Des. Codes Cryptogr. (2017) 82:231–247
DOI 10.1007/s10623-016-0279-3

Almost cover-free codes and designs

Arkadii D'yachkov[1] · Ilya Vorobyev[1] · Nikita Polyanskii[1] · Vladislav Shchukin[1]

Received: 20 September 2015 / Revised: 5 August 2016 / Accepted: 31 August 2016 /
Published online: 16 September 2016

Abstract An s-subset of codewords of a binary code X is said to be (s, ℓ)-*bad* in X if the code X contains a subset of ℓ other codewords such that the conjunction of the ℓ codewords is covered by the disjunctive sum of the s codewords. Otherwise, the s-subset of codewords of X is called (s, ℓ)-*good* in X. A binary code X is said to be a cover-free (CF) (s, ℓ)-code if the code X does not contain (s, ℓ)-bad subsets. In this paper, we introduce a natural *probabilistic* generalization of CF (s, ℓ)-codes, namely: a binary code X is said to be an almost CF (s, ℓ)-code if the relative number of its (s, ℓ)-good s-subsets is close to 1. We develop a random coding method based on the ensemble of binary constant weight codes to obtain lower bounds on the capacity of such codes. Our main result shows that the capacity for almost CF (s, ℓ)-codes is essentially greater than the rate for ordinary CF (s, ℓ)-codes.

Keywords Almost cover-free codes · Designs · Capacity · Random coding bound

Mathematics Subject Classification 94B25 · 94B65

This is one of several papers published in *Designs, Codes and Cryptography* comprising the "Special Issue on Coding and Cryptography".

The material in this work was presented in part at the 2015 IEEE International Symposium on Information Theory. This paper is the full paper of [8] and provides significant technical contributions over [8], e.g., proofs of all lemmas. Refer to Sect. 4.

✉ Nikita Polyanskii
nikitapolyansky@gmail.com

Arkadii D'yachkov
agd-msu@yandex.ru

[1] Department of Probability Theory, Faculty of Mechanics and Mathematics, Lomonosov Moscow State University, Moscow, Russian Federation 119992

1 Statement of problem and results

1.1 Notations and conventions

In what follows, the symbol $\triangleq$ denotes definitional equalities. For any positive integer n put $[n] \triangleq \{1, 2, \ldots, n\}$. Let N and t be positive integers. The standard symbol $\lfloor a \rfloor$ ($\lceil a \rceil$) will be used to denote the largest (least) integer $\leq a$ ($\geq a$). Introduce a binary $N \times t$ matrix $X = \|x_i(j)\|$ having N rows $\boldsymbol{x}_i \triangleq \big(x_i(1), x_i(2), \ldots, x_i(t)\big)$, $i \in [N]$, and t columns $\boldsymbol{x}(j) \triangleq \big(x_1(j), x_2(j), \ldots, x_N(j)\big)$, $j \in [t]$. Any such matrix X is called a *binary code of length N and size $t = \lfloor 2^{RN} \rfloor$* (briefly, an (N, R)-*code*), where a fixed parameter $R > 0$ is called the *rate* of code X. A column $\boldsymbol{x}(j) \in \{0, 1\}^N$ is called a *jth codeword*. The number of 1's in column $x(j)$, i.e., $|\boldsymbol{x}(j)| \triangleq \sum\limits_{i=1}^{N} x_i(j)$, is called the *weight* of $x(j)$, $j \in [t]$.

For binary vectors $\mathbf{u} \triangleq (u_1, \ldots, u_N) \in \{0, 1\}^N$ and $\mathbf{v} \triangleq (v_1, \ldots, v_N) \in \{0, 1\}^N$, we will use the standard notation of component-wise *disjunction* $\mathbf{u} \bigvee \mathbf{v}$ and *conjunction* $\mathbf{u} \bigwedge \mathbf{v}$. We say that $\mathbf{u}$ is *covered* by $\mathbf{v}$ ($\mathbf{v} \succeq \mathbf{u}$) if $\mathbf{u} \bigvee \mathbf{v} = \mathbf{v}$.

1.2 Almost cover-free codes

Let s and ℓ be positive integers such that $s + \ell \leq t$ and $\mathscr{P}_s(t) \triangleq \{\mathscr{S} : \mathscr{S} \subset [t],\ |\mathscr{S}| = s\}$ is the collection of all s-subsets of the set $[t]$. Note that $|\mathscr{P}_s(t)| = \binom{t}{s}$.

Definition 1 Let $X = (\boldsymbol{x}(1), \boldsymbol{x}(2), \ldots, \boldsymbol{x}(t))$ be an arbitrary binary code of length N and size t. A set $\mathscr{S} \in \mathscr{P}_s(t)$ is said to be (s, ℓ)-*bad* in X if there exists a set $\mathscr{L}$, $\mathscr{L} \subset [t] \setminus \mathscr{S}$ of size $|\mathscr{L}| = \ell$ such that

$$\bigvee_{j \in \mathscr{S}} \boldsymbol{x}(j) \ \succeq \ \bigwedge_{j \in \mathscr{L}} \boldsymbol{x}(j).$$

Otherwise, the set $\mathscr{S} \in \mathscr{P}_s(t)$ is called an (s, ℓ)-*good* set in X.

