ϵ-Almost Selectors and their Applications to Multiple-Access Communication

Annalisa De Bonis and Ugo Vaccaro

***Abstract*—Consider a group of stations connected through a multiple-access channel, with the constraint that if at a time instant exactly one station transmits a message, then the message is successfully received by any other station, whereas if two or more stations simultaneously transmit their messages then a conflict occurs and all messages are lost. Let us assume that n is the number of stations and that an (arbitrary) subset A of them, $|A| \leq k \leq n$, is *active*, that is, there are at most k stations that have a message to send over the channel. In the classical *Conflict Resolution Problem*, the issue is to schedule the transmissions of each station to let every active station use the channel alone (i.e., without conflict) at least once, and this requirement must be satisfied whatever might be the set of active stations A. The parameter to optimize is, usually, the worst case number of transmissions that any station has to attempt before all message transmissions are successful. In this paper we study the following question: is it possible to obtain a significant improvement on the protocols that solve the classical Conflict Resolution Problem if we allow the protocols to fail over a "small" fraction of all possible subsets of active stations? In other words, is it possible to significantly reduce the number of transmissions that must be attempted if the set of active stations is chosen uniformly at random and the conflict resolution algorithm is only required to work correctly with "high" probability? In this paper we will show that this is indeed the case. Our main technical tool is a generalization of selectors, a recently introduced combinatorial structure that has found applications in several areas. As it turned out for selectors, we believe that our new combinatorial structures are likely to be useful also outside the present context.**

***Index Terms*—: Multiple-access channel, conflict resolution, superimposed codes, selectors.**

I. The communication model

Our scenario consists of a multiaccess system where n stations have access to the channel and at most a certain number k of stations might be *active* at the same time, i.e., might transmit simultaneously over the channel. An active station successfully transmits if and only if it transmits singly on the channel. We follow the model and assumptions laid out in the seminal paper by Massey and Mathys [31]. 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 $1, 2, \ldots, t$. In a distributed model, a scheduling algorithm can be represented by a set of n Boolean vectors of length t, identified by integers from 1 through n, each of which corresponds to a distinct station, with the meaning that station j is scheduled to transmit at step i if and only if the i-th entry of its associated Boolean vector j is 1. In fact station j really transmits at step i if and only if it is an active station and is scheduled to transmit at that step.

A *conflict resolution* algorithm for the above described multiaccess system is a scheduling protocol that allows active stations to transmit successfully. A *non adaptive* conflict resolution algorithm is a protocol that schedules all transmissions in advance, i.e., for each step $i = 1, \ldots, t$ establishes which stations should transmit at step i without looking at what happened over the channel at the previous steps. A non adaptive conflict resolution algorithm is conveniently represented by the Boolean matrix having as columns the n Boolean vectors associated with the scheduling of the transmissions of the n stations. Entry (i, j) of such a matrix is 1 if and only if station j is scheduled to transmit at step i. The parameter of interest to be minimized is the number of rows of the matrix which represents the number of time slots over which the conflict resolution algorithm schedules the transmissions of the n stations so that up to k active stations transmit with success.

The multiple-access channel without feedback

When stations receive no *feedback* from the channel then the conflict resolution algorithm must schedule transmissions in such a way that each active station transmits singly to the channel at some step, i.e., in such a way that no other active station is scheduled to transmit at that same step. In this case non adaptive algorithms are an obliged choice since at each given step the conflict resolution algorithm has to schedule transmissions without knowing which stations succeeded to transmit their messages in the previous steps. A conflict resolution algorithm for this model is represented by a Boolean matrix M with the property that for any k columns of M and for any column $\mathbf{c}$ chosen among these k columns, there exists a row in correspondence of which $\mathbf{c}$ has a 1-entry and the remaining $k-1$ columns have 0-entries. In other words, for any choice of k out of n columns of M, the submatrix formed by these k columns contains all rows of the identity matrix I_k. Matrices that satisfy this property have been very well studied in the literature and are known under different names, such

The authors are with Dipartimento di Informatica, Università di Salerno, Fisciano (SA), Italy e-mail: debonis@dia.unisa.it, uvaccaro@unisa.it .

as superimposed codes [28], $(k-1)$-cover free families [24], $(k-1)$-disjunct codes [20], and strongly selective families [13], [14]. Extensions of such matrices that find applications in several different contexts have also been considered (see [18], [19] and references therein quoted). The best constructions for these combinatorial structures [35] imply the existence of conflict resolution algorithms that solve all conflicts among up to k active stations in

$$O\left(k^2 \log \frac{n}{k}\right)^{1} \tag{1}$$

number of steps. Remarkably, it is known that the number of rows of these matrices is lower bounded by

$$\Omega\left(\frac{k^2}{\log k} \log \frac{n}{k}\right), \tag{2}$$

[2], [22], [25], [36], and therefore any conflict resolution algorithm for this model should use at least this number of steps.

The authors of [17] introduced the following combinatorial structure that generalizes disjunct codes by introducing a parameter m which fixes the minimum number of distinct rows of the identity matrix I_k that should appear in any submatrix of k columns.

Definition 1: [17] Given integers k, m, and n, with $1 \leq m \leq k \leq n$, we say that a Boolean matrix M with t rows and n columns is a (k, m, n)-selector if for any choice of k out of n columns of M the submatrix formed by these k columns contains m rows of the identity matrix I_k. The integer t is the size of the (k, m, n)-selector. The minimum size of a (k, m, n)-selector is denoted by $t_s(k, m, n)$.

A (k, m, n)-selector provides us with a non adaptive algorithm that allows at least m out of exactly k active stations to transmit successfully. Protocols based on (k, m, n)-selectors employ a number of steps which decreases with the maximum number of active stations that might not succeed in transmitting their messages, as shown by the following bound on the minimum number of rows of (k, m, n)-selectors [17].

$$t_s(k, m, n) \leq \frac{ek^2}{k-m+1} \ln\left(\frac{n}{k}\right) + \frac{ek(2k-1)}{k-m+1}. \tag{3}$$

Notice that if there are less than k active stations then protocols based on (k, m, n)-selectors guarantee a number of successful transmissions smaller than m since it might happen that some (eventually all) of the m out of k stations which are scheduled to transmit singly to the channel are not active. In particular, if there are $j \leq k$ active stations the protocol guarantees $m - (k - j)$ active stations to transmit successfully, and therefore, independently from the actual number of active stations, such a protocol schedules transmissions so that at most $k-m$ active stations do not succeed to transmit their messages.

The multiple-access channel with feedback

In addition to the situation when stations receive no feedback from the channel, we consider also a communication model in which any transmitting station receives feedback on whether its transmission has been successful or not [29]. In such a model an active station has the capability to become *inactive* (i.e., to refrain from transmitting) after it has transmitted successfully. As in the previous model, a non adaptive conflict resolution algorithm should guarantee that for each active station there is a step at which it transmits singly. However, in this scenario an active station transmits singly to the channel also at time slots where it is scheduled to transmit simultaneously with some of the other k stations that were initially active, provided that these stations transmitted successfully at one of the previous steps. A Boolean matrix M represents a non adaptive conflict resolution algorithm for this more relaxed model *if and only if* any subset S of k columns of M satisfies the following property.

(*) *There are k row indices $i_1, i_2, \ldots, i_k$, with $i_1 < i_2 < \ldots < i_k$, and a permutation $[j_1, \ldots, j_k]$ of the indices of the columns in S, such that the submatrix of M formed by rows with indices $i_1, \ldots, i_k$, taken in this order, and columns with indices $j_1, \ldots, j_k$, taken in this order, is a $k \times k$ lower unitriangular matrix, i.e., a $k \times k$ matrix in which all entries in the diagonal are 1 and all those above the diagonal are 0.*

We will refer to a matrix in which all subsets of k column indices satisfy property (*) as a KG(k, n)-code and will denote the minimum length of KG(k, n)-codes by $t_{KG}(k, n)$. The name KG(k, n)-code comes from the initials of Komlós and Greenberg [29] who were the first to prove the following upper bound on the minimum length of KG(k, n)-codes.

$$t_{KG}(k, n) = O\left(k \log \frac{n}{k}\right) \tag{4}$$

Interestingly, the above bound is tight with the lower bound

$$\Omega\left(k \log \frac{n}{k}\right), \tag{5}$$

first proved by the author of [15], and, subsequently and independently, in [14], [16]. The authors of [17], [26] suggested a simpler construction which achieves the same asymptotic efficiency as the protocols in [29] and consists in concatenating $(2^i, 2^{i-1}, n)$-selectors starting from $i = \lceil \log k \rceil$ through $i = 1$, i.e., with the rows of the $(2^{\lceil \log k \rceil}, 2^{\lceil \log k \rceil - 1}, n)$-selector being placed at the top of the matrix and those of the $(2, 1, n)$-selector being placed at the bottom. Such a construction has also much smaller constants hidden in the big-Oh notation expressing the length of KG(k, n)-codes.

Our results

In this paper we study non adaptive conflict resolution protocols for both the two above described multiple-access models, that is, for multiple access channels with and without feedback. Our goal is to investigate what happens in terms of time efficiency if we allow the protocols to (*possibly*) fail over a "small" fraction of all possible subsets of active stations. In other words, we want to give an answer to the following question: is it possible to significantly reduce the number of steps used by the conflict resolution protocol if we tolerate that the protocol does not guarantee to behave correctly for a small fraction of all possible subsets of active

[1] Unless differently specified, all logarithms in this paper are of base 2

stations? In order to study this question, we introduce two new combinatorial structures that consist in a generalization of selectors and KG(k,n)-codes, respectively. In these matrices only a ratio $(1-\epsilon)$ of all k-column subsets is guaranteed to satisfy the desired properties, and therefore the corresponding conflict resolution protocols are guaranteed to work correctly only for $(1-\epsilon)\binom{n}{k}$ possible subsets of active stations.

Our paper is organized as follows. In Section II we first introduce a new version of (k,m,n)-selectors that correspond to protocols, for the multiple-access channel without feedback, that schedule transmissions so that, for at least a fraction $(1-\epsilon)$ of all possible subsets of k active stations, one has that at least m out of k stations are scheduled to transmit singly to the channel. Then, we introduce a new version of KG(k,n)-codes that furnish scheduling protocols for the multiple-access channel with feedback that allow to solve all conflicts for at least a fraction $(1-\epsilon)$ of all possible subsets of k active stations. In Section III, we recall the basic notion of hypergraph along with the related concepts of cover and partial cover which are at the core of our constructions. Our main technical results are contained in Section IV where we give constructions and non-existence results for the combinatorial structures introduced in Section II. We rely on these combinatorial results to derive upper and lower bounds on the number of time slots used by the corresponding conflict resolution algorithms. Our main results are summarized by the following theorems[2]. The proofs of Theorems 1, 2, 3, and 4 are deferred to Section IV after we have introduced the appropriate combinatorial tools in Section II. The proof of Theorem 5 is given in Section VI.

Theorem 1: Let k, m, and n be integers such that $1 \leq m \leq k \leq n$, and let ϵ be a real number such that $0 < \epsilon < 1$. There exists a conflict resolution algorithm for a multiple-access channel without feedback that schedules the transmissions of n stations in such a way that for at least a $(1-\epsilon)$ ratio of all possible subsets of k active stations, one has that at least m out k active stations transmit successfully. The number t of time slots used by the conflict resolution algorithm is

$$t \leq \frac{ek}{k-m+1}\left(1+\ln\frac{\binom{k}{k-m+1}}{\epsilon}\right).$$

At a first glance, it might seem strange that the upper bound in Theorem 1, as well as that in Theorem 2 below, appears not to be dependent on n. The dependency on n is in fact somehow hidden in the parameter ϵ, whose value, in practical situations, will depend on n in a way determined by the application at hand. We remark that for $\epsilon = \frac{1}{\binom{n}{k}} - \delta$, for any $\delta > 0$, the conflict resolution algorithm of Theorem 1 allows at least m active stations to transmit successfully whichever the subset of k active stations is. Observe that, for ϵ approaching $\frac{1}{\binom{n}{k}}$, the upper bound of Theorem 1 approaches upper bound (3).

Theorem 2: Let k and n be integers such that $1 \leq k \leq n$, and let ϵ be a real number such that $0 < \epsilon < 1$. 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 for at least a $(1-\epsilon)$ ratio of all possible subsets of k stations, one has that if the set of (up to k) active stations is entirely contained in one of those k-subsets then all active stations transmit successfully. The number t of time slots used by the conflict resolution algorithm is

$$t < 2e\lceil \log k \rceil \ln\left(\frac{\log k}{\epsilon}\right) + O(k).$$

Similarly to what happens with the algorithm of Theorem 1, we have that for $\epsilon = \frac{1}{\binom{n}{k}} - \delta$, for any $\delta > 0$, the conflict resolution algorithm of Theorem 2 allows all active stations to transmit with success, whichever the subset of up to k active stations is. For $\epsilon = \frac{1}{\binom{n}{k}}$, the upper bound of Theorem 2 is $O\left(k(\log k)(\log \frac{n}{k})\right)$.

