $G(n,p)$ has diameter d asymptotically almost surely.

Actually, Bollobás proved more: he showed that if $p^d n^{d-1} = \log(n^2/c)$ for some positive constant c, then the distribution of the number of pairs of vertices in $G(n,p)$ that are of distance $d+1$ from each other tends to a Poisson distribution and therefore the probability that $G(n,p)$ has diameter d tends to $e^{-c/2}$ while the probability that the diameter is $d+1$ tends to $1 - e^{-c/2}$.

The main lemma which essentially deals with the required bounds on $|B_{G(n,p)}(x,r)|$ and $|B_{G(n,p)}(x,r) \triangle B_{G(n,p)}(y,r)|$ is Lemma 5 in the paper of Bollobás. He carries out all calculations only for $p^d n^{d-1} = \log(n^2/c)$, but when p is bounded away from this threshold value, all calculations are even easier. Let us summarize how Bollobás obtains his bounds. For a fixed vertex x, we write $B(x,j) = B_{G(n,p)}(x,j)$ and $S(x,j) = B(x,j) \setminus B(x,j-1)$.

- To obtain bounds on the size of $B(x,r)$ for a fixed vertex x of $G(n,p)$ Bollobás exposes the edges of $G(n,p)$ in rounds. In the jth round he exposes edges between $S(x,j-1)$ and $V(G(n,p)) \setminus B(x,j-1)$ to determine $S(x,j)$ and thus $B(x,j)$. Note that if we condition on $B(x,j-1)$, then the size of $S(x,j)$ is a binomial random variable as for each vertex y of $V(G(n,p)) \setminus B(x,j-1)$ the probability of $y \in S(x,j)$ is $1 - (1-p)^{|S(x,j)|}$ and these events are independent. Careful analysis of concentration inequalities for binomial random variables gives a very small error probability of $|B(x,r)|$ not behaving as expected. Then the union bound is used to get the result below for all vertices.

- Similarly to the steps above, fixing two vertices x, y of $G(n,p)$ one can expose the edges in the following rounds. In round j one exposes edges between $S(x,j-1)$ and $V(G(n,p)) \setminus B(x,j-1)$ and also between $S(y,j-1)$ and $V(G(n,p)) \setminus B(y,j-1)$. Vertices in $B(x,j) \cap B(y,j) \setminus (B(x,j-1) \cap B(y,j-1))$ may come either from $B(x,j-1) \setminus B(y,j-1)$ or $B(y,j-1) \setminus B(x,j-1)$ or from $V(G(n,p)) \setminus (B(x,j-1) \cup B(y,j-1))$. Thus the number of new vertices in the intersection will be the sum of 3 binomially distributed random variable. Concentration inequalities and the union bound is used again to obtain bounds for all pairs x, y.

Lemma 3.5 (Bollobás, Lemma 5 in [3]). *Let d denote the diameter of $G(n,p)$ with $2 \leq d$ and $(\log n)/d - 3\log\log n \to \infty$. Then the following holds for $G(n,p)$ asymptotically almost surely: for any $1 \leq j \leq d-2$ and any pair x, y of vertices we have*

$$|B_{G(n,p)}(x,j)| = (1 + o(1))(np)^j \quad \text{and} \quad |B_{G(n,p)}(x,j) \cap B_{G(n,p)}(y,j)| = O(n^{2j-1}p^{2j}).$$

Lemma 3.5 enables us to obtain bounds on $M(G(n,p), r)$ when $r \leq d-2$.

Theorem 3.6. *Suppose the functions $d = d(n) \geq 3$ and $0 < p = p(n) < 1$ satisfy*

$$(\log n)/d - 3\log\log n \to \infty, \quad p^d n^{d-1} - 2\log n \to \infty, \quad p^{d-1} n^{d-2} - 2\log n \to -\infty.$$

Then for any $1 \leq r \leq d-2$ we have that

$$(1 - o(1))n^{1-r}p^{-r}\frac{\log n}{\log(n^{1-r}p^{-r})} \leq M(G(n,p), \leq r) \leq (1+o(1))n^{1-r}p^{-r}\ln n$$

holds asymptotically almost surely.

