Generalized Selectors and Locally Thin Families with Applications to Conflict Resolution in Multiple Access Channels Supporting Simultaneous Successful Transmissions

Annalisa De Bonis

Università di Salerno, Fisciano, Italy
debonis@dia.unisa.it

Abstract

We consider the *Conflict Resolution Problem* in the context of a multiple-access system in which several stations can transmit their messages simultaneously to the channel. We assume that there are n stations and that at most k, $k \leq n$, stations are *active* at the same time, i.e, are willing to transmit a message over the channel. If in a certain instant at most d, $d \leq k$, active stations transmit to the channel then their messages are successfully transmitted, whereas if more than d active stations transmit simultaneously then their messages are lost. In this latter case we say that a *conflict* occurs. The present paper investigates *non-adaptive* conflict resolution algorithms working under the assumption that active stations receive a *feedback* from the channel that informs them on whether their messages have been successfully transmitted. If a station becomes aware that its message has been correctly sent over the channel then it becomes immediately *inactive*, that is, stops transmitting. The measure to optimize is the number of time slots needed to solve conflicts among all active stations. The fundamental question is how much this measure decreases with the number d of messages that can be simultaneously transmitted with success. In this paper we prove that it is possible to achieve a speedup linear in d by providing a conflict resolution algorithm that uses a $1/d$ ratio of the number of time slots used by the optimal conflict resolution algorithm for the particular case $d = 1$ [20]. Moreover, we derive a lower bound on the number of time slots needed to solve conflicts non-adaptively which is within a $\log(k/d)$ factor from the upper bound. To the aim of proving these results, we introduce a new combinatorial structure that consists in a generalization of Komlós and Greenberg codes [20]. Constructions of these new codes are obtained via a new kind of selectors [11], whereas the non-existential result is implied by a non-existential result for a new generalization of the locally thin families of [1, 10]. We believe that the combinatorial structures introduced in this paper and the related results may be of independent interest.

1998 ACM Subject Classification C.2.2 Network Protocols

Keywords and phrases Multiple-Access channels, Multi Access Communication, Conflict Resolutions, New Combinatorial Structures, Selectors

Digital Object Identifier 10.4230/LIPIcs.OPODIS.2016.22

1 Introduction

Conflict resolution is a fundamental problem in multiple-access communication and has been widely investigated in the literature both for its practical implications and for the many theoretical challenges it poses [6]. Commonly, this problem is studied under the assumption of the so called *collision model* in which simultaneous transmission attempts by two or more stations result in the destruction of all messages. However, as already observed in [16] and

20th International Conference on Principles of Distributed Systems (OPODIS 2016).
Editors: Panagiota Fatourou, Ernesto Jiménez, and Fernando Pedone; Article No. 22; pp. 22:1–22:16

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

more recently in [5], this restrictive multiaccess model does not capture the features of many important multiuser communication systems in which several messages can be successfully transmitted at the same time. Examples of communication systems allowing multiple-packet reception include Code Division Multiple Access (CDMA) systems in which several stations share the same frequency band, and Multiple-Input Multiple-Output (MIMO) systems, that enhance the capacity of a radio link by using multiple antennas at the transmitter and the receiver. These systems are largely used in the phone standards, satellite communication systems, and in wireless communication networks. Multiple-packet reception is also achieved through coding techniques specifically designed for copying with collisions. Recently, the authors of [4] introduced a coding technique, for the finite-field additive radio network model, that enables broadcast in a network with a bounded number of transmitters. These codes have the property that, when codewords from at most a certain number of different transmitting nodes are summed up, then the receiving nodes are able to recover the original transmissions.

Given the growing relevance of systems allowing multiple-packet reception in modern communication technologies, it is crucial to consider multiple-access models that better capture the phenomenon occurring when several packets can be sent simultaneously over the channel. The following quotation from [5] well emphasizes the importance of these communication models: "Traditionally, practical design and theoretical analysis of random multiple access protocols have assumed the classical collision channel model – namely, a transmitted packet is considered successfully received as long as it does not overlap or 'collide' with another. Although this model is analytically amenable and reflected the state of technology when networking was an emerging field, the classical collision model does not represent the capabilities of today's transceivers. In particular, present transceiver technologies enable users to correctly receive multiple simultaneously transmitted data packets. With proper design, this capability – commonly referred to as multiple packet reception (MPR) [17, 16] – can significantly enhance network performance.".

Communication models allowing multiple simultaneous successful transmissions have received great attention in the literature in recent times [5, 12, 14, 18, 22, 24, 25]. The following fundamental question arises when studying conflict resolution in the above described models: How fast does the number of time slots needed to solve conflicts decrease with the number d of messages that can be simultaneously transmitted with success? In this paper we show that it is possible to achieve a speedup linear in d when dealing with multiple-access systems with feedback, i.e., systems in which whenever an active station transmits to the channel, it receives a feedback that informs the station on whether its transmission has been successful.

1.1 The Model and Related Work

We consider a multiple-access system in which n stations have access to the channel and at most $k \leq n$ stations are willing to transmit a message at the same time. We call these stations *active* stations. If at most $d \leq k$ active stations transmit to the channel then these stations succeed to transmit their messages, whereas if more than d stations transmit then all messages are lost. In this latter case, we say that a *conflict* occurs. We assume that time is divided into time slots and that transmissions occur during these time slots. We also assume that all stations have a global clock and that active stations start transmitting at the same time slot. A scheduling algorithm for such a multiaccess system is a protocol that schedules the transmissions of the n stations over a certain number t of time slots (*steps*) identified by integers $1, 2, \ldots, t$. Whenever an active station transmits to the channel, it

receives a feedback from the channel that informs the station on whether its transmission has been successful. As soon as an active station becomes aware that its message has been successfully transmitted, it becomes *inactive* and does not transmit in the following time slots, even though it is scheduled to transmit by the protocol. For the particular case $d = 1$, our model corresponds to the multiple-access model with feedback considered by Komlós and Greenberg in [20].

In this paper we focus on non-adaptive scheduling algorithms, that is, algorithms that schedule all transmissions in advance so that all stations transmit according to a predetermined protocol known to them from the very beginning. Please notice that the knowledge of the feedback cannot affect the schedule of transmissions but can only signal a station to become inactive after it has successfully transmitted. A non-adaptive scheduling algorithm is represented by a $t \times n$ Boolean matrix where each column is associated with a distinct station and a station j is scheduled to transmit at step i if and only if entry (i, j) of the matrix is 1. In fact station j really transmits at step i if and only it is an active station and is scheduled to transmit at that step.

A *conflict resolution algorithm* is a scheduling protocol that schedules transmissions in such a way that all active stations transmit with success, i.e., for each active station there is a time slot in which it is scheduled to transmit on the channel and at most $d - 1$ other active stations are allowed to transmit in that time slot. The conflict resolution protocols considered in this paper are non-adaptive. The parameter we are interested in minimizing is the number of rows of the matrix which corresponds to the number of time slots over which the conflict resolution algorithm schedules the transmissions.

