Theoretical Computer Science 806 (2020) 219–235

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

A new kind of selectors and their applications to conflict resolution in wireless multichannels networks

Annalisa De Bonis *, Ugo Vaccaro

Dipartimento di Informatica, Università di Salerno, I-84084 Fisciano (SA), Italy

ARTICLE INFO

Article history:
Available online 27 March 2019

Keywords:
Multichannels networks
Multi-access communication
Conflict resolution algorithms
Selectors
Superimposed codes
Frameproof codes

ABSTRACT

We investigate the benefits of using multiple channels in wireless networks, under the full-duplex multi-packet reception model of communication. The main question we address is the following: Is a speedup linear in the number of channels available, for some interesting communication primitive? We provide a positive answer to this interrogative for the *Information Exchange Problem*, in which up to k arbitrary nodes have information they intend to share with the entire network. In particular, we give non-adaptive deterministic protocols for both the scenario in which the channels provide the transmitting stations with the feedback on whether their transmissions have been successful and for the scenario in which channels provide no such feedback. To this aim, we devise and exploit a new combinatorial structure that generalizes well known combinatorial tools, widely used in the area of data-exchange in multiple-access channels (i.e., strongly selective families, selectors, and related mathematical objects). For our new combinatorial structures we provide both existential results and randomized algorithms to generate them. We also prove non-existence results showing that our protocol for the model with feedback is optimal, whereas that for the no-feedback scenario uses a number of time slots that exceeds the lower bound by a $\log k$ factor. Leveraging on properties of error correcting codes, we show that for an infinite set of the relevant parameters n and k, one can construct our combinatorial structure for the no-feedback scenario in polynomial time and of minimum length.

1. Introduction

Recent advances in technology and protocols have made wireless devices and services that operate on multiple channels increasingly available (e.g., WiFi [1] and Bluetooth [7]). Furthermore, the introduction of *realistic* full-duplex relaying technologies (see [4,6,30,36,39], the consortia [26,23], and references therein quoted), that allow terminals to transmit and receive *simultaneously* over a same frequency band, opens up a new spectrum of potentialities to greatly increase the throughput of a communication system. However, technology alone is not necessarily a panacea. More often than not, technological advances need to be supported by the introduction of new conceptual tools in order to transit from a mere potentiality into a full actuality. Under these premises, in this paper we seek an answer to the following question: How much faster can we disseminate data in a network if we have full-duplex multiple communication channels at our disposal (as opposed to relying on single communication channels)?

* Corresponding author.
E-mail addresses: adebonis@unisa.it (A. De Bonis), uvaccaro@unisa.it (U. Vaccaro).

https://doi.org/10.1016/j.tcs.2019.03.034
0304-3975/

1.1. The communication model

Our scenario consists of a single-hop communication network in which n stations have a common access to a multichannel $\mathcal{C}$, comprising of $q-1 \geq 1$ individual channels labeled with the integers $1, 2, \ldots, q-1$ (think of a wireless network, with stations able to transmit and listen over $q-1$ different frequencies). We assume that at most a certain number $k \leq n$ of stations might be *active* at the same time, i.e., might want to transmit simultaneously over the channels. Notice that the up to k active stations are not known beforehand. Each station may choose which of the $q-1$ channels to use for transmitting. An active station *successfully* transmits if and only if it transmits *singly* on the channel of its choice. We assume that the channels are without *collision detection*, i.e., the stations are not able to distinguish between the case when the channel is silent and the case when a collision occurs on the channel. 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 this paper, a *conflict resolution* algorithm for the above described multiaccess system is a scheduling algorithm that allows all 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 $i = 1, \ldots, t$, it establishes which stations should transmit at step i without looking at what happened over the channels at the previous steps. A non-adaptive scheduling algorithm is equivalent to a set of n vectors, identified by integers from 1 through n (each of which corresponding to a distinct station), of length t over the alphabet $\{0, 1, \ldots, q-1\}$, with the meaning that station j is scheduled to transmit at step i over the channel $s \in \{1, \ldots, q-1\}$ if and only if the i-th entry of its associated vector j is equal to s. If the i-th entry is equal to 0, the station has to stay silent at step i. A non-adaptive scheduling algorithm is conveniently represented by the q-ary matrix having as columns the n q-ary vectors associated with the scheduling of the transmissions of the n stations. Obviously, entry (i, j) of such a matrix is equal to $s \neq 0$ if and only if station j is scheduled to transmit at step i over channel $s \in \{1, \ldots, q-1\}$, and it is equal to $s = 0$ if and only if station j remains silent at step i. The parameter of interest to be minimized is the number of rows of the matrix. In fact, that number represents the number of time slots in which the conflict resolution algorithm schedules the transmissions of the n stations, so that up to k active stations transmit with success.

In order to fully define the communication model, we must also specify *whether or not* each transmitting station is able to discriminate between the event that its transmission has been successful, from the event that its transmission has been unsuccessful (i.e., another station has attempted a simultaneous transmission over the same channel at the same time instant, thus causing a collision). In other words, we must differentiate the situation in which transmitting stations receive feedback from the channels about whether or not their transmissions have been successful, from the situation in which channels do not provide this feedback.

1.1.1. The multiple-access channel without feedback

When stations receive no *feedback* from the channel, the conflict resolution algorithm must necessarily schedule transmissions in such a way that each active station transmits singly over at least one of the $q-1$ channels at some step, i.e., for each active station there is a time slot at which it is scheduled to transmit on some channel s and no other active station is scheduled to transmit on channel s at that same time slot. A conflict resolution algorithm for this model is equivalent to a q-ary 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 an entry equal to some $s \in \{1, \ldots, q-1\}$ and the remaining $k-1$ columns have entries different from s. When $q = 2$, matrices that satisfy this property have been very well studied in the literature where they are known under different names, such as superimposed codes [32], $(k-1)$-cover free families [24], k-disjunct codes [19], and strongly selective families [12,13]. To the best of our knowledge, in the literature there are no papers studying the Information Exchange Problem under the model described in this section in the general case when q is an arbitrary integer larger than or equal to 2. A multi-access model very similar to ours has been considered by the authors of [9] in the context of the problem of waking-up a single-hop radio network of n stations. The connections between this paper and ours are discussed in Section 3.

1.1.2. The multiple-access channel with feedback

In addition to the situation when stations receive no feedback from the channels, we consider also the communication model in which any transmitting station receives feedback on whether its transmission has been successful or not (see [34] for the case when only one channel is available). In such a model, an active station has the capability to become *inactive* (i.e., to refrain from transmitting) after it has transmitted successfully for the first time. We remark that, in this model, a multiple-access channel provides only the transmitting stations with the feedback on whether their transmissions have been successful or not but it does not give any feedback to the other stations on whether a collision has occurred in the case when no message is received on that channel. 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 over some of the $q-1$ available channels. However, in this scenario, if a certain active station j transmits with success at time slot t, then after time slot t, station j does not transmit anymore even if the protocol schedules it to transmit. Consequently, after time slot t, a *still* active station results in transmitting singly to a channel also in the case when it is scheduled to transmit simultaneously on that same channel with station j. The characterization of q-ary matrices that represent non-adaptive conflict resolution algorithms for this different model of communication is less simple than that for the case without feedback, and will be presented in the

technical part of this paper. To the best of our knowledge, the only paper that studies the information exchange problem in a multiple-access model similar to the one described in this section, is [42]. We will discuss the relations between our results and those in that paper in Section 3.

2. Our results

In this paper we study non-adaptive conflict resolution protocols for both the above described multiple-access models, that is, for multiple-access channels with and without feedback. Our main goal is to provide an answer to the following question: Is a speedup linear in the number of channels available, for some interesting communication primitive? We provide a positive answer to this interrogative for the basic *Information Exchange Problem* [31,35], in which up to k arbitrary nodes have information they intend to share with the entire network. In other words, we show that the time needed to disseminate these up to k information items decreases linearly with the number of available channels. In order to study this question, we introduce two new combinatorial structures that consist in a generalization of selectors [17,11] (and therefore of superimposed codes, $(k-1)$-cover free families, k-disjunct codes, strongly selective families) and a generalization of Komlós and Greenberg codes [34].

Since the number n of stations in the network and the number k of active stations can be arbitrary, while the number $q-1$ of channels is usually severely limited by technological factors [1,7], in the scheduling protocols described in this paper we assume that $q \leq k$. However, for completeness we shall derive our combinatorial results in their full mathematical generality, that is, also for the case $q > k$. As it will become clear as the paper unfolds, scheduling protocols also for the case $q > k$ can be obtained in a relatively straightforward way, once one has constructed the appropriate combinatorial tools.

Our main results are summarized by the following two theorems.

Theorem 1. *Let k, m, q, and n be integers such that $1 \leq m \leq k \leq n$, and $2 \leq q \leq k$. There exists a non-adaptive deterministic conflict resolution algorithm for a multiple-access channel $\mathcal{C}$* without feedback *(comprising of $q-1 \geq 1$ individual channels) that schedules the transmissions of n stations in such a way that, for all possible subsets of up to k active stations, one has that at most $k-m$ out of k active stations fail to transmit successfully. The number t of time slots used by the conflict resolution algorithm is*

$$t \leq \frac{4(k-1)}{(k-m+1)(q-1)}\left(\ln\binom{n}{k-1} + \ln k + \ln\binom{k}{k-m+1} + 1\right) = O\left(\frac{k^2}{(k-m+1)q}\log\frac{n}{k}\right). \tag{1}$$

When $m = k$ (as in the Information Exchange Problem) the number of time slots t used by the conflict resolution algorithm is

$$t = O\left(\frac{k^2}{q}\log\frac{n}{k}\right), \tag{2}$$

and for any *conflict resolution algorithm for the Information Exchange Problem it holds that the number of time slots t is such that*

$$t = \Omega\left(\frac{k^2}{q\log k}\log\frac{n}{k}\right). \tag{3}$$

The proof technique of (1) and (2) is based on the Lovász Local Lemma. By exploiting the recent constructive version of the Lovász Local Lemma given by Moser and Tardos [38], we design a randomized algorithm that generates the protocols whose existence is implied by Theorem 1. To prove the lower bound (3) we uncover a relation between the combinatorial structures we introduce in this paper and the well known superimposed codes of [32].

In case of a single channel *with feedback*, it is well known that one can solve conflicts among transmitting stations by protocols that require a number of time steps substantially smaller than that required for a channel without feedback, and this thanks to the seminal paper [34]. In our present case of several channels with feedback, we can prove an analogous result. More precisely, we have that the following (optimal) result holds.

Theorem 2. *Let k, q, and n be integers such that $2 \leq q \leq k \leq n$. There exists a non-adaptive deterministic conflict resolution algorithm for a multiple-access channel $\mathcal{C}$* with feedback *(comprising of $q-1 \geq 1$ individual channels) that schedules the transmissions of n stations in such a way that, for all possible subsets of k stations, one has that all active stations transmit successfully. The number t of time slots used by the conflict resolution algorithm is*

$$t = O\left(\frac{k}{q}\log\frac{n}{k}\right). \tag{4}$$

Moreover, for any *conflict resolution algorithm in this scenario it holds that the number of time slots t is such that*

$$t = \Omega\left(\frac{k}{q}\log\frac{n}{k}\right). \tag{5}$$

The randomized algorithm that generates the protocols in Theorem 1 can be exploited also to design an algorithm that generates the conflict resolution protocol for the multiple-access model with feedback of Theorem 2.

We remark that our upper bounds for the multiple-access model without feedback hold for the parameters q, k, and n being known in advance, whereas the upper bound of Theorem 2 holds also in the case when there is no a priori knowledge of the maximum number k of active stations. Indeed, in the multiple-access model with feedback, conflicts among an unknown number of active stations can be resolved by using a standard argument in the area. More explicitly, one runs iteratively the conflict resolution algorithm given for the case when an upper bound on the number of active stations is known in advance. At each iteration (stage), one doubles the number of stations that are *assumed* to be active. In other words, at stage i one runs a conflict resolution algorithm for a number k_i of supposedly active stations equal to 2^i. Since at stage $\lceil \log k \rceil$ we run an algorithm for a number of active stations larger than or equal to k, one is guaranteed that all active stations transmit with success within this stage. It is easy to see that the complexity of this protocol is asymptotically equivalent to that of the protocol that works having the value of the number of active stations in input.