Proof. To obtain the lower bound we have to plug the values $M = n$ and $m = (1+o(1))(np)^r$ into Theorem 3.3. Note that the value of m follows from Lemma 3.5. Then Theorem 3.3 yields

$$M(G(n,p), r) \geq \frac{\log n}{\log e^{\frac{n}{(1+o(1))(np)^r}}}\frac{n}{(1+o(1))(np)^r} = (1 - o(1))n^{1-r}p^{-r}\frac{\log n}{\log(n^{1-r}p^{-r})}.$$

To obtain the upper bound we will place balls of radius r at centers $u_1, u_2, ..., u_t \in V(G(n,p))$ chosen uniformly at random and independently of each other. We introduce the random indicator variables Y_{v_i,v_j} of the event that the vertices v_i, v_j are not separated by the balls $B_{G(n,p)}(u_1, r), \ldots, B_{G(n,p)}(u_t, r)$. We write $Y = \sum_{1 \leq i < j \leq n} Y_{v_i,v_j}$ to denote the random variable of the total number of unseparated vertices. Observe that by Lemma 3.5 we have $|B_{G(n,p)}(v_i, r) \setminus B_{G(n,p)}(v_j, r)| = (1+o(1))(np)^r$ and thus

$$\mathbb{E}(Y_{v_i,v_j}) = \left(1 - \frac{|B_{G(n,p)}(v_i, r) \triangle B_{G(n,p)}(v_j, r)|}{n}\right)^t = \left(1 - \frac{(2+o(1))(np)^r}{n}\right)^t.$$

Therefore if $t \geq n^{1-r}p^{-r}\ln n$ holds, then we have

$$\mathbb{E}(Y) \leq \binom{n}{2}\left(1 - \frac{(2+o(1))(np)^r}{n}\right)^t \leq \frac{1}{2}\exp(2\ln n - 2tn^{r-1}p^r) \leq \frac{1}{2}.$$

Therefore there exists a particular choice of $u_1, u_2, ..., u_t$ such that the number of non-separated pairs of vertices is 0. $\square$

One would expect that $|B_{G(n,p)}(x, d-1)| = (1+o(1))(pn)^{d-1}$ should hold. It does for most values of p, but for a tiny range it fails since if $n^{-(d-2)/(d-1)} \leq p \leq n^{-(d-2)/(d-1)}(2\log n)^{1/(d-1)}$, then $(pn)^{d-1}$ is larger than n and still the diameter of $G(n,p)$ is d. For this range of p we introduce the function $f(n) = f_d(n)$ defined by the equation $(np)^{d-1} = n \cdot f(n)$. We will subdivide the interval of p for which the diameter of $G(n,p)$ is d into four smaller intervals. When $f(n) \to \infty$, then balls of radius $d-1$ consist of almost all vertices, and therefore we will be interested in their complement and in the fact that whether these complements are still bigger than the balls of radius $d-2$. If this is so, then it is still worth picking randomly centered balls of radius $d-1$, than those of radius $d-2$.

Lemma 3.7. *(i) If $(np)^{d-1} = o(n)$, then a.a.s*

$$|B(x, d-1)| = (1+o(1))(np)^{d-1}, |B(x_1, d-1)\triangle B(x_2, d-1)| = (2+o(1))(np)^{d-1}$$

hold for all vertices $x, x_1, x_2 \in V(G(n,p))$.

(ii) If $(np)^{d-1} = (c + o(1))n$ holds for some positive constant c, then a.a.s.

$$|\overline{B(x, d-1)}| = (1 - e^{-c} + o(1))n, |B(x_1, d-1)\triangle B(x_2, d-1)| = (1 - e^{-2c} + o(1))n/e^{f(n)}$$

hold for all vertices $x, x_1, x_2 \in V(G(n, p))$.
 (iii) If $f(n) \le (1 - \varepsilon)\frac{1}{d-1}\log n$ holds, then a.a.s.

$$|\overline{B(x, d-1)}| = (1 + o(1))n/e^{f(n)}, |B(x_1, d-1)\triangle B(x_2, d-1)| = (2 + o(1))n/e^{f(n)}$$

hold for all vertices $x, x_1, x_2 \in V(G(n, p))$.
 (iv) If $f(n) \ge (1 + \varepsilon)\frac{1}{d-1}\log n$ holds, then a.a.s. we have

$$|\overline{B(x, d-1)}| < |B(x, d-2)|$$

for all $x \in V(G(n, p))$.