For the particular case $d = 1$, Komlós and Greenberg [20] gave a non-adaptive protocol that uses $O\left(k \log \frac{n}{k}\right)$ time slots to solve all conflicts among up to k active stations. Later on, the authors of [11, 21] proved the same upper bound by providing a simple construction based on selectors [11]. The above upper bound has been shown to be the best possible in [9], and later on, independently by the authors of [8, 10]. The lower bound in [8, 9, 10] improved on the $\Omega\left(\frac{k}{\log k} \log n\right)$ lower bound in [19], which additionally holds for adaptive algorithms that however are not the topic of this paper.

In [12] it has been studied the *no-feedback* version of the multiple-access problem considered in the present paper, i.e., the scenario in which at most d out of up to k active stations can transmit their messages simultaneously with success and the stations receive no feedback from the channel. That paper provides both upper and lower bounds on the minimum number of time slots needed to solve conflicts in the no-feedback model. The lower bound has been later improved by a combinatorial result given in [13]. Interestingly, the upper bound of [12] and the lower bound of [13] exceed, respectively, our upper and lower bounds for multiple-access systems with feedback by a $\frac{k}{d}$ factor.

1.2 Our results

In this paper we investigate the conflict resolution problem under the multiaccess model described in the previous section. To this aim, we introduce a new generalization of Komlós and Greenberg codes [20]. We prove that these new codes are equivalent to scheduling algorithms that allow up to k active stations to transmit with success in our setting, thus showing that upper and lower bounds on the minimum length of these codes translate into upper and lower bounds on the minimum number of time slots needed to solve conflicts. We present upper and lower bounds of the minimum length of these codes that differ asymptotically by a $\log(k/d)$ factor. These bounds are a consequence of the corresponding

bounds for other two new combinatorial structures also introduced in this paper. In particular, the proposed construction of generalized Komlós and Greenberg codes is based on a new version of (k, m, n)-selectors [11] having an additional parameter d. We give an existential result for this version of selectors based on the Lovász Local Lemma and present a Moser-Tardos type randomized algorithm to generate selectors meeting the proved upper bound. The lower bound follows from a non-existential result for a new combinatorial structure that can be regarded as an extension of the selective families of [3, 7] and the $\leq k$-locally thin codes of [10]. We call these new structures $(\leq k, d, n)$-locally thin codes.

Our main results are summarized by the following theorems.

▶ **Theorem 1.** *Let k, d, and n be integers such that $1 \leq d \leq k \leq n$. There exists a conflict resolution algorithm for a multiple-access channel with feedback that schedules the transmissions of n stations in such a way that all active stations transmit with success, provided that the number of active stations is at most k and that the channel allows up to d stations to transmit their messages simultaneously with success. The number of time slots t used by this algorithm is*

$$t = O\left(\frac{k}{d}\log\frac{n}{k}\right).$$

▶ **Theorem 2.** *Let k, d, and n be positive integers such that $3(d+1) \leq k \leq n$. Let $\mathcal{A}$ be* any *conflict resolution algorithm for a multiple-access channel with feedback that schedules the transmissions of n stations in such a way that all active stations transmit with success, provided that the number of active stations is at most k and that the channel allows up to d stations to transmit their messages simultaneously with success. The number of time slots t needed by $\mathcal{A}$ is*

$$t = \Omega\left(\frac{k}{d\log(k/d)}\log\frac{n}{k(d+1)}\right).$$

We remark that the asymptotic upper bound of Theorem 1 holds also in the case when there is no a priori knowledge of the number k of active stations. In this case, conflicts are resolved by running the conflict resolution algorithm of Theorem 1 iteratively (in stages), each time doubling the number of stations that are assumed to be active. In other words, at stage i the conflict resolution algorithm of Theorem 1 is run for a number k_i of supposedly active stations equal to 2^i. At stage $\lceil \log k \rceil$, the algorithm of Theorem 1 is run for a number of active stations larger than or equal to k and we are guaranteed that all active stations transmit with success within that stage.

Our paper is organized as follows. In Section 2 we introduce the fundamental combinatorial tools. We first introduce the new generalization of Komlós and Greenberg codes and prove that these new codes are equivalent to conflict resolution algorithms for our problem. Then, we introduce our generalized version of selectors and describe how to obtain a conflict resolution protocol by exploiting these combinatorial structures. We conclude Section 2 by giving the definition of $(\leq k, d, n)$-locally thin codes and show that our generalized version of Komlós and Greenberg codes is indeed a $(\leq k, d, n)$-locally thin code, thus proving that any non-existential result for $(\leq k, d, n)$-locally thin codes implies a non existential result for the conflict resolution protocols in our model. In Section 3 we provide existential results for generalized selectors and exploit these existential results to obtain the upper bound of Theorem 1. In Section 3 we also give a randomized algorithm to generate selectors meeting the proved upper bound. In Section 4 we give a lower bound on the minimum length of

$(\leq k, d, n)$-locally thin codes that implies the lower bound of Theorem 2. In that section, we also present a non-existential result for a combinatorial structure satisfying a weaker property than that of $(\leq k, d, n)$-locally thin codes and that can be regarded as a generalization of the k-locally thin families of [1]. Besides its combinatorial interest, this result implies a lower bound on the number of times slots needed to solve conflicts when the number of active stations is known to be *exactly* equal to k.

Due to space limit, some of the proofs are omitted in the present version of the paper.

2 Combinatorial Structures

In the following, for a positive integer m, we denote by $[m]$ the set $\{1, 2, \ldots, m\}$. Given a matrix M, we denote the set of its columns and the set of its column indices by M itself. The rows of a $t \times n$ matrix are numbered from the top to the bottom with integers from 1 to t. The n stations are identified by integers from 1 through n and for a given subset $S \subseteq [n]$ and an n-column matrix M, we denote by $M[S]$ the submatrix formed by the columns of M with indices in S.

2.1 Generalized Komlós and Greenberg Codes

▶ **Definition 3.** Let k, d, and n be integers such that $1 \leq d \leq k \leq n$. A $t \times n$ Boolean matrix is said to be a KG (k, d, n)-code of length t if for any submatrix M' of k columns of M there exists a non-empty set of row indices $\{i_1, \ldots, i_\ell\} \subseteq [t]$, with $i_1 < i_2 < \ldots < i_\ell$, such that the following property holds.

> There exists a partition $\{M'_1, \ldots, M'_\ell\}$ of the set of columns of M' such that, for $j = 1, \ldots, \ell$, $1 \leq |M'_j| \leq d$ and the i_j-th row of M' has all entries at the intersection with the columns of M'_j equal to 1 and those at the intersection with the columns in $M'_{j+1}, \ldots, M'_\ell$ equal to 0.

We will denote by $t_{KG}(k, d, n)$ the minimum length of a KG (k, d, n)-code.

The following theorem states that a KG (k, d, n)-code is indeed equivalent to a scheduling protocol for our multiple-access model with feedback that allows all up to k active stations to transmit with success.

▶ **Theorem 4.** *Let k, d, and n be integers such that $1 \leq d \leq k \leq n$, and let $\mathcal{A}$ be a scheduling algorithm for a multiple-access channel with feedback that allows up to d stations to transmit their messages simultaneously with success. $\mathcal{A}$ is a conflict resolution algorithm that schedules the transmissions of n stations in such a way that all of the up to k active stations transmit with success, if and only if the Boolean matrix corresponding to $\mathcal{A}$ is a KG (k, d, n)-code.*

Proof. The proof is omitted and will be given in the extended version of the paper. ◀

In the following two sections we introduce our generalized versions of selectors and locally thin codes and unveil their relationships with KG (k, d, n)-codes and, consequently, with our conflict resolution problem.