Our paper is organized as follows. In Section 3 we discuss the related literature. In Section 4 we first introduce, in Section 4.1, a new version of selectors that, in our *no-feedback scenario*, correspond to protocols that schedule transmissions so that at least a certain number m out of k active stations transmit with success. Then, Section 4.1 presents the constructions and the non-existence result that imply Theorem 1. In Section 4.2, we introduce a new version of Komlós and Greenberg codes that, in the multiple-access model *with feedback*, furnishes a scheduling protocol that solves all conflicts for all possible subsets of up to k active stations. For these codes, in Section 4.2, we provide the construction and the lower bound that imply Theorem 2. In Section 5, we highlight a few connections between the combinatorial structures introduced in this paper and the well known Frameproof Codes [5]. Finally, in Section 6, we show the fact that, for an infinite set of the relevant parameters n and k, one can construct our combinatorial structures for the Information Exchange Problem in the no-feedback scenario in polynomial time and of optimal (minimum) length.

3. Related work

Information dissemination in single-channel, single-hop networks is a well known and widely studied problem. We refer the reader to the survey papers [10,28] for a presentation of this area. The study of distributed algorithms in multi-channels wireless networks is relatively recent (see [8,16,18,29,31,47–49] and references therein quoted). However, most of these papers considered either randomized algorithms, or different communication primitives, or different communication models. To the best of our knowledge, the present paper seems to be the first to study the Information Exchange Problem in the no-feedback communication model described in Section 1.1.1. The authors of [42] studied the Information Exchange Problem in the same multiple-access model with feedback described in Section 1.1.2. However, they assume that the number of available channels may depend on the total number n of stations and that no a priori information about n is given. That paper provides randomized algorithms that solve the Information Exchange Problem when the number of available multiple-access channels is linear in the total number of stations. We remark that all algorithms given in the present paper are deterministic and work with an arbitrary number $q-1$ of channels. Another paper that considers a communication model closely related to ours, though in the context of a different problem, is [9]. This paper studies the problem of waking-up a single-hop radio network of n stations each of which is connected to a certain number of channels. Up to a certain number k of these n stations become active spontaneously at arbitrary times and the goal of waking-up the network is accomplished when one of the active stations transmits singly to one of the available channels. Notice that in the Information Exchange Problem we consider, the goal is accomplished when *all* active stations have successfully transmitted their packets, and it is assumed that all up to k stations become active in the same round.

4. Combinatorial tools and their applications

Let M be an arbitrary t-rows n-column matrix, where we denote by $\mathbf{c}_1, \ldots, \mathbf{c}_n$ the $t \times 1$ columns of M. We assume that the entries of M are over the alphabet $\{0, 1, \ldots, q-1\}$, where q is an integer larger than or equal to 2. For each column $\mathbf{c}$ of M and row index $i \in \{1, \ldots, t\} = [t]$, we denote by $\mathbf{c}(i)$ the i-th component of vector $\mathbf{c}$. Given arbitrary column vectors $\mathbf{c}, \mathbf{c}_1, \ldots, \mathbf{c}_g$, we say that $\mathbf{c}$ is *covered* by $\mathbf{c}_1, \ldots, \mathbf{c}_g$ if, for any $i \in [t]$, $\mathbf{c}(i) \neq 0$ implies that there exists at least a vector $\mathbf{c}' \in \{\mathbf{c}_1, \ldots, \mathbf{c}_g\}$ (depending on i), such that it is $\mathbf{c}(i) = \mathbf{c}'(i)$.

4.1. Combinatorial tools for the multiple-access model without feedback

In this section we introduce the combinatorial objects that correspond to conflict resolution algorithms for the no-feedback scenario, and give constructions and non-existence results for these objects.

Definition 1. Given positive integers q, k, and n, with $q \geq 2$ and $2 \leq k \leq n$, we say that a $t \times n$ matrix M with entries in $\{0, 1, \ldots, q-1\}$ is a q-ary (k, n)-*strongly selective* code of length t if for any k distinct columns $\mathbf{c}_{j_1}, \ldots, \mathbf{c}_{j_k}$ of M, one has that no column in $\{\mathbf{c}_{j_1}, \ldots, \mathbf{c}_{j_k}\}$ is covered by the remaining $k-1$ columns. In other words, for any $\ell \in \{1, \ldots, k\}$, there exists a row index i_ℓ such that $\mathbf{c}_{j_\ell}(i_\ell) \neq 0$ and $\mathbf{c}_{j_f}(i_\ell) \neq \mathbf{c}_{j_\ell}(i_\ell)$ for each $f \in \{1, \ldots, k\} \setminus \{\ell\}$. The minimum length of a q-ary (k, n)-strongly selective code is denoted by $t(q, k, n)$.

For $q = 2$, the codes defined above correspond to the strongly selective families of [13] or equivalently to the $(k-1)$-cover free families of [24] and the $(k-1)$-superimposed codes of [20]. We can prove the following result.

Theorem 3. *Given positive integers q, k, and n, with $2 \leq q \leq k \leq n$, there exists a q-ary (k, n)-strongly selective code of length*

$$t \leq \frac{4(k-1)}{q-1}\left(\ln\binom{n}{k-1} + 2\ln k + 1\right) = O\left(\frac{k^2}{q}\log\frac{n}{k}\right).$$

In Section 1.1.1 we have observed that a conflict resolution algorithm for the no-feedback scenario is equivalent to a q-ary matrix with the property that for any k columns and for any column $\mathbf{c}$ chosen among these k columns, there exists a row in correspondence of which $\mathbf{c}$ has an entry equal to some $s \in \{1, \ldots, q-1\}$ and the remaining $k-1$ columns have entries different from s. It is immediate to see that a q-ary matrix satisfies this property if and only if it is a q-ary (k, n)-strongly selective code. Consequently, we have that Theorem 3 implies the upper bound (2) of Theorem 1.

In order to prove Theorem 3 (and also Theorem 1 in its full generality), we introduce the following generalization of q-ary (k, n)-strongly selective codes.

Definition 2. Given positive integers q, m, k, and n, with $m \leq k$, $2 \leq k \leq n$, and $q \geq 2$, we say that a $t \times n$ matrix M with entries in $\{0, 1, \ldots, q-1\}$ is a q-ary (k, m, n)-selector of size t if for any k distinct columns $\mathbf{c}_{j_1}, \ldots, \mathbf{c}_{j_k}$ of M, one has that among $\mathbf{c}_{j_1}, \ldots, \mathbf{c}_{j_k}$ there are at least m columns $\mathbf{c}$ such that $\mathbf{c}$ is not covered by the remaining $k-1$ columns in $\{\mathbf{c}_{j_1}, \ldots, \mathbf{c}_{j_k}\} \setminus \{\mathbf{c}\}$. The minimum size of a q-ary (k, m, n)-selector is denoted by $t_{sel}(q, k, m, n)$.

For $q = 2$, q-ary (k, m, n)-selectors correspond to the (k, m, n)-selectors introduced in [17]. It is clear that a q-ary (k, n)-strongly selective code corresponds to a q-ary (k, k, n)-selector. In Section 4.2, we will see that q-ary (k, m, n)-selectors prove also useful as buildings blocks for obtaining conflict resolution protocols in the multiple-access scenario with feedback.

The following theorem discloses the relevance of q-ary (k, m, n)-th selectors with respect to our multi-access problem.

Theorem 4. *Given positive integers q, m, k, and n, with $m \leq k$, $2 \leq k \leq n$, and $q \geq 2$, a scheduling algorithm is a conflict resolution algorithm for a multiple-access channel $\mathcal{C}$ without feedback (comprising of $q - 1 \geq 1$ individual channels) that schedules the transmissions of n stations in such a way that all but at most $k - m$ of the up to k active stations transmit successfully,* if and only if *the corresponding q-ary matrix is a q-ary (k, m, n)-selector.*

Proof. Let M be a q-ary (k, m, n)-selector. Let us assume that $j \leq k$ be the number of active stations, and let M' be any k-column submatrix of M containing all j columns associated with the j active stations. By Definition 2, one has that M' contains at least m columns each of which is not covered by the remaining $k-1$ columns in M'. Therefore, at least $m - (k - j)$ of the j columns of M' associated with the j active stations are not covered by the remaining $k-1$ columns in M'. Each of the active stations associated with these columns transmits singly, and therefore successfully, to one of the $q-1$ channels at some time slot. It follows that a q-ary (k, m, n)-selector corresponds to a conflict resolution algorithm for the no-feedback scenario that allows all but at most $j - (m - (k - j)) = k - m$ stations to transmit with success.

Conversely, let M be an n-column matrix, with entries in $\{0, 1, \ldots, q-1\}$, associated with a scheduling algorithm for the no-feedback scenario that schedules the transmissions of the n stations so that at most $k - m$ out of the up to k active stations transmit with success. Let us consider an arbitrary k-column submatrix M' of M and let S be the set of k stations associated with the columns of M'. Suppose that S is the set of active stations. Since an active station transmits with success only if it transmits singly on one of the $q-1$ channels at some step, we have that the stations of S that transmit with success are those associated with the columns of M' that are not covered by the remaining columns of M'. Consequently, if the number of these columns is less than m then the number of active stations that transmit with success is less than m, and consequently, there are more than $k - m$ active stations that do not transmit with success. Therefore, the conflict resolution algorithm associated with M must satisfy the property that any k-column submatrix M' of M contains at least m columns such that each of these columns is not covered by the remaining $k-1$ columns of M'. In other words, M is a q-ary (k, m, n)-selector. □

Notice that by setting $m = k$ in Theorem 4, we get that q-ary (k, n)-strongly selective codes are equivalent to conflict resolution protocols for our no-feedback scenario that allow all up to k active stations to transmit with success, as already noticed in the discussion following Theorem 3.

To avoid trivialities, throughout the paper we assume that $k = o(\sqrt{n})$. This is not only reasonable on practical ground, but also because the opposite case is, in a sense, trivial also from the mathematical point of view. Indeed, it is known that if k is such that $\frac{15+\sqrt{23}}{34}k^2 > n$ then no better binary (k, n)-strongly selective code exits than the $n \times n$ identity matrix I_n [41]. It is not hard to see that for the same range of values of k and n, the length t of any q-ary (k, n)-strongly selective code is such that $t \geq \lceil n/(q-1) \rceil$, and a construction meeting this bound is easy to devise.

In the following, given an event E, we will denote by $\bar{E}$ its complementary event. In order to prove our existential results, we need to recall the celebrated Lovász Local Lemma for the symmetric case (see [2]), as stated below.

Lemma 1. *Let b and d be integers with $0 \leq d \leq b$, and let $E_1, E_2, \ldots, E_b$ be b events in an arbitrary probability space. Suppose that each event E_i is mutually independent of at least $b-d-1$ events in $\{E_1, E_2, \ldots, E_b\} \setminus \{E_i\}$, and that $\Pr[E_i] \leq P$ for all $1 \leq i \leq b$. If $(d+1)eP \leq 1$, then $\Pr[\cap_{i=1}^{n} \bar{E}_i] > 0$, where $e = 2.71828\ldots$ is the base of the natural logarithm.*

Using Lovász Local Lemma we will prove the following basic result.

Theorem 5. *Given positive integers q, m, k, and n, with $m \leq k \leq n$ and $q \geq 2$, there exists a q-ary (k, m, n)-selector of size*

$$t \leq \frac{4(k-1)}{(k-m+1)v}\left(\ln\binom{n}{k-1} + \ln k + \ln\binom{k}{k-m+1} + 1\right)$$
$$= O\left(\frac{k^2}{(k-m+1)v}\log\frac{n}{k}\right),$$

where $v = q-1$ if $q \leq k$, and $v = k$ if $q > k$.