Proof. Let us condition on the event that is stated in Lemma 3.5. More precisely, when proving statements about $B(x, d-1)$, we pick a vertex x, expose edges in rounds among vertices $S(x, j-1)$ and $V(G(n, p)) \setminus B(x, j-1)$. We condition on the event E_x that $|S(x, d-2)| = (1+o(1))(np)^{d-2}$. Bollobás proved that E_x holds with probability $1-n^{-K}$ for some large constant K and thus E_x holds with high probability for all vertices simultanously. Similarly, when proving statements on $|B(x_1, d-1) \cap B(x_2, d-1)|$ we pick two vertices x_1 and x_2, expose edges in rounds among vertices in $S(x_1, j-1)$ and $V(G(n, p)) \setminus B(x_1, j-1)$ and also among vertices in $S(x_2, j-1)$ and $V(G(n, p)) \setminus B(x_2, j-1)$. We condition on the event E_{x_1, x_2} that $|S(x_1, d-2)| = (1+o(1))(np)^{d-2} = |S(x_2, d-2)|$ and $|B(x_1, d-2) \cap B(x_2, d-2)| = O((np)^{d-3})$. Again, these events hold with probability $1 - n^{-K}$ for some large constant K.

Before considering the four cases let us observe that $\Omega(n^{1/2}) = (np)^{d-2} = o(n)$. We will also use the inequalities $pa - p^2\binom{a}{2} \le 1 - (1 - p)^a \le pa$.

Consider first the cases (i), (ii) when $(np)^{d-1} = o(n)$ or $(np)^{d-1} = (c + o(1))n$ holds for some positive constant c. In this case for a vertex $y \notin B(x, d-2)$ we have $\mathbb{P}(y \in S(x, d-1)) = 1 - (1 - p)^{|S(x, d-2)|}$. This is $(1 + o(1))p^{d-1}n^{d-2}$ and thus $\mathbb{E}(|S(x, d-1)|) = (1 + o(1))(np)^{d-1}$ if $(np)^{d-1} = o(n)$, while $(n - o(n))(1 - (1 - p)^{|S(x, d-2)|}) = (1 - e^{-c} + o(1))n$ if $(np)^{d-1} = (c + o(1))n$. As $(np)^{d-1} \ge (np)^{d-2} = \Omega(n^{1/2})$ holds, by Chernoff's inequality we obtain that $\mathbb{P}(|S(x, d-1) - \mathbb{E}(|S(x, d-1)|)| \ge n^{1/3})$ is exponentially small in a positive power of n. The union bound gives that asymptotically almost surely $|B(x, d-1)| = (1 + o(1))|S(x, d-1)| = \mathbb{E}(|S(x, d-1)|)$ holds for all vertices.

Let us condition on the sets $B(x_1, d-2), S(x_1, d-2), B(x_2, d-2)$ and $S(x_2, d-2)$. Then a vertex y can become an element of $B(x_1, d-1) \cap B(x_2, d-1)$ in the following ways:
 (a) either $y \notin B(x_1, d-2) \cup B(x_2, d-2)$ and (a1) it is connected to a vertex in $B(x_1, d-2) \cap B(x_2, d-2)$ or (a2) is connected to at least one vertex both from $S(x_1, d-2)$ and $S(x_2, d-2)$ or
 (b) $y \in S(x_1, d-2) \setminus B(x_2, d-2)$ and y is connected to a vertex in $S(x_2, d-2)$ or vice versa. The number of such vertices is binomially distributed in all three cases.

In case (a1), using the result of Lemma 3.5 on $|B(x_1, d-2) \cap B(x_2, d-2)|$ the expected value of such vertices is $O(p^{2(d-2)+1}n^{2(d-2)}) = O(\frac{p}{(pn)^2}((pn)^{d-1})^2) = O(\frac{pn}{(pn)^2}(pn)^{d-1}) = o((pn)^{d-1})$.