2.2 Generalized Selectors

The following definition introduces a new combinatorial structure that will be employed as a building block to construct KG (k, d, n)-codes. This new structure generalizes the notion of (k, m, n)-selectors introduced in [11] and corresponds to this notion for $d = 1$.

▶ **Definition 5.** Let k, m, d, and n be integers such that $1 \leq d \leq m \leq k \leq n$. A $t \times n$ Boolean matrix is said to be a (k, m, d, n)-selector if any k-column submatrix M' of M contains a set R of rows such that each row in R has Hamming weight comprised between 1 and d, and the Boolean sum of all rows of R has Hamming weight at least m. The number of rows t of the (k, m, d, n)-selector is the size of the selector. The minimum size of (k, m, d, n)-selectors is denoted by $t_{sel}(k, m, d, n)$.

A (k, m, d, n)-selector defines a scheduling algorithm for our multiaccess model that, in the presence of up to k active stations, allows all but at most $k - m$ of these stations to transmit with success. Indeed, all active stations that are scheduled to transmit in the time slots corresponding to the rows in R, transmit with success, since for each of those time slots there are at most d stations scheduled to transmit in that time slot. Notice that an active station might be scheduled to transmit in more than one of those time slots but it will become inactive as soon as it transmits with success. Let $p \leq k$ be the total number of active stations. Since the Boolean sum of the rows in R has Hamming weight at least m, then at least $m - (k - p)$ 1-entries in that Boolean sum are associated with active stations, and consequently, at least $m - (k - p)$ active stations transmit with success and at most $p - (m - (k - p)) = k - m$ active stations do not succeed to transmit their messages.

In the following we will show how to use (k, m, d, n)-selectors to obtain a KG (k, d, n)-code. The idea of this construction is similar to the one employed in [11, 21] to obtain a KG $(k, 1, n)$-code by using (k, m, n)-selectors as building blocks. From now on, unless specified differently, "log" will denote the logarithm in base 2. For the moment, let us assume for the sake of the simplicity that k and d be powers of 2. Our construction works as follows. We concatenate the rows of $(2^{v+1}, 2^v, d, n)$-selectors, for $v = \log d, \ldots, \log k - 1$, with the rows of the $(k, k/2, d, n)$-selector being placed at the top and those of the $(2d, d, d, n)$-selector being placed at the bottom. Then we add an all-1 row at the bottom of the matrix. Let M be the resulting matrix. Notice that the protocol defined by M consists in running the protocols defined by the $(2^{v+1}, 2^v, d, n)$-selectors, starting from the protocol associated with the $(k, k/2, d, n)$-selector through the one associated with the $(2d, d, d, n)$-selector. In the last time slot, corresponding to the bottommost row of M, the protocol schedules all stations to transmit. We will show that M defines a scheduling algorithm the allows up to k active stations to transmit with success, which, by Theorem 4, is equivalent to showing that M is a KG (k, d, n)-code. Let us assume that there are at most $k \leq n$ active stations. We observed that a (k, m, d, n)-selector provides a scheduling algorithm that schedules the transmissions so that at most $k - m$ active stations do not succeed to transmit their messages. Therefore, after running the scheduling protocol for $v = \log k - 1$, i.e., the protocol associated with the $(k, k/2, d, n)$-selector, the algorithm is left with at most $k/2$ active stations. Then the algorithm runs the protocol for $v = \log k - 2$, i.e., the protocol associated with the $(k/2, k/4, d, n)$-selector. This protocol allows all but at most $k/4$ of the remaining active stations to transmit with success. Extending this reasoning to an arbitrary $v \in \{\log d, \ldots, \log k - 1\}$, we have that after running the protocol associated with the $(2^{v+1}, 2^v, d, n)$-selector, there are at most 2^v stations that are still active. Therefore, after running the protocol associated with the $(2d, d, d, n)$-selector, there are most d active stations and no conflict can occur in the last time slot. In the last time slot all stations are scheduled to transmit, and consequently, all remaining active stations transmit with success in that time slot. For arbitrary values of k and d (not necessarily powers of 2), we replace in the above construction $\log k$ and $\log d$ by $\lceil \log k \rceil$ and $\lfloor \log d \rfloor$, respectively. The above construction implies the following upper bound on the minimum length $t_{KG}(k, d, n)$

of a KG (k, d, n)-code:

$$t_{KG}(k, d, n) = O\Big(\sum_{i=\lfloor \log d \rfloor}^{\lceil \log k \rceil - 1} t_{sel}(2^{i+1}, 2^i, d, n)\Big). \tag{1}$$

2.3 Generalized Locally Thin Codes

In this section we define a novel combinatorial structure that is strictly related to our problem in that non-existential results for this structure translate into non-existential results for KG (k, d, n)-codes.

▶ **Definition 6.** Let k, d, and n be integers such that $1 \leq d \leq k \leq n$. A $t \times n$ Boolean matrix M is said to be a $(\leq k, d, n)$-locally thin code of length t if the submatrix formed by any subset of s, $d \leq s \leq k$, columns of M contains a row with a number of 1's comprised between 1 and d. We will denote by $t_{LT}(\leq k, d, n)$ the minimum length of a $(\leq k, d, n)$-locally thin code.

Let M be a $(\leq k, d, n)$-locally thin code and let $\mathbf{F}$ be the family of the sets whose characteristic vectors are the columns of M. The family $\mathbf{F}$ has the property that for any subfamily $\mathbf{F'} \subseteq \mathbf{F}$ with $d \leq |\mathbf{F'}| \leq k$, there exists an element $x \in [t]$ such that $1 \leq |\{F \in \mathbf{F'} : x \in F\}| \leq d$. For $d = 1$, these families correspond to the selective families of [3, 7] and to the $\leq k$-locally thin families of [10]. The authors of [8, 9, 10] proved an $\Omega(k \log(n/k))$ lower bound on the minimum size of the ground set of $\leq k$-locally thin families which is tight with the upper bound on the length of KG $(k, 1, n)$-code [20].

The following theorem establishes a relation between $(\leq k, d, n)$-locally thin codes and KG (k, d, n)-codes.

▶ **Theorem 7.** *Let k, d, and n be integers such that $1 \leq d \leq k \leq n$. Any KG (k, d, n)-code is a $(\leq k, d, n)$-locally thin code.*

Proof. Let M be a KG (k, d, n)-code and suppose by contradiction that M is not a $(\leq k, d, n)$-locally thin code. This implies that there exists a subset of s, $d \leq s \leq k$, columns of M such that the submatrix M_s formed by these s columns contains no row with a number of 1's comprised between 1 and d. Let M' be a k-column submatrix of M such that $M' \supseteq M_s$. Since M is a KG (k, d, n)-code, Definition 3 implies that there exists a non-empty set of row indices $\{i_1, \ldots, i_\ell\} \subseteq [t]$, with $i_1 < i_2 < \ldots < i_\ell$, such that the following property holds:

> There exists a partition $\{M'_1, \ldots, M'_\ell\}$ of the set of columns of M' such that, for $j = 1, \ldots, \ell$, $1 \leq |M'_j| \leq d$ and the i_j-th row of M' has all entries at the intersection with the columns of M'_j equal to 1 and those at the intersection with the columns in $M'_{j+1}, \ldots, M'_\ell$ equal to 0.