Proof. The proof is based on the probabilistic method. Let $M = [M(i,j)]$ be a random q-ary $t \times n$ matrix such that all entries in M are chosen independently with probabilities

$$Pr\{M(i,j) = s\} = p, \quad \text{for each } s \in \{1, \ldots, q-1\},$$

and

$$Pr\{M(i,j) = 0\} = 1 - (q-1)p.$$

For a given set Υ of k columns of M, let us denote by E_Υ the event that the matrix formed by the columns of Υ does not satisfy the property of q-ary (k, m, n)-selectors, i.e., Υ does not contain a subset of at least m columns each of which is not covered by the remaining columns of Υ. It is immediate to see that event E_Υ occurs if and only if for at least $k-m+1$ columns of Υ, say $\mathbf{c}_{\ell_1}, \ldots, \mathbf{c}_{\ell_g} \in \Upsilon$, $g \geq k-m+1$, one has that each column $\mathbf{c}_{\ell_r} \in \{\mathbf{c}_{\ell_1}, \ldots, \mathbf{c}_{\ell_g}\}$ is covered by the remaining $k-1$ columns in $\Upsilon \setminus \{\mathbf{c}_{\ell_r}\}$. For a subset of $k-m+1$ columns $\Psi = \{\mathbf{c}_{\ell_1}, \ldots, \mathbf{c}_{\ell_{k-m+1}}\} \subseteq \Upsilon$, let E_Ψ denote the event that each column in Ψ is covered by the remaining $k-1$ columns in Υ. In other words, there exists no row index $i \in [t]$ such that for some $\mathbf{c} \in \Psi$, it holds $\mathbf{c}(i) = s \in \{1, \ldots, q-1\}$ and $\mathbf{c}'(i) \neq s$ for all $\mathbf{c}' \in \Upsilon \setminus \{\mathbf{c}\}$.

For a column $\mathbf{c} \in \Psi$ and a row index $i \in [t]$, we denote by $E_{\mathbf{c},i}$ the event that $\mathbf{c}(i) = s$ for some $s \neq 0$ and $\mathbf{c}'(i) \neq s$ for all $\mathbf{c}' \in \Upsilon \setminus \{\mathbf{c}_r\}$. Notice that for a fixed row index $i \in [t]$, events in $\{\bar{E}_{\mathbf{c},i} : \mathbf{c} \in \Psi\}$ are not mutually independent. To see this, let us consider $a+1$ arbitrary columns $\mathbf{c}, \mathbf{c}_1, \ldots \mathbf{c}_a \in \Psi$. We will show that if events $\bar{E}_{\mathbf{c}_1,i}, \ldots, \bar{E}_{\mathbf{c}_a,i}$ all occur then the event $\bar{E}_{\mathbf{c},i}$ is less likely to occur. Notice that $E_{\mathbf{c},i}$ occurs if only if the i-th entry of $\mathbf{c}(i)$ is different from zero e different from any of the values occurring in the i-th entries of the remaining $k-1$ columns of Υ. Now, suppose that for a certain number f, $f \leq a$, of columns $\mathbf{c}'_1, \ldots, \mathbf{c}'_f \in \{\mathbf{c}_1, \ldots, \mathbf{c}_a\}$ one has that $E_{\mathbf{c}'_1,i}, \ldots, E_{\mathbf{c}'_f,i}$ all occur. If these f events occur then each of $\mathbf{c}'_1(i), \ldots, \mathbf{c}'_f(i)$ is different from zero and is distinct from each of the i-th entries of the remaining $k-1$ columns in Υ, thus also implying that $\mathbf{c}'_1(i), \ldots, \mathbf{c}'_f(i)$ are pairwise distinct. Consequently, in comparison with the case when all events $E_{\mathbf{c}_1,i}, \ldots, E_{\mathbf{c}_a,i}$ occur, in this case there are f more values that $\mathbf{c}(i)$ is not allowed to assume in order for event $E_{\mathbf{c},i}$ to occur. From the above discussion it follows that if none of events $E_{\mathbf{c}_1,i}, \ldots, E_{\mathbf{c}_a,i}$ then the event $E_{\mathbf{c},i}$ is more likely to occur. Consequently, for any subset of columns $\{\mathbf{c}_1, \ldots \mathbf{c}_a\} \subseteq \Psi$, $1 \leq a \leq k-m$, and any column $\mathbf{c} \in \Psi \setminus \{\mathbf{c}_1, \ldots \mathbf{c}_a\}$, it holds that

$$P\{\bar{E}_{\mathbf{c},i} | \bar{E}_{\mathbf{c}_1,i} \cap \ldots \cap \bar{E}_{\mathbf{c}_a,i}\} \leq P\{\bar{E}_{\mathbf{c},i}\}. \tag{6}$$

The following is a consequence of (6).

$$\begin{aligned} &P\{\bar{E}_{\mathbf{c}_{\ell_1},i} \cap \ldots \cap \bar{E}_{\mathbf{c}_{\ell_{k-m+1}},i}\} \\ &= P\{\bar{E}_{\mathbf{c}_{\ell_1},i}\} \cdot P\{\bar{E}_{\mathbf{c}_{\ell_2},i} | \bar{E}_{\mathbf{c}_{\ell_1},i}\} \cdot \ldots \cdot P\{\bar{E}_{\mathbf{c}_{\ell_{k-m+1}},i} | \bar{E}_{\mathbf{c}_{\ell_1},i} \cap \ldots \cap \bar{E}_{\mathbf{c}_{\ell_{k-m}},i}\} \\ &\leq P\{\bar{E}_{\mathbf{c}_{\ell_1},i}\} \cdot \ldots \cdot P\{\bar{E}_{\mathbf{c}_{\ell_{k-m+1}},i}\}. \end{aligned} \tag{7}$$

For a fixed column $\mathbf{c} \in \Psi$ and a fixed row index $i \in [t]$, we have that $P(E_{\mathbf{c},i}) = (q-1)p(1-p)^{k-1}$, and consequently, it holds that

$$P(\bar{E}_{\mathbf{c},i}) = 1 - (q-1)p(1-p)^{k-1}. \tag{8}$$

By (7) and (8), we have that the probability that all events $\bar{E}_{\mathbf{c}_{\ell_1},i}, \ldots, \bar{E}_{\mathbf{c}_{\ell_{k-m+1}},i}$ occur is at most $(1-(q-1)p(1-p)^{k-1)})^{k-m+1}$, from which it follows that

$$P\{E_\Psi\} \leq [1-(q-1)p(1-p)^{k-1}]^{(k-m+1)t}. \tag{9}$$

It also follows that the probability that Υ contains a subset of $k-m+1$ columns such that each of these $k-m+1$ is covered by the remaining $k-1$ columns in Υ is

$$Pr\{E_\Upsilon\} \leq \binom{k}{k-m+1}[1-(q-1)p(1-p)^{k-1}]^{t(k-m+1)}. \tag{10}$$

Observe that there are at most $k\binom{n-1}{k-1}$ k-column subsets containing one or more columns of Υ, and consequently, event E_Υ is independent from all but at most $k\binom{n-1}{k-1}$ events in $\{E_{\Upsilon'} : \Upsilon' \subseteq M, |\Upsilon'| = k\}$. Therefore, applying Lemma 1 with

$$P = \binom{k}{k-m+1}[1-(q-1)p(1-p)^{k-1}]^{t(k-m+1)}, \qquad d = k\binom{n-1}{k-1},$$

one has that the matrix M has positive probability of being a q-ary (k,m,n)-selector if

$$e\binom{k}{k-m+1}[1-(q-1)p(1-p)^{k-1}]^{t(k-m+1)}\left(k\binom{n-1}{k-1}+1\right) \leq 1. \tag{11}$$

Inequality (11) is satisfied if

$$t \geq \frac{\ln\left(k\binom{n-1}{k-1}+1\right)+\ln\binom{k}{k-m+1}+1}{-(k-m+1)\ln\left[1-(q-1)p\,(1-p)^{k-1}\right]}. \tag{12}$$

From the above discussion, it follows that M has a strictly positive probability of being a q-ary (k,m,n)-selector if t satisfies the above inequality (12). The well known inequality

$$-\ln(1-x) > x, \text{ for } 0 < x < 1, \tag{13}$$

implies that the righthand side of inequality (12) is strictly smaller than

$$\frac{\ln\left(k\binom{n-1}{k-1}+1\right)+\ln\binom{k}{k-m+1}+1}{(k-m+1)(q-1)p\,(1-p)^{k-1}} \leq \frac{\ln\left(k\binom{n}{k-1}\right)+\ln\binom{k}{k-m+1}+1}{(k-m+1)(q-1)p\,(1-p)^{k-1}}, \tag{14}$$

where latter inequality is a consequence of the fact that $k\binom{n-1}{k-1}+1 \leq k\binom{n}{k-1}$. Therefore, in order for a q-ary (k,m,n)-selector of length t to exist it is sufficient that

$$t \leq \frac{\ln\left(k\binom{n}{k-1}\right)+\ln\binom{k}{k-m+1}+1}{(k-m+1)(q-1)p\,(1-p)^{k-1}}. \tag{15}$$

To the aim of obtaining the upper bounds in the statement of the theorem, we discuss separately the cases $2 \leq q \leq k$ and $q > k$, since we need to choose the value of p accordingly.

Case $2 \leq q \leq k$**.**
In this case we set $p = \frac{1}{k}$. Then, by replacing p with $\frac{1}{k}$ in inequality (15) we get that there exists a q-ary (k,m,n)-selector of length t with

$$t \leq \frac{\ln\left(k\binom{n}{k-1}\right)+\ln\binom{k}{k-m+1}+1}{(k-m+1)(q-1)\frac{1}{k}\left(1-\frac{1}{k}\right)^{k-1}}. \tag{16}$$

Since $k \geq 2$ and $\left(1-\frac{1}{k}\right)^k$ increases with k, one has that $\left(1-\frac{1}{k}\right)^k \geq 1/4$ and it follows that the righthand side of the above inequality is at most

$$\frac{4\left(\ln\left(k\binom{n}{k-1}\right)+\ln\binom{k}{k-m+1}+1\right)}{(k-m+1)(q-1)\frac{1}{k}\left(1-\frac{1}{k}\right)^{-1}} = \frac{4(k-1)\left(\ln\left(k\binom{n}{k-1}\right)+\ln\binom{k}{k-m+1}+1\right)}{(k-m+1)(q-1)}.$$

Therefore, one has that there exists a q-ary (k,m,n)-selector of length t being at most

$$\frac{4(k-1)\left(\ln\binom{n}{k-1}+\ln k+\ln\binom{k}{k-m+1}+1\right)}{(k-m+1)(q-1)}.$$

In order to obtain the asymptotic bound in the statement of the theorem we exploit the following well known upper bound on the binomial coefficient.

$$\binom{a}{b} \leq \left(\frac{ea}{b}\right)^b. \tag{17}$$

Notice that upper bound (17) follows from inequality $\binom{a}{b} \leq \frac{a^b}{b!}$ and from the Taylor series of e^b which implies $b! \geq \frac{b^b}{e^b}$.

Applying inequality (17) to $\binom{n}{k-1}$ and $\binom{k}{k-m+1}$ in (16), we get

$$\begin{aligned} t &\leq \frac{4(k-1)\left(\ln\left(\frac{en}{k-1}\right)^{k-1} + \ln k + \ln\left(\frac{ek}{k-m+1}\right)^{k-m+1} + 1\right)}{(k-m+1)(q-1)} \\ &= O\left(\frac{k^2}{(k-m+1)(q-1)} \ln \frac{n}{k}\right). \end{aligned}$$

Case $q > k$.