11

In case (a2), the expected value of such vertices is $(1-(1-p)^{|S(x_1,d-2)|})(1-(1-p)^{|S(x_2,d-2)|})$ which is $O(\frac{1}{n}((pn)^{d-1})^2) = o((pn)^{d-1})$ provided $(np)^{d-1} = o(n)$. If $(np)^{d-1} = (c+o(1))n$, then this value is $((1-e^{-c})^2 + o(1))n$.

Finally, in case (b), the expected value of such vertices is $O((pn)^{d-2}p^{d-1}n^{d-2}) = o((pn)^{d-1})$. In all three cases Chernoff's inequality and the union bound can be applied to obtain the same bounds not only for one, but all pairs of vertices.

Altogether, we obtain that $|B(x_1,d-1)\triangle B(x_2,d-1)| = (2+o(1))(np)^{d-1}$ holds almost surely for all x_1, x_2 if $(np)^{d-1} = o(n)$ and that $|B(x_1,d-1)\triangle B(x_2,d-1)| = [2(1-e^{-c}) - (1-e^{-c})^2 + o(1)]n = (1-e^{-2c}+o(1))n$ if $(np)^{d-1} = (c+o(1))n$ holds.

Let us now turn our attention to (iii) and (iv). In these cases we are interested in $|\overline{B(x,d-1)}|$. Let us repeat that $f(n)$ is defined by $f(n) = f_d(n) = (np)^{d-1} = n \cdot f(n)$ and that $f(n)$ tends to infinity as otherwise we were in case (i) or (ii). If we condition on $B(x,d-2)$ and $S(x,d-2)$, then the size of $\overline{B(x,d-1)}$ is binomially distributed with $\mathbb{E}(|\overline{B(x,d-1)}|) = (n-|B(x,d-2)|)(1-p)^{|S(x,d-2)|} = (1+o(1))e^{\log n - (1+o(1))f(n)}$. If $f(n) \geq (1+\varepsilon)\frac{1}{d-1}\log n$ holds for some positive ε, then this expression is $o((np)^{d-2})$. Chernoff's inequality and the union bound ensures that this holds for all x and we are done by (iv).

Using Chernoff's inequality we obtain that as long as $f(n) \leq (1-\varepsilon)\frac{1}{d-1}\log n$ holds, we have a.a.s $|\overline{B(x,d-1)}| = (1+o(1))n/e^{f(n)} > |B(x,d-2)|$ for any vertex x. As $I := \overline{B(x_1,d-1)} \cap \overline{B(x_2,d-1)} = \overline{B(x_1,d-1) \cup B(x_2,d-1)}$ we have that for any $y \in V(G(n,p))\backslash(B(x_1,d-2)\cup B(x_2,d-2))$ the probability that y is in I is $(1-p)^{|S(x_1,d-2)\cup S(x_2,d-2)|}$ and thus $\mathbb{E}(|I|) = (1+o(1))n/e^{2f(n)}$. Note that as $f(n)$ tends to infinity, we have $n/e^{2f(n)} = o(n/e^{f(n)})$. The usual Chernoff's inequality argument shows that $|\overline{B(x_1,d-1)} \cap \overline{B(x_2,d-1)}| = o(n/e^{f(n)})$ and thus $|\overline{B(x_1,d-1)}\triangle\overline{B(x_2,d-1)}| = (2+o(1))n/e^{f(n)}$ holds a.a.s for any $x_1, x_2 \in V(G(n,p))$. $\qquad\square$

The results of Lemma 3.7 enable us to imitate the proof of Theorem 3.6 to obtain bounds on $M(G(n,p))$. Part (iv) of Lemma 3.7 shows that in that range of p, one cannot improve the bounds of Theorem 3.6 by using random balls of radius $d-1$. In all other cases, replacing the values of $|B(x_1,d-1)\triangle B(x_2,d-1)| = |\overline{B(x_1,d-1)}\triangle\overline{B(x_2,d-1)}|$ does indeed improve on bounds obtained by placing balls of radius $d-2$. However, their formulation would be a little lengthy due to the fact that sometimes one has to calculate with the size of balls of radius $d-1$ and sometimes with the size of their complements. Therefore here we state only the simplest case (corresponding to case (i) in Lemma 1.7.) and just mention that in all cases the lower and upper bounds that can be obtained differ only by a factor of a power of $\log n$.