Let $M'_{f_1}, \ldots, M'_{f_b} \in \{M'_1, \ldots, M'_\ell\}$, with $f_1 < f_2 < \ldots < f_b$, be the members of the partition having non-empty intersection with M_s, i.e., $M_s \cap M_{f_q} \neq \emptyset$, for each $q \in \{1, \ldots, b\}$, and $M_s \cap (M'_{f_1} \cup \ldots \cup M'_{f_b}) = M_s$. By our assumption that M_s does not contain any row with Hamming weight comprised between 1 and d, it follows that, for each row index $i \in [t]$, the i-th row of M_s has either Hamming weight 0 or Hamming weight larger than d. In the former case, the i-th row has a 0 in correspondence of at least one column in each of $M'_{f_1}, \ldots, M'_{f_b}$, whereas in the latter case the row has entries equal to 1 in correspondence of columns belonging to at least two of $M'_{f_1}, \ldots, M'_{f_b}$, since these submatrices contain at most d columns. Let us consider the row of M_s with index i_{f_1}. By Definition 3, one has that this

row has the entries at the intersection with the columns in M'_{f_1} equal to 1 and those at the intersection with the columns in $M'_{f_2} \cup \ldots \cup M'_{f_b}$ equal to 0. However, from what we have just observed, the i_{f_1}-th row of M_s has either a 0 in correspondence of at least one column in each of $M'_{f_1}, \ldots, M'_{f_b}$, or has entries equal to 1 in correspondence of columns belonging to at least two of $M'_{f_1}, \ldots, M'_{f_b}$. In the former case, the i_{f_1}-th row of M_s has an entry equal to 0 also at the intersection with some column in M'_{f_1}, whereas in the latter case the i_{f_1}-th row has a 1-entry in correspondence of some column in at least one of $M'_{f_2}, \ldots, M'_{f_b}$. In both cases, the i_{f_1}-the row of M_s does not satisfy the property of Definition 3, thus contradicting the fact that M is a KG (k, d, n)-code. ◀

3 Existential Results

First we prove an existential result for (k, m, d, n)-selectors for $k > 2(m-1)$. In order to prove this result, we need to recall the celebrated Lovász Local Lemma for the symmetric case (see [2]), as stated below.

▶ **Lemma 8.** *Let $E_1, E_2, \ldots, E_b$ be events in an arbitrary probability space. Suppose that each event E_i is mutually independent of a set of all other events E_j except for at most D, and that $\Pr[E_i] \leq P$ for all $1 \leq i \leq b$. If $eP(D+1) \leq 1$, then $\Pr[\cap_{i=1}^{n} \bar{E}_i] > 0$.*

By exploiting the above result we prove the following theorem.

▶ **Theorem 9.** *Let k, m, d, and n be positive integers such that $1 \leq d \leq m$ and $2(m-1) < k \leq n$. The minimum size $t_{sel}(k, m, d, n)$ of a (k, m, d, n)-selector is*

$$t_{sel}(k,m,d,n) \leq \begin{cases} 16\left(\ln k + (k-1)\ln\left(\frac{en}{k-1}\right) + (k-m+1)\ln\left(\frac{ek}{k-m+1}\right) + 1\right) & \text{if } 1 \leq d \leq 2 \\ \frac{\ln k + (k-1)\ln\left(\frac{en}{k-1}\right) + (k-m+1)\ln\left(\frac{ek}{k-m+1}\right) + 1}{\frac{d(k-m+1)}{4k} - \ln(4/3)} & \text{otherwise,} \end{cases}$$

where e denotes the Neper's constant $e = 2,71828\ldots$.

Proof. We will prove the existence of a (k, m, d, n)-selector with size smaller than or equal to the stated upper bound. The proof is based on Lemma 8. Let M be a $t \times n$ random binary matrix M where each entry is 1 with probability p and 0 with probability $1-p$. For a given k-column submatrix M' of M, let us denote by $E_{M'}$ the event that M' does not satisfy the property of Definition 5. Notice that, since there are at most $k\binom{n-1}{k-1}$ k-column submatrices containing one or more columns of M', it holds that $E_{M'}$ is independent from all but at most $k\binom{n-1}{k-1}$ events in $\{E_{\tilde{M}} : \tilde{M} \subseteq M, |\tilde{M}| = k\} \setminus \{E_{M'}\}$. In order to apply Lemma 8, we need to derive an upper bound P on the probability of each given event $E_{M'} \in \{E_{\tilde{M}} : \tilde{M} \subseteq M, |\tilde{M}| = k\}$.

In the following we will say that a row is w-good if its Hamming weight is comprised between 1 and d. The probability $Pr\{E_{M'}\}$ is the probability that the submatrix M' contains no subset R of w-good rows such that the Boolean sum of the rows in R has Hamming weight larger than or equal to m. To this aim, we notice that this event holds if and only if there exists a set A of $k-m+1$ column indices such that all w-good rows of M' have all zeros at the intersection with the columns with indices in A. For a fixed subset A of $k-m+1$ indices of columns of M', we denote by $\hat{E}_A$ the event that all w-good rows of M' have zeros at the intersection with the columns with indices in A. Hence, we have that

$$Pr\{E_{M'}\} = Pr\Big\{ \bigcup_{\substack{A \subseteq M': \\ |A| = k-m+1}} \hat{E}_A \Big\} \leq \sum_{\substack{A \subseteq M': \\ |A| = k-m+1}} Pr\{\hat{E}_A\}. \tag{2}$$

For a fixed A, event $\hat{E}_A$ holds if and only if, for any row index $i = 1, \ldots, t$, one has either that the i-th row of M' is not w-good or that the i-th row of M' is w-good and has all zeroes at the intersection with the columns with indices in A. Therefore, one has that

$$\begin{aligned} Pr\{\hat{E}_A\} &= Pr\Big\{\bigcap_{i=1}^{t}\{\{\text{the } i\text{-th row of } M' \text{ is not } w\text{-good}\} \\ &\qquad \cup\{\text{the } i\text{-th row of } M' \text{ is } w\text{-good and } M'(i,j)=0 \text{ for all } j\in A\}\}\Big\} \\ &\leq \prod_{i=1}^{t}\Big(Pr\{\text{the } i\text{-th row of } M' \text{ is not } w\text{-good}\} \\ &\qquad + Pr\{\text{the } i\text{-th row of } M' \text{ is } w\text{-good and } M'(i,j)=0 \text{ for all } j\in A\}\Big) \\ &= (P_1+P_2)^t, \end{aligned} \tag{3}$$

where

$$P_1 = Pr\{\text{the } i\text{-th row of } M' \text{ is not } w\text{-good}\} \tag{4}$$

and

$$P_2 = \{\text{the } i\text{-th row of } M' \text{ is } w\text{-good and } M'(i,j)=0 \text{ for all } j\in A\}, \tag{5}$$

for any fixed $i \in [t]$. Notice indeed that P_1 and P_2 do not depend on i.