Theorem 3: Let k and n be integers such that $2 \leq k \leq n$, and ϵ be a real number such that $\frac{1}{\binom{n}{k}} \leq \epsilon \leq \left(\frac{k}{2n}\right)^2$. Let t denote the minimum number of time slots needed by *any* conflict resolution algorithm for a multiple-access channel without feedback that schedules the transmissions of n stations in such a way that for at least a $(1-\epsilon)$ ratio of all possible subsets of k stations, one has that if the subset of (up to k) active stations is entirely contained in one of these subsets then all active stations transmit with success. It holds that

$$t = \Omega\left(\frac{\left(\log\frac{1}{\epsilon}\right)^2}{\left(\log\frac{n}{k}\right)\left(\log\log\frac{1}{\epsilon} - \log\log\frac{n}{k}\right)}\right).$$

By setting $m = k$ in the upper bound of Theorem 1, we obtain an $O\left(k\log\frac{1}{\epsilon}\right)$ upper bound on the minimum number of time slots of an optimal scheduling algorithm that behaves like the one in the hypothesis of Theorem 3. In Section IV-B, we will show that, for $\epsilon < (1/k)^c$, for any constant $c > 0$, this upper bound differs from the lower bound of Theorem 3 by an $O\left(\frac{k(\log n)(\log k)}{\log\frac{1}{\epsilon}}\right)$ factor. One can see that the smaller is ϵ, the smaller is the difference between these upper and lower bounds. When ϵ approaches $\frac{1}{\binom{n}{k}}$, this gap approaches $O(\log k)$. Interestingly, this gap is the same as the one between upper bound (1) and lower bound (2), which are the best upper and lower bounds for conflict resolution algorithms that solve conflicts among all possible subsets of k active stations in the multiple-access model without feedback. Designing such conflict resolution algorithms, or equivalently, constructing the associated binary codes, in a way that all conflicts are solved within an optimal number of time slots $\Theta\left(\frac{k^2}{\log k}\log n\right)$ (or, more modestly, just showing their existence), for *all* values of the involved parameters, is a problem that is still unresolved since decades [21].

Theorem 4: Let k and n be integers such that $2 \leq k \leq n$, and ϵ be a real number such that $\frac{1}{\binom{n}{k}} \leq \epsilon \leq \left(\frac{k}{2n}\right)^2$. Let t denote the minimum number of time slots needed by *any* conflict resolution algorithm for a multiple-access channel with feedback that schedules the transmissions of n stations in such a way that for at least a $(1-\epsilon)$ ratio of all possible subsets of k stations, one has that if the subset of (up to k) active stations is entirely contained in one of these subsets then all active stations transmit with success. It holds that

$$t = \Omega\left(\log\frac{1}{\epsilon}\right).$$

[2]However, see Section V for some improvements of the constructions

For $\epsilon = \frac{1}{\binom{n}{k}}$, the lower bound of Theorem 4 is $\Omega(k \log \left(\frac{n}{k}\right))$, which is equal to the lower bound (5), holding for conflict resolution algorithms that solve conflicts among all possible subsets of k active stations in the multiple-access model with feedback [14], [15], [16]. In Section IV-B, we will show that the upper bound of Theorem 2 differs from this lower bound by a factor of order $O\left(\log k + \frac{k}{\log \frac{1}{\epsilon}}\right)$.

Our protocol for a multiple-access channel with feedback, as well as those of [17], [26], can be made to work also in the case in which the parameter k is not known a priori, as asserted by the following theorem.

Theorem 5: Let k^* and n be integers such that k^* is not known in advance and $1 \leq k^* \leq n$, and let ϵ be a real number such that $0 < \epsilon < 1$. 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 for at least $(1-\epsilon)\binom{n}{k^*}$ possible subsets of k^* stations, one has that if the set of active stations is one of those subsets then all active stations transmit successfully and the algorithm uses a number t of time slots with

$$t < 8e \log k^* \ln \left(\frac{\log k^*}{\epsilon}\right) + O\left(\min\{k^{*2}, n\}\right).$$

Related work

Communication over a multiple access channel raises many challenging algorithmic and combinatorial problems. We refer the reader to the relevant chapters of the monographs [7], [20] and to the survey papers [10], [27] for a thorough presentation of the area to which our paper belongs. We discuss here only the results that are strictly related to ours.

In the case the error parameter ϵ is sufficiently small, our questions reduce, essentially, to the construction of good superimposed codes [28] and KG(k, n) codes [29]. The first problem has been recently subject of a breakthrough, namely in [35] the first *efficient* algorithm to construct superimposed codes of length $O(k^2 \log n)$ is presented. Regarding KG(k, n) codes, it is known since the seminal paper [29] that KG(k, n) codes of length $O(k \log(n/k))$ exist, and this bound was shown to be (asymptotically) optimal in [15], and in [14], [16]. Much simpler construction of KG(k, n) codes of length $O(k \log(n/k))$ were shown in [17], [26]. However, to date, there is no polynomial time algorithm to construct KG(k, n) codes of optimal length $O(k \log(n/k))$.

To the best of our knowledge, the problem of extending superimposed codes and KG(k, n) codes in order to deal with a controllable probability of error (in the sense we have previously explained) is new. There is a recent line of work contained in the papers [1], [4], [5], [9], [23], [32], [33] that considers extensions of matrices, strictly related to superimposed codes, in the same spirit as ours, that is, by requiring that the properties that such matrices must satisfy have to hold only for a fixed fraction of all possible k-tuples of columns. However, the results in the papers [1], [4], [5], [9], [23], [32], [33] have no direct implications on our results.

II. Combinatorial tools

For a positive integer n, we will denote by $[n]$ the set $\{1, \ldots, n\}$ and by $[n]_k$ the family of all k element subsets of $[n]$. The following definition introduces a new notion of selectors in which only a fraction $(1 - \epsilon)$ of all k-column subsets is guaranteed to satisfy the condition described in Definition 1.

Definition 2: Given integers k, m, and n, with $1 \leq m \leq k \leq n$, and a real number $0 < \epsilon < 1$, we say that a Boolean matrix M with t rows and n columns is an *ϵ-almost (k, m, n)-selector* if at least $(1 - \epsilon)\binom{n}{k}$ distinct subsets of k out of n columns of M are such that the k columns in each of those subsets form a submatrix that contains m rows of the identity matrix I_k. The integer t is the size of the ϵ-almost (k, m, n)-selector.

The ϵ-almost version of $(k+1, k+1, n)$-selectors are in fact the ϵ-almost version of k-disjunct codes and correspond to the notion of type 2 $(k-1, \epsilon)$-disjunct codes introduced in [32]. In the context of conflict resolution in the presence of a multiple-access channel without feedback, an ϵ-almost (k, m, n)-selector is equivalent to a protocol that schedules transmissions so that, for at least a fraction $(1 - \epsilon)$ of all possible subsets of exactly k active stations, one has that at least m active stations transmit singly to the channel. By the same argument used in Section I, one can see that, independently from the actual number of active stations, which is possibly smaller than k, the protocol corresponding to an ϵ-almost (k, m, n)-selector schedules transmissions so that there are at most $k - m$ active stations that do not transmit successfully, provided that the active stations are contained in some of the k-station subsets corresponding to one of the k-column subsets that satisfy the property of (k, m, n)-selectors, i.e., form a submatrix that contains m rows of the identity matrix. We will use this observation later on in the proof of Theorem 8. One is expected to trade off the weaker selection capacity of ϵ-almost (k, m, n)-selectors, which translates into the possibility of failing to resolve conflicts on a limited set of inputs, for a better efficiency of conflict resolution algorithms. In fact, we will show that, for large enough values of ϵ, there are ϵ-almost (k, m, n)-selectors with a number of rows significantly smaller than that of their "exact" counterpart.

The following definition introduces the analogous notion of KG(k, n)-codes in which only a fraction $(1 - \epsilon)$ of all k-column subsets is guaranteed to satisfy property (*).

Definition 3: Given integers k and n, with $1 \leq k \leq n$, and a real number $0 < \epsilon < 1$, we say that a Boolean matrix M with t rows and n columns is an ϵ-almost *KG(k, n)-code* if property (*) is satisfied by at least $(1-\epsilon)\binom{n}{k}$ k-column subsets S of M. The integer t is the length of the code.

In the context of conflict resolution for a multiple-access channel with feedback, ϵ-almost KG(k, n)-codes are equivalent to conflict resolution algorithms that guarantee all active stations to transmit successfully only in the presence of some subsets of active stations, i.e., if and only if the subset of active stations corresponds to a subset of one of the up to $(1-\epsilon)\binom{n}{k}$ k-column subsets that satisfy property (*).

A. A first, simple construction

As a warm-up, and in order to gain intuition, let us consider the following elementary construction of ϵ-almost (k,m,n)-selectors. We start by constructing a (classical) selector with the same parameters k and m, and with a number b of columns possibly smaller than n. If the parameter b, $k \leq b \leq n$, is such that

$$\binom{b}{k} \geq (1-\epsilon)\binom{n}{k}, \tag{6}$$

then it is immediate to see that the n column matrix, obtained by taking the columns of this (k,m,b)-selector and filling the remaining entries arbitrarily, is indeed an ϵ-almost (k,m,b)-selector. Let b be the smallest integer for which inequality (6) holds. The well known inequality

$$\binom{a}{c} \geq (a/c)^c. \tag{7}$$

implies that $\binom{b}{k} \geq (b/k)^k$ and therefore inequality (6) is satisfied if $(b/k)^k \geq (1-\epsilon)\binom{n}{k}$, which holds for $b \geq k(1-\epsilon)^{1/k}\binom{n}{k}^{1/k}$. Thus, by replacing n with

$$b = \left\lceil k(1-\epsilon)^{1/k}\binom{n}{k}^{1/k} \right\rceil \tag{8}$$

in (3), we have that the following fact holds.

Fact 1 *Given integers k, m, and n, with $1 \leq m \leq k \leq n$, and a real number $0 < \epsilon < 1$, there exists an ϵ-almost (k,m,n)-selector of size $t = O\left(\frac{k}{k-m+1}\ln\left((1-\epsilon)\binom{n}{k}\right) + \frac{k^2}{k-m+1}\right)$.*

The same simple idea can be applied to the construction of ϵ-almost KG(k,n)-codes. We consider the n column matrix obtained by taking the columns of an optimal KG(k,b)-code and other $(n-b)$ columns whose entries can be fixed arbitrarily. By setting b as in the above construction of ϵ-almost (k,m,n)-selectors, one has that the resulting matrix is indeed an ϵ-almost KG(k,n)-code. Then, the following fact is an immediate consequence of upper bound (4) on the minimum length of KG(k,n)-codes.

Fact 2 *For any integers k and n, with $1 \leq k < n$, and a real number $0 < \epsilon < 1$, there exists an ϵ-almost KG(k,n)-code of size t, with $t = O\left(\ln\left((1-\epsilon)\binom{n}{k}\right) + k\right)$.*

The rest of the paper is devoted to improve the simple upper bounds of Fact 1 and Fact 2 and that will represent our main technical contribution. We start with obtaining a construction for ϵ-almost (k,m,n)-selectors which improves considerably on the construction given in the present section. Our improved result relies on the fact that an ϵ-almost (k,m,n)-selector can be seen as a *partial cover* of a properly defined hypergraph. In the following section, we briefly recall the definitions of hypergraph, cover and partial cover of an hypergraph, along with the related terminology.

III. Hypergraphs and covers

Given a finite set X and a family $\mathcal{F}$ of subsets of X, a hypergraph is a pair $\mathcal{H} = (X, \mathcal{F})$. The set X will be denoted by $V(\mathcal{H})$ and its elements will be called vertices of $\mathcal{H}$, while the family $\mathcal{F}$ will be denoted by $E(\mathcal{H})$ and its elements will be called hyperedges of $\mathcal{H}$. A hypergraph is said to be *uniform* if all edges contain the same number of vertices and it is said to be *regular* if all vertices have the same degree, i.e, belong to the same number of edges. A vertex $v \in V(\mathcal{H})$ is said to *cover* an edge $E \in E(\mathcal{H})$ if $v \in E$. A *cover* of $\mathcal{H}$, also called *integral cover*, is a subset $T \subseteq V(\mathcal{H})$ such that for any hyperedge $E \in E(\mathcal{H})$ we have $T \cap E \neq \emptyset$, i.e., the vertices of T covers all edges in $E(\mathcal{H})$. A *fractional cover* of $\mathcal{H}$ is an assignment of vertex-weights $\{t_v \geq 0 : v \in V(\mathcal{H})\}$ such that the constraint $\sum_{v\in E} t_v \geq 1$ holds for all edges in $E(\mathcal{H})$. The size of the fractional cover is defined as $\sum_{v\in V(\mathcal{H})} t_v$. The minimum size of a cover of $\mathcal{H}$ will be denoted by $\tau(\mathcal{H})$,whereas the minimum size of a fractional cover of $\mathcal{H}$ will be denoted $\tau^*(\mathcal{H})$. Notice that the assignment of vertex-weights $\{t_v = \frac{1}{\min_{E\in\mathcal{F}}|E|} : v \in V(\mathcal{H})\}$ is a fractional cover of $\mathcal{H}$. Therefore, one has

$$\tau^*(\mathcal{H}) \leq \frac{|V(\mathcal{H})|}{\min_{E\in\mathcal{F}}|E|}. \tag{9}$$

Below, we recall the notions of partial cover ($(1-\epsilon)$-cover) and fractional partial cover (fractional $(1-\epsilon)$-cover). For $0 < \epsilon < 1$, a $(1-\epsilon)$-cover T of a hypergraph $\mathcal{H} = (V(\mathcal{H}), E(\mathcal{H}))$ is a collection of vertices which cover at least $(1-\epsilon)|E(\mathcal{H})|$ edges, i.e., $|\{E \in E(\mathcal{H}) : T \cap E \neq \emptyset\}| \geq (1-\epsilon)|E(\mathcal{H})|$. A fractional $(1-\epsilon)$-cover of a hypergraph $\mathcal{H} = (V(\mathcal{H}), E(\mathcal{H}))$ is an assignment of vertex-weights $\{t_v \geq 0 : v \in V(\mathcal{H})\}$ such that the constraint $\sum_{v\in E} t_v \geq 1$, holds for at least $(1-\epsilon)|E(\mathcal{H})|$ edges $E \in E(\mathcal{H})$. The size of the fractional partial cover is defined as $\sum_{v\in V(\mathcal{H})} t_v$. We denote by $\tau_\epsilon(\mathcal{H})$ and by $\tau^*_\epsilon(\mathcal{H})$ the minimum sizes of the integral $(1-\epsilon)$-cover and the fractional $(1-\epsilon)$-cover, respectively. The following theorem [34] bounds from above the minimum sizes of integral $(1-\epsilon)$-covers of regular and uniform hypergraphs.