Let the symbol $\mathbf{B}(s, \ell, X)$ ($\mathbf{G}(s, \ell, X)$) denote the collection of all (s, ℓ)-bad ((s, ℓ)-good) sets $\mathscr{S} \in \mathscr{P}_s(t)$ in X. Obviously, $|\mathbf{B}(s, \ell, X)| + |\mathbf{G}(s, \ell, X)| = \binom{t}{s}$.

Proposition 1 *For $s \geq 2$ and $\ell \geq 1$, any $(s, \ell+1)$-good ((s, ℓ)-bad) set in X is (s, ℓ)-good ($(s, \ell+1)$-bad) set in X, i.e., the following inclusions hold:* $\mathbf{B}(s, \ell, X) \subset \mathbf{B}(s, \ell+1, X)$ *and* $\mathbf{G}(s, \ell+1, X) \subset \mathbf{G}(s, \ell, X)$.

Definition 2 Let ε, $0 \leq \varepsilon \leq 1$, be a fixed parameter. A code X is said to be an *almost cover-free* (s, ℓ)-*code* of *error probability* ε or, briefly, *CF* (s, ℓ, ε)-*code* if

$$\frac{|\mathbf{B}(s, \ell, X)|}{\binom{t}{s}} \leq \varepsilon \quad \Longleftrightarrow \quad |\mathbf{G}(s, \ell, X)| \geq (1 - \varepsilon) \binom{t}{s}. \tag{1}$$

Example 1 Consider 5×5 code X:

$$X = \begin{vmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{vmatrix} \tag{2}$$

Then $\mathbf{G}(2, 2, X) = \big\{\{1; 2\},\ \{1; 3\},\ \{1; 4\},\ \{1; 5\},\ \{2; 3\}\big\}$ and X is a CF $(2, 2, \frac{1}{2})$-code.

From Definition 2 and Proposition 1 it follows

Proposition 2 *Any CF* $(s, \ell+1, \varepsilon)$*-code is a CF* (s, ℓ, ε)*-code.*

Monotonicity with respect to parameter s is provided by

Proposition 3 *If* X *is a CF* (s, ℓ, ε)*-code of size* t *and length* N*, then there exists a CF* $(s-1, \ell, \varepsilon)$*-code* X' *of size* $t-1$ *and length* N*.*

Proof (Proposition 3) Let

$$\mathbf{B}(s, \ell, X, i) \triangleq \{\mathscr{S} : i \in \mathscr{S} \in \mathbf{B}(s, \ell, X)\}$$

denote the collection of all (s, ℓ)-bad sets $\mathscr{S}$ in X containing the element $i \in [t]$. Note that the cardinalities $|\mathbf{B}(s, \ell, X, i)|$, $0 \le |\mathbf{B}(s, \ell, X, i)| \le \binom{t-1}{s-1}$, $i \in [t]$, satisfy the equality:

$$\sum_{i=1}^{t} |\mathbf{B}(s, \ell, X, i)| = s \cdot |\mathbf{B}(s, \ell, X)| \le s \binom{t}{s} \varepsilon,$$

where the last inequality follows from (1). This means that there exists $j \in [t]$, such that $|\mathbf{B}(s, \ell, X, j)| \le \binom{t-1}{s-1} \varepsilon$. Now we check that the code X' obtained from X by deleting the column $\boldsymbol{x}(j)$ is a CF $(s-1, \ell, \varepsilon)$-code of size $t-1$ and length N. Suppose the number of $(s-1, \ell)$-bad sets in X' exceeds $\binom{t-1}{s-1} \varepsilon$, and assume for simplicity that $j = t$. If a set $S' \subset [t-1]$ is $(s-1, \ell)$-bad in X', then the union $S' \cup \{t\}$ is an (s, ℓ)-bad set in X and belongs to $\mathbf{B}(s, \ell, X, t)$. We derive a contradiction since $|\mathbf{B}(s, \ell, X, t)| \le \binom{t-1}{s-1} \varepsilon$. □

For $\varepsilon = 0$, the concept of CF (s, ℓ, ε)-code can be considered as a natural probabilistic generalization of the combinatorial concept of *CF* (s, ℓ)*-code* that is defined in [4,9] as the *incidence matrix* of a *family of finite sets in which no intersection of* ℓ *sets is covered by the union of* s *others*. For the case $\ell = 1$, CF codes and their applications were introduced in [11]. For $\ell \ge 2$, CF (s, ℓ)-codes along with their applications to *key distribution patterns* were first suggested in [13].