By (2) and (3), we have that

$$Pr\{E_{M'}\} \leq \binom{k}{k-m+1}(P_1+P_2)^t. \tag{6}$$

Applying Lemma 8 with $P = \binom{k}{k-m+1}(P_1+P_2)^t$ and $D = k\binom{n-1}{k-1}$, one has that M has positive probability of being a (k,m,n)-strongly selective code if

$$e\binom{k}{k-m+1}(P_1+P_2)^t\left(k\binom{n-1}{k-1}+1\right) \leq 1. \tag{7}$$

Inequality (7) holds for t satisfying the following inequality

$$t \geq \frac{1+\ln\binom{k}{k-m+1}+\ln\left(k\binom{n-1}{k-1}+1\right)}{-\ln(P_1+P_2)}. \tag{8}$$

Therefore, there exists a (k,m,d,n)-selector of size t, for any t satisfying the above inequality.

The proof of the following claim is omitted and will be given in the extended version of the paper.

Claim 1. Let k, m, d, and n be positive integers such that $1 \leq d \leq m$ and $2(m-1) < k \leq n$. If we choose

$$p = \begin{cases} \frac{d}{2k} & \text{if } d \in \{1,2\} \\ \frac{d}{4k} & \text{if } d \geq 3 \end{cases}$$

then it holds

$$-\ln(P_1+P_2) \geq \begin{cases} \frac{1}{16} & \text{if } d \in \{1,2\} \\ (k-m+1)\left(\frac{d}{4k}\right)-\ln\left(\frac{4}{3}\right) & \text{if } d \geq 3. \end{cases}$$

In order for a (k, m, d, n)-selector of size t to exist it is sufficient that t satisfies inequality (8). Claim 1 implies that the righthand side of (8) is at most

$$16\left(1+\ln\binom{k}{k-m+1}+\ln\left(k\binom{n-1}{k-1}+1\right)\right),$$

if $1 \leq d \leq 2$, and it is at most

$$\frac{1+\ln\binom{k}{k-m+1}+\ln\left(k\binom{n-1}{k-1}+1\right)}{\frac{d(k-m+1)}{4k}-\ln(\frac{4}{3})},$$

if $d \geq 3$. Therefore, the upper bounds on $t_{sel}(k, m, d, n)$ in the statement of the theorem are implied by these two upper bounds on the righthand side of (8), along with inequality $\left(k\binom{n-1}{k-1}+1\right) \leq k\binom{n}{k-1}$ and the following well known upper bound on the binomial coefficient:

$$\binom{z}{y} \leq \left(\frac{ez}{y}\right)^y. \qquad \blacktriangleleft$$

Theorem 9 implies that bound (1) on $t_{KG}(k, d, n)$ in Section 2 is

$$O\Big(\sum_{i=\lfloor \log d\rfloor}^{\lceil \log k\rceil -1}\frac{2^{i+1}}{d}\log\frac{n}{2^{i+1}}\Big)=O\left(\frac{k}{d}\log\frac{n}{k}\right).$$

Therefore, the following theorem holds.

▶ **Theorem 10.** *Let k, d, and n be positive integers such that $d \leq k \leq n$. The minimum length $t_{KG}(k, d, n)$ of a KG (k, d, n)-code is*

$$t_{KG}(k, d, n) = O\left(\frac{k}{d}\log\frac{n}{k}\right).$$

Theorem 1 follows from Theorems 4 and 10. In virtue of Theorem 7, we have that Theorem 10 implies an existential results for $(\leq k, d, n)$-locally thin code. For $d = 1$, this existential result attains the same asymptotic upper bound as the one in [8].

Below, we provide a randomized algorithm that generates (k, m, d, n)-selectors meeting the upper bound of Theorem 9. Algorithm 1 is obtained by specializing a technique introduced by Moser and Tardos [23] to generate the structures whose existence is guaranteed by the Lovász Local Lemma. Theorem 1.2 of Moser and Tardos [23] implies that the expected number of times the resampling step (line 14 in Algorithm 1) is repeated is at most $\frac{n}{k^2}$. As a consequence, for fixed k, Algorithm 1 runs in expected polynomial time.

4 Non existential results

The following theorem states a lower bound on the minimum length of $(\leq k, d, n)$-locally thin codes. The proof of this theorem exploits and generalizes an interesting lower bound proof technique used by the authors of [1].

▶ **Theorem 11.** *Let k, d, and n be positive integers such that $3(d + 1) \leq k \leq n$. The minimum length $t_{LT}(\leq k, d, n)$ of a $(\leq k, d, n)$-locally thin code is*

$$t_{LT}(\leq k, d, n) > \frac{\lfloor \frac{k}{d+1}\rfloor}{\log\left(e\lfloor \frac{k}{d+1}\rfloor\right)}\log\left(\frac{n}{k(d+1)}\right).$$

Algorithm 1: Algorithm that generates (k, m, d, n)-selectors.

Input: Integers k, m and n, where $1 \leq d \leq m$ and $2(m-1) < k \leq n$.

Output: M : a (k, m, d, n)-selector.

1 Let $t := \begin{cases} 16\left(\ln k + (k-1)\ln\left(\frac{en}{k-1}\right) + (k-m+1)\ln\left(\frac{ek}{k-m+1}\right) + 1\right) & \text{if } 1 \leq d \leq 2 \\ \frac{\ln k + (k-1)\ln\left(\frac{en}{k-1}\right) + (k-m+1)\ln\left(\frac{ek}{k-m+1}\right) + 1}{\frac{d(k-m+1)}{4k} - \ln(4/3)} & \text{otherwise,} \end{cases}$;

2 Let $p := \begin{cases} \frac{d}{2k} & \text{if } 1 \leq d \leq 2 \\ \frac{d}{4k} & \text{otherwise.} \end{cases}$;

3 Construct a $t \times n$ matrix M where each entry $M(i,j)$ is chosen independently at random with $Pr\{M(i,j) = 1\} = p$ and $Pr\{M(i,j) = 0\} = 1 - p$.;

4 **repeat**

5 Set *flag*:= true;

6 **for** *each set C of k columns of M* **do**

7 **if** *C does not satisfy the property of Definition 5* **then**

8 Set *flag*:= false;

9 Set *missing-column-set* := C;

10 **break**;

11 **end**

12 **end**

13 **if** *flag = false* **then**

14 Choose all the entries in the k columns of *missing-column-set* independently at random, with each of those entries being 1 with probability p and 0 with probability $1 - p$.;

15 **end**

16 **until** *flag = true*;

17 Output M;

Proof. Let us write k as $k = (d+1)\lfloor\frac{k}{d+1}\rfloor + q$, with $0 \leq q \leq d$ and let $u = \lfloor\frac{k}{d+1}\rfloor$. We denote by α a positive rational number $\alpha = \frac{a}{b}$ satisfying the following inequalities

$$\frac{1}{u} \leq \alpha < \frac{1}{2}. \tag{9}$$

Let us denote by $n_{LT}(\leq k, d, t)$ the maximum value of n for which there exists a $(\leq k, d, n)$-locally thin code of length t. We will prove that

$$n_{LT}(\leq k, d, t) < k(d+1) \cdot 2^{h(\alpha)t}. \tag{10}$$

First we will show that for any $\alpha < \frac{1}{2}$ it holds

$$\frac{\alpha}{e} \leq 2^{-\frac{h(\alpha)}{\alpha}} < \frac{\alpha}{2}, \tag{11}$$

where $h(\alpha)$ denotes the binary entropy of α. Notice that, since we can choose $\alpha = \frac{1}{u}$, the upper bound (10) on $n_{LT}(\leq k, d, t)$, along with the lefthand side of (11), implies that

$$n_{LT}(\leq k, d, t) < k(d+1)\left(e\lfloor\frac{k}{d+1}\rfloor\right)^{t/\lfloor\frac{k}{d+1}\rfloor},$$

from which the lower bound on $t_{LT}(\leq k, d, n)$ in the statement of the theorem follows.