Theorem 6: [34] If $\mathcal{H}$ is a regular and uniform hypergraph then it holds that

$$\tau_\epsilon(\mathcal{H}) \leq \tau^*_\epsilon(\mathcal{H})\left(1 + \ln\left(\frac{1}{\epsilon}\right)\right).$$

IV. Combinatorial results

In this section we present constructions and lower bounds for the combinatorial structures introduced in Section II.

A. Constructing ϵ-almost (k,m,n)-selectors via partial coverings

We first give an upper bound on the length of ϵ-almost (k,m,n)-selectors and then exploit this result to derive an upper bound on the length of ϵ-almost KG(k,n)-codes. The upper bound on the minimum size of ϵ-almost (k,m,n)-selectors is achieved by constructing a hypergraph $\mathcal{H}$ in such a way that any partial cover of $\mathcal{H}$ that covers a properly defined ratio of the edges in $E(\mathcal{H})$ is indeed an ϵ-almost (k,m,n)-selector. Theorem 6 is then used to derive the desired upper bound on the minimum size of such a partial cover.

Theorem 7: Given integers k, m, and n, with $1 \leq m \leq k \leq n$, and a real number $0 < \epsilon < 1$, there exists an ϵ-almost (k,m,n)-selector of size t such that

$$t \leq \frac{ek}{k-m+1}\left(1 + \ln\frac{\binom{k}{k-m+1}}{\epsilon}\right).$$

Proof: We aim at constructing a hypergraph $\mathcal{H}$ in such a way that, for a given $0 < \epsilon' < 1$, any $(1-\epsilon')$-partial cover of $\mathcal{H}$ is an ϵ-almost (k,m,n)-selector. Then, we will exploit Theorem 6 to derive the desired upper bound on the selector size.

The hypergraph $\mathcal{H}$ is defined as in the proof of Theorem 1 of [17]. We denote by X the set of all binary vectors $\mathbf{x} = (x_1, \ldots, x_n)$ of length n containing n/k 1's. For any integer i, $1 \leq i \leq k$, let $\mathbf{a}_i$ be the binary vector of length k having all components equal to zero but that in position i, that is, $\mathbf{a}_1 = (1,0,\ldots,0)$, $\mathbf{a}_2 = (0,1,\ldots,0)$, $\ldots$, $\mathbf{a}_k = (0,0,\ldots,1)$. Moreover, for any set of indices $S = \{i_1, \ldots, i_k\}$, with $1 \leq i_1 < i_2 < \ldots < i_k \leq n$, and for any binary vector $\mathbf{a} = (a_1, \ldots, a_k) \in \{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$, let $E_{\mathbf{a},S}$ be the set of binary vectors $E_{\mathbf{a},S} = \{\mathbf{x} = (x_1, \ldots, x_n) \in X : x_{i_1} = a_1, \ldots, x_{i_k} = a_k\}$. For any set $A \subseteq \{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ of size r, $r = 1, \ldots, k$, and any set $S \subseteq \{1, \ldots, n\}$, with $|S| = k$, we define the hyperedge $E_{A,S} = \bigcup_{\mathbf{a} \in A} E_{\mathbf{a},S}$ and, for any $r = 1, \ldots, k$, we define the set of hyperedges $\mathbf{F}_r = \{E_{A,S} : A \subset \{\mathbf{a}_1, \ldots, \mathbf{a}_k\}, |A| = r, \text{ and } S \subseteq \{1, \ldots, n\}, |S| = k\}$ and the hypergraph $\mathcal{H}_r = (X, \mathbf{F}_r)$.

Let $\ell = \lceil (1-\epsilon)\binom{n}{k} \rceil$ and let ϵ' be a real number such that $0 < \epsilon' < 1$ and

$$\left\lceil \binom{k}{k-m+1}\binom{n}{k}(1-\epsilon') \right\rceil \geq \binom{n}{k}\left[\binom{k}{k-m+1} - 1\right] + \ell. \quad (10)$$

We will prove that *any* $(1-\epsilon')$-cover of $\mathcal{H}_{k-m+1}$ is an ϵ-almost (k,m,n)-selector. Let T be *any* $(1-\epsilon')$-cover of $\mathcal{H}_{k-m+1}$. First notice that $|E(\mathcal{H}_{k-m+1})| = |\mathbf{F}_{k-m+1}| = \binom{k}{k-m+1}\binom{n}{k}$, therefore inequality (10) implies that any $(1-\epsilon')$-cover of $\mathcal{H}_{k-m+1}$ covers a number of edges larger than or equal to $\binom{n}{k}\left[\binom{k}{k-m+1} - 1\right] + \ell$. Moreover, by construction, for a fixed subset $S \in [n]_k$ there are exactly $\binom{k}{k-m+1}$ edges $E_{A,S}$ in $\mathcal{H}_{k-m+1}$, each for any of the possible subsets $A \subseteq \{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ of size $k-m+1$. It follows that for at least ℓ subsets $S^* \in [n]_k$, T covers all $\binom{k}{k-m+1}$ edges E_{A,S^*}. Let us denote by $\mathcal{G} \subseteq [n]_k$ the set of these ℓ subsets. It is possible to prove that the submatrix of T formed by the columns in any subset $S^* \in \mathcal{G}$ contains m rows of the identity matrix I_k. Indeed, assume by contradiction that there exists a set of indices $S^* = \{i_1, \ldots, i_k\} \in \mathcal{G}$ such that the submatrix of T obtained by considering only the columns of T with indices $i_1, \ldots, i_k$ contains *at most* $m-1$ distinct rows of I_k. It follows that there exists a subset A of $k-m+1$ vectors of $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ that does not contain any of these $m-1$ vectors, and consequently, E_{A,S^*} is not one of the edges covered by T, i.e., $T \cap E_{A,S^*} = \emptyset$. This contradicts the fact that for any $S^* \in \mathcal{G}$ and any subset A of $k-m+1$ vectors of $\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$, T covers the edge E_{A,S^*}. Since $|\mathcal{G}| = \ell \geq (1-\epsilon)\binom{n}{k}$ then T is an ϵ-almost (k,m,n)-selector. By Theorem 6 there exists a $(1-\epsilon')$-cover of size

$$\tau_{\epsilon'}(\mathcal{H}_{k-m+1}) \leq \tau^*_{\epsilon'}(\mathcal{H}_{k-m+1})\left(1 + \ln\left(\frac{1}{\epsilon'}\right)\right). \quad (11)$$

One can see that if we choose

$$\epsilon' = \frac{\epsilon}{\binom{k}{k-m+1}}, \quad (12)$$

then inequality (10) is satisfied with equality. To estimate the upper bound on $\tau_{\epsilon'}(\mathcal{H}_{k-m+1})$ in (11), we need to compute the minimum size $\tau^*_{\epsilon'}(\mathcal{H}_{k-m+1})$ of the fractional $(1-\epsilon')$-cover of $\mathcal{H}_{k-m+1}$. Notice that it holds that $\tau^*_{\epsilon'}(\mathcal{H}_{k-m+1}) \leq \tau^*(\mathcal{H}_{k-m+1})$ and inequality (9) implies

$$\tau^*(\mathcal{H}_{k-m+1}) \leq \frac{|X|}{\min_{E \in \mathbf{F}_{k-m+1}} |E|} = \frac{\binom{n}{n/k}}{(k-m+1)\binom{n-k}{n/k-1}}.$$

As shown in the proof of Theorem 1 of [17], inequality $\frac{\binom{n}{n/k}}{\binom{n-k}{n/k-1}} \leq ek$ holds, from which it follows

$$\tau^*_{\epsilon'}(\mathcal{H}_{k-m+1}) \leq \frac{ek}{k-m+1}. \quad (13)$$

The bound stated by the theorem follows from (11), (12) and (13). ∎

By the discussion following Definition 2 in Section II, where we have seen that an ϵ-almost (k,m,n)-selector furnishes a protocol that schedules transmissions so that, for at least a fraction $(1-\epsilon)$ of all possible subsets of exactly k active stations, one has that at least m active stations transmit singly to the channel, one sees that Theorem 1 is in fact a direct consequence of Theorem 7.

By setting $m = k+1$ in the upper bound on the length of ϵ-almost $(k+1,m,n)$-selectors derived from Theorem 7, we obtain the following upper bound on the minimum length of ϵ-almost k-disjunct codes of size n.

Corollary 1: Given integers k and n, with $1 \leq k < n$, and a real number $0 < \epsilon < 1$, there exists an ϵ-almost k-disjunct code with length $t \leq e(k+1)\left(1 + \ln\frac{(k+1)}{\epsilon}\right)$.

Remark 1: Reference *[32]* presented a construction for ϵ-almost k-disjunct codes with length

$$O\left((k-1)^{3/2}\ln n \frac{\sqrt{\frac{\ln(2k)}{\epsilon}}}{\ln(k-1) - \ln\ln\frac{2k}{\epsilon} + \ln(4a)}\right),$$

for $\epsilon > 2ke^{-a(k-1)}$ and a being a constant larger than 1. We notice that the constraint that $\epsilon > 2ke^{-a(k-1)}$ does not allow constructions of ϵ-almost k-disjunct codes in which the number $\epsilon\binom{n}{k}$ of "bad" k-column subsets is small enough, as in our constructions. In fact, $\epsilon > 2ke^{-a(k-1)}$ implies that $\epsilon\binom{n}{k} > 2ke^a\left(\frac{n}{ke^a}\right)^k$. Moreover, for $\epsilon > 2ke^{-a(k-1)}$ the bound of our Corollary 1 is $O(k^2)$ which is at least as good as the bound in [32] when $\epsilon \leq c\left(\frac{\ln n\sqrt{\ln(2k)}}{\sqrt{k-1}(\ln(k-1) - \ln\ln\frac{2k}{\epsilon} + \ln(4a))}\right)^2$, for any positive constant c, that is almost always the case.

Remark 2: It is interesting also to remark that for $\epsilon = \frac{1}{\binom{n}{k}} - \delta$, for any $\delta > 0$, the ϵ-almost disjunct codes reduce to the classical disjunct codes. Moreover, the bound one obtains from our Corollary 1 with that value of ϵ reduces to $t = O(k^2 \log\frac{n}{k})$, which is equal to the upper bound on the length of classical disjunct codes constructed in [35].

Remark 3: It might be also worthwhile to consider the case $m = 1$ of Theorem 7. In this case one obtains a binary matrix M with the following property: there are at least $(1-\epsilon)\binom{n}{k}$ subsets of k columns such that for each of those subsets there exists a row in which the entries of M at the intersection of that row and of the k columns in that subset are all 0, with the exception of just one entry of value 1. From the point of view of a protocol for conflict resolution, matrix M would correspond to a protocol that would allow the successful transmission of at least one station out of *exactly* k conflicting ones, provided that the subset of k conflicting stations is among the $(1-\epsilon)\binom{n}{k}$ "good" k-station subsets. By Theorem 7, the length t of such a protocol would be such that

$$t \leq e\left(1 + \ln\frac{1}{\epsilon}\right). \tag{14}$$

At a first look, bound (14) seems too good to be true. Additionally, a deep result of *[3]* implies that in the case k is even, then any binary matrix that satisfies the above property for all subsets of k columns (and not for just $(1-\epsilon)\binom{n}{k}$ of them) *must* have a number t of rows such that

$$t \geq \frac{k}{2}\log n. \tag{15}$$

However, a moment of reflection shows that (14) is less surprising that it looks, since, in general, ϵ might depend on n and k. For instance, if we insist that all subsets of k columns are "good", that is, $\epsilon = \frac{1}{\binom{n}{k}} - \delta$, for any $\delta > 0$, then bound (14) (optimally) matches the lower bound (15). On the other hand, a comparison of bounds (15) and (14) shows that by introducing a (controllable) possibility of error in the protocol, one can obtain quite significant improvements.