Let $t(N, s, \ell)$ be the maximal size of CF (s, ℓ)-codes of length N and let $N(t, s, \ell)$ be the minimal length of CF (s, ℓ)-codes of size t. Then the number

$$R(s, \ell) \triangleq \varlimsup_{N \to \infty} \frac{\log_2 t(N, s, \ell)}{N} = \varlimsup_{t \to \infty} \frac{\log_2 t}{N(t, s, \ell)} \tag{3}$$

is called [4] the *rate* of CF (s, ℓ)-codes. In the recent papers [5,6], one can find a detailed survey of the best known lower and upper bounds on the rate $R(s, \ell)$.

Let us introduce the following definition which uses standard information-theoretic terminology [2], introduce

Definition 3 Let R, $R > 0$, be a fixed parameter. Taking into account inequality (1), define the *error* for almost CF (s, ℓ)-codes:

$$\varepsilon(s, \ell, R, N) \triangleq \min_{X : t = \lfloor 2^{RN} \rfloor} \left\{ \frac{|\mathbf{B}(s, \ell, X)|}{\binom{t}{s}} \right\}, \tag{4}$$

where the minimum is taken over all (N, R)-codes X. The function

$$\mathbf{E}(s, \ell, R) \triangleq \varlimsup_{N \to \infty} \frac{-\log_2 \varepsilon(s, \ell, R, N)}{N}, \tag{5}$$

is said to be the error *exponent* for almost CF (s, ℓ)-codes, the number

$$C(s, \ell) \triangleq \sup\{R : \mathbf{E}(s, \ell, R) > 0\} \tag{6}$$

is said to be the *capacity* for almost CF (s, ℓ)-codes and rate $R(s, \ell)$ defined by (3) is called the *zero-error* capacity for almost CF (s, ℓ)-codes.

In other words, the capacity is the supremum over all rates $R > 0$ such that there is a sequence of (N_n, R)-codes $\{X_n\}$ with $N_n \to \infty$ and the fraction of (s, ℓ)-bad sets in X_n exponentially converges to zero with respect to N_n. As for the zero-error capacity, we require the property that the corresponding fraction of (s, ℓ)-bad sets is zero for all codes X_n. The issue is, does such relaxation allow to improve the parameters of CF codes.

For the particular case $\ell = 1$, Definitions 1–3 were suggested in our paper [7], in which we introduce the concept of almost disjunctive list-decoding codes. The best presently known constructions of such codes were proposed in [3]. Bounds on the rate for these constructions were computed in the recent paper [1].

Definitions 1–3 and Proposition 1–3 lead to

Theorem 1 (Monotonicity properties) *For* $s \geq 1,\ \ell \geq 2,\ R > 0$, *the following inequalities hold*

$$\begin{aligned} R(s+1, \ell) &\leq R(s, \ell) \leq R(s, \ell - 1), \\ C(s+1, \ell) &\leq C(s, \ell) \leq C(s, \ell - 1), \\ \mathbf{E}(s+1, \ell, R) &\leq \mathbf{E}(s, \ell, R) \leq \mathbf{E}(s, \ell - 1, R). \end{aligned}$$

1.3 Almost cover-free designs

By $\hat{\mathscr{P}}_s(\ell, t)$ denote the collection of *supersets* p, $\mathsf{p} \triangleq (P_1, P_2, \ldots, P_s)$, $P_i \subset \mathscr{P}_\ell(t), i \in [s]$, where each p consists of s disjoint sets $P \subset [t]$ of size $|P| = \ell$, i.e.:

$$\hat{\mathscr{P}}_s(\ell, t) \triangleq \left\{ \begin{array}{c} \mathsf{p} = (P_1, P_2, \ldots, P_s) : \\ P_i \subset [t],\ |P_i| = \ell, \\ P_i \cap P_j = \varnothing \text{ for } i \neq j, i, j \in [s], \end{array} \right\}.$$

Example 2 Let $t = 7$, $s = 2$ and $\ell = 2$. Then $(\{1; 2\}, \{5; 6\})$ belongs to $\hat{\mathscr{P}}_2(2, 7)$, while $(\{1; 2; 3\}, \{4; 7\})$ is not included to $\hat{\mathscr{P}}_2(2, 7)$.

Obviously, the collection $\hat{\mathscr{P}}_s(\ell, t)$ has cardinality

$$|\hat{\mathscr{P}}_s(\ell, t)| = \frac{1}{s!}\binom{t}{s\ell}\binom{s\ell}{(s-1)\ell}\cdots\binom{2\ell}{\ell}. \tag{7}$$

For a superset $\mathsf{p} \in \hat{\mathscr{P}}_s(\ell, t)$ and a code X, introduce the binary vector $\mathbf{r}(\mathsf{p}, X) \triangleq (r_1, r_2, \ldots, r_N) \in \{0, 1\}^N$:

$$\mathbf{r}(\mathsf{p}, X) \triangleq \bigvee_{P \in \mathsf{p}} \bigwedge_{j \in P} \boldsymbol{x}(j). \tag{8}$$