In this case we set $p = \frac{1}{q-1}$. Notice that in this case the matrix M does not contain any 0-entries. By inequality (15) with p replaced by $\frac{1}{q-1}$, we have that there exists a q-ary (k, m, n)-selector of length t with

$$t \leq \frac{\ln\left(k\binom{n}{k-1}\right) + \ln\binom{k}{k-m+1} + 1}{(k-m+1)\left(1-\frac{1}{q-1}\right)^{k-1}}. \tag{18}$$

Since $k \leq q-1$, one has that $\left(1-\frac{1}{q-1}\right)^{k-1} \geq \left(1-\frac{1}{k}\right)^{k-1}$. We already observed, when discussing the case $q \leq k$, that $\left(1-\frac{1}{k}\right)^k \geq 1/4$, from which it follows that the righthand side of the above inequality (18) is at most

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

Therefore, it follows that, for $q > k$, there exists a q-ary (k, m, n)-selector of length

$$t \leq \frac{4(k-1)\left(\ln\binom{n}{k-1} + \ln k + \ln\binom{k}{k-m+1} + 1\right)}{k(k-m+1)}.$$

The above upper bound along with upper bound (17) on the binomial coefficient implies that there exists a q-ary (k, m, n)-selector of length

$$\begin{aligned} t &\leq \frac{4(k-1)\left(\ln\left(\frac{en}{k-1}\right)^{k-1} + \ln k + \ln\left(\frac{ek}{k-m+1}\right)^{k-m+1} + 1\right)}{(k-m+1)k} \\ &= O\left(\frac{k^2}{(k-m+1)k} \ln \frac{n}{k}\right). \quad \square \end{aligned}$$

Theorem 5, along with Theorem 4, implies the upper bound (1) of Theorem 1.

By setting $m = k$ in the upper bound of Theorem 5 for $q \leq k$, we get the upper bound of Theorem 3, and, by setting $m = k$ also in the statement of Theorem 4, we get the upper bound (2) of Theorem 1.

Moser and Tardos [38] gave a randomized algorithm to generate the structures whose existence is guaranteed by the Lovász Local Lemma. They considered a set of mutually independent random variables $\mathcal{X}$ in a fixed probability space Ω and a family of "bad" events $\{E_1, \ldots, E_b\}$ in Ω satisfying the hypothesis of the Lovász Local Lemma, with each event E_i being determined by the values of some subset $S \subseteq \mathcal{X}$ of these variables. The minimal subset $S \subseteq \mathcal{X}$ of variables that determine E_i is denoted by $vbl(E_i)$. An evaluation of the variables in $\mathcal{X}$ is said to *violate* event E_i if it makes E_i happen. Moser and Tardos [38] gave an efficient randomized algorithm to find an evaluation of the variables in $\mathcal{X}$ that does not violate any event in the family. The idea of the algorithm is the following. The algorithm starts with a random evaluation $v(X)$ of each variable $X \in \mathcal{X}$. If some event in $\{E_1, \ldots, E_b\}$ is violated by this evaluation then the algorithm picks one of the violated events E_i and sample a new random assignment of values for the variables in $vbl(E_i)$. Each variable in $vbl(E_i)$ is sampled independently and according to its distribution. The values of the other variables in $\mathcal{X}$ do not change. Moser and Tardos refer to this operation as a *resampling* of the event E_i. After resampling E_i, the algorithm checks whether there are violated events and if so performs the resampling of one of these events. The algorithm continues in this way until there

are no violated events. The following theorem [38] states that the above algorithm performs a small expected number of resampling steps before it reaches an evaluation of the variables not violating any of the events in the family. Notice that the original theorem in [38] is stated for events $E_1, \ldots, E_b$ that satisfy the hypothesis of the fully general version of the Lovász Local Lemma (asymmetric Lovász Local Lemma). Here, we present a formulation of that theorem for events that satisfy the symmetric version of the Lovász Local Lemma, i.e., Lemma 1.

Theorem 6. *[38] Let $\mathcal{X}$ be a finite set of mutually independent random variables in an arbitrary probability space Ω. Let b and d be integers with $0 \leq d \leq b$, and let $\{E_1, \ldots, E_b\}$ be a finite set of events determined by the variables in $\mathcal{X}$. If $E_1, \ldots, E_b$ satisfy the hypothesis of Lemma 1 then the randomized algorithm described above performs an expected number of resampling steps which is at most $\frac{b}{d}$ before it finds an evaluation of the variables in $\mathcal{X}$ not violating any of the events in $\{E_1, \ldots, E_b\}$.*

By exploiting the technique in [38], we design Algorithm 1 that constructs a q-ary (k, m, n)-selector meeting the upper bound of Theorem 5 and obtain the following result.

Theorem 7. *Given positive integers q, m, k, and n, with $m \leq k \leq n$ and $q \geq 2$, Algorithm 1 generates the q-ary (k, m, n)-selector whose existence is stated by Theorem 5. Moreover, the repeat-until loop in this algorithm is iterated an expected number of times which is at most $\frac{n}{k^2}$.*

Proof. Let t be bounded as in the statement of Theorem 5, and let M be the $t \times n$ random q-ary matrix $M = [M(i, j)]$ constructed in the proof of Theorem 5. Moreover, as in that proof, for a given set Υ of k columns of M, we denote by E_Υ the event that the matrix formed by the columns of Υ does not satisfy the property of q-ary (k, m, n)-selectors in Definition 2, i.e., Υ does not contain a subset of at least m columns each of which is not covered by the remaining columns of Υ. Observe that events in $\{E_\Upsilon : \Upsilon$ is a subset of k columns of $M\}$ depend on the $n{\cdot}t$ random variables $M(i, j)$'s. The proof of Theorem 5 shows that the events in $\{E_\Upsilon : \Upsilon$ is a subset of k columns of $M\}$, satisfy the hypothesis of Lemma 1 with $d = k\binom{n-1}{k-1}$ and $b = \binom{n}{k}$.

Now let us turn our attention to Algorithm 1. Notice that Algorithm 1 performs the same steps as those of the algorithm in [38]. To see this, set $\mathcal{X}$ in that algorithm to be the set of random variables $\{M(i, j) : i = 1, \ldots, t$ and $j = 1, \ldots, n\}$, and the bad events to be the $\binom{n}{k}$ events in the set $\{E_\Upsilon : \Upsilon$ is a subset of k columns of $M\}$. Moreover, for each set Υ of k columns of M, the minimal subset of random variables $vbl(E_\Upsilon)$ that determine E_Υ is the set of entries $M(i, j)$'s belonging to the k columns in Υ. Algorithm 1 starts by choosing the values of each entry $M(i, j)$ independently at random according to the probability distribution used to generate M in the proof of Theorem 5. Then, in the repeat-until loop (lines 4-16), the algorithm iteratively checks whether there is a subset C of k columns of M that do not satisfy the property of Definition 2, i.e., such that the corresponding event E_C is violated. If so, Algorithm 1 performs the resampling of E_C in line 14 by sampling a new assignment of values for the entries $M(i, j)$'s belonging to the k columns in C, i.e., for the random variables in $vbl(E_C)$. Algorithm 1 exits the repeat-until loop when there are no more violated events, i.e., no subset of k-columns violates the condition of Definition 2.

In order to complete the proof of the theorem, we need only to show that the set of bad events $\{E_\Upsilon : \Upsilon$ is a subset of k columns of $M\}$ in Algorithm 1 satisfy the hypothesis of Lemma 1. To see this, we observe that Algorithm 1 sets t as in the statement of Theorem 5 and samples the random variables $M(i, j)$'s independently and according to the probability distribution in the proof of Theorem 5. Therefore, from what we have observed at the beginning of this proof, it follows that the events in the set $\{E_\Upsilon : \Upsilon$ is a subset of k columns of $M\}$ satisfy the hypothesis of Lemma 1. Consequently, we can exploit Theorem 6 to infer that Algorithm 1 reaches an evaluation of the random variables $M(i, j)$'s not violating any event in $\{E_\Upsilon : \Upsilon$ is a subset of k columns of $M\}$ after an expected number of resampling steps which is at most $\frac{b}{d} = \frac{\binom{n}{k}}{k\binom{n-1}{k-1}} = \frac{n}{k^2}$. □

Notice that, for fixed k, Theorem 7 implies that Algorithm 1 runs in expected polynomial time.

Notice also that, by setting $m = k$ in Algorithm 1, we obtain a randomized algorithm that generates the q-ary (k, n)-strongly selective codes whose existence is stated by Theorem 3.

We now turn our attention to non-existential results. We prove the following theorem showing that our conflict resolution algorithm is not far from being optimal.

Theorem 8. *Given positive integers q, k, and n, with $q \geq 2$ and $2 \leq k \leq n$, the minimum length of any q-ary (k, n)-strongly selective code is*

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

where the hidden constant is larger than or equal to $1/2$.

Algorithm 1: Algorithm for q-ary (k, m, n)-selectors.

Input: Integers t, k, m, q and n, where $q \geq 2$, $1 \leq m \leq k$, and $2 \leq k \leq n$.
Output: M: a q-ary (k, m, n)-selector.

Let $t := \frac{4(k-1)}{(k-m+1)v}\left(\ln\binom{n}{k-1} + \ln k + \ln\binom{k}{k-m+1} + 1\right)$, where $v = q-1$ if $2 \leq q \leq k$, and $v = k$ if $q > k$;
Let $p := \begin{cases} \frac{1}{k} & \text{if } 2 \leq q \leq k \\ \frac{1}{q-1} & \text{if } q > k \end{cases}$;
Construct a $t \times n$ matrix M where each entry $M(i, j)$ is chosen independently at random from $\{0, 1, \ldots, q-1\}$ with $Pr\{M(i, j) = s\} = p$, for $s \in \{1, \ldots, q\}$, and $Pr\{M(i, j) = 0\} = 1 - (q-1)p$;
repeat
    Set *flag* := true;
    for *each set C of k columns of M* **do**
        if *C does not satisfy the property of Definition 2* **then**
            Set *flag*:= false;
            Set *missing-column-set* := C;
            break;
        end
    end
    if *flag* = *false* **then**
        Choose entries $M(i, j)$'s in the k columns of *missing-column-set* independently at random from $\{0, 1, \ldots, q-1\}$ with $Pr\{M(i, j) = s\} = p$, for $s \in \{1, \ldots, q\}$, and $Pr\{M(i, j) = 0\} = 1 - (q-1)p$;
    end
until *flag* = *true*;
Output M;

Proof. Let M be a $t \times n$ q-ary (k, n)-strongly selective code and let M_B be the $(q-1) \cdot t \times n$ Boolean matrix obtained by replacing each entry by a Boolean column of length $q-1$, with each 0-entry being replaced by the all-zero column and each non-zero entry being replaced by a Boolean column with a single entry equal to 1, according to the following mapping

$$0 \to \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad 1 \to \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad 2 \to \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \quad q-1 \to \begin{pmatrix} 0 \\ 0 \\ \vdots \\ q-1 \end{pmatrix}. \tag{19}$$

By Definition 1, one has that, for any k distinct columns $\mathbf{c}_{j_1}, \ldots, \mathbf{c}_{j_k}$ of M, each column $\mathbf{c}_{j_\ell} \in \{\mathbf{c}_{j_1}, \ldots, \mathbf{c}_{j_k}\}$ is such that there exists a row index i_ℓ for which $\mathbf{c}_{j_\ell}(i_\ell) = s$ for some $s \neq 0$ and $\mathbf{c}_{j_f}(i_\ell) \neq s$ for each $f \in \{1, \ldots, k\} \setminus \{\ell\}$. As a consequence, entry $\mathbf{c}_{j_\ell}(i_\ell)$ is expanded into a Boolean column of M_B with all entries equal to 0 but the s-th one, whereas each entry in $\{\mathbf{c}_{j_1}(i_\ell), \ldots, \mathbf{c}_{j_k}(i_\ell)\} \setminus \{\mathbf{c}_\ell(i_\ell)\}$ is expanded into a Boolean column with the s-th entries equal to 0. From the above argument it follows that M_B is a (classical) binary (k, n)-strongly selective code. The stated lower bound now follows from the very well known $\Omega(\frac{k^2}{\log k} \log \frac{n}{k})$ lower bound [3,20,27,40] on the length of binary (k, n)-strongly selective codes. The constant hidden in that Ω-notation is larger than or equal to 1/2 [20]. □

In view of the equivalence between protocols that solve conflicts in channels with no feedback and q-ary (k, n)-strongly selective codes, it is clear that Theorem 8 implies (3) of Theorem 1.