Theorem 8 below establishes an upper bound on the minimum length of ϵ-almost KG(k,n)-codes. The main idea of the construction attaining this upper bound is to concatenate optimal ϵ_i-almost $(2^i, 2^{i-1}, n)$-selectors, starting from $i = \lceil\log k\rceil$ through $i = 1$, i.e., with the rows of the $\epsilon_{\lceil\log k\rceil}$-almost $(2^{\lceil\log k\rceil}, 2^{\lceil\log k\rceil-1}, n)$-selector being placed at the top of the matrix and those of the ϵ_1-almost $(2,1,n)$-selector being placed at the bottom. The values of the ϵ_i's are given in the proof of Theorem 8 and are a consequence of the analysis therein. Then the all-1 row is placed at the bottom of the resulting matrix. The intuition behind this construction is the following. Consider a set X of $2^{\lceil\log k\rceil}$ columns of this matrix and let X_i, for $i = 1,\ldots,\lceil\log k\rceil$, denote the set of the columns obtained by restricting the columns in X to the entries corresponding to the rows of the ϵ_i-almost $(2^i, 2^{i-1}, n)$-selector. Suppose that, for $i = 1,\ldots,\lceil\log k\rceil$, each of the 2^i-column subsets of X_i is one the $(1-\epsilon_i)\binom{n}{2^i}$ 2^i-column subsets that satisfies the property of $(2^i, 2^{i-1}, n)$-selector, i.e., the submatrix formed by these 2^i columns contains 2^{i-1} rows of the identity matrix I_{2^i}. By the discussion in Section II, we have that the protocol based on the ϵ_i-almost $(2^i, 2^{i-1}, n)$-selector schedules transmissions in such a way that if the set of active stations is entirely contained in one of the 2^i-column subsets of X_i then there are at most 2^{i-1} active stations that do not succeed to transmit their messages. For $i = 1,\ldots,\lceil\log k\rceil$, let t_i denote the number of rows of the $(2^i, 2^{i-1}, n)$-selector. Suppose that the up to k active stations are associated with columns belonging to X. By the above argument, one has that after the first $t_{\lceil\log k\rceil}$ steps there are at most $2^{\lceil\log k\rceil-1}$ active stations. After the next $t_{\lceil\log k\rceil-1}$ time slots, at most $2^{\lceil\log k\rceil-2}$ of these up to $2^{\lceil\log k\rceil-1}$ stations are still active. Continuing in this way, one can see that, for each $i = 1,\ldots,\lceil\log k\rceil$, after the first $t_{\lceil\log k\rceil} + t_{\lceil\log k\rceil-1} + \ldots + t_i$ time slots, the algorithm is left with at most 2^{i-1} active stations. At the end, there is at most a single station which is still active, and therefore, can transmit without conflict with other active stations.

Theorem 8: Given integers k and n, with $1 \leq k \leq n$, and a real number $0 < \epsilon < 1$, there exists an ϵ-almost *KG*(k,n)-code with length

$$t < 2e\lceil\log k\rceil \ln\left(\frac{\lceil\log k\rceil}{\epsilon}\right) + O(k).$$

Proof: We will prove that there exists a $t \times n$ Boolean matrix $\tilde{M}$, with t being upper bounded as in the statement of the theorem, such that $\tilde{M}$ contains at least $(1-\epsilon)\binom{n}{k}$ subsets of k columns that satisfy property (*). This is equivalent to prove that there are at least $(1-\epsilon)\binom{n}{k}$ subsets S of k stations such that the scheduling algorithm corresponding to $\tilde{M}$ allows any subset of active stations in S to transmit successfully. Let us recall that Theorem 7 implies that for any integer $1 \leq i \leq \lceil\log k\rceil$ and any real number ϵ_i with $0 < \epsilon_i < 1$, there exists an ϵ_i-almost $(2^i, 2^{i-1}, n)$-selector of size

$$t_i < 2e\left(1 + \ln\frac{\binom{2^i}{2^{i-1}+1}}{\epsilon_i}\right). \tag{16}$$

For $i = 1,\ldots,\lceil\log k\rceil$, let us denote by $\tilde{M}_{(\epsilon_i,2^i,n)}$ such an ϵ_i-almost $(2^i, 2^{i-1}, n)$-selector, and let

$$\epsilon_i = \begin{cases} \frac{\epsilon}{\binom{k}{2^i}\lceil\log k\rceil} & \text{for } i = 1,\ldots,\lceil\log k\rceil - 1, \\ \frac{\epsilon}{\lceil\log k\rceil} & \text{for } i = \lceil\log k\rceil. \end{cases} \tag{17}$$

We set $\tilde{M}$ to be the matrix obtained by concatenating the rows of matrices $\tilde{M}_{(\epsilon_i,2^i,n)}$, for $i = \lceil\log k\rceil,\ldots,1$, taken in this order, that is with the rows of $\tilde{M}_{(\epsilon_{\lceil\log k\rceil},2^{\lceil\log k\rceil},n)}$ being placed at the top of the matrix and those of $\tilde{M}_{(\epsilon_1,2,n)}$ being placed at the bottom, and by appending the all-1 row at the bottom. We will show that there are at least $(1-\epsilon)\binom{n}{k}$ subsets S of k stations such that the conflict resolution algorithm based on $\tilde{M}$ allows any subset of active stations in S to transmit successfully, thus proving that $\tilde{M}$ is an ϵ-almost KG(k,n)-code.

For $i = 1,\ldots,\lceil\log k\rceil$, let us denote by $\mathcal{G}_i$ the family of the subsets of 2^i columns of $\tilde{M}_{(\epsilon_i,2^i,n)}$ that satisfy the property of $(2^i, 2^{i-1}, n)$-selectors, i.e., the family of 2^i-column subsets for which it holds that the submatrix of $\tilde{M}_{(\epsilon_i,2^i,n)}$ formed by the 2^i columns in any member of the family contains 2^{i-1} rows of the identity matrix I_{2^i}. We will prove that the following claim holds.

Claim Let S be a subset of k columns of $\tilde{M}$ such that $S \subseteq S'$ for some $S' \in \mathcal{G}_{\lceil\log k\rceil}$. If, for $i = 1,\ldots,\lceil\log k\rceil - 1$, it holds that each of the 2^i-column subsets of S are members of $\mathcal{G}_i$, then $\tilde{M}$ allows any subset A of initially active stations such that $A \subseteq S$ to transmit successfully.

Assume that the hypothesis of the claim is true. In Section II we observed that any protocol based on an ϵ-almost (k,m,n)-selector schedules transmissions in such a way that at the end there are at most $k-m$ active stations that did not succeed to transmit their messages, provided that all active stations are contained in one of the $(1-\epsilon)\binom{n}{k}$ subsets of k stations corresponding to the k-column subsets that satisfy the property of (k,m,n)-selectors. Consequently, if the set of active stations has size at most 2^i and is subset of some set $S_i \in \mathcal{G}_i$, then the protocol based on $\tilde{M}_{(\epsilon_i,2^i,n)}$ schedules transmissions so that at the end we are left with at most 2^{i-1} active stations. Recall, that an active station becomes inactive immediately after it has transmitted successfully. Let us consider the conflict resolution algorithm based on $\tilde{M}$. Since we are assuming that initially the number of active stations is at most $k \leq 2^{\lceil \log k \rceil}$ and that the subset of up to k active stations is entirely contained in some member of $G_{\lceil \log k \rceil}$, then, by the above argument, one has that after the first $t_{\lceil \log k \rceil}$ steps there are at most $2^{\lceil \log k \rceil - 1}$ active stations. The hypothesis of the claim implies that these up to $2^{\lceil \log k \rceil - 1}$ active stations are contained in some set $S_{\lceil \log k \rceil - 1} \in \mathcal{G}_{\lceil \log k \rceil - 1}$ and therefore, there are at most $2^{\lceil \log k \rceil - 2}$ stations that are still active after the next $t_{\lceil \log k \rceil - 1}$ steps. Continuing in this way, one can see that, for each $i = 1, \ldots, \lceil \log k \rceil$, after the first $t_{\lceil \log k \rceil} + t_{\lceil \log k \rceil - 1} + \ldots + t_i$ steps, the algorithm is left with at most 2^{i-1} active stations. At the end, after $\sum_{i=1}^{\lceil \log k \rceil} t_i$ steps, there is at most a single station which is still active. Obviously, this last station can transmit without conflict with other active stations. The all-1 row at the bottom of the matrix takes care of this last active station. This concludes the proof of the claim.

In the following, we will prove that the hypothesis of the claim holds for at least $(1-\epsilon)\binom{n}{k}$ subsets S of k columns of $\tilde{M}$, thus showing that $\tilde{M}$ is an ϵ-almost KG(k,n)-code.

For $i = 1, \ldots, \lceil \log k \rceil - 1$, any of the $\epsilon_i \binom{n}{2^i}$ subsets of 2^i columns not belonging to $\mathcal{G}_i$ could be contained in up to $\binom{n-2^i}{k-2^i}$ k-column subsets of $\tilde{M}$. It follows that the number of k-column subsets S of $\tilde{M}$ that do not satisfy the condition of the claim is at most

$$\sum_{i=1}^{\lceil \log k \rceil - 1} \epsilon_i \binom{n}{2^i}\binom{n-2^i}{k-2^i} = \sum_{i=1}^{\lceil \log k \rceil - 1} \epsilon_i \binom{k}{2^i}\binom{n}{k}. \quad (18)$$

Now we need to estimate the total number of k-column subsets S such that $S \subseteq S'$ for some $S' \in \mathcal{G}_{\lceil \log k \rceil}$. To this aim we use the well known fact (see [8], Cor. 4, p. 13) that for any family $\mathbf{F}$ of m-subsets on the ground set $[n]$, $m \leq n$, and for any positive integer $s \leq m$, we have that

$$\begin{aligned} &|\{F' \subset [n] : |F'| = s \text{ and } F' \subset F \text{ for some } F \in \mathbf{F}\}| \\ &\quad \geq \frac{\binom{n}{s}}{\binom{n}{m}} \cdot |\mathbf{F}|. \end{aligned} \quad (19)$$

By replacing m with $2^{\lceil \log k \rceil}$ and s with k, we have that the total number of k-column subsets S such that $S \subseteq S'$ for some $S' \in \mathcal{G}_{\lceil \log k \rceil}$ is at least $|\mathcal{G}_{\lceil \log k \rceil}| \frac{\binom{n}{k}}{\binom{n}{2^{\lceil \log k \rceil}}} \geq (1 - \epsilon_{\lceil \log k \rceil})\binom{n}{2^{\lceil \log k \rceil}} \frac{\binom{n}{k}}{\binom{n}{2^{\lceil \log k \rceil}}} = (1 - \epsilon_{\lceil \log k \rceil})\binom{n}{k}$. By (18) one has that at most $\sum_{i=0}^{\lceil \log k \rceil - 1} \epsilon_i \binom{k}{2^i}\binom{n}{k}$ of these subsets do not satisfy the hypothesis of the claim. It follows that the number of k-column subsets for which the claim holds is at least

$$\begin{aligned} &(1-\epsilon_{\lceil \log k \rceil})\binom{n}{k} - \sum_{i=1}^{\lceil \log k \rceil - 1} \epsilon_i \binom{k}{2^i}\binom{n}{k} \\ &\quad = \binom{n}{k} - \sum_{i=1}^{\lceil \log k \rceil - 1} \epsilon_i \binom{k}{2^i}\binom{n}{k} - \epsilon_{\lceil \log k \rceil}\binom{n}{k}. \end{aligned}$$

Therefore, the total number of k-column subsets that do not satisfy property (*) is at most

$$\sum_{i=1}^{\lceil \log k \rceil - 1} \epsilon_i \binom{k}{2^i}\binom{n}{k} + \epsilon_{\lceil \log k \rceil}\binom{n}{k}.$$

By setting ϵ_i as in (17), we obtain that the above bound is equal to $\epsilon\binom{n}{k}$ thus proving that $\tilde{M}$ is an ϵ-almost KG(k,n)-code.