One can see that the ith component of $\mathbf{r}(\mathsf{p}, X)$ can be written in the form:

$$r_i = \begin{cases} 1, & \text{if } \exists\, P \in \mathsf{p} \text{ such that } x_i(j) = 1 \text{ for all } j \in P, \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

Definition 4 Let $X = (\boldsymbol{x}(1), \boldsymbol{x}(2), \ldots, \boldsymbol{x}(t))$ be an arbitrary binary code of length N and size t. A superset p, $\mathsf{p} \in \hat{P}_s(\ell, t)$, is said to be an (s, ℓ)-*bad* superset in X, if there exists another superset $\mathsf{p}' \in \hat{P}_s(\ell, t)$, $\mathsf{p} \neq \mathsf{p}'$, such that $\mathbf{r}(\mathsf{p}, X) = \mathbf{r}(\mathsf{p}', X)$. Otherwise, the superset p is said to be an (s, ℓ)-*good* superset in X.

Let $\hat{\mathbf{B}}(s, \ell, X)$ and $\hat{\mathbf{G}}(s, \ell, X)$ denote the collection of all (s, ℓ)-bad and (s, ℓ)-good supersets, respectively. Obviously, $|\hat{\mathbf{B}}(s, \ell, X)| + |\hat{\mathbf{G}}(s, \ell, X)| = |\hat{\mathscr{P}}_s(\ell, t)|$.

Definition 5 Let ε, $0 \leq \varepsilon \leq 1$, be a fixed parameter. A code X is said to be an *almost cover-free* (s, ℓ)-*design* of *error probability* ε or, briefly, CF (s, ℓ, ε)-design if

$$\frac{|\hat{\mathbf{B}}(s, \ell, X)|}{|\hat{\mathscr{P}}_s(\ell, t)|} \leq \varepsilon \iff |\hat{\mathbf{G}}(s, \ell, X)| \geq (1 - \varepsilon)\, |\hat{\mathscr{P}}_s(\ell, t)|. \tag{10}$$

Example 3 For the code X described in Example 1, the collection of $(2, 2)$-bad supersets

$$\hat{\mathbf{B}}(s, \ell, X) = \big\{(\{1; 2\}, \{4; 5\}),\ (\{1; 3\}, \{4; 5\}),\ (\{1; 4\}, \{2; 3\}),\ (\{1; 5\}, \{2; 3\})\big\}.$$

It follows that X is a CF $(2, 2, \frac{4}{15})$-design.

Definition 6 Let R, $R > 0$, be a fixed parameter. Taking into account inequality (10) define the *error* for almost CF (s, ℓ)-designs:

$$\hat{\varepsilon}(s, \ell, R, N) \triangleq \min_{X : t = \lceil 2^{RN} \rceil} \left\{ \frac{|\hat{\mathbf{B}}(s, \ell, X)|}{|\hat{\mathscr{P}}_s(\ell, t)|} \right\}, \tag{11}$$

where the minimum is taken over all (N, R)-codes X. The function

$$\hat{\mathbf{E}}(s, \ell, R) \triangleq \varlimsup_{N \to \infty} \frac{-\log_2 \hat{\varepsilon}(s, \ell, R, N)}{N}, \tag{12}$$

is said to be the error *exponent* for almost CF (s, ℓ)-designs, the number

$$\hat{C}(s, \ell) \triangleq \sup\{R : \hat{\mathbf{E}}(s, \ell, R) > 0\}$$

is called the *capacity* for almost CF (s, ℓ)-designs.

For the particular case $\ell = 1$, Definitions 4–6 were already introduced in [12] to describe the model called *planning screening experiments*. In [12], it was proved that the capacity of almost CF $(s, 1)$-designs $\hat{C}(s, 1) = 1/s$. One can see that Definitions 4–6 represent a natural generalization of almost CF $(s, 1)$-designs. We conjecture that the capacity $\hat{C}(s, \ell) = 1/(s\,\ell)$.

In Sect. 2, we establish

Theorem 2 *The capacities and the error exponents satisfy the inequality*

$$C(s, \ell) \leq \hat{C}(s, \ell) \leq 1/(s\,\ell), \quad \mathbf{E}(s, \ell, R) \leq \hat{\mathbf{E}}(s, \ell, R).$$

During the proof of Theorem 2 we show that a CF (s, ℓ, ε)-code is a CF $(s, \ell, \hat{\varepsilon})$-design such that $\hat{\varepsilon} \leq f(s, \ell)\varepsilon$, where $f(s, \ell)$ is an independent function of length of the code.

However, in spite of the greater capacity, *using* of CF (s, ℓ, ε)-designs for the superset identification problem formally defined in Sect. 1.5 is *practically unacceptable*, since it requires much greater complexity, which is evidently equal to the complexity of exhaustive search. For fixed s, ℓ and $t \to \infty$, the order of the complexity is $|\hat{\mathscr{P}}_s(\ell, t)| \sim t^{s\,\ell}$. It will be shown in Sect. 1.5 that CF (s, ℓ, ε)-codes are efficient CF (s, ℓ, ε)-designs and for such codes the algorithm of identification supersets $\mathsf{p} \in \hat{\mathscr{P}}_s(\ell, t)$, is essentially faster than the trivial one, and its complexity is proportional to t^{ℓ}.