Let us prove inequalities (11). By the definition of binary entropy, one has that

$$h\left(\frac{a}{b}\right) = \frac{a}{b}\log\frac{b}{a} + \left(\frac{b-a}{b}\right)\log\left(\frac{b}{b-a}\right) = \frac{a}{b}\log\frac{b}{a} + \frac{1}{b}\cdot\log\left(1+\frac{a}{b-a}\right)^{b-a}. \quad (12)$$

Since $\left(1+\frac{a}{b-a}\right)^{b-a}$ increases with b, one has that $2^a < \left(1+\frac{a}{b-a}\right)^{b-a} \leq e^a$, where the left inequality follows from the righthand side of (9) that implies $b > 2a$. Therefore, by (12), it holds

$$\frac{a}{b}\log\left(\frac{2b}{a}\right) < h\left(\frac{a}{b}\right) \leq \frac{a}{b}\log\left(\frac{eb}{a}\right). \quad (13)$$

By replacing $\frac{a}{b}$ with α, inequalities (13) can be rewritten as $\alpha\log\left(\frac{2}{\alpha}\right) < h(\alpha) \leq \alpha\log\left(\frac{e}{\alpha}\right)$, from which we have that inequalities (11) hold.

Now we prove that $n_{LT}(\leq k, d, t) < k(d+1)\cdot 2^{h(\alpha)t}$. The proof is by induction on t.

For $t = 1$, any $t \times n$ Boolean matrix M has a single row that either contains at least $\frac{n}{2}$ entries equal to 0 or at least $\frac{n}{2}$ entries equal to 1. Consequently, if we assume by contradiction that $|M| = n \geq k(d+1)\cdot 2^{h(\alpha)t} \geq k(d+1)$ then the single row of M would either contain at least $k(d+1)/2$ occurrences of 0 or at least $k(d+1)/2$ occurrences of 1. This implies that there exist $k(d+1)/2 \geq k$ entries that are either all equal to 0 or all equal to 1 thus contradicting the hypothesis that M is a $(\leq k, d, n)$-locally thin code.

Let us consider $t > 1$ and let us assume by induction hypothesis that $n_{LT}(\leq k, d, t-1) < k(d+1)\cdot 2^{h(\alpha)(t-1)}$. Let M be a $t \times n$ $(\leq k, d, n)$-locally thin code of length t and let us assume by contradiction that $n \geq k(d+1)\cdot 2^{h(\alpha)t}$. We consider the following two cases.

Case 1. There exists an integer i in $[t]$ such that there are at least $2^{-h(\alpha)}n$ columns of M with the i-th entry equal to 0. In this case, if we remove the i-th entry from each of these columns, we have that the resulting columns form a matrix $\tilde{M}$ that is a $(\leq k, d, n)$-locally thin code of length $t-1$. Since we are assuming that $n \geq k(d+1)\cdot 2^{h(\alpha)t}$, it holds $|\tilde{M}| \geq 2^{-h(\alpha)}k(d+1)2^{h(\alpha)t} = k(d+1)\cdot 2^{h(\alpha)(t-1)}$. By induction hypothesis, $\tilde{M}$ cannot be a $(\leq k, d, n)$-locally thin code of length $t-1$, thus contradicting the fact that M is $(\leq k, d, n)$-locally thin code.

Case 2. For each element $i \in [t]$, there are less than $2^{-h(\alpha)}n$ columns of M with the i-th entry equal to 0. This implies that for a fixed i and for u randomly chosen columns $\mathbf{c}_1, \ldots, \mathbf{c}_u$ of M, the probability that $\mathbf{c}_1, \ldots, \mathbf{c}_u$ all have the i-th entry equal to 0 is less than $2^{-uh(\alpha)}$. By the lefthand side of (9) this probability is at most $2^{-\frac{h(\alpha)}{\alpha}}$, which by the righthand side of (11) is less than $\frac{\alpha}{2}$. Therefore, the expected number of 0-entries in the Boolean sum $\bigvee_{j=1}^{u}\mathbf{c}_j$ is less than $\frac{t\alpha}{2}$. Let X denote the number of 0-entries in the Boolean sum of u randomly chosen columns. We have shown that $E[X] < \frac{t\alpha}{2}$. Markov's inequality implies that, for any non-negative random variable Y and for any $b > 0$, it holds $Pr\{Y \geq b\} \leq \frac{E[Y]}{b}$. By our upper bound on $E[X]$ and by Markov's inequality, one has $\Pr\{\bigvee_{j=1}^{u}\mathbf{c}_j$ has at least $t\alpha$ 0-entries$\} < \frac{t\alpha}{2}\cdot\frac{1}{t\alpha} = \frac{1}{2}$. It follows that $Pr\{\bigvee_{j=1}^{u}\mathbf{c}_j$ has Hamming weight larger than $t - t\alpha\} > \frac{1}{2}$. Let $m = 2(d+1)\lceil 2^{h(\alpha)t}\rceil$ and let $\mathcal{B}_1, \ldots, \mathcal{B}_m$ be m randomly chosen subsets of u columns of M such that $\mathcal{B}_j \cap \mathcal{B}_\ell = \emptyset$, for $j \neq \ell$. Such subsets $\mathcal{B}_1, \ldots, \mathcal{B}_m$ can be generated by randomly permuting the columns of M, and then picking a set of $m \cdot u$ consecutive columns in the resulting matrix. In order to obtain $\mathcal{B}_1, \ldots, \mathcal{B}_m$, this set of columns is partitioned into m disjoint subsets each consisting of u consecutive columns. We have shown that $\bigvee_{\mathbf{c}\in\mathcal{B}_\ell}\mathbf{c}$ has Hamming weight larger than $t - t\alpha$ with probability larger than $\frac{1}{2}$, and consequently, the expected number

of subfamilies $\mathcal{B}_j$'s among $\mathcal{B}_1, \ldots, \mathcal{B}_m$ such that $\bigvee_{F \in \mathcal{B}_j} F$ has Hamming weight larger than or equal to $t - t\alpha$ is at least $\frac{m}{2}$. By linearity of expectation, there is a random choice of $\mathcal{B}_1, \ldots, \mathcal{B}_m$ such that there are at least $f \geq \frac{m}{2}$ subfamilies $\mathcal{B}'_1, \ldots, \mathcal{B}'_f$ among $\mathcal{B}_1, \ldots, \mathcal{B}_m$ for which one has that $\bigvee_{\mathbf{c} \in \mathcal{B}'_\ell} \mathbf{c}$, for $\ell = 1, \ldots, f$, has Hamming weight larger than or equal to $t - t\alpha$. However, one has that the number of pairwise distinct binary vector of length t with Hamming weight larger than or equal to $t - t\alpha$ is