Now we derive the claimed upper bound on the number of rows of $\tilde{M}$. Upper bound (16) and equation (17) imply that the length of $\tilde{M}$ is

$$\begin{aligned} t &< 2e \sum_{i=1}^{\lceil \log k \rceil} \left(1 + \ln \frac{\binom{2^i}{2^{i-1}+1}}{\epsilon_i}\right) + 1 \\ &= 2e\lceil \log k \rceil + 2e \sum_{i=1}^{\lceil \log k \rceil - 1} \ln \frac{\binom{2^i}{2^{i-1}+1}\binom{k}{2^i}\lceil \log k \rceil}{\epsilon} \\ &\quad + 2e \ln \frac{\binom{2^{\lceil \log k \rceil}}{2^{\lceil \log k \rceil - 1}+1}\lceil \log k \rceil}{\epsilon} + 1. \end{aligned} \quad (20)$$

We upper bound the righthand side of (20) by noticing that $\binom{2^i}{2^{i-1}+1} < \binom{2^i}{2^{i-1}}$ and applying the well known inequality $\binom{a}{c} \leq (ea/c)^c$ to $\binom{2^i}{2^{i-1}}$ and $\binom{k}{2^i}$, thus obtaining

$$\begin{aligned} t &< 2e\lceil \log k \rceil \\ &\quad + 2e \sum_{i=1}^{\lceil \log k \rceil - 1} \ln \left(\frac{(2e)^{2^{i-1}}(ek/2^i)^{2^i}\lceil \log k \rceil}{\epsilon}\right) \\ &\quad + 2e \ln \left(\frac{(2e)^{2^{\lceil \log k \rceil - 1}}\lceil \log k \rceil}{\epsilon}\right) + 1 \\ &= 2e\lceil \log k \rceil \left(1 + \ln\left(\frac{\lceil \log k \rceil}{\epsilon}\right)\right) \\ &\quad + 2e \ln(2e) \sum_{i=1}^{\lceil \log k \rceil} 2^{i-1} + 2e \ln e \sum_{i=1}^{\lceil \log k \rceil - 1} 2^i \\ &\quad + 2e \sum_{i=1}^{\lceil \log k \rceil - 1} 2^i \ln(k/2^i) + 1 \\ &< 2e\lceil \log k \rceil \left(1 + \ln\left(\frac{\lceil \log k \rceil}{\epsilon}\right)\right) \\ &\quad + 2e(2k-1)\ln(2e^2) \end{aligned}$$

$$+4ek\sum_{i=1}^{\lceil\log k\rceil-1}\frac{1}{2^i}\ln 2^i+1. \quad (21)$$

The summation in (21) follows from observing that

$$\begin{aligned}\sum_{i=1}^{\lceil\log k\rceil-1}2^i\ln\left(k/2^i\right) &< 2k\sum_{i=1}^{\lceil\log k\rceil-1}\frac{2^i}{2^{\lceil\log k\rceil}}\ln\left(\frac{2^{\lceil\log k\rceil}}{2^i}\right)\\ &= 2k\sum_{i=1}^{\lceil\log k\rceil-1}\frac{1}{2^i}\ln 2^i\end{aligned}$$

By using the well known inequality $\sum_{i=0}^{\infty} ix^i \leq \frac{x}{(1-x)^2}$, holding for $|x|<1$, to bound from above $\sum_{i=0}^{\lceil\log k\rceil-1}\frac{i}{2^i}$ in (21), it follows $t < 2e\lceil\log k\rceil\left(1+\ln\left(\frac{\lceil\log k\rceil}{\epsilon}\right)\right)+2e(2k-1)\ln(2e^2)+8ek\ln 2+1$, that concludes the proof of the theorem. ∎

The discussion following Definition 3 in Section II implies that, in the context of conflict resolution for a multiple-access channel with feedback, an ϵ-almost KG(k,n)-codes are equivalent to conflict resolution algorithms that, for at least a ratio $(1-\epsilon)$ of all possible subsets of k stations, guarantee all active stations to transmit successfully if they are entirely contained in one of those k-subsets. In view of this equivalence, Theorem 2 is a direct consequence of Theorem 8.

B. Non-existence results

In this section we derive lower bounds on the minimum number of rows of ϵ-almost KG(k,n)-codes and ϵ-almost (k,k,n)-selectors.

Theorem 9: Given integers k and n, with $2\leq k\leq n$, and a real number ϵ, with $\frac{1}{\binom{n}{k}}\leq\epsilon\leq\left(\frac{k}{2n}\right)^2$, one has that the minimum size of an ϵ-almost (k,k,n)-selector is

$$\Omega\left(\frac{\left(\log\frac{1}{\epsilon}\right)^2}{\left(\log\frac{n}{k}\right)\left(\log\log\frac{1}{\epsilon}-\log\log\frac{n}{k}\right)}\right).$$

Proof: Let M be an ϵ-almost (k,k,n)-selector with t rows. Suppose that there exists an integer f, with $2\leq f\leq k$, such that for each submatrix $\tilde{M}$ of f columns of M, one has that $\tilde{M}$ contains the f rows of the identity matrix I_f. If such an integer f exists, then M is a (classical) (f,f,n)-selector and we can obtain a lower bound on the number of rows of M by exploiting the $\Omega(\frac{k^2}{\log k}\log\frac{n}{k})$, $k\geq 2$, lower bound on the minimum number of rows of (k,k,n)-selectors (or equivalently $(k-1)$-disjunct codes) in (2). By replacing k with f in that lower bound, we get

$$t=\Omega\left(\frac{f^2}{\log f}\log\frac{n}{f}\right). \quad (22)$$

The rest of the proof is devoted to finding a value of f such that each f-column submatrix of M contains the f rows of the identity matrix I_f, so that we can apply lower bound (22) on t. Obviously, in order to get a good lower bound, we aim at finding a value of f as large as possible. To this aim, we compute a lower bound on the smallest integer g, $3\leq g\leq k$ such that there is at least a g-column submatrix M' of M that does not contain one or more rows of I_g. In order to derive a lower bound on g, we observe that, for any k-column submatrix M'' containing all g columns of M', one has that M'' does not contain one or more rows of the identity matrix I_k. Indeed, if otherwise, there would be k pairwise distinct rows of M'' having a single entry equal to 1, and in g of these rows the 1-entry would be at the intersection with the columns of M'. It is immediate to see that if restrict these g rows to the entries at intersection with the columns of M', we obtain the g rows of the identity matrix I_g, thus contradicting the assumption that M' does not contain one or more rows of I_g. The number of k-column submatrices containing all columns of M' is $\binom{n-g}{k-g}$, and by the above discussion, one has that each of these submatrices does not contain one or more rows of I_k. On the other hand, by definition of ϵ-almost (k,k,n)-selector, there are at most $\epsilon\binom{n}{k}$ submatrices of k columns of M such that each of those submatrices does not contain one or more rows of the identity matrix I_k. As a consequence, the following inequality must be satisfied:

$$\binom{n-g}{k-g}\leq\epsilon\binom{n}{k}. \quad (23)$$

Notice that we can restrict our attention to values of g smaller than $\lceil k/2\rceil$ because for $g\geq\lceil k/2\rceil$, it is immediate to see that the lower bound in the statement of the theorem holds. Indeed, for $g\geq\lceil k/2\rceil$, M is an ϵ-almost (f,f,n)-selector for any $f<\lceil k/2\rceil$, and by setting $f=\lceil k/2\rceil-1$ in lower bound (22), we get $t=\Omega\left(\frac{k^2}{\log k}\log\frac{n}{k}\right)$, i.e., the same lower bound holding for the case when all k-column submatrices contain all rows of the identity matrix I_k. The hypothesis $\epsilon\geq\frac{1}{\binom{n}{k}}$, along with the well known inequality

$$\binom{a}{c}\leq(ea/c)^c, \quad (24)$$

implies that $\log\frac{1}{\epsilon}\leq\log\left(\frac{en}{k}\right)^k$. The hypothesis $\epsilon\leq\left(\frac{k}{2n}\right)^2$ implies that the lower bound in the statement of the theorem increases with $\frac{1}{\epsilon}$, and consequently, one has that the above lower bound $t=\Omega\left(\frac{k^2}{\log k}\log\frac{n}{k}\right)$ implies the lower bound in the statement of the theorem. Therefore, from now on, we assume $g<\lceil k/2\rceil$. By this assumption, we have that

$$\begin{aligned}\frac{\binom{n}{k}}{\binom{n-g}{k-g}} &= \frac{n}{k}\cdot\frac{n-1}{k-1}\cdot\ldots\cdot\frac{n-g+1}{k-g+1}<\left(\frac{n}{k-g}\right)^g\\ &< \left(\frac{2n}{k}\right)^g. \quad (25)\end{aligned}$$

Inequalities (23) and (25) imply that

$$\frac{1}{\epsilon}<\left(\frac{2n}{k}\right)^g,$$

from which we obtain that

$$g>\frac{\log\frac{1}{\epsilon}}{\log\left(\frac{2n}{k}\right)}. \quad (26)$$

Observe that, by the hypothesis of the theorem, it holds $\epsilon\leq\left(\frac{k}{2n}\right)^2$, and this inequality along with (26) implies $g\geq 3$, as desired. From (26) it follows that the smallest integer g such that M is not a (g,g,n)-selector is at least as large as

$\left\lceil \frac{\log \frac{1}{\epsilon}}{\log\left(\frac{2n}{k}\right)} \right\rceil$. Consequently, for $f = \left\lceil \frac{\log \frac{1}{\epsilon}}{\log\left(\frac{2n}{k}\right)} \right\rceil - 1$, one has that M is an (f, f, n)-selector and, by replacing k with f in lower bound (22), we get the following lower bound on the number t of rows of M:

$$t = \Omega\left(\frac{\left(\log \frac{1}{\epsilon}\right)^2}{\left(\log \frac{2n}{k}\right)^2 \left(\log\log \frac{1}{\epsilon} - \log\log \frac{2n}{k}\right)} \log\left(\frac{n \log \frac{2n}{k}}{\log \frac{1}{\epsilon}}\right)\right).$$

Notice that the assumption $g < \lceil k/2 \rceil$, along with (26), implies that

$$\frac{\log \frac{1}{\epsilon}}{\log\left(\frac{2n}{k}\right)} < \frac{k}{2}. \tag{27}$$

By inequality (27), we have that $\log\left(\frac{n \log \frac{2n}{k}}{\log \frac{1}{\epsilon}}\right)$ is larger than $\log \frac{2n}{k}$, and consequently, the above lower bound on t is

$$\Omega\left(\frac{\left(\log \frac{1}{\epsilon}\right)^2}{\left(\log \frac{2n}{k}\right)\left(\log\log \frac{1}{\epsilon} - \log\log \frac{2n}{k}\right)}\right),$$

which implies the lower bound in the statement of the theorem. ∎

The discussion in Section II implies that an ϵ-almost (k, k, n)-selector is equivalent to a scheduling algorithm that schedules the transmissions in such a way that, for at least $(1-\epsilon)\binom{n}{k}$ subsets of k stations, one has that if the subset of up to k active stations is entirely contained in one of these k-station subsets then all active stations transmit with success. Therefore, a lower bound on the minimum number of rows of ϵ-almost (k, k, n)-selectors translates into a lower bound on the minimum number of time slots needed by such a scheduling algorithm. The lower bound of Theorem 3 is an immediate consequence of the lower bound of Theorem 9.

Notice that Theorem 9 holds for $\frac{1}{\binom{n}{k}} \leq \epsilon \leq \left(\frac{k}{2n}\right)^2$. This hypothesis does not impose a significant constraint on ϵ since for $\epsilon = \frac{1}{\binom{n}{k}} - \delta$, for *any* $\delta > 0$, ϵ-almost (k, k, n)-selectors are indeed (classical) ϵ-almost (k, k, n)-selectors and the minimum number of rows of these codes is lower bounded by (2). Moreover, having $\epsilon > \left(\frac{k}{2n}\right)^2$ would lead to an ϵ-almost (k, k, n)-selector with very little structure, since more than $\left(\frac{k}{2n}\right)^2 \binom{n}{k} \geq \frac{1}{4}\left(\frac{n}{k}\right)^{k-2}$ of its k-column submatrices would not satisfy the desired property.

Notice that, for $\epsilon \geq \frac{1}{\binom{n}{k}}$, it holds that $\log\log \frac{1}{\epsilon} - \log\log \frac{n}{k} \leq \log\log \binom{n}{k} - \log\log \frac{n}{k}$, which, by (24), is at most $\log k + \log\log\left(e\frac{n}{k}\right) - \log\log \frac{n}{k} = O(\log k)$. Therefore, the lower bound of Theorem 9 is

$$t = \Omega\left(\frac{\left(\log \frac{1}{\epsilon}\right)^2}{\left(\log \frac{n}{k}\right)(\log k)}\right). \tag{28}$$

Observe that above lower bound increases as ϵ becomes smaller and approaches $\Omega\left(\frac{\left(\log \binom{n}{k}\right)^2}{(\log \frac{n}{k})(\log k)}\right)$, for ϵ approaching $\frac{1}{\binom{n}{k}}$. In virtue of (6), this lower bound is $\Omega\left(\frac{k^2}{\log k} \log \frac{n}{k}\right)$, which is the same as the lower bound in (2), holding for the case when all k-column submatrices contain all rows of the identity matrix I_k.

Below, we compare lower bound (28) with the following upper bound on the minimum number of rows of a (k, k, n)-selector, obtainable by setting $m = k$ in the upper bound of Theorem 7:

$$O\left(k \log \frac{k}{\epsilon}\right). \tag{29}$$

The ratio between upper bound (29) and lower bound (28) is

$$O\left(\frac{k\left(\log \frac{k}{\epsilon}\right)\left(\log \frac{n}{k}\right)(\log k)}{\left(\log \frac{1}{\epsilon}\right)^2}\right). \tag{30}$$

For $\epsilon \leq (1/k)^c$, for any constant $c > 0$, it is $\log \frac{k}{\epsilon} \leq \left(\frac{c+1}{c}\right) \log \frac{1}{\epsilon}$, and consequently (30) is

$$O\left(\frac{k(\log \frac{n}{k})(\log k)}{\left(\log \frac{1}{\epsilon}\right)}\right).$$

If we replace k with $k+1$ in the statement of Theorem 9, we obtain a lower bound on the minimum length of ϵ-almost k-disjunct codes. The above discussion implies that, for $\frac{1}{\binom{n}{k}} \leq \epsilon \leq \left(\frac{1}{k}\right)^c$, for any positive constant c, the upper bound of Corollary 1 differs by an $O\left(\frac{k(\log \frac{n}{k})(\log k)}{\left(\log \frac{1}{\epsilon}\right)}\right)$ factor from this lower bound.