1.4 Lower bounds on $R(s, \ell)$, $C(s, \ell)$

The best presently known upper and lower bounds on the rate $R(s, \ell)$ of cover-free (s, ℓ)-codes were presented in [5,6]. If $\ell \geq 1$ is fixed and $s \to \infty$, then these bounds have the following asymptotic form:

$$R(s, \ell) \geq \frac{(\ell+1)^{\ell+1}}{e^{\ell+1}} \frac{\log_2 s}{s^{\ell+1}} (1 + o(1)),$$
$$R(s, \ell) \leq \frac{(\ell+1)^{\ell+1}}{2e^{\ell-1}} \frac{\log_2 s}{s^{\ell+1}} (1 + o(1)). \tag{13}$$

In the present paper, we suggest a modification of the random coding method developed in [5] and [7], which permits us to obtain a lower bound on the capacity $C(s, \ell)$. Let $h(x)$

$$h(x) \triangleq -x \log_2 x - (1-x) \log_2 (1-x), \ 0 < x < 1,$$

denote the binary entropy function. In Sect. 3, we prove

Theorem 3 (Random coding lower bound $\underline{C}(s, \ell)$)

(1) *For $\ell \geq 2$, the capacity for almost CF codes satisfies the inequality*

$$C(s, \ell) \geq \underline{C}(s, \ell) \triangleq \frac{1}{\ell} \max_{0 \leq Q \leq 1} \mathscr{D}(\ell, Q, \hat{q}), \tag{14}$$

where the function $\mathscr{D}(\ell, Q, \hat{q})$ is defined in the parametric form

$$\mathscr{D}(\ell, Q, \hat{q}) \triangleq \ell h(Q) + (1-Q)\ell \log_2 z - (1-\hat{q}) \log_2 (1-(1-z)^{\ell}) + \ell \left(\frac{(1-Q)}{z}(1-z) - \left(\frac{(1-Q)}{z} - \hat{q} \right) (1-z)^{\ell} \right) \log_2 (1-z), \tag{15}$$

and parameters z and $\hat{q}$ are uniquely determined by the following equations

$$Q = \frac{(1-z)(1-(1-z)^{\ell}) - (1-\hat{q}) z (1-z)^{\ell}}{1-(1-z)^{\ell}}, \quad \hat{q} = 1 - (1-Q)^{s}. \tag{16}$$

(2) *For $\ell \geq 2$ and $s \to \infty$, the lower asymptotic bound on $C(s, \ell)$ is*

$$C(s, \ell) \geq \frac{\log_2 e}{s^{\ell}} \cdot \frac{\ell^{\ell-1}}{e^{\ell}} (1 + o(1)). \tag{17}$$

1.5 Boolean model for nonadaptive search of supersets

Definition 7 [4]. A binary $N \times t$ matrix X is called a *cover-free (s, ℓ)-design* or, briefly, *CF (s, ℓ)-design* if for any $\mathsf{p}', \mathsf{p}'' \in \hat{\mathscr{P}}_s(\ell, t)$, $\mathsf{p}' \neq \mathsf{p}''$, the vectors $\mathbf{r}(\mathsf{p}', X) \neq \mathbf{r}(\mathsf{p}'', X)$.

Remind a well known application of CF codes. Suppose a set of t samples is given. We identify it with the set $[t]$. In the present paper we consider a generalization of *the Boolean search model for sets* [11] which is called *the Boolean search model for supersets* [4]. Assume that a *positive superset* $\mathsf{p} \in \hat{\mathscr{P}}_s(\ell, t)$ is fixed. Our aim is to detect it using the minimal number of group tests, where *each test checks whether a testing group contains at least one set $P \in \mathsf{p}$ or not*. Now assume that we use N tests. They can be encoded by a code $X = \|x_i(j)\|$. A column $\boldsymbol{x}(j)$ corresponds to the jth sample, and a row $\boldsymbol{x}_i$ corresponds to the ith test. We put $x_i(j) \triangleq 1$ iff the jth sample is included into the ith testing group. Otherwise, $x_i(j) \triangleq 0$.

The *outcomes* (9) of all N tests form the binary vector $\mathbf{r}(\mathsf{p}, X)$ (8), where $\mathsf{p} \in \hat{\mathscr{P}}_s(\ell, t)$ is the (*unknown*) positive superset. Thus, the code X should be designed in such a way that we should be able to detect a superset p given the vector $\mathbf{r}(\mathsf{p}, X)$. Obviously, it is possible if and only if X is a CF (s, ℓ)-design. Note that we deal with the *nonadaptive* search model arised from the needs of molecular biology and first suggested in [14].