4.2. Combinatorial tools for the multiple-access model with feedback

In order to prove Theorem 2, dealing with conflict resolution in channels with feedback, we need the combinatorial objects defined below.

Definition 3. Given positive integers q, k, and n, with $q \geq 2$ and $2 \leq k \leq n$, we say that a $t \times n$ matrix M with entries in $\{0, 1, \ldots, q-1\}$ is a q-ary KG(k, n)-code of length t if for any $t \times k$ submatrix M' of M the following property holds.

> There exists a non-empty set of row indices $\{i_1, \ldots, i_\ell\} \subseteq [t]$, with $i_1 < i_2 < \ldots < i_\ell$, and a partition $\{M'_1, \ldots, M'_\ell\}$ of the set of columns of M' such that for each column $\mathbf{c}$ of M'_j, $j = 1, \ldots, \ell$, one has that $\mathbf{c}(i_j) \neq 0$ and that all other columns in $M'_j, \ldots, M'_\ell$ have the i_j-th entry different from $\mathbf{c}(i_j)$.

The minimum length of a q-ary KG(k, n)-code is denoted by $t_{KG}(q, k, n)$.

One can see that in the case $q = 2$, q-ary KG(k, n)-codes coincide with the classical combinatorial structure introduced by Komlós and Greenberg in [34]. The relevance of Definition 3 to our discussion is explained in Theorem 9, showing that a q-ary KG(k, n)-code is indeed equivalent to a scheduling protocol for the multiple-access channel with feedback that allows

all up to k out of n active stations to transmit with success. Intuitively, the row indices $i_1 < i_2 < \ldots < i_\ell$ in Definition 3 correspond to the time slots at which the active stations transmit with success provided that all columns associated with the up to k active stations are contained in M'. Indeed, a scheduling algorithm allows all active stations to transmit with success if and only if the set of up to k active stations can be partitioned into a certain number ℓ of subsets $S_1, \ldots, S_\ell$ such that, for $j = 1, \ldots, \ell$, all stations in S_j are scheduled to transmit singly to one of the $q-1$ channels at some time slot i_j, with $i_1 < i_2 < \ldots < i_\ell$. This corresponds to saying that at time slot i_1 each active station in S_1 is scheduled to transmit to a channel that has not been assigned to any other active stations, and that, for $j = 2, \ldots, \ell$, at time slot i_j, each station in S_j is scheduled to transmit to a channel that has not been assigned to any active station in $S_j \cup S_{j+1} \cup \ldots \cup S_\ell$. Notice that, at time slot i_j, the stations contained in the sets $S_1, \ldots, S_{j-1}$, have already transmitted with success, and therefore, at time slot i_j all (still) active stations are contained in $S_j \cup S_{j+1} \cup \ldots \cup S_\ell$. The following theorem shows that a scheduling algorithm satisfies the above said property if and only if the corresponding q-ary matrix satisfies the property in Definition 3 with the sets of columns $M'_1, \ldots, M'_\ell$ being such that each station in S_j is associated with a distinct column in M'_j.

Theorem 9. *Given positive integers q, k, and n, with $q \geq 2$ and $2 \leq k \leq n$, a scheduling algorithm is a conflict resolution algorithm for a multiple-access channel $\mathcal{C}$ with feedback (comprising of $q-1 \geq 1$ individual channels) that schedules the transmissions of n stations in such a way that all of the up to k active stations transmit successfully,* if and only if *the corresponding q-ary matrix is a q-ary KG(k, n)-code.*

Proof. Given a matrix M, we denote the set of its columns and the set of its column indices by M itself. Consequently, given a set S of stations, the intersection $M \cap S$ is the set of column indices of M associated with the stations in S and, given a submatrix M' of M, the intersection $M' \cap S$ denotes the subset of S formed by stations associated with columns belonging to M'.

Let us first prove the "if part". Let M be a q-ary KG(k, n)-code and let S be the set of up to k active stations. We will prove that, in our multiple-access model with feedback, the scheduling algorithm associated with M allows all active stations in S to transmit with success. To this aim, let us consider any set S' of exactly k stations such that $S \subseteq S'$ and let $M' = M \cap S'$ be the k-column submatrix of M formed by the columns associated with the k stations in S'. By Definition 3, the following property holds.

> There exists a non-empty set of row indices $\{i_1, \ldots, i_\ell\} \subseteq [t]$, with $i_1 < i_2 < \ldots < i_\ell$, and a partition $\{M'_1, \ldots, M'_\ell\}$ of the set of columns of M' such that for each column $\mathbf{c}$ of M'_j, one has that $\mathbf{c}(i_j) \neq 0$ and that all other columns in $M'_j, \ldots, M'_\ell$ have the i_j-th entry different from $\mathbf{c}(i_j)$.

For $j = 1, \ldots, \ell$, let us denote by $S_j \subseteq S$ the set of active stations that are associated with columns in M'_j, i.e., $S_j = S \cap M'_j$. Notice that for some $j \in \{1, \ldots, \ell\}$, it might eventually be $S_j = \emptyset$. The above mentioned property of q-ary KG(k, n)-codes implies that for each column $\mathbf{c}$ in M'_1, one has $\mathbf{c}(i_1) \neq 0$ and $\mathbf{c}(i_1) \neq \mathbf{c}'(i_1)$, for any other column $\mathbf{c}' \in M'_1 \cup \ldots \cup M'_\ell$, and consequently, at time slot i_1 the active station in S_1 associated with $\mathbf{c}$ transmits singly on channel $\mathbf{c}(i)$. Since $\mathbf{c}$ is an arbitrary column of M_1 and each station of S_1 is associated with a distinct column of M_1, we have that each station of S_1 transmits with success at time slot i_1. Therefore, after time slot i_1 the *still* active stations are those in $S_2, \ldots, S_\ell$, and the conflict resolution algorithm needs to solve conflicts that arise among these stations only. At time slot i_2 the property of q-ary KG(k, n)-codes guarantees that for each column $\mathbf{c}$ in M'_2, one has $\mathbf{c}(i_2) \neq 0$ and $\mathbf{c}(i_2) \neq \mathbf{c}'(i_2)$, for any other column $\mathbf{c}' \in M'_2 \cup \ldots \cup M'_\ell$. This implies that at time slot i_2 the active station associated with $\mathbf{c}$ is assigned a channel, namely $\mathbf{c}(i_2)$, which is different from any of the channels assigned to the other stations in $S_2 \cup \ldots \cup S_\ell$. It follows that the station associated with $\mathbf{c}$ transmits singly on the channel $\mathbf{c}(i_2)$. Since $\mathbf{c}$ is an arbitrary column of M_2 and each station of S_2 is associated with a distinct column of M_2, we have that each station of S_2 transmits with success at time slot i_2. Inductive reasoning shows that, for $j = 1, \ldots, \ell$, all stations in $S_1 \cup \ldots \cup S_j$ are inactive after time slot i_j, as a consequence of the fact that each of them transmitted with success in one of time slots $i_1, \ldots, i_j$.

Now we prove the "only if part". To this aim, let us consider a conflict resolution algorithm for our multiple-access model with feedback that allows each of the up to k active stations to transmit successfully. Let M be the q-ary matrix associated with this algorithm. We will show that M is a q-ary KG(k, n)-code. To this aim, let S be an arbitrary set of exactly k stations and let us assume that S is the set of active stations. We partition S into the sets $S_1, \ldots, S_\ell$ defined as follows. S_1 is the subset of active stations that transmit with success at the same time slot, before all active stations in $S_2 \cup \ldots \cup S_\ell$, S_2 is the subset of active stations that transmit with success at the same time slot, before all active stations in $S_3 \cup \ldots \cup S_\ell$, and so on. For $j = 1, \ldots, \ell$, let i_j denote the time slot at which the stations in S_j transmit with success. By definition of S_j it is $i_1 < i_2 < \ldots < i_\ell$. Let M' be the k-column submatrix of M formed by the columns associated with the stations in S, i.e., $M' = M \cap S$, and M'_j denote the submatrix of M' formed by the columns associated with the stations in S_j, i.e., $M'_j = M' \cap S_j$, for $j = 1, \ldots, \ell$. The submatrices $M'_1, \ldots, M'_\ell$ form a partition of M'. In order for the active stations in S_j to transmit with success at time slot i_j, each station in S_j must transmit over a channel that is not assigned to any other active station. In other words, each station in S_j must be assigned a non-zero integer between 1 and $q-1$ which is not assigned to any other station in $S_j \cup \ldots \cup S_\ell$. This translates into saying that for each column $\mathbf{c}$ of M'_j and for any other column $\mathbf{c}' \in M'_j \cup \ldots \cup M'_\ell$, one has $\mathbf{c}(i_j) \neq 0$ and $\mathbf{c}(i_j) \neq \mathbf{c}'(i_j)$. It follows that for the k-column submatrix M' of M, there exist indices $i_1, \ldots, i_\ell$ and matrices $M'_1, \ldots, M'_\ell$ satisfying the property of Definition 3. Since M' is an arbitrary k-column submatrix of M, in that it

corresponds to an arbitrary set S of k stations, this property is satisfied for any k-column submatrix of M, and consequently, M is a KG(k,n)-code. □

The following result, together with the above Theorem 9, prove formula (4) of Theorem 2.

Theorem 10. *Given positive integers q, k, and n, with $2 \leq q \leq k \leq n$, there exists a q-ary KG(k,n)-code of length t with*

$$t = O\left(\frac{k}{q}\log\frac{n}{k}\right). \tag{20}$$

Proof. In this proof we exploit that the fact, stated in Theorem 4, that a q-ary (k,m,n)-selector is equivalent to a scheduling algorithm that allows all but at most $k-m$ active stations to transmit with success.

In order to obtain a q-ary KG(k,n)-code, we use the same idea as the one exploited in [17,35] to construct a binary KG(k,n)-code. The idea consists in concatenating q-ary $(2^j, 2^{j-1}, n)$-selectors, one on the top of the another, starting from $j = \lceil \log k \rceil$ through $j = 1$, with the rows of the $(2^j, 2^{j-1}, n)$-selector placed on the top of the $(2^{j-1}, 2^{j-2}, n)$-selector. An all-one row is then placed at very end (bottom) of the resulting matrix. Let M be the final matrix so obtained. The number of rows of M is $\sum_{j=1}^{\lceil \log k \rceil} t_{sel}(q, 2^j, 2^{j-1}, n)$. By the upper bound for $q \leq k$ of Theorem 5, this number of rows is

$$O\left(\sum_{j=1}^{\lceil \log k \rceil} \frac{2^{j+1}}{q}\log\frac{n}{2^j}\right) = O\left(\frac{k}{q}\log\frac{n}{k}\right).$$

To the aim of proving that M is a q-ary KG(k,n)-code, we show that M defines a scheduling algorithm for our channel with feedback that allows each of the up to k active stations to transmit with success. In virtue of Theorem 9, this implies that M is a q-ary KG(k,n)-code.