The following theorem provides a lower bound on the minimum length of KG(k, n)-codes.

Theorem 10: Given integers k and n, with $2 \leq k \leq n$, and a real number ϵ, with $\frac{1}{\binom{n}{k}} \leq \epsilon \leq \left(\frac{k}{2n}\right)^2$, one has that the minimum size of an ϵ-almost KG(k, n)-code is

$$\Omega\left(\log \frac{1}{\epsilon}\right).$$

Proof: The proof of the lower bound in the statement of the theorem is similar to that of Theorem 9. Let M be an ϵ-almost KG(k, n)-code with t rows. Suppose that there exists an integer f, with $2 \leq f \leq k$, such that each subset S of f column indices of M satisfies property (*) of Section I, with k being replaced by f. In this case, M is a KG(f, n)-code and we can compute a lower bound on the number t of rows of M by exploiting the $\Omega(k \log \frac{n}{k})$ lower bound in (5), i.e., the lower bound on the minimum length of (classical) KG(k, n)-codes with $k \geq 2$. By replacing k with f in this lower bound, we get

$$t = \Omega\left(f \log \frac{n}{f}\right). \tag{31}$$

In order to obtain the lower bound in the statement of the theorem, we compute a value of f for which one is guaranteed that each subset of f columns of M satisfies property (*), so that we can apply lower bound (31) on t. Let g denote the smallest integer g such that there exists at least a set S' of g column indices that does not satisfy property (*), with k being replaced by g in that property. Notice that if a set S'' of k-column indices of M contains all g indices in S', then one has that S'' also does not satisfy property (*). Indeed, suppose by contradiction that S'' satisfies property (*), i.e., there are k row indices $i_1, i_2, \ldots, i_k$, with $i_1 < i_2 < \ldots < i_k$, and a permutation $[j_1, \ldots, j_k]$ of the indices in S'', such that

the submatrix formed by rows with indices $i_1, \ldots, i_k$, taken in this order, and columns with indices $j_1, \ldots, j_k$, taken in this order, form a $k \times k$ lower unitriangular matrix. Now let us remove from $[j_1, \ldots, j_k]$ the indices that do not belong to S', and let us denote by $[j'_1, \ldots, j'_g]$ the resulting permutation of the column indices in S'. For each column index j_z removed from $[j_1, \ldots, j_k]$, we remove the corresponding index i_z from the ordered sequence of row indices $i_1, \ldots, i_k$ and denote by $i'_1, \ldots, i'_g$ the ordered sequence of the remaining row indices. The submatrix formed by rows with indices $i'_1, \ldots, i'_g$, taken in this order, and columns with indices $j'_1, \ldots, j'_g$, taken in this order, form a $g \times g$ lower unitriangular matrix, thus contradicting the assumption that S' does not satisfy property (*). The above discussion implies that there are at least $\binom{n-g}{k-g}$ subsets of k column indices of M that do not satisfy property (*) and, since M is an ϵ-almost KG(k,n)-code, one has that the number of such subsets does not exceed $\epsilon\binom{n}{k}$. As in the proof of Theorem 9, we can restrict our attention to values of g smaller than $\lceil k/2 \rceil$. Indeed, for $g \geq \lceil k/2 \rceil$, M is an ϵ-almost KG(f,n)-code for any $f < \lceil k/2 \rceil$, and by setting $f = \lceil k/2 \rceil - 1$ in lower bound (31), one obtains the lower bound $t = \Omega\left(k \log \frac{n}{k}\right)$, i.e., the same lower bound as the one holding for (classical) KG(k,n)-codes. In virtue of (24), this bound implies the lower bound in the statement of the theorem for any $\epsilon \geq \frac{1}{\binom{n}{k}}$. Therefore, in the rest of the proof we will assume $g < \lceil k/2 \rceil$. By the same calculations as in the proof of Theorem 9, we obtain that the inequality $\binom{n-g}{k-g} \leq \epsilon\binom{n}{k}$ implies that $g \geq \left\lceil \frac{\log \frac{1}{\epsilon}}{\log \frac{2n}{k}} \right\rceil$. Consequently, for $f = \left\lceil \frac{\log \frac{1}{\epsilon}}{\log \frac{2n}{k}} \right\rceil - 1$, M is a KG(f,n)-code and lower bound (31) on t holds and implies the following lower bound

$$t = \Omega\left(\frac{\log \frac{1}{\epsilon}}{\log \frac{2n}{k}} \log\left(\frac{n \log \frac{2n}{k}}{\log \frac{1}{\epsilon}}\right)\right). \tag{32}$$

As in the proof of Theorem 9, one has that inequality (27) holds. From that inequality, it follows that $\frac{\log \frac{2n}{k}}{\log \frac{1}{\epsilon}} \geq \frac{2}{k}$, and consequently, it is $\log\left(\frac{n \log \frac{2n}{k}}{\log \frac{1}{\epsilon}}\right) \geq \log \frac{2n}{k}$, which, along with (32), implies the lower bound in the statement of the theorem. ∎

Similarly to what we have done in the discussion following Theorem 9, we can argue that the hypothesis on ϵ in Theorem 10 does not represent a significant constraint.

In Section II, we have seen that ϵ-almost KG(k,n)-codes are equivalent to conflict resolution algorithms, for the multiple-access channel with feedback, that for at least a $(1-\epsilon)$ fraction of all possible subsets of k stations, allow all (up to k) active stations to transmit with success, provided that they are contained in one of those k-station subsets. The lower bound of Theorem 4 is an immediate consequence of this fact and of Theorem 10.

Notice that the upper bound of Theorem 8 differs asymptotically from the lower bound of Theorem 10 by a factor equal to

$$\frac{(\log k)\left(\log\left(\frac{\log k}{\epsilon}\right)\right) + k}{\log \frac{1}{\epsilon}}$$

$$\frac{(\log k)(\log\log k)}{\log \frac{1}{\epsilon}} + \log k + \frac{k}{\log \frac{1}{\epsilon}},$$

and therefore, we have that the gap between our upper and lower bounds on the length of ϵ-almost KG(k,n)-code is $O\left(\log k + \frac{k}{\log \frac{1}{\epsilon}}\right)$.

V. Improvements

A possible drawback of our solutions is that our scheduling algorithms do not offer any guarantee for (at most) a fraction ϵ of all $\binom{n}{k}$ k-subsets of conflicting stations. However, if one concatenates a (k,m,n)-selector to an ϵ-almost (k,k,n)-selector, then one obtains a scheduling algorithm that allows at least m active stations to transmit successfully even if the k-subset of active stations falls within these $\epsilon\binom{n}{k}$ subsets. The following theorem follows from setting $m = k$ in the upper bound of Theorem 7 and from upper bound (3) on the size of (k,m,n)-selectors.

Theorem 11: Given integers k, m, and n, with $1 \leq m \leq k \leq n$, and a real number $0 < \epsilon < 1$, there exists an n-column matrix such that

1) each subset of k columns forms a $k \times k$ submatrix that contains at least m rows of the identity matrix I_k and
2) for at least $(1-\epsilon)\binom{n}{k}$ subsets of k columns one has that each of these subsets forms a $k \times k$ submatrix that contains all rows of the identity matrix I_k

and has a number of rows

$$t \leq ek\left(1 + \ln \frac{k}{\epsilon}\right) + \frac{ek^2}{k-m+1} \ln\left(\frac{n}{k}\right) + \frac{ek(2k-1)}{k-m+1}.$$

For $m = k+1-\frac{k \ln(n/k)}{c \ln(k/\epsilon)}$, where c is an arbitrary constant such that $1 \leq c \leq \frac{k \ln\left(\frac{n}{k}\right)}{\ln\left(\frac{k}{\epsilon}\right)}$, the scheduling algorithm of Theorem 11 uses the same asymptotic number of time slots as the algorithm based on ϵ-almost (k,k,n)-selectors. Indeed, by setting $m = k+1-\frac{k \ln(n/k)}{c \ln(k/\epsilon)}$ in the statement of Theorem 11, we get the following result.

Corollary 2: Let k and n be integers such that $1 \leq k \leq n$, and let ϵ be a real number such that $\frac{k^k}{n^{k-1}} \leq \epsilon \leq \frac{k^2}{n}$. There exists a conflict resolution algorithm for a multiple-access channel without feedback that schedules the transmissions of n stations in such a way that for at least a $(1-\epsilon)$ ratio of all possible subsets of k active stations, the algorithm allows all k conflicting stations to transmit successfully, whereas for the remaining subsets of k stations it allows at least $k+1-\frac{k \ln(n/k)}{c \ln(k/\epsilon)}$ stations to transmit successfully, where c is an arbitrary constant such that $1 \leq c \leq \frac{k \ln\left(\frac{n}{k}\right)}{\ln\left(\frac{k}{\epsilon}\right)}$. The number of time slots used by the conflict resolution algorithm is

$$t \leq e(c+1)k \ln\left(\frac{k}{\epsilon}\right) + ec(2k-1)\frac{\ln(k/\epsilon)}{\ln(n/k)} + ek.$$

The above corollary implies that for any $\frac{k^k}{n^{k-1}} \leq \epsilon \leq \frac{k^2}{n}$ and for any arbitrary constant c, with $1 \leq c \leq \frac{k \ln\left(\frac{n}{k}\right)}{\ln\left(\frac{k}{\epsilon}\right)}$, there exists a scheduling algorithm such that

- it uses $t \leq e(c+1)k \ln\left(\frac{k}{\epsilon}\right) + ec(2k-1)\frac{\ln(k/\epsilon)}{\ln(n/k)} + ek$ time slots, i.e., the *same* asymptotic number of time slots

as the scheduling algorithm based on ϵ-almost (k,k,n)-selectors, and

- in addition to solving *all conflicts* among a ratio $(1-\epsilon)$ of all possible subsets of k active stations, it allows at least $k\left(\frac{c-1}{c}\right)+1$ stations to transmit successfully *whichever* the subset of k active stations is.

Notice that for $\epsilon=\frac{k^2}{n}$, the above scheduling algorithm uses $t=O(k\log(n/k))$ time slots, i.e., the same asymptotic number of time slots used by the scheduling algorithm based on (classical) KG(k,n)-codes, which, however, solves conflicts only under the assumption that the stations receive feedback from the channel.

We can apply a similar idea to the multiple-access channel with feedback to obtain a scheduling algorithm that solves all conflicts among (up to) k active stations if the subset of active stations is contained in one of the $(1-\epsilon)\binom{n}{k}$ "good" k-subsets, and among a smaller subset of active stations otherwise. Let ϵ_i be defined as in (17) in the proof of Theorem 11. The idea is to concatenate ϵ_i-almost $(2^i,2^{i-1},n)$-selectors, as in the construction of Theorem 8, for values of i larger than an *appropriately* chosen $\hat{i}$ (that will be explicitly determined later on), and (classical) $(2^i,2^{i-1},n)$-selectors for smaller i's. More precisely, the desired matrix is obtained by first concatenating the ϵ_i-almost $(2^i,2^{i-1},n)$-selectors from the top to the bottom in order of decreasing i, starting with $i=\lceil\log k\rceil$ and ending with $i=\hat{i}+1$. Then, the rows of the (classical) $(2^i,2^{i-1},n)$-selectors are appended to the bottom of the above described matrix. The rows of the (classical) $(2^i,2^{i-1},n)$-selectors are also arranged in decreasing order of the parameter i, starting with those of the $(2^{\hat{i}},2^{\hat{i}-1},n)$-selector through those of the $(2,1,n)$-selector. An additional all-1 row is appended at bottom of the matrix to allow the last active station to transmit. Let t_u denote the number of rows in the concatenation of the ϵ_i-almost $(2^i,2^{i-1},n)$-selectors and t_ℓ denote the number of rows in the concatenation of (classical) $(2^i,2^{i-1},n)$-selectors. Notice that t_u is upper bounded by the number of rows of the ϵ-almost KG(k,n)-code of Theorem 8, and consequently, it holds that

$$t_u\le 2e\lceil\log k\rceil\ln\left(\frac{\lceil\log k\rceil}{\epsilon}\right)+O(k). \tag{33}$$

For the purpose of obtaining a code whose number of rows is asymptotically the same as the one of Theorem 8, t_ℓ should be upper bounded by the righthand side of (33), too.