Let X be a binary $N \times t$ matrix and $\mathsf{p}^{(\mathrm{un})} \in \hat{\mathscr{P}}_s(\ell, t)$ be an *unknown* superset. Any fixed set $P' \subset [t]$, $|P'| \le \ell$, is called *acceptable* for the *known* vector $\mathbf{r}^{(\mathrm{kn})} \triangleq \mathbf{r}(\mathsf{p}^{(\mathrm{un})}, X)$ if the conjunction $\bigwedge\limits_{j \in P'} \boldsymbol{x}(j)$ is covered by $\mathbf{r}^{(\mathrm{kn})}$. In the given model, an effective decoding algorithm is based on the following

Proposition 4 [4] *If X is an CF (s, ℓ)-code, then any superset $\mathsf{p} \in \hat{\mathscr{P}}_s(\ell, t)$ is composed of all acceptable sets for $\mathbf{r}^{(\mathrm{kn})}$. For fixed s, ℓ and $t \to \infty$, this means that the decoding complexity is proportional to $\binom{t}{\ell} \sim t^{\ell}$ and does not depend on s.*

Note that in the general case of CF (s, ℓ)-design and the trivial decoding algorithm, we need to check all possible supersets $\mathsf{p} \in \hat{\mathscr{P}}_s(\ell, t)$. If s and ℓ are fixed and $t \to \infty$, then the number of such comparisons (decoding complexity) is proportional to $|\hat{\mathscr{P}}_s(\ell, t)| \sim t^{s\ell}$. Thus, CF (s, ℓ)-codes form a class of CF (s, ℓ)-designs for which the decoding algorithm based on Proposition 4 is strongly better than the trivial one.

Let $\ell \ge 1$ be fixed and $s \to \infty$. Taking into account (13), we conclude that for sufficiently large t the *use of optimal CF (s, ℓ)-codes* gives the bounds on $\log_2 t/N$:

$$\log_2 t/N \le \frac{\log_2 s}{s^{\ell+1}} \cdot \frac{(\ell+1)^{\ell+1}}{2e^{\ell-1}}\,(1 + o(1)),$$

$$\log_2 t/N \ge \frac{\log_2 s}{s^{\ell+1}} \cdot \frac{(\ell+1)^{\ell+1}}{e^{\ell+1}}\,(1 + o(1)).$$

The capacity $C(s, \ell)$ can be interpreted as the theoretical tightest upper bound on the information rate $\log_2 t/N$ of CF (s, ℓ, ε)-codes with error probability $\varepsilon \to 0$. Therefore, the bound (17) means that for $\ell \ge 2, s \to \infty$ and sufficiently large t, *using of optimal CF (s, ℓ, ε)-codes* guarantees the inequality:

$$\log_2 t/N \ge \frac{\log_2 e}{s^{\ell}} \cdot \frac{\ell^{\ell-1}}{e^{\ell}}(1 + o(1)).$$

2 Proof of Theorem 2

Proof For any superset $\mathsf{p} \in \hat{P}_s(\ell, t)$, $\mathsf{p} = \{P_1, P_2, \ldots, P_s\}$, define the collection $T(\mathsf{p})$:

$$T(\mathsf{p}) \triangleq \left\{ \mathscr{S} \in \mathscr{P}_s(t) : \quad \mathscr{S} = \{a_1, a_2, \ldots, a_s\}, \quad a_i \in P_i \in \mathsf{p},\ i \in [s] \right\}.$$

One can check that $|T(\mathsf{p})| = \ell^s$. Observe that if all sets $\mathscr{S} \in T(\mathsf{p})$ are (s, ℓ)-good in X, then the superset p is also a (s, ℓ)-good superset in X.

Assume that a code X is a CF (s, ℓ, ε)-code. It means that the number (1) of bad (s, ℓ)-sets doesn't exceed $\varepsilon \cdot \binom{t}{s}$. Given a bad (s, ℓ)-set $B \in \mathscr{P}_s(t)$ for the code X, one can check that the number of $\mathsf{p} \in \hat{P}_s(\ell, t)$ such that $B \in T(\mathsf{p})$ is at most $\binom{t-s}{s(\ell-1)}\binom{s(\ell-1)}{(s-1)(\ell-1)} \cdots \binom{2(\ell-1)}{\ell-1}$. This implies that the number of bad (s, ℓ)-supersets is at most $\varepsilon \cdot \binom{t}{s}\binom{t-s}{s(\ell-1)}\binom{s(\ell-1)}{(s-1)(\ell-1)} \cdots \binom{2(\ell-1)}{\ell-1}$ or $\varepsilon \cdot \ell^s \cdot |\hat{\mathscr{P}}_s(\ell, t)|$, where the cardinality $|\hat{\mathscr{P}}_s(\ell, t)|$ is computed in (7). Thus, X is also a

CF $(s, \ell, \varepsilon \cdot \ell^s)$-design. In other words, we proved the relations $C(s, \ell) \le \hat{C}(s, \ell)$ and $\mathbf{E}(s, \ell, R) \le \hat{\mathbf{E}}(s, \ell, R)$.