Let $S_{\lceil \log k \rceil}$ be an arbitrary set of $2^{\lceil \log k \rceil}$ stations and let us assume that the up to k active stations are contained in $S_{\lceil \log k \rceil}$. Since M contains at the top the rows of a q-ary $(2^{\lceil \log k \rceil}, 2^{\lceil \log k \rceil - 1}, n)$-selector and $k \leq 2^{\lceil \log k \rceil}$, then Theorem 4 with k being replaced by $2^{\lceil \log k \rceil}$ and m by $2^{\lceil \log k \rceil - 1}$, implies that all but at most $2^{\lceil \log k \rceil} - 2^{\lceil \log k \rceil - 1} = 2^{\lceil \log k \rceil - 1}$ active stations transmit with success. Let $S_{\lceil \log k \rceil - 1} \subseteq S_{\lceil \log k \rceil}$ denote the subset of these up to $2^{\lceil \log k \rceil - 1}$ *still* active stations. Immediately after the rows of the q-ary $(2^{\lceil \log k \rceil}, 2^{\lceil \log k \rceil - 1}, n)$-selector, in M we find placed the rows of a q-ary $(2^{\lceil \log k \rceil - 1}, 2^{\lceil \log k \rceil - 2}, n)$-selector. Theorem 4, with k replaced by $2^{\lceil \log k \rceil - 1}$ and m by $2^{\lceil \log k \rceil - 2}$, implies that all but at most $2^{\lceil \log k \rceil - 1} - 2^{\lceil \log k \rceil - 2} = 2^{\lceil \log k \rceil - 2}$ active stations transmit with success, and therefore, after running the scheduling algorithm associated with the rows of the q-ary $(2^{\lceil \log k \rceil - 1}, 2^{\lceil \log k \rceil - 2}, n)$-selector, one is left with at most $2^{\lceil \log k \rceil - 2}$ active stations. Let $S_{\lceil \log k \rceil - 2} \subseteq S_{\lceil \log k \rceil - 1}$ denote the subset of these up to $2^{\lceil \log k \rceil - 2}$ *still* active stations. Extending this reasoning to an arbitrary $j \in \{1, \ldots, \lceil \log k \rceil\}$, we define S_j as the subset of active stations that have not been scheduled to transmit successfully in any of the time slots corresponding to the rows positioned above those of the q-ary $(2^j, 2^{j-1}, n)$-selector. For $j = 0$, S_0 is the set of active stations that have not been scheduled to transmit successfully in any of the time slots corresponding to the rows of all concatenated selectors. Iteratively applying the argument used for $j = \lceil \log k \rceil - 1$ and $j = \lceil \log k \rceil - 2$, one has that S_j contains at most 2^j active stations. Therefore, S_0 either is empty or contains the unique active station that has not transmitted successfully in any of the time slots corresponding to the rows of the q-ary selectors. This station transmits with success at the very last time slot when all stations are scheduled to transmit by the all-one row but no other station is active. □

The q-ary KG(k,n)-code in Theorem 10 is built by concatenating the q-ary (k,m,n)-selectors of Theorem 5, with $k = 2^j$ and $m = 2^{j-1}$, for $j = 1, \ldots, \lceil \log k \rceil$. Therefore, by Theorem 7, it is possible to generate this q-ary KG(k,n)-code by running Algorithm 1 with $k = 2^j$ and $m = 2^{j-1}$, for $j = 1, \ldots, \lceil \log k \rceil$.

We now prove the lower bound (5) of Theorem 2, showing the *optimality* of our construction. We introduce the following definition.

Definition 4. Given positive integers q, k, and n, with $q \geq 2$ and $2 \leq k \leq n$, we say that a $t \times n$ matrix M with entries in $\{0, 1, \ldots, q-1\}$ is a q-ary (k,n)-locally thin code of length t if for any submatrix M' of up to k columns of M there is a row index i such that, for some $s \in \{1, \ldots, q-1\}$, one has that s occurs *exactly* in one entry of the i-th row of M'. The minimum length of a q-ary (k,n)-locally thin code is denoted by $t_{LT}(q,k,n)$.

In the case $q = 2$, locally thin codes have been studied in [13–15,33] (each paper independently rediscovered the same result, but with a different terminology).

Theorem 11. *Given positive integers q, k, and n, with $q \geq 2$ and $2 \leq k \leq n$, the minimum length $t_{LT}(q,k,n)$ of a q-ary (k,n)-locally thin code is*

$$t_{LT}(q,k,n) = \Omega\left(\frac{k}{q}\log\frac{n}{k}\right).$$

Proof. Let M be a $t \times n$ q-ary (k,n)-locally thin code and let M_B be the $(q-1)\cdot t \times n$ Boolean matrix obtained, as in the proof of Theorem 8, by replacing each entry $M(i,j)$ of M by a Boolean column of length $q-1$ according to the mapping (19) in the proof of Theorem 8. By Definition 4, one has that, for any distinct columns $\mathbf{c}_{j_1},\ldots,\mathbf{c}_{j_p}$, $p \le k$, of M, there exists a row index i and a column $\mathbf{c} \in \{\mathbf{c}_{j_1},\ldots,\mathbf{c}_{j_p}\}$ such that, $\mathbf{c}(i) = s$ for some $s \neq 0$ and $\mathbf{c}_{j_f}(i) \neq s$ for each $\mathbf{c}_{j_f} \in \{\mathbf{c}_{j_1},\ldots,\mathbf{c}_{j_p}\} \setminus \{\mathbf{c}\}$. It follows that entry $\mathbf{c}(i)$ is expanded into a Boolean column of M_B with all entries equal to 0 but the s-th one, whereas each entry in $\{\mathbf{c}_{j_1}(i),\ldots,\mathbf{c}_{j_p}(i)\} \setminus \{\mathbf{c}(i)\}$ is expanded into a Boolean column with the s-th entry equal to 0. From the above argument it follows that M_B is a binary (k,n)-locally thin code. The lower bound in the statement of the theorem is a consequence of the $\Omega\left(k\log\frac{n}{k}\right)$ lower bound of [13–15,33] on the length of binary (k,n)-locally thin codes. □

The relevance of locally thin codes to our discourse is explained by the following result.

Lemma 2. *Given positive integers q, k, and n, with $q \ge 2$ and $2 \le k \le n$,* any *q-ary KG(k,n)-code is a q-ary (k,n)-locally thin code.*

Proof. Again, given a matrix M, we denote the set of its columns and the set of its column indices by M itself.

Let M be a q-ary KG(k,n)-code and assume by contradiction that M is not a q-ary (k,n)-locally thin code. By the contradiction hypothesis, M contains a submatrix $\tilde{M}$ of $p \le k$ columns such that, for each element $s \in \{1,\ldots,q-1\}$, any row of $\tilde{M}$ contains either no entry or at least two entries equal to s. Let M' be a k-column submatrix of M that contains all columns of $\tilde{M}$. By Definition 3, the following property holds.

> There exists a non-empty set of row indices $\{i_1,\ldots,i_\ell\} \subseteq [t]$, with $i_1 < i_2 < \ldots < i_\ell$, and a partition $\{M'_1,\ldots,M'_\ell\}$ of the set of columns of M' such that, for each column $\mathbf{c}$ of M'_j, one has that $\mathbf{c}(i_j) \neq 0$ and all other columns in $M'_j,\ldots,M'_\ell$ have the i_j-th entry different from $\mathbf{c}(i_j)$.

Let us consider a set of row indices $\{i_1,\ldots,i_\ell\}$ and a partition $\{M'_1,\ldots,M'_\ell\}$ of the set of columns of M' satisfying the above property. We denote by $h \in \{1,\ldots,\ell\}$ the smallest index for which $M'_h \cap \tilde{M} \neq \emptyset$. By the contradiction hypothesis, there is no element $s \neq 0$ that occurs exactly once in the i_h-th row of $\tilde{M}$. Consequently, there are two possible cases: either all entries in the i_h-th row of $M'_h \cap \tilde{M}$ are equal to 0, or $M'_h \cap \tilde{M}$ contains at least a column $\mathbf{c}$ such that $\mathbf{c}(i_h) \neq 0$ and, in this latter case, the contradiction hypothesis implies there exists a column $\mathbf{c}' \neq \mathbf{c}$ in $\tilde{M}$ such that $\mathbf{c}(i_h) = \mathbf{c}'(i_h)$. Notice that, $\tilde{M} \subseteq M'_h \cup \ldots \cup M'_\ell$, and consequently, $\mathbf{c}'$ belongs to one of submatrices $M'_h,\ldots,M'_\ell$. In the former case, M'_h contains at least a column $\mathbf{c}$ such that $\mathbf{c}(i) = 0$, whereas in the latter case, M'_h contains at least a column $\mathbf{c}$ such that $\mathbf{c}(i_h) \neq 0$ and $\mathbf{c}(i_h) = \mathbf{c}'(i_h)$, for some other column $\mathbf{c}'$ belonging to one of submatrices $M'_h,\ldots,M'_\ell$. In both cases, we have a contradiction to the fact that $\{i_1,\ldots,i_\ell\}$ and $\{M'_1,\ldots,M'_\ell\}$ satisfy the above property of Definition 3. □

The following theorem is an immediate consequence of Theorem 11 and Lemma 2.

Theorem 12. *Given positive integers q, k, and n, with $q \ge 2$ and $2 \le k \le n$, the minimum length $t_{KG}(q,k,n)$ of a q-ary KG(k,n)-code is*

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

Clearly, above Theorems 9 and 12 imply the lower bound (5) of Theorem 2.

5. Relationships with frameproof codes

The combinatorial structures of Definition 1 have similarities with the q-ary *k-frameproof codes* introduced by Boneh and Shaw [5] in the context of fingerprinting for digital data, i.e., cryptographic schemes intended to protect copyrighted material, such as digital data, software packages, pay-per-view television broadcasts, etc. To this aim the distributors of the protected material mark each copy with a distinct codeword before distributing the copy. Alternatively, codes may be associated with the keys used to recover the content of the protected material. This marking allows the distributors to detect any unauthorized copy and trace it back to the user who released the unauthorized copy. Several variants of traceability codes have been studied in the literature, and consequently, different definitions of frameproof codes have been given. In this section we are concerned with the frameproof codes defined by Fiat and Tassa [25], and extensively studied in [43–46]. This version of q-ary k-frameproof codes are useful in preventing any coalition of at most k users to frame a user not in the coalition by generating the codeword held by that user. In the following we give a formal definition of this variant of q-ary k-frameproof codes.

A q-ary k-frameproof code of size n and length t is a $t \times n$ q-ary matrix having the property that for any k arbitrary distinct columns $\mathbf{c}_{j_1},\ldots,\mathbf{c}_{j_k}$ and for any other column $\mathbf{c}$, there exists a row index i such that $\mathbf{c}(i) \notin \{\mathbf{c}_{j_1}(i),\ldots,\mathbf{c}_{j_k}(i)\}$. We denote the minimum length of a q-ary k-frameproof code of size n by $t_{FP}(q,k,n)$. It is obvious that a q-ary $(k+1,n)$-strongly selective code of length t is a k-frameproof code of size n and length t. On the other hand, a q-ary k-frameproof code is

not necessarily a q-ary $(k+1,n)$-strongly selective code since the above said entry $\mathbf{c}(i)$ is possibly equal to 0. However, the minimum length of $(k+1,n)$-strongly selective codes is at most twice the minimum length of k-frameproof codes of size n. Indeed, if one is given a k-frameproof code M of size n then one can construct a $(k+1,n)$-strongly selective code twice as long as M by concatenating the rows of M to those of the matrix obtained by replacing each entry $M(i,j)$ in M by $q-1-M(i,j)$. Therefore, our lower bound on the minimum length of strongly selective codes of Theorem 8 implies the following lower bound on the minimum length $t_{FP}(q,k,n)$ of k-frameproof codes of size n.

Theorem 13. *Given positive integers q, k, and n, with $q \geq 2$ and $2 \leq k \leq n$, the minimum length of a q-ary k-frameproof code of size n is*

$$t_{FP}(q,k,n) = \Omega\left(\frac{k^2}{q\log k}\log\frac{n}{k}\right),$$

where the hidden constant is larger than or equal to $1/4$.

In [43] it has been proved an upper bound on the size of k-frameproof codes (see Thm. 3.7 of [43]) that translates into the following lower bound on the minimum length of k-frameproof codes of size n.

$$t_{FP}(q,k,n) > k\left(\frac{\log(n/k+1)}{\log q} - 1\right). \tag{21}$$

We observe that our lower bound of Theorem 13 differs asymptotically from the lower bound (21) by a factor of $\frac{k\log q}{q\log k}$ that, for $q<k$, is larger than 1. Therefore, for $q<k$ our lower bound of Theorem 13 outperforms the lower bound (21).