In order to obtain an upper bound on t_ℓ, we apply upper bound (3) with k replaced by 2^i and m by 2^{i-1}, for $i=1,\ldots,\hat{i}$, thus obtaining

$$\begin{aligned}
t_\ell &\le \sum_{i=1}^{\hat{i}}\frac{2^{2i}e}{2^i-2^{i-1}+1}\ln\left(\frac{n}{2^i}\right)+\frac{2^ie(2^{i+1}-1)}{2^i-2^{i-1}+1}\\
&< \sum_{i=1}^{\hat{i}}2e2^i\ln\left(\frac{n}{2^i}\right)+4e2^i\\
&= 2e(\ln n+2)\sum_{i=1}^{\hat{i}}2^i-2e\ln 2\sum_{i=1}^{\hat{i}}i2^i\\
&= 2e(\ln n+2)(2^{\hat{i}+1}-1)-2e\ln 2\sum_{i=1}^{\hat{i}}i2^i. \qquad (34)
\end{aligned}$$

By applying the following well known equality

$$\sum_{i=1}^{a}ix^i=\frac{x-x^{a+2}}{(x-1)^2}+\frac{(a+1)x^{a+1}}{x-1} \tag{35}$$

with $x=2$ and $a=\hat{i}$ to the summation in the righthand side of (34), we obtain

$$\begin{aligned}
t_\ell &\le 2e(\ln n+2)(2^{\hat{i}+1}-1)\\
&\quad -2e\ln 2(2+(\hat{i}-1)2^{\hat{i}+1}) \qquad (36)\\
&< 2e(\ln n+2)2^{\hat{i}+1}-2e(\ln 2)(\hat{i}-1)2^{\hat{i}+1} \qquad (37)\\
&= e2^{\hat{i}+2}\ln\left(\frac{e^2n}{2^{\hat{i}-1}}\right). \qquad (38)
\end{aligned}$$

If we set $\hat{i}=\lceil\log f(k)\rceil$, where $f(k)$ is an arbitrary non decreasing function such that $f(k)\le k$, we obtain that (38) is less than $2e\lceil\log k\rceil\ln\left(\frac{\log k}{\epsilon}\right)$, for any $\epsilon\le\log k\left(\frac{f(k)}{2e^2n}\right)^{\frac{4f(k)}{\lceil\log k\rceil}}$. Therefore, we have that $t_\ell\le t_u$, and as a consequence, the total number of rows in the resulting matrix is $t_u+t_\ell+1\le 2t_u+1$. In conclusion, we have proved the following result.

Theorem 12: There exists a scheduling algorithm for the multiple-access channel with feedback such that

- it uses $t=4e\lceil\log k\rceil\ln\left(\frac{\lceil\log k\rceil}{\epsilon}\right)+O(k)$ time slots, i.e., the same asymptotic number of time slots of the scheduling algorithm based on ϵ-almost KG(k,n)-codes, and
- it allows *any* subset of active stations to transmit successfully, provided that there are no more than $f(k)$ active stations, and for at least a ratio $(1-\epsilon)$ of all subsets of k stations, it solves all conflicts among up to k active stations belonging to one of those k-subsets,

where $f(k)$ is an arbitrary non decreasing function such that $f(k)\le k$ and $\epsilon\le\log k\left(\frac{f(k)}{2e^2n}\right)^{\frac{4f(k)}{\lceil\log k\rceil}}$.

Notice that for $\epsilon=\log k\left(\frac{f(k)}{2e^2n}\right)^{\frac{4f(k)}{\lceil\log k\rceil}}$ the above scheduling algorithm uses $t=O(f(k)\log(n/f(k)))$ time slots, which for $f(k)=o(k)$ is asymptotically smaller than the number of time slots (4) used by the scheduling algorithm based on (classical) KG(k,n)-codes. On the other hand, as $f(k)$ approaches k, the values of the ratio ϵ for which Theorem 12 holds get smaller and smaller, and consequently, the number of time slots used by the conflict resolution algorithm approaches the number of time slots used by the scheduling algorithm based on (classical) KG(k,n)-codes, as one would expect.

VI. Solving conflicts among an unknown number of active stations

In this section we consider the case when there is no a priori knowledge on the maximum number of stations that can be active at the same time. Notice that in the multiaccess model without feedback the problem is void of interest, since any conflict resolution algorithm for this case needs to use exactly n time slots. Indeed, the algorithm must reserve a different time slot for each of the n stations to transmit singly over the

channel since it has to cope with the eventuality that all n stations are active at the same time. For that reason, in this section we focus on conflicts resolution strategies for multiple-access channels with feedback. In this model, the problem of efficiently solving conflicts within an a priori unknown number of active stations poses non trivial challenges. Nevertheless, we will show that we can solve the problem with protocols that use, essentially, the same number of time slots as the protocols that work *assuming* the knowledge of an upper bound on the number of active stations.

Differently from the case when the number of active stations is bounded by the known parameter k, in the case presently considered, the maximum number of time slots needed to solve all conflicts *does not* correspond to the number of rows of the underlying combinatorial structure. This is due to the fact that the combinatorial structure is designed so as to deal with any possible value of the actual number of active stations, whereas conflicts are solved as soon as all active stations transmit with success and no further transmission needs to occur.

The idea at the basis of the conflict resolution algorithm of Theorem 5 is that of making successive guesses on the number of active stations and trying to solve conflicts for each of the guessed values. The algorithm succeeds in solving all conflicts if the guessed value is an upper bound on the actual number of active stations. Before proving Theorem 5 we prove the following preliminary result that applies the above explained idea.

Theorem 13: Let k^* and n be integers such that k^* is not known in advance and $1 \leq k^* \leq n$, and let ϵ be a real number such that $0 < \epsilon < 1$. 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 for at least $(1-\epsilon)\binom{n}{2^{2^{\lceil \log\log k^* \rceil}}}$ possible subsets of $2^{2^{\lceil \log\log k^* \rceil}}$ stations, one has that if the set of k^* active stations is entirely contained in one of those subsets then all active stations transmit successfully and the algorithm uses a number of time slots t, with

$$t < 8e \log k^* \ln\left(\frac{\log k^*}{\epsilon}\right) + O\left(\min\{{k^*}^2, n\}\right).$$

Proof: Let $\mathcal{A}_k$ be the conflict resolution algorithm whose existence is stated by Theorem 2 and let t_k be the number of time slots used by $\mathcal{A}_k$. Recall that $\mathcal{A}_k$ schedules the transmissions in such a way that for at least a $(1-\epsilon)$ ratio of all possible subsets of k stations, one has that if the set of active stations is entirely contained in one of those k-subsets then all active stations transmits successfully. Let us denote by $\mathcal{Q}_k$ the family of those $(1-\epsilon)\binom{n}{k}$ subsets of k stations. For $k = n$, we define $\mathcal{Q}_n = \{[n]\}$. Notice that $\mathcal{A}_k$ works under the hypothesis that the parameter k is known in advance.

We design a conflict resolution algorithm $\mathcal{A}$ which works with an unknown number k^* of active stations as follows. The conflict resolution algorithm $\mathcal{A}$ applies algorithm A_k iteratively, with increasing values of the parameter k. Each application of this algorithm conceptually corresponds to a guess k of the value of the parameter k^*: if the guessed value k is larger than or equal than the real number k^* of active stations then, when running algorithm $\mathcal{A}_k$, all active stations transmit with success provided that they belong to one of the k-subsets in $\mathcal{Q}_k$.

More precisely, algorithm $\mathcal{A}$ works as follows. For $i = 0, 1, 2, \ldots, \lceil \log\log n \rceil - 1$ the active stations transmit, iteratively, according to algorithm $\mathcal{A}_k$ with $k = 2^{2^i}$ and then apply a conflict resolution algorithm which consists in scheduling transmissions by assigning to each of the n stations a different time slot, thus eventually using up to n additional time slots. Notice that for $i = \lceil \log\log k^* \rceil$ one has that the value of $k = 2^{2^i}$ is at least k^*, and consequently, all active stations are scheduled to transmit with success by $\mathcal{A}_k$ provided that they are entirely contained in one of the k-subsets in $\mathcal{Q}_k$. Therefore, if $\lceil \log\log k^* \rceil \leq \lceil \log\log n \rceil - 1$, algorithm $\mathcal{A}$ runs algorithm $\mathcal{A}_k$ with $k = 2^{2^{\lceil \log\log k^* \rceil}}$, and we are guaranteed that after $\sum_{i=0}^{\lceil \log\log k^* \rceil} t_{2^{2^i}}$ all active stations transmit with success provided that they are entirely contained in one of the k-subsets in $\mathcal{Q}_k$. On the other hand, if $\lceil \log\log k^* \rceil = \lceil \log\log n \rceil$, algorithm $\mathcal{A}$ does not run algorithm $\mathcal{A}_k$ with $k = 2^{2^{\lceil \log\log k^* \rceil}}$ and we have no guarantee that any of algorithms $\mathcal{A}_1, \ldots, \mathcal{A}_{\lceil \log\log n \rceil - 1}$ schedules transmissions so as to ensure the desired behavior. However, for any possible value of k^*, algorithm $\mathcal{A}$ allows all active stations to transmit with success by assigning to each of the n stations a different time slot.

Notice that, differently from what happens in the proof of Theorem 8, here the parameter k of the concatenated codes grows at a super-exponential rate. This is a consequence of the mathematics in the analysis below, showing that an exponential growth of the parameter k, as the one in the proof of Theorem 8, would indeed lead to a larger number of rows in the resulting matrix. In particular, one would obtain a $\Theta(\log^2 k^*)$ factor in place of the $\log k^*$ factor that appears in the first term of the claimed upper bound.

In the following we estimate the number of time slots within which active stations transmit with success, provided that they belong to one of the k-subsets in $\mathcal{Q}_k$, where k is either equal to $2^{2^{\lceil \log\log k^* \rceil}}$ or equal to n. Recall that $\mathcal{Q}_n$ is $\{[n]\}$. From the above discussion, such a number of time slots is equal to

$$\sum_{i=0}^{\lceil \log\log k^* \rceil - 1} t_{2^{2^i}} + \begin{cases} n & \text{if } \lceil \log\log k^* \rceil = \lceil \log\log n \rceil, \\ t_{2^{2^{\lceil \log\log k^* \rceil}}} & \text{otherwise.} \end{cases} \tag{39}$$

Let us estimate

$$\sum_{i=0}^{\lceil \log\log k^* \rceil - 1} t_{2^{2^i}} \tag{40}$$

in (39). From Theorem 2 we have

$$\sum_{i=0}^{\lceil \log\log k^* \rceil - 1} t_{2^{2^i}}$$

$$< \sum_{i=0}^{\lceil \log\log k^* \rceil - 1} \left(2e \lceil \log 2^{2^i} \rceil \ln\left(\frac{\log 2^{2^i}}{\epsilon}\right) + O(2^{2^i}) \right)$$

$$
\begin{aligned}
&= 2e \sum_{i=0}^{\lceil \log\log k^* \rceil - 1} 2^i \left(\ln 2 \cdot \log\log 2^{2^i} + \ln\left(\frac{1}{\epsilon}\right)\right) \\
&\quad + \sum_{i=0}^{\lceil \log\log k^* \rceil - 1} O(2^{2^i}) \\
&= 2e \ln 2 \sum_{i=0}^{\lceil \log\log k^* \rceil - 1} 2^i i \\
&\quad + 2e \ln\left(\frac{1}{\epsilon}\right) \sum_{i=0}^{\lceil \log\log k^* \rceil - 1} 2^i + \sum_{i=0}^{\lceil \log\log k^* \rceil - 1} O(2^{2^i}) \\
&= 2e \ln 2 \sum_{i=0}^{\lceil \log\log k^* \rceil - 1} 2^i i \\
&\quad + 2e \ln\left(\frac{1}{\epsilon}\right) (2^{\lceil \log\log k^* \rceil} - 1) + O\left(\sum_{i=0}^{\lceil \log\log k^* \rceil - 1} 2^{2^i} \right).
\end{aligned}
$$

We exploit equality (35) with $x = 2$ and $a = \lceil \log\log k^* \rceil - 1$ to compute the summation $\sum_{i=0}^{\lceil \log\log k^* \rceil - 1} 2^i i$, and we upper bound the quantity $\sum_{i=0}^{\lceil \log\log k^* \rceil - 1} 2^{2^i}$ by $\sum_{j=0}^{\log k^*} 2^j = 2k^* - 1$, thus obtaining that (40) is smaller than

$$
\begin{aligned}
& 2e \ln 2 \left(2 + (\lceil \log\log k^* \rceil - 2) 2^{\lceil \log\log k^* \rceil}\right) + \\
& + 2e \ln\left(\frac{1}{\epsilon}\right) (2^{\lceil \log\log k^* \rceil} - 1) + O(k^*) \\
< \; & 2e \ln 2 \left(2 + (\log\log k^* - 1) 2^{\log\log k^* + 1}\right) \\
& + 2e \ln\left(\frac{1}{\epsilon}\right) (2^{\log\log k^* + 1} - 1) + O(k^*) \\
= \; & 2e \ln 2 \, (2 + (\log\log k^* - 1) 2 \log k^*) \\
& + 2e \ln\left(\frac{1}{\epsilon}\right) (2 \log k^* - 1) + O(k^*) \\
= \; & 4e \ln 2 (1 - \log k^*) + 4e \log k^* \ln\left(\frac{\log k^*}{\epsilon}\right) \\
& - 2e \ln\left(\frac{1}{\epsilon}\right) + O(k^*) \\
= \; & 4e \log k^* \ln\left(\frac{\log k^*}{\epsilon}\right) + 4e \ln 2 \\
& - 2e \ln\left(\frac{1}{\epsilon}\right) - 4e \ln k^* + O(k^*). \qquad (41)
\end{aligned}
$$