Now, fix $R > 0$ and $\varepsilon > 0$ and suppose that the code X is a CF (s, ℓ, ε)-design of length N and size $t \triangleq \lfloor 2^{RN} \rfloor$. Observe that for any two different good (see Def. 4) supersets $\mathsf{p}, \mathsf{p}' \in \hat{\mathbf{G}}(s, \ell, X)$, $\mathsf{p} \ne \mathsf{p}'$, two vectors $\mathbf{r}(\mathsf{p}, X)$ and $\mathbf{r}(\mathsf{p}', X)$ defined by (8) are distinct, i.e., $\mathbf{r}(\mathsf{p}, X) \ne \mathbf{r}(\mathsf{p}, X)$. From the definition (10) of CF (s, ℓ, ε)-design, we get

$$(1-\varepsilon)|\hat{\mathscr{P}}_s(\ell, t)| = (1-\varepsilon)\frac{1}{s!}\binom{t}{s\ell}\binom{s\ell}{(s-1)\ell}\cdots\binom{2\ell}{\ell} \le 2^N. \tag{18}$$

The comparison of the left and right-hand sides of (18) and the equality $t = \lfloor 2^{RN} \rfloor$ lead to the bound

$$\hat{\varepsilon}(s, \ell, R, N) \ge 1 - 2^N \cdot \left(|\hat{\mathscr{P}}_s(\ell, t)|\right)^{-1} = 1 - 2^{-N\lfloor (s\ell \cdot R - 1) + o(1) \rfloor}, \quad N \to \infty.$$

This inequality means that the condition $R < 1/(s\ell)$ is necessary for $\hat{\mathbf{E}}(s, \ell, R) > 0$. In other words, we prove that $\hat{C}(s, \ell) \le \frac{1}{s\ell}$. □

3 Proof of Theorem 3

Proof (Statement 1) For a code X, the number $|\mathbf{B}(s, \ell, X)|$ of (s, ℓ)-bad sets in the code X can be represented in the form:

$$|\mathbf{B}(s, \ell, X)| \triangleq \sum_{\mathscr{S} \in \mathscr{P}_s(t)} \psi(X, \mathscr{S}), \tag{19}$$

where

$$\psi(X, \mathscr{S}) \triangleq \begin{cases} 1 & \text{if the set } \mathscr{S} \in \mathbf{B}(s, \ell, X), \\ 0 & \text{otherwise.} \end{cases}$$

Let $Q, 0 < Q < 1$, and $R, 0 < R < 1$, be fixed parameters. Define the ensemble $\{N, t, Q\}$ of binary $(N \times t)$-matrices $X = (\boldsymbol{x}(1), \boldsymbol{x}(2), \ldots \boldsymbol{x}(t))$, where columns $\boldsymbol{x}(i)$, $i \in [t]$, $t \triangleq \lfloor 2^{RN} \rfloor$, are chosen independently and equiprobably from the set consisting of $\binom{N}{\lfloor QN \rfloor}$ columns of the fixed weight $\lfloor QN \rfloor$. Fix two subsets $\mathscr{S}, \mathscr{L} \subset [t]$ such that $|\mathscr{S}| = s$, $|\mathscr{L}| = \ell$ and $\mathscr{S} \cap \mathscr{L} = \varnothing$. From (19) it follows that for $\{N, t, Q\}$, the expectation $\overline{|\mathbf{B}(s, \ell, X)|}$ of the number $|\mathbf{B}(s, \ell, X)|$ is

$$\overline{|\mathbf{B}(s, \ell, X)|} = |\mathscr{P}_s(t)| \Pr\{\mathscr{S} \in \mathbf{B}(s, \ell, X)\}.$$

Thus, the expectation of the error probability for almost CF (s, ℓ)-codes is

$$\mathscr{E}^{(N)}(s, \ell, R, Q) \triangleq |\mathscr{P}_s(t)|^{-1}\, \overline{|\mathbf{B}(s, \ell, X)|} = \Pr\{\mathscr{S} \in \mathbf{B}(s, \ell, X)\}, \tag{20}$$

where $t = \lfloor 2^{RN} \rfloor$. The evident *random coding upper bound* on the error probability (4) for cover-free (s, ℓ)-codes is formulated for any $0 < Q < 1$ as the following inequality:

$$\varepsilon(s, \ell, R, N) \triangleq \min_{X: t = \lfloor 2^{RN} \rfloor} \left\{ \frac{|\mathbf{B}(s, \ell, X)|}{|\mathscr{P}_s(t)|} \right\} \le \mathscr{E}^{(N)}(s, \ell, R, Q). \tag{21}$$

The expectation $\mathscr{E}^{(N)}(s, \ell, R, Q)$ defined by (20) can be represented as follows

$$\mathscr{E}^{(N)}(s, \ell, R, Q) = \mathscr{P}_2^{(N)}(s, Q, k) \cdot \sum_{k=\lfloor QN \rfloor}^{\min\{N, s\lfloor QN \rfloor\}} \Pr\left\{ \mathscr{S} \in \mathbf{B}(s, \ell, X) \Big/ \left| \bigvee_{i \in \mathscr{S}} \boldsymbol{x}(i) \right| = k \right\}$$

$$\leq \sum_{k=\lfloor QN \rfloor}^{\min\{N, s\lfloor QN \rfloor\}} \mathscr{P}_2^{(N)}(s, Q, k) \cdot \min\left\{1; \binom{t}{\ell} \mathscr{P}_1^{(N)}(\ell, Q, k)\right\}, \tag{22}$$

where we apply the total probability formula and the standard union bound for the conditional probability:

$$\Pr\left\{ \bigcup_i C_i \Big/ C \right\} \leq \min\left\{1; \sum_i \Pr\{C_i / C\}\right\}$$

and for $\lfloor QN \rfloor \leq k \leq \min\{N, s\lfloor QN \rfloor\}$, introduce the notations for conditional probability

$$\mathscr{P}_1^{(N)}(\ell, Q, k) \triangleq \Pr\left\{ \bigvee_{i \in \mathscr{S}} \boldsymbol{x}(i) \succeq \bigwedge_{j \in \mathscr{L}} \boldsymbol{x}(j) \Big/ \left| \bigvee_{i \in \mathscr{S}} \boldsymbol{x}(i) \right| = k \right\} \tag{23}$$

and for probability of the event "the weight of disjunctive sum of s distinct codewords in the ensemble $\{N, t, Q\}$ equals k":

$$\mathscr{P}_2^{(N)}(s, Q, k) \triangleq \Pr\left\{ \left| \bigvee_{i \in \mathscr{S}} \boldsymbol{x}(i) \right| = k \right\}. \tag{24}$$

Let $k \triangleq \lfloor qN \rfloor$ and the functions

$$\mathscr{D}(\ell, Q, q) \triangleq \lim_{N \to \infty} \frac{-\log_2\left(\mathscr{P}_1^{(N)}(\ell, Q, k)\right)}{N}, \tag{25}$$

$$\mathscr{A}(s, Q, q) \triangleq \lim_{N \to \infty} \frac{-\log_2\left(\mathscr{P}_2^{(N)}(s, Q, k)\right)}{N} \tag{26}$$

denote the exponents of the logarithmic asymptotic behavior for the probabilities (23) and (24) in the ensemble $\{N, t, Q\}$, respectively. Define $\hat{q} \triangleq 1 - (1 - Q)^s$. Now we formulate two auxiliary lemmas, the proof of which can be found in Sect. 4. □

Lemma 1 *The function $\mathscr{A}(s, Q, q)$ of the parameter q, $Q < q < \min\{1, sQ\}$, defined by (26) can be represented in the parametric form*

$$\mathscr{A}(s, Q, q) \triangleq (1 - q)\log_2(1 - q) + q \log_2\left(\frac{Qy^s}{1 - y}\right)$$

$$+ sQ \log_2\left(\frac{1 - y}{y}\right) + s \cdot h(Q), \tag{27}$$

$$q = Q\frac{1 - y^s}{1 - y}, \qquad 0 < y < 1. \tag{28}$$

In addition, $\mathscr{A}(s, Q, q)$ as a function of q attains its unique minimal value which is equal to 0 at $q = \hat{q} \triangleq 1 - (1 - Q)^s$.

Lemma 2 *For $\ell \geq 2$, the value of the function $\mathscr{D}(\ell, Q, q)$ defined by (25) at point $q = \hat{q}$ can be represented as in (15).*

The inequality (22) and the random coding bound (21) imply that the error probability exponent (12) satisfies the inequality

$$\mathbf{E}(s, \ell, R) \geq \underline{\mathbf{E}}(s, \ell, R) \triangleq \max_{0 \leq Q \leq 1} E(s, \ell, R, Q), \tag{29}$$

$$E(s, \ell, R, Q) \triangleq \min_{Q < q < \min\{1, sQ\}} \left\{ \mathscr{A}(s, Q, q) + [\mathscr{D}(\ell, Q, q) - \ell R]^+ \right\}, \tag{30}$$

where $[x]^+$ denotes the positive part of the real function x. Lemma 1 states that $\mathscr{A}(s, Q, q) > 0$ if $q \neq \hat{q}$. In particular, the condition $q \neq \hat{q}$ implies $E(s, \ell, R, Q) > 0$. Therefore, if $\ell R < \mathscr{D}(\ell, Q, \hat{q})$ then $E(s, \ell, R, Q) > 0$, what, in turn, means (see (6) and (29)) that

$$C(s, \ell) \geq \underline{C}(s, \ell) \triangleq \frac{1}{\ell} \max_{0 \leq Q \leq 1} \mathscr{D}(\ell, Q, \hat{q}