Theorem 3.8. *Suppose the functions $d = d(n) > 3$ and $0 < p = p(n) < 1$ satisfy $(\log n)/d - 3\log\log n \to \infty$, $p^d n^{d-1} - 2\log n \to \infty$. Furthermore, if $(np)^{d-1} = o(n)$ holds, then a.a.s we have*

$$(1-o(1))\frac{n}{(np)^{d-1}}\frac{\log n}{\log n/(np)^{d-1}} \leq M(G(n,p)) \leq (1+o(1))\frac{n}{(np)^{d-1}}\ln n.$$

4 Graphs of bounded maximum degree

The aim of this section is to prove Theorem 1.2, that is for any $G \in \mathcal{G}_{\Delta,n}$ we have $A(G) = \Theta_\Delta(\log n)$ where $\mathcal{G}_{\Delta,n}$ denotes the set of finite graphs on n vertices with maximum degree at most Δ. Note that the lower bound $\log n \leq A(G)$ holds for any graph as the sequence of answers must differ for all vertices of G. The idea behind Theorem 1.2 is the following: for every $G \in \mathcal{G}_\Delta$, $v \in V$ and $r \geq 1$ we have $|B_G(v, r+1)|/|B_G(v, r)| \leq \Delta - 1$, and this enables us to imitate the halving argument of adaptive binary search. This is formulated in the following lemma.

Lemma 4.1. *Let G be a connected graph with maximum degree Δ. For every $X \subseteq V(G)$ with $|X| \geq \Delta + 2$, there exists a vertex $v(X) \in V(G)$ and a natural number $r(X)$ satisfying*

$$\frac{1}{\Delta+2}|X| \leq |B(v(X), r(X)) \cap X| \leq \frac{\Delta+1}{\Delta+2}|X|.$$

Proof. For all $w \in V(G)$ and $X \subseteq V(G)$ we define the following function:

$$f(w, X) := \min\left\{r : |B(w,r) \cap X| \geq \frac{\Delta+1}{\Delta+2}|X|\right\}.$$

Since G is connected, $f(w, X)$ is well defined for all $w \in V(G)$. Also observe that $f(w, X) \geq 2$, since $\frac{\Delta+1}{\Delta+2}|X| \geq \Delta + 1$.

Define $m(X) = \min\{f(w, X) : w \in V(G)\}$, and let $w(X)$ be any vertex satisfying $f(w(X), X) = m(X)$. We claim that we can choose one of the neighbours of $w(X)$ as our vertex $v(X)$. Suppose, for the sake of contradiction, that none of the neighbors of $w(X)$ fulfill the statement of the lemma. By the definition of $m(X)$, we know that $|B(u, m(X)-1) \cap X| \leq \frac{1}{\Delta+2}|X|$ holds for all vertices $u \in V(G)$. But we also know that

$$B(w(X), m(X)) \subseteq \bigcup_{\{u, w(X)\} \in E(G)} B(u, m(X) - 1).$$

This gives us

$$X \cap B(w(X), m(X)) \quad \subseteq \quad X \cap \bigcup_{\{u, w(X)\} \in E(G)} B(u, m(X) - 1).$$

This gives a contradiction since $|X \cap B(w(X), m(X))| \geq \frac{\Delta+1}{\Delta+2}|X|$ holds by the definition of $m(X)$, whereas $|X \cap \bigcup_{\{u, w(X)\} \in E(G)} B(u, m(X) - 1)| \leq \sum_{\{u, w(X)\} \in E(G)} |B(u, m(X) - 1)| \leq \frac{\Delta}{\Delta+2}|X|$. $\square$

Now we are ready to prove Theorem 1.2.