Now let us now estimate $t_{2^{2^{\lceil \log\log k^* \rceil}}}$. Notice that

$$
\begin{aligned}
& t_{2^{2^{\lceil \log\log k^* \rceil}}} \\
& = 2e \lceil \log 2^{2^{\lceil \log\log k^* \rceil}} \rceil \ln\left(\frac{\log 2^{2^{\lceil \log\log k^* \rceil}}}{\epsilon}\right) \\
& \quad + O\left(2^{2^{\lceil \log\log k^* \rceil}}\right) \\
& < 2e 2^{\log\log k^* + 1} \ln\left(\frac{2^{\log\log k^* + 1}}{\epsilon}\right) \\
& \quad + O\left(2^{2^{\lceil \log\log k^* \rceil}}\right) \\
& = 4e \log k^* \ln\left(\frac{2 \log k^*}{\epsilon}\right) + O\left(2^{2^{\lceil \log\log k^* \rceil}}\right) \\
& = 4e \log k^* \ln\left(\frac{\log k^*}{\epsilon}\right) + 4e (\ln 2) \log k^* \\
& \quad + O\left(2^{2^{\lceil \log\log k^* \rceil}}\right). \qquad (42)
\end{aligned}
$$

Therefore, by upper bound (41) on (40), and by (42), expression (39) is less than

$$
\begin{cases}
4e \log k^* \ln\left(\dfrac{\log k^*}{\epsilon}\right) & \\
\quad + O(k^*) + n & \text{if } \lceil \log\log k^* \rceil = \lceil \log\log n \rceil, \\
8e \log k^* \ln\left(\dfrac{\log k^*}{\epsilon}\right) & \\
\quad + O(k^*) + O\left(2^{2^{\lceil \log\log k^* \rceil}}\right) & \text{otherwise.}
\end{cases}
$$

The above upper bound is at most

$$
\begin{aligned}
& 8e \log k^* \ln\left(\frac{\log k^*}{\epsilon}\right) \\
& + \begin{cases} O(n) & \text{if } \lceil \log\log k^* \rceil = \lceil \log\log n \rceil, \\ O\left(2^{2^{\lceil \log\log k^* \rceil}}\right) & \text{otherwise.} \end{cases} \qquad (43)
\end{aligned}
$$

For $\lceil \log\log k^* \rceil = \lceil \log\log n \rceil$, it is immediate to see that the above upper bound implies the bound in the statement of the theorem, since in this case it holds $\min\{{k^*}^2, n\} = n$. In order to prove the stated upper bound for $\lceil \log\log k^* \rceil < \lceil \log\log n \rceil$, we need to show that in this case it holds that $\min\{{k^*}^2, n\} \geq 2^{2^{\lceil \log\log k^* \rceil}}$. If $\min\{{k^*}^2, n\} = {k^*}^2$ then that inequality is obviously satisfied. If $\min\{{k^*}^2, n\} = n$, then it holds that $\lceil \log\log n \rceil = \lceil \log\log k^* \rceil + 1$, from which it follows that $n > 2^{2^{\lceil \log\log n \rceil - 1}} = 2^{2^{\lceil \log\log k^* \rceil}}$. Therefore, one has that (43) implies the stated upper bound also for $\lceil \log\log k^* \rceil < \lceil \log\log n \rceil$ ∎

Notice that the algorithm of Theorem 13 schedules transmissions so that there are at least $(1-\epsilon)\binom{n}{2^{2^{\lceil \log\log k^* \rceil}}}$ subsets of $2^{2^{\lceil \log\log k^* \rceil}}$ stations such that if the k^* active stations belong to one of those subsets, then they all transmit with success. In the proof of Theorem 13 we have denoted by $\mathcal{Q}_{2^{2^{\lceil \log\log k^* \rceil}}}$ the family of such subsets of stations.

The rest of this section is devoted to prove Theorem 5. To this aim, we just need to show that the algorithm of Theorem 13 guarantees, for at least $(1-\epsilon)\binom{n}{k^*}$ possible subsets of k^* stations, that if the set of active stations is one of those k^*-subsets then all active stations transmit successfully. In other words, we need to prove that there are at least $(1-\epsilon)\binom{n}{k^*}$ *distinct* k^*-subsets such that each of them is contained in some member of $\mathcal{Q}_{2^{2^{\lceil \log\log k^* \rceil}}}$. In order to prove that, we apply inequality (19) in the proof of Theorem 8 with $m = 2^{2^{\lceil \log\log k^* \rceil}}$, $s = k^*$, and $\mathbf{F} = \mathcal{Q}_{2^{2^{\lceil \log\log k^* \rceil}}}$. Hence, we obtain that there exists a family of k^*-subsets such that each member is contained in one member of $\mathcal{Q}_{2^{2^{\lceil \log\log k^* \rceil}}}$ and has size at least

$$
\begin{aligned}
& \frac{\binom{n}{k^*}}{\binom{n}{2^{2^{\lceil \log\log k^* \rceil}}}} \cdot \left| \mathcal{Q}_{2^{2^{\lceil \log\log k^* \rceil}}} \right| \\
& \geq \frac{\binom{n}{k^*}}{\binom{n}{2^{2^{\lceil \log\log k^* \rceil}}}} \cdot (1-\epsilon) \binom{n}{2^{2^{\lceil \log\log k^* \rceil}}} \\
& = (1-\epsilon)\binom{n}{k^*}.
\end{aligned}
$$

This completes the proof of Theorem 5.

FINAL REMARKS

In this paper we have initiated the study of conflict resolution protocols that trade off a possibility of failing to resolve conflicts on a limited set of inputs for a better efficiency of the protocols. By using tools from hypergraph theory we have seen that is indeed possible to design protocols that are significantly shorter than the classical conflict resolution algorithms, while maintaining the probability of working incorrectly quite small. It would be interesting to apply and extend this point of view to more general scenarios. Another interesting and more specific problem that could be investigated, would be to find efficient algorithms (in the sense of [35]) to construct the ϵ-almost selectors introduced in the present paper. In principle, one could obtain the codes introduced in this paper by standard random generation techniques, and therefore design a simple random sampling procedure that iteratively generates random codes, at each step eventually discarding the generated code, in case it does not satisfies the desired property, and performing a new resampling step. However, this would lead to an inefficient algorithm. On the other hand, the *classical* case when no error is tolerated, i.e., $\epsilon < \frac{1}{\binom{n}{k}}$, suffers from the same problem, in that also in this case the question of devising an efficient deterministic strategy to construct (k, m, n)-selectors and $\mathrm{KG}(k, n)$-codes is wide open.

ACKNOWLEDGEMENTS

We are grateful to the Associate Editor and to the reviewers for the careful reading of our paper and for the many helpful comments.

REFERENCES

[1] M. Aldridge, L. Baldassini, O. Johnson: Group testing algorithms: bounds and simulations. IEEE Trans. Inf. Theory, vol. 60, no. 6, pp. 3671–3687 (2014).

[2] N. Alon and V. Asodi: Learning a hidden subgraph. SIAM J. Discrete Mathematics 18, pp. 697–712 (2005).

[3] N. Alon, E. Fachini, and J. Körner: Locally thin set families. Combinatorics, Probability and Computing 9, pp. 481–488 (2000).

[4] G. K. Atia, V. Saligrama: Boolean compressed sensing and noisy group testing. IEEE Trans. Inf. Theory, vol. 58, no. 3, pp. 1880–1901 (2012).

[5] A. Barg and A. Mazumdar: Almost disjunct matrices from codes and designs, arXiv:1510.02873 (2015).

[6] L.A. Bassalygo and V.V. Rykov: Multiple-access hyperchannel. Problems of Information Transmission 49 (2013), no. 4, pp. 299–307.

[7] E. Biglieri, L. Györfi: Multiple Access Channels. NATO Security through Science Series - D: Information and Communication Security, 10 (2007).

[8] B. Bollobás: Combinatorics - Set Systems, Hypergraphs, Families of Vectors, and Combinatorial Probability. Cambridge University Press (1986).

[9] C. L. Chan, S. Jaggi, V. Saligrama, S. Agnihotri: Non-adaptive group testing: explicit bounds and novel algorithms. IEEE Transactions on Information Theory, vol. 60, no. 5, pp. 3019–3035 (2014).

[10] B.S. Chlebus: Randomized communication in radio networks, a chapter in Handbook on Randomized Computing. P.M. Pardalos, S. Rajasekaran, J.H. Reif, and J.D.P. Rolim, (Eds.), Kluwer Academic Publishers, 2001, vol. I, pp. 401–456.

[11] B. S. Chlebus, L. Gąsieniec, D. R. Kowalski, A. Shvartsman: A robust randomized algorithm to perform independent tasks. J. Discrete Algorithms 6(4), pp. 651–665 (2008).

[12] B. S. Chlebus, D. R. Kowalski, A. Pelc, M. A. Rokicki: Efficient Distributed Communication in Ad-Hoc Radio Networks. In: L. Aceto, M. Henzinger, J. Sgall (eds) ICALP 2011, Part II. LNCS, vol. 6756, pp. 613–624. Springer, Heidelberg (2011).

[13] M. Chrobak, L. Gąsieniec, W. Rytter: Fast Broadcasting and Gossiping in Radio Networks, Fast broadcasting and gossiping in radio networks. Journal of Algorithms 43(2), pp. 177–189 (2002).

[14] A.E.F. Clementi, A. Monti and R. Silvestri: Distributed broadcast in radio networks of unknown topology. Theoretical Computer Science, 302(1-3), pp. 337–364 (2003).

[15] G.D. Cohen: Applications of coding theory to communication combinatorial problems. Discrete Mathematics, 83, Issues 2-3, pp. 237–248 (1990).

[16] M. Csűrös and M. Ruszinkó: Single-user tracing and disjointly superimposed codes. IEEE Transactions on Information Theory, 51, no. 4, pp. 1606–1611 (2005).

[17] A. De Bonis, A, Gąsieniec, U. Vaccaro: Optimal two-stage algorithms for group testing problems. SIAM J. Computing 34, no. 5, pp. 1253–1270 (2005).

[18] A. De Bonis, U. Vaccaro: Constructions of generalized superimposed codes with applications to group testing and conflict resolution in multiple access channels. Theoretical Computer Science 306(1-3), pp. 223–243 (2003).

[19] A. De Bonis, U. Vaccaro: Optimal algorithms for two group testing problems, and new bounds on generalized superimposed codes. IEEE Transactions on Information Theory 52(10), pp. 4673–4680 (2006).

[20] D.Z. Du, F.K. Hwang: Combinatorial Group Testing and Its Applications. World Scientific, River Edge, NJ (2000).

[21] Du, D.Z., Hwang, F. K.: Pooling Design and Nonadaptive Group Testing. Series on Appl. Math. vol. 18. World Scientific (2006).

[22] A.G. D'yachkov and V.V. Rykov: Bounds on the length of disjunct codes. Problemy Peredachi Informatsii 18, No. 3, pp. 7–13 (1982) [Probl. Inf. Trans. (Engl. Transl.), 1982, vol. 18, no. 3, pp. 166-171].

[23] A. G. D'yachkov, I.V. Vorobyev, N. A. Polyanskii, and V. Yu. Shchukin: Almost cover-free codes and designs. arXiv:1410.8566 (2014).

[24] P. Erdös, P. Frankl, and Z. Füredi: Families of finite sets in which no set is covered by the union of r others. Israel J. of Mathematics 51, pp. 75–89 (1985).

[25] Z. Füredi: On r-cover-free families. Journal of Combinatorial Theory, Ser. A 73 , pp. 172–173 (1996).

[26] D. R. Kowalski: On selection problem in radio networks. Proceedings of the twenty-fourth annual ACM symposium on Principles of distributed computing (PODC '05), pp. 158–166, ACM Press (2005).

[27] S. Gyory: Coding for a multiple access OR channel: A survey. Discrete Applied Mathematics, vol. 156, pp. 1407–1430 (2008).

[28] W.H. Kautz, R.C. Singleton: Nonrandom binary superimposed codes. IEEE Transactions on Information Theory 10, pp. 363–377 (1964).

[29] J. Komlós and A.G. Greenberg: An asymptotically fast nonadaptive algorithm for conflict resolution in multiple-access channels, IEEE Transactions on Information Theory 31, No. 2, pp. 302–306 (1985).

[30] L. Lovàsz: On the ratio of optimal integral and fractional covers. Discrete Mathematics 13, pp. 383–390 (1975).

[31] J.L. Massey and P. Mathys: The collision channel without feedback. IEEE Transactions on Information Theory, vol. 31, pp. 192–204 (1985).

[32] A. Mazumdar, On almost disjunct matrices for group testing. Algorithms and Computation, Springer, pp. 649–658 (2012).

[33] A. Mazumdar, Nonadaptive group testing with random set of defectives via constant-weight codes. arXiv:1503:03596 (2015).

[34] Ö. Sümer: Partial covering of hypergraphs. In: Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '05), pp. 572–581 (2005).
[35] E. Porat and A. Rothschild: Explicit nonadaptive combinatorial group testing schemes. IEEE Transactions on Information Theory 57(12), pp. 7982–7989 (2011).
[36] M. Ruszinkó: On the upper bound of the size of the r-cover-free families. Journal of Combinatorial Theory, Ser. A 66(2), pp. 302–310 (1994).