By replacing k with $k+1$ and setting $m=k+1$ in the statement of Theorem 5, we obtain an upper bound on the minimum length of $(k+1,n)$-strongly selective codes that translates in the following upper bound on the minimum length of k-frameproof codes of size n.

Theorem 14. *Given positive integers q, k, and n, with $q \geq 2$ and $1 \leq k \leq n$, there exists a q-ary k-frameproof code of size n and length*

$$t_{FP}(q,k,n) \leq \begin{cases} \frac{4k}{q-1}\left(\ln\binom{n}{k} + 2\ln(k+1) + 1\right) = O\left(\frac{k^2}{q}\log\frac{n}{k}\right), & \text{if } q \leq k, \\ 4\left(\ln\binom{n}{k} + 2\ln(k+1) + 1\right) = O\left(k\log\frac{n}{k}\right) & \text{otherwise.} \end{cases}$$

Theorem 14 provides an upper bound for k-frameproof codes, over an alphabet of arbitrary size $q \geq 2$, that differs asymptotically by a $\log k$ factor from the lower bound of Theorem 13, if $q \leq k$, and by a $\log q$ factor from lower bound (21), if $q > k$.

We remark that the frameproof codes of Theorem 14 can be generated in pseudo-polynomial time by Algorithm 1, by replacing k with $k+1$ and setting $m=k+1$.

To our knowledge, the best existential result for k-frameproof codes has been provided by the authors of [45] who obtained the following upper bound on the minimum length of k-frameproof codes (see Cor. 19 of [45]) holding for an asymptotically large size of the alphabet (i.e., $q \to \infty$).

$$t_{FP}(q,k,n) \leq \frac{k}{\log\left(\frac{q}{k}(1+o(1))\right)}\log\left(\frac{nk!}{k!-1}\right). \tag{22}$$

The upper bound (22) can be compared with the upper bound stated by Theorem 14 for the case $q > k$. Though the bound (22) outperforms our upper bound for the case $q > k$, we remark that our bound holds for arbitrary values of q and k such that $q > k \geq 1$, whereas (22) holds only for $q \to \infty$.

6. Optimal length q-ary strongly selective codes via Reed-Solomon codes

Our upper bound of Theorem 3 and lower bound of Theorem 8 on the minimum length of q-ary (k,n)-strongly selective codes differ by a $\log k$ multiplicative factor. Clearly, the same gap transfers to our estimation of the length of optimal schedules for conflict resolution in our multichannels scenario with *no feedback*. Closing this gap would give a breakthrough result in some long-standing combinatorial open problem, i.e., the exact estimation of the optimal length of classical binary superimposed codes and of frameproof codes. In this section we show that this gap can indeed be closed for an infinite number of values of the involved parameters. To this aim we give a construction of q-ary (k,n)-strongly selective codes based on the well known Reed-Solomon codes, whose definition we recall below.

Definition 5. Let q be a prime power and let us denote by $\mathbb{F}_q$ the finite field with q elements. For any integer m we denote by $\mathbb{F}_q^m$ the vector space comprising all m-components vectors with components in $\mathbb{F}_q$. Let $\alpha_1, \ldots, \alpha_N$ be distinct elements from $\mathbb{F}_q$ and let us choose integers N and a such that $a \leq N \leq q$. We define an encoding function for Reed-Solomon codes as $RS : \mathbb{F}_q^a \rightarrow F_q^N$ as follows. A message (vector) $\mathbf{m} = (m_0, m_1, \ldots, m_{a-1}) \in \mathbb{F}_q^a$, with each $m_i \in \mathbb{F}_q$, is first mapped to a polynomial $f_{\mathbf{m}}(x) = \sum_{i=0}^{a-1} m_i x^i$, of degree $a-1$, $a > 2$. The encoding $RS(\mathbf{m})$ of $\mathbf{m}$ (i.e., the codeword associated to $\mathbf{m}$) is the value of the polynomial $f_{\mathbf{m}}(x)$ *at all* the α_i's:

$$RS(\mathbf{m}) = (f_{\mathbf{m}}(\alpha_1), f_{\mathbf{m}}(\alpha_2), ..., f_{\mathbf{m}}(\alpha_N)).$$

The parameter N in the above definition is commonly taken equal to $q-1$ and consequently the polynomials $f_{\mathbf{m}}(x)$ are evaluated at all points in $\mathbb{F}_q \setminus \{0\}$. Reed-Solomon codes are linear codes (i.e., are vector spaces over $\mathbb{F}_q$), with each codeword being of length N and the number of distinct codewords being equal to the number of distinct polynomials $f_{\mathbf{m}}(x) = \sum_{i=0}^{a-1} m_i x^i$ over $\mathbb{F}_q$, that is, equal to q^a. It is well known that Reed-Solomon codes meet the Singleton bound in that they have minimum Hamming distance equal to $N-a+1$ (i.e., each pair of codewords differ in at least $N-a+1$ coordinates, and this is the best one can achieve among the class of codes of codeword length N and cardinality q^a). For undefined terms and concepts of Error Correcting Codes see [37].

The authors of [32] exploited Reed-Solomon codes to obtain *binary* (k,n)-strongly selective codes with $k = \lfloor \frac{N-1}{a-1} \rfloor$ and length $(q-1)N$. We generalize their construction to obtain (k,n)-strongly selective codes on an alphabet of *arbitrary* cardinality v, $2 \leq v \leq q$. Recall that, from basic results in the theory of finite fields, the cardinality q of the finite field $\mathbb{F}_q$ on which Reed-Solomon codes are defined can be any prime power [37].

Our construction proceeds as follows. We start with a q-ary Reed-Solomon code of length $N = q-1$ and minimum distance $N-a+1$. It will be convenient to consider the q-ary matrix M, of dimension $N \times q^a$, where the columns of M are the q^a codewords of the Reed-Solomon code. Subsequently, we replace each entry in M by a column vector of length $\lfloor \frac{q-1}{v-1} \rfloor$ having a single non-zero entry. Namely, an entry equal to $x \in \{0, 1, \ldots, q-1\}$ is replaced by the $\lfloor \frac{q-1}{v-1} \rfloor$-entry column vector $\mathbf{c}_x^T = (\mathbf{c}_x(0), \ldots, \mathbf{c}_x(\lfloor \frac{q-1}{v-1} \rfloor))^T$ with

$$\mathbf{c}_x(i) = \begin{cases} x \bmod (v-1) + 1, & \text{if } i = \lfloor \frac{x}{v-1} \rfloor, \\ 0 & \text{otherwise.} \end{cases}$$

Let S denote the v-ary matrix that results from replacing each entry $x \in \{0, 1, \ldots, q-1\}$ in M by the vector $\mathbf{c}_x$. The matrix S has $t = N\lfloor \frac{q-1}{v-1} \rfloor$ rows, and each column of S has *exactly* N non-zero entries, given that each of the N entries in the original column of M has been replaced by a column vector with exactly one non-zero entry.

We now show that any column of S differs from any other column of S in at least $a-1$ *non-zero* entries. To this aim, we observe that, for any two distinct symbols $x, x' \in \{0, 1, \ldots, q-1\}$, we have that at least one of the following conditions holds:

- $x \bmod (v-1) + 1 \neq x' \bmod (v-1) + 1$,
- $\lfloor \frac{x}{v-1} \rfloor \neq \lfloor \frac{x'}{v-1} \rfloor$.

Consequently, vectors $\mathbf{c}_x$ and $\mathbf{c}_{x'}$ differ in at least one entry. More precisely, if the second of the above two conditions is satisfied then there exist two indices i and i', namely $i = \lfloor \frac{x}{v-1} \rfloor$ and $i' = \lfloor \frac{x'}{v-1} \rfloor$, such that $0 = \mathbf{c}'_x(i) \neq \mathbf{c}_x(i)$ and $0 = \mathbf{c}_x(i') \neq \mathbf{c}_{x'}(i')$, whereas if the second condition is not satisfied then the first condition must necessarily hold, and consequently, there exists an index i, namely $i = \lfloor \frac{x}{v-1} \rfloor = \lfloor \frac{x'}{v-1} \rfloor$, such that $\mathbf{c}_x(i) \neq 0$, $\mathbf{c}_{x'}(i) \neq 0$, and $\mathbf{c}_x(i) \neq \mathbf{c}_{x'}(i)$. In both of the above cases, one has that $\mathbf{c}_x$ has a non-zero entry which is different from the corresponding entry of $\mathbf{c}_{x'}$, and $\mathbf{c}_{x'}$ has a non-zero entry which is different from the corresponding entry of $\mathbf{c}_x$. It follows that if two codewords of the Reed-Solomon code differ in d entries for some integer d (that is, if two columns of M differ in d entries), then each of the two corresponding columns in S has d non-zero entries at which it differs from the other column.

From the above discussion we have that, for any pair of columns in S, the number of non-zero entries at which they differ is at least as large as the minimum distance of the Reed-Solomon code, namely $N-a+1$. This, along with the fact that each column in S has exactly N non zero entries, implies that for any two columns $\mathbf{c}_1, \mathbf{c}_2 \in S$ there are at most $N - (N-a+1) = a-1$ indices i's such that $\mathbf{c}_1(i) = \mathbf{c}_2(i) \neq 0$. As a consequence, for any column $\mathbf{c} \in S$ and for any other k columns $\mathbf{c}_1, \ldots, \mathbf{c}_k \in S$, $\mathbf{c}$ has at most $a-1$ non-zero entries in common with each of $\mathbf{c}_1, \ldots, \mathbf{c}_k$, and therefore, there exist at most $k(a-1)$ indices i's such that $\mathbf{c}(i) \neq 0$ and $\mathbf{c}_j(i) = \mathbf{c}(i)$ for some $j \in \{1, \ldots, k\}$. If we choose $k = \left\lfloor \frac{N-1}{a-1} \right\rfloor = \left\lceil \frac{N}{a-1} \right\rceil - 1$ then it holds $k(a-1) < N$ and, by the above argument, there is an index i such that $\mathbf{c}(i) \neq 0$ and $\mathbf{c}_1(i) \neq \mathbf{c}(i), \mathbf{c}_2(i) \neq \mathbf{c}(i), \ldots, \mathbf{c}_k(i) \neq \mathbf{c}(i)$. In other words, S is a v-ary (k,n)-strongly selective code, where $n = q^a$, and its length t and "selection" capability k are equal to

$$t = N \left\lfloor \frac{q-1}{v-1} \right\rfloor \qquad k = \left\lceil \frac{N}{a-1} \right\rceil - 1. \tag{23}$$

Now, let us consider the parameter a be constant, and let instead the size $q = N + 1$ of the field F_q vary among the set of prime powers (there are many). Moreover, we also constrain the value of k in such a way that $k \leq bn^{1/a}$, for some constant b. The relations between t, n, k and the alphabet cardinality of the v-ary (k, n)-strongly selective code that we have constructed above, is evaluated as follows:

$$\begin{aligned}
t &\leq \frac{(q-1)^2}{v-1} && \text{(since } N = q-1 \text{ and from the first identity in (23))}\\
&= \frac{1}{v-1}\frac{(q-1)^2}{(a-1)^2}(a-1)^2 &&\\
&\leq (a-1)^2\frac{(k+1)^2}{v-1} && \text{(from the second identity in (23))}\\
&\leq (a-1)a\frac{(k+1)^2}{v-1} &&\\
&\leq (a-1)\frac{(k+1)^2}{(v-1)(\log k - \log b)}\log n && \text{(because we are in the regime } k \leq bn^{1/a}\text{).}
\end{aligned}$$

Since a and b are constant, while q is an *arbitrary* prime power, we have found an infinite set of values of the parameters $n = q^a$ and k for which there exist v-ary (k, n)-strongly selective codes of length t matching (asymptotically) the lower bound $\Omega\left(\frac{k^2}{(v-1)\log k}\log\frac{n}{k}\right)$ proved in Theorem 8. The following theorem is therefore a consequence of the above construction.