Proof of Theorem 1.2: Let $G \in \mathcal{G}_{\Delta,n}$ be a connected graph and set $X_0 = V(G)$. For $i > 0$ we repeatedly apply Lemma 4.1 to X_{i-1} as long as $|X_{i-1}| \geq \Delta + 2$ holds. We ask the query $B(v(X_{i-1}), r(X_{i-1}))$ and let

$$X_i := \begin{cases} X_{i-1} \cap B(v(X_{i-1}), r(X_{i-1})) & \text{if the answer is yes;} \\ X_{i-1} \setminus B(v(X_{i-1}), r(X_{i-1})) & \text{if the answer is no.} \end{cases}$$

Clearly the unknown vertex is contained in X_i for all values of i. If $|X_{i-1}| \leq \Delta + 2$, then we ask its vertices one by one (more precisely the balls of radius 0 around its vertices). The length of this process is at most

$$\log_{\frac{\Delta+3}{\Delta+2}} n + \Delta + 1.$$

This completes the proof of the theorem. $\qquad\square$

By the same argument, one can prove a version of Theorem 1.2 where the upper bound $\Delta(n)$ on the maximum degree in the graph is a function which grows with n. Namely, we obtain a $\log_{\frac{\Delta(n)+3}{\Delta(n)+2}} n + \Delta(n) + 2$ upper bound on the length of an adaptive search.

References

[1] N. Alon and J. Spencer. *The probabilistic method.* Wiley-Interscience., 1991.

[2] Y. Ben-Haima, S. Gravierb, A. Lobsteinc, and J. Moncel. Adaptive identification in graphs. *J. Combin. Theory Ser. A*, 115(7):1114–1126, 2008.

[3] B. Bollobás. The diameter of random graphs. *Trans. Amer. Math. Soc.*, 267:41–52, 1981.

[4] I. Charon, G. Cohen, O. Hudry, and A. Lobstein. New identifying codes in the binary hamming space. *European J. Combin.*, 31:491–501, 2010.

[5] G. Cohen, I. Honkala, S. Litsyn, and A. Lobstein. *Covering Codes.* North-Holland Publishing Co., 1997.

[6] G. Exoo, V. Junnila, T. Laihonen, and S. Ranto. Upper bounds for binary identifying codes. *Advances in Applied Mathematics*, 42(3):277–289, 2009.

[7] G. Exoo, V. Junnila, T. Laihonen, and S. Ranto. Improved bounds on identifying codes in binary Hamming spaces. *European J. Combin.*, 31(3):813–827, 2010.

[8] G. Exoo, T. Laihonen, and S. Ranto. New bounds on binary identifying codes. *Discrete Applied Mathematics*, 156(12):2250–2263, 2008.

[9] F. Foucaud and G. Perarnau. Bounds for identifying codes in terms of degree parameters. *Electron. J. Combin.*, 19, 2012.

[10] A. Frieze, R. Martin, J. Moncel, M. Ruszinkó, and C.Smyth. Codes identifying sets of vertices in random networks. *Discrete Math.*, 307:1094–1107, 2007.

[11] F. Harary and R. Melter. On the metric dimension of a graph. *Ars Combinatoria*, 2:191–195, 1976.

[12] S. Janson and T. Laihonen. On the size of identifying codes in binary hypercubes. *J. Combin. Theory Ser. A*, 116(5):1087–1096, 2009.

[13] V. Junnila. Adaptive identication of sets of vertices in graphs. *Discrete Math. Theor. Comput. Sci.*, 14(1):69–86, 2012.

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

[15] G. Katona. On separating systems of a finite set. *J. Combin. Theory*, 1(2):174–194, 1966.

[16] G. Katona. Combinatorial search problems, problem booklet of the 5th emléktábla workshop, http://www.renyi.hu/∼emlektab/index_booklet.html. 2013.

[17] V. L. Klee and D. G. Larmann. Diameters of random graphs. *Canad. J. Math.*, 33:618–640, 1981.

[18] M. Krivelevich, B. Sudakov, and V. H. Vu. Covering codes with improved density. *IEEE Trans. Inform. Theory*, 49:1812–1815, 2003.

[19] A. Lobstein. http://www.infres.enst.fr/∼ lobstein/debutbibidetlocdom.pdf.

[20] J. Moncel. Codes identifiants dans les graphes. *These de Doctorat, Université de Grenoble, France*, 2005.

[21] P. Slater. Leaves of trees. *Congr. Numer.*, 14:549–559, 1975.