$$\sum_{s=t-t\alpha}^{t} \binom{t}{s} = \sum_{s=0}^{t\alpha} \binom{t}{s} \leq 2^{th(\alpha)}, \tag{14}$$

where the last inequality follows from the well known inequality $\sum_{i=0}^{b} \binom{g}{i} \leq 2^{gh(b/g)}$, holding for $b/g \leq 1/2$, [15] . Since it is $m = 2(d+1)\lceil 2^{h(\alpha)t} \rceil$, then there are at most $\frac{m}{2(d+1)}$ pairwise distinct vectors of Hamming weight larger than or equal to $t - t\alpha$. We have shown that there exist $f \geq \frac{m}{2}$ subfamilies $\mathcal{B}'_1, \ldots, \mathcal{B}'_f$ such that $\bigvee_{\mathbf{c} \in \mathcal{B}'_\ell} \mathbf{c}$, for $\ell = 1, \ldots, f$, has Hamming weight larger than or equal to $t - t\alpha$. As a consequence, for at least a binary vector $\mathbf{c}_v$, there are $d+1$ sets $\mathcal{B}'_{j_1} \ldots, \mathcal{B}'_{j_{d+1}} \subseteq \{\mathcal{B}'_1, \ldots, \mathcal{B}'_f\}$ such that $\bigvee_{\mathbf{c} \in \mathcal{B}'_{j_q}} \mathbf{c} = \mathbf{c}_v$, for $q = 1, \ldots, d+1$. In other words, $\mathbf{c}_v$ occurs at least $d+1$ times among the Boolean sums $\bigvee_{\mathbf{c} \in \mathcal{B}'_1} \mathbf{c}, \ldots, \bigvee_{\mathbf{c} \in \mathcal{B}'_f} \mathbf{c}$. Therefore, the submatrix formed by the $(d+1)u = (d+1)\lfloor \frac{k}{d+1} \rfloor \leq k$ columns of $\mathcal{B}'_{j_1} \ldots, \mathcal{B}'_{j_{d+1}}$ is such that each row is either an all-zero row or has at least $d+1$ entries equal to 1, thus contradicting the assumption the M is a $(\leq k, d, n)$-locally thin code. ◀

Theorem 2 is an immediate consequence of Theorems 7 and 11.

The technique used to prove the lower bound of Theorem 11 allows also to obtain a lower bound on the length of codes satisfying a weaker property than that of $(\leq k, d, n)$-locally thin codes. We refer to these codes as (k, d, n)-*locally thin codes*. A $t \times n$ Boolean matrix M is a (k, d, n)-locally thin code of length t if and only if any submatrix formed by *exactly* k columns of M contains at least a row whose Hamming weight is comprised between 1 and d. If we interpret the columns of such a code as the characteristic vectors of n sets on the ground set $[t]$, then these sets have the property that for any k of them there exists an $i \in [t]$ that is contained in at at least one of these k sets and in no more than d of them. For $d = 1$, these families correspond to the k-locally thin code of [1].

▶ **Theorem 12.** *Let k, d, and n be positive integers such that $4(d+1) \leq k \leq n$. The minimum length $t_{LT}(k, d, n)$ of a (k, d, n)-locally thin code is*

$$t_{LT}(k, d, n) > \frac{\left(\lfloor \frac{k}{d+1} \rfloor - 1\right)}{\log\left(e\left(\lfloor \frac{k}{d+1} \rfloor - 1\right)\right)} \log\left(\frac{n}{k(d+1)}\right).$$

Proof. The proof is similar to that of Theorem 11 with the difference that here we write k as $k = (d+1)\lfloor \frac{k}{d+1} \rfloor + q = (d+1)\left(\lfloor \frac{k}{d+1} \rfloor - 1\right) + d + 1 + q$, with $0 \leq q \leq d$, and set $u = \lfloor \frac{k}{d+1} \rfloor - 1$. Moreover, here in order to prove the lower bound for Case 2, we need to prove the existence of a submatrix of *exactly* k columns that does not contain any row of Hamming weight comprised between 1 and d. To this aim, let us consider the subsets of u columns $\mathcal{B}'_{j_1}, \ldots, \mathcal{B}'_{j_{d+1}}$ whose existence has been proved in the proof of Theorem 11. The subsets $\mathcal{B}'_{j_1}, \ldots, \mathcal{B}'_{j_{d+1}}$ are such that the Boolean sums $\bigvee_{\mathbf{c} \in \mathcal{B}'_{j_1}} \mathbf{c}, \ldots, \bigvee_{\mathbf{c} \in \mathcal{B}'_{j_{d+1}}} \mathbf{c}$ have Hamming weight at least $t - t\alpha$, and the submatrix formed by the columns in $\mathcal{B}'_{j_1} \cup \ldots \cup \mathcal{B}'_{j_{d+1}}$ contains no row of Hamming weight comprised between 1 and d. The number of columns in this submatrix is $(d+1)u = (d+1)\left(\lfloor \frac{k}{d+1} \rfloor - 1\right) \leq k$. We will show that it is possible to add

columns to this submatrix so as to obtain a submatrix with exactly k-columns and with no row of Hamming weight comprised between 1 and d. To this aim, let us consider the columns of M that do not belong to any of $\mathcal{B}'_{j_1}, \ldots, \mathcal{B}'_{j_{d+1}}$. Let us denote by $i_1, \ldots, i_z$ the indices of the 0-zero entries in the Boolean sums $\bigvee_{\mathbf{c}\in\mathcal{B}'_{j_1}} \mathbf{c}, \ldots, \bigvee_{\mathbf{c}\in\mathcal{B}'_{j_{d+1}}} \mathbf{c}$. By the way $\mathcal{B}'_{j_1} \ldots, \mathcal{B}'_{j_{d+1}}$ have been defined in the proof of Theorem 11, one has $z \leq t\alpha$. We will prove that there are at least $d+1+q$ columns whose restrictions to the entries with indices $i_1, \ldots, i_z$ are identical. This implies that the $t \times k$ submatrix formed by $d+1+q$ of these columns and the columns in $\mathcal{B}'_{j_1} \ldots, \mathcal{B}'_{j_{d+1}}$ is such that each row is either an all-zero row or has at least $d+1$ entries equal to 1, thus contradicting the fact that M is a $(\leq k, d, n)$-locally thin code.

The rest of the proof is devoted to prove that there are at least $d+1+q$ distinct columns of M not in $\mathcal{B}'_{j_1} \ldots, \mathcal{B}'_{j_{d+1}}$ whose restrictions to the entries with indices $i_1, \ldots, i_z$ are identical. We observe that the number of columns of M that do not belong to any of $\mathcal{B}'_{j_1} \ldots, \mathcal{B}'_{j_{d+1}}$ is $n-(d+1)\left(\lfloor\frac{k}{d+1}\rfloor - 1\right)$. By the contradiction assumption, it holds $n \geq k(d+1)\cdot 2^{h(\alpha)t}$, and consequently, the above said number of columns is at least $k(d+1)\cdot 2^{h(\alpha)t} - (d+1)\left(\lfloor\frac{k}{d+1}\rfloor - 1\right)$ which, by the righthand side of (11), is larger than $k(d+1)\left(\frac{2}{\alpha}\right)^{t\alpha} - (d+1)\left(\lfloor\frac{k}{d+1}\rfloor - 1\right)$. Since $k(d+1)\cdot\left(\frac{2}{\alpha}\right)^{t\alpha} - (d+1)\left(\lfloor\frac{k}{d+1}\rfloor - 1\right) > kd\cdot\left(\frac{2}{\alpha}\right)^{t\alpha} > 2d\cdot 2^{t\alpha}$, it follows that there are more than $2d\cdot 2^{t\alpha}$ columns of M not in $\mathcal{B}'_{j_1} \ldots, \mathcal{B}'_{j_{d+1}}$. Among these columns there are at most $2^z \leq 2^{t\alpha}$ columns whose restrictions to indices $i_1, \ldots, i_z$ are pairwise distinct. As a consequence, there are at least $2d+1 \geq d+1+q$ columns of M not in $\mathcal{B}'_{j_1} \ldots, \mathcal{B}'_{j_{d+1}}$ whose restrictions to indices $i_1 \ldots, i_z$ are identical. ◀

For k even, the authors of [1] proved an $\Omega(k \log n)$ lower bound on the minimum size of the ground set of k-locally thin families, whereas, for arbitrary values of k, they gave an $\Omega\left(\frac{k}{\log k}\log n\right)$ lower bound. This latter lower bound is asymptotically the same as the lower bound obtained by setting d equal to 1 in the lower bound of Theorem 12.

Notice that Theorem 12 gives a lower bound on the minimum number of time slots needed to solve all conflicts in our model when the number of active stations is *exactly* k.

5 Conclusions

We have presented upper and lower bounds on the minimum number of time slots needed to solve conflicts among up to k active stations in a multiple-access system with feedback where at most d stations can transmit simultaneously with success over the channel. Interestingly, we have proved that it is possible to resolve conflicts in a number of time slots linearly decreasing with the number d of messages that can be simultaneously transmitted with success. Indeed, we have provided a conflict resolution algorithm that uses a $1/d$ ratio of the number of time slots used by the optimal conflict resolution algorithm for the particular case $d = 1$ [20].

The upper and lower bounds given in this paper differ asymptotically by a $\log(k/d)$ factor. An interesting open problem is to close this gap by improving on the lower bound on the minimum length of KG $(\leq k, d, n)$-codes.

References

1 Noga Alon, Emanuela Fachini, and János Körner. Locally thin set families. *Combinatorics, Probability and Computing*, 9(06):481–488, 2000.

2 Noga Alon and Joel Spencer. *The Probabilistic Method. Interscience series in discrete mathematics and optimization.* Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., Hoboken, NJ„ third edition, 2008.

3 Stefano Basagni, Danilo Bruschi, and Imrich Chlamtac. A mobility-transparent deterministic broadcast mechanism for ad hoc networks. *IEEE/ACM transactions on networking*, 7(6):799–807, 1999.

4 Keren Censor-Hillel, Bernhard Haeupler, Nancy Lynch, and Muriel Médard. Bounded-contention coding for the additive network model. *Distributed Computing*, 28(5):297–308, 2015.

5 Douglas S. Chan, Toby Berger, and Lang Tong. Carrier sense multiple access communications on multipacket reception channels: theory and applications to IEEE 802.11 wireless networks. *IEEE Transactions on Communications*, 61(1):266–278, 2013.

6 Bogdan S. Chlebus. Randomized communication in radio networks. In P. M. Pardalos, S. Rajasekaran, J. Reif, and J. D. P. Rolim, editors, *Handbook of Randomized Computing*, volume 1, pages 401–456. Kluwer Academic Publishers, 2001.

7 Bogdan S. Chlebus, Leszek Gąsieniec, Alan Gibbons, Andrzej Pelc, and Wojciech Rytter. Deterministic broadcasting in unknown radio networks. In *Proceedings of the Eleventh Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA'00, pages 861–870, Philadelphia, PA, USA, 2000. Society for Industrial and Applied Mathematics.

8 Andrea E. F. Clementi, Angelo Monti, and Riccardo Silvestri. Selective families, superimposed codes, and broadcasting on unknown radio networks. In *Proceedings of the 12th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 709–718. Society for Industrial and Applied Mathematics, 2001.

9 Gérard D. Cohen. Applications of coding theory to communication combinatorial problems. *Discrete Mathematics*, 83(2):237–248, 1990.

10 Miklós Csűrös and Miklós Ruszinkó. Single-user tracing and disjointly superimposed codes. *IEEE transactions on information theory*, 51(4):1606–1611, 2005.

11 Annalisa De Bonis, Leszek Gąsieniec, and Ugo Vaccaro. Optimal two-stage algorithms for group testing problems. *SIAM Journal on Computing*, 34(5):1253–1270, 2005.

12 Annalisa De Bonis and Ugo Vaccaro. Constructions of generalized superimposed codes with applications to group testing and conflict resolution in multiple access channels. *Theoretical Computer Science*, 306(1):223–243, 2003.

13 Annalisa De Bonis and Ugo Vaccaro. Optimal algorithms for two group testing problems, and new bounds on generalized superimposed codes. *IEEE transactions on information theory*, 52(10):4673–4680, 2006.

14 Aditya Dua. Random access with multi-packet reception. *IEEE Transactions on Wireless Communications*, 7(6):2280–2288, 2008.

15 Jörg Flum and Martin Grohe. *Parameterized Complexity Theory.* Springer, 2006.

16 Sylvie Ghez, Sergio Verdu, and Stuart C. Schwartz. Stability properties of slotted aloha with multipacket reception capability. *IEEE Transactions on Automatic Control*, 33(7):640–649, 1988.

17 Sylvie Ghez, Sergio Verdú, and Stuart C. Schwartz. Optimal decentralized control in the random access multipacket channel. *IEEE Transactions on Automatic Control*, 34(11):1153–1163, 1989.

18 Jasper Goseling, Michael Gastpar, and Jos H. Weber. Random access with physical-layer network coding. *IEEE Transactions on Information Theory*, 61(7):3670–3681, 2015.

19 Albert G. Greenberg and Schmuel Winograd. A lower bound on the time needed in the worst case to resolve conflicts deterministically in multiple access channels. *Journal of the ACM (JACM)*, 32(3):589–596, 1985.

20 Janos Komlos and Albert Greenberg. An asymptotically fast nonadaptive algorithm for conflict resolution in multiple-access channels. *IEEE Transactions on Information Theory*, 31(2):302–306, 1985.

21 Dariusz R. Kowalski. On selection problem in radio networks. In *Proceedings of the 24th Annual ACM Symposium on Principles of Distributed Computing*, pages 158–166. ACM, 2005.

22 Avery Miller. On the complexity of neighbourhood learning in radio networks. *Theoretical Computer Science*, 608:135–145, 2015.

23 Robin A. Moser and Gábor Tardos. A constructive proof of the general lovász local lemma. *Journal of the ACM (J. ACM)*, 57(2):11–15, 2010.

24 Alexander Russell, Sudarshan Vasudevan, Bing Wang, Wei Zeng, Xian Chen, and Wei Wei. Neighbor discovery in wireless networks with multipacket reception. *IEEE Transactions on Parallel and Distributed Systems*, 26(7):1984–1998, 2015.

25 Boris Tsybakov. Packet multiple access for channel with binary feedback, capture, and multiple reception. *IEEE Transactions on Information Theory*, 50(6):1073–1085, 2004.