Theorem 15. *Let a be an integer constant with $a > 2$, and b be a positive constant. Moreover, let $q > a$ be a prime power such that $\left\lceil\frac{q-1}{a-1}\right\rceil - 1 \leq bn^{1/a}$. For $k = \left\lceil\frac{q-1}{a-1}\right\rceil - 1$, $n = q^a$, and for any integer v, $2 \leq v \leq q$, there exists a v-ary (k, n)-strongly selective code of asymptotically optimal length*

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

It has not escaped our attention that, for $v = 2$, Theorem 15 implies that, for an infinite set of the parameters k and n, it is possible to construct in *polynomial* time, binary (k, n)-strongly selective codes of *optimal* length $\Theta\left(\frac{k^2}{\log k}\log n\right)$.

Constructing binary (k, n)-strongly selective codes of optimal length $\Theta\left(\frac{k^2}{\log k}\log n\right)$, for *all* values of the involved parameters (or, more modestly, just showing their existence) is a problem that has been open for decades. Recently, a general unconstrained tight bound $\Theta\left(\frac{k^2}{\log k}\log n\right)$ was claimed in [21], however this claim has now been retracted [22].

7. Conclusions

In this paper we have studied the Information Exchange Problem in wireless networks under the assumption that stations have a common access to a multichannel comprising of a certain number of individual multi-access channels. In our knowledge this is the first paper that presents protocols for the Information Exchange Problem that work with an arbitrary number of individual channels (i.e., not dependent on the other parameters of the problem). Our results show that the time needed to disseminate the information held by the active stations decreases linearly with the number of available channels. In order to study this question, we have introduced two new combinatorial structures that consist in a generalization of selectors [11,17] and a generalization of Komlós and Greenberg codes [34]. Our upper and lower bounds for these structures show that our protocols are not far from being optimal. We have also presented an algorithm that generates the combinatorial structures underlying our protocols. For a constant number of active stations this algorithm runs in polynomial time.

An interesting future research goal is to give an answer to the question of whether the results obtained in this paper for the Information Exchange Problem can be extended also to the scenario where there is no global clock. In other words, it would be interesting to see how the time needed to disseminate information in the network depends on the number of available channels in an asynchronous communication model.

Acknowledgements

The authors are grateful to the reviewers for the careful reading of the paper and for the many helpful comments.

References

[1] I.802.11 Wireless LAN MAC and physical layer specification, available at http://www.ieee802.org/.
[2] N. Alon, J.H. Spencer, The Probabilistic Method, third edition, Wiley-Interscience Series in Discr. Math. and Optimization, John Wiley & Sons, Inc., 2008.
[3] N. Alon, V. Asodi, Learning a hidden subgraph, SIAM J. Discrete Math. 18 (4) (2005) 697–712.
[4] I. Avgouleas, V. Angelakis, N. Pappas, Utilizing multiple full-duplex relays in wireless systems with multiple packet reception, in: 19th IEEE International Workshop on Computer Aided Modeling and Design of Communication Links and Networks, CAMAD 2014, 2014, pp. 193–197.
[5] D. Boneh, J. Shaw, Collusion-secure fingerprinting for digital data, IEEE Trans. Inf. Theory 44 (5) (1998) 1897–1905.
[6] D. Bharadia, E. McMilin, S. Katti, Full duplex radios, ACM SIGCOMM 2013, Comput. Commun. Rev. 43 (4) (2013) 375–386.
[7] Bluetooth Consortium, Bluetooth specification, available at https://www.bluetooth.com/.
[8] I. Chlamtac, A. Faragó, An optimal channel access protocol with multiple reception capacity, IEEE Trans. Comput. 43 (1994) 480–484.
[9] B.S. Chlebus, G. De Marco, D.R. Kowalski, Scalable wake-up of multi-channel single-hop radio networks, Theor. Comput. Sci. 615 (2016) 23–44.
[10] B.S. Chlebus, Randomized communication in radio networks, in: P.M. Pardalos, S. Rajasekaran, J.H. Reif, J.D.P. Rolim (Eds.), Handbook on Randomized Computing, vol. I, Kluwer Academic Publishers, 2001, pp. 401–456.
[11] B.S. Chlebus, D.R. Kowalski, Almost optimal explicit selectors, in: M. Liskiewicz, R. Reischuk (Eds.), 15th International Symposium on Fundamentals of Computation Theory, FCT 2005, in: Lecture Notes in Computer Science, vol. 3623, 2005, pp. 270–280.
[12] M. Chrobak, L. Gąsieniec, W. Rytter, Fast broadcasting and gossiping in radio networks, J. Algorithms 43 (2) (2002) 177–189.
[13] A.E.F. Clementi, A. Monti, R. Silvestri, Distributed broadcast in radio networks of unknown topology, Theor. Comput. Sci. 302 (1–3) (2003) 337–364.
[14] G.D. Cohen, Applications of coding theory to communication combinatorial problems, Discrete Math. 83 (2–3) (1990) 237–248.
[15] M. Csűrös, M. Ruszinkó, Single-user tracing and disjointly superimposed codes, IEEE Trans. Inf. Theory 51 (4) (2005) 1606–1611.
[16] S. Daum, F. Kuhn, C. Newport, Efficient symmetry breaking in multi-channel radio networks, in: M.K. Aguilera (Ed.), 26th International Symposium on Distributed Computing, DISC 2012, in: Lecture Notes in Computer Science, vol. 7611, 2012, pp. 238–252.
[17] A. De Bonis, L. Gąsieniec, U. Vaccaro, Optimal two-stage algorithms for group testing problems, SIAM J. Comput. 34 (5) (2005) 1253–1270.
[18] S. Dolev, S. Gilbert, M. Khabbazian, C.C. Newport, Leveraging channel diversity to gain efficiency and robustness for wireless broadcast, in: D. Peleg (Ed.), 25th International Symposium on Distributed Computing, DISC 2011, in: Lecture Notes in Computer Science, vol. 6950, 2011, pp. 252–267.
[19] D.Z. Du, F.K. Hwang, Combinatorial Group Testing and Its Applications, World Scientific, River Edge, NJ, 2000.
[20] A.G. D'yachkov, V.V. Rykov, Bounds on the length of disjunctive codes, Probl. Inf. Transm. 18 (3) (1982) 166–171.
[21] A.G. D'yachkov, I.V. Vorobév, N.A. Polyansky, V.Yu. Shchukin, Bounds on the rate of disjunctive codes, Probl. Inf. Transm. 50 (1) (2014) 27–56.
[22] A.G. D'yachkov, I.V. Vorobév, N.A. Polyansky, V.Yu. Shchukin, Erratum to: "Bounds on the rate of disjunctive codes", Probl. Inf. Transm. 50 (2014) 27, Probl. Inf. Transm. 52 (2) (2016) 200.
[23] Duplo Consortium, http://www.fp7-duplo.eu/.
[24] P. Erdös, P. Frankl, Z. Füredi, Families of finite sets in which no set is covered by the union of r others, Isr. J. Math. 51 (1–2) (1985) 75–89.
[25] A. Fiat, T. Tassa, Dynamic traitor tracing, in: M. Weiner (Ed.), Advances in Cryptology, CRYPTO '99, in: Lecture Notes in Computer Science, vol. 1666, 1999, pp. 354–371.
[26] Flexicon Consortium, http://flexicon.ee.columbia.edu/.
[27] Z. Füredi, On r-cover-free families, J. Comb. Theory, Ser. A 73 (1) (1996) 172–173.
[28] S. Győry, Coding for a multiple access OR channel: a survey, Discrete Appl. Math. 156 (9) (2008) 1407–1430.
[29] M. Halldórsson, Y. Wang, D. Yu, Leveraging multiple channels in ad hoc networks, in: 2015 ACM Symposium on Principles of Distributed Computing, PODC 2015, 2015, pp. 431–440.
[30] K. Haneda, M. Valkama, T. Riihonen, E. Antonio-Rodriguez, D. Korpi, Design and implementation of full-duplex transceivers, in: F. Luo, J. Zhang (Eds.), Signal Processing for 5G: Algorithms and Implementations, Wiley, 2016, pp. 402–428.
[31] S. Holzer, Y. Pignolet, J. Smula, R. Wattenhofer, Time-optimal information exchange on multiple channels, in: The Seventh ACM SIGACT/SIGMOBILE International Workshop on Foundations of Mobile Computing, FOMC 2011, 2011, pp. 69–76.
[32] W.H. Kautz, R.C. Singleton, Nonrandom binary superimposed codes, IEEE Trans. Inf. Theory 10 (4) (1964) 363–377.
[33] L.S. Khasin, Conflict resolution in a multiple access channel, Probl. Pereda. Inf. 25 (4) (1989) 63–68.
[34] J. Komlós, A.G. Greenberg, An asymptotically fast non-adaptive algorithm for conflict resolution in multiple-access channels, IEEE Trans. Inf. Theory 31 (2) (1985) 302–306.
[35] D.R. Kowalski, On selection problem in radio networks, in: PODC '05, ACM Press, 2005, pp. 158–166.
[36] H. Krishnaswamy, G. Zussman, One chip, twice the bandwidth, IEEE Spectr. 53 (7) (2016) 38–54.
[37] F.J. MacWilliams, N.J.A. Sloane, The Theory of Error-Correcting Codes, North Holland Publishing Co., 1977.
[38] R.A. Moser, G. Tardos, A constructive proof of the general Lovász local lemma, J. ACM 57 (2) (2010) 11–15.
[39] N. Pappas, M. Kountouris, A. Ephremides, A. Traganitis, Relay-assisted multiple access with full-duplex multi-packet reception, IEEE Trans. Wirel. Commun. 14 (7) (2015) 3544–3558.
[40] M. Ruszinkó, On the upper bound of the size of the r-cover-free families, J. Comb. Theory, Ser. A 66 (2) (1994) 302–310.
[41] C. Shangguan, G. Ge, New bounds on the number of tests for disjunct matrices, IEEE Trans. Inf. Theory 62 (12) (2016) 7518–7521.
[42] W. Shi, Q.-S. Hua, D. Yu, Y. Wang, F.C.M. Lau, Efficient information exchange in single-hop multi-channel radio networks, in: X. Wang, R. Zheng, T. Jing, K. Xingvol (Eds.), 7th International Conference on Wireless Algorithms, Systems, and Applications, WASA 2012, in: Lecture Notes in Computer Science, vol. 7405, 2012, pp. 438–449.
[43] J.N. Staddon, D.R. Stinson, R. Wei, Combinatorial properties of frameproof and traceability codes, IEEE Trans. Inf. Theory 47 (3) (2001) 1042–1049.
[44] D.R. Stinson, R. Wei, K. Chen, On generalized separating hash families, J. Comb. Theory, Ser. A 115 (1) (2008) 105–120.
[45] D.R. Stinson, G.M. Zaverucha, Some improved bounds for secure frameproof codes and related separating hash families, IEEE Trans. Inf. Theory 54 (6) (2008) 2508–2514.
[46] D.R. Stinson, R. Wei, L. Zhu, New constructions for perfect hash families and related structures using combinatorial designs and codes, J. Comb. Des. 8 (3) (2000) 189–200.
[47] Y. Yan, D. Yu, Y. Wang, J. Yu, F.C. Lau, Bounded information dissemination in multi-channel wireless networks, J. Comb. Optim. 31 (3) (2016) 996–1012.
[48] D. Yu, Y. Wang, Y. Yan, J. Yu, F.C. Lau, Speedup of information exchange using multiple channels in wireless ad hoc networks, in: 2015 IEEE Conference on Computer Communication, INFOCOM, 2015, pp. 2029–20137.
[49] Y. Wang, Y. Wang, D. Yu, J. Yu, F.C. Lau, Information exchange with collision detection on multiple channels, J. Comb. Optim. 31 (2016) 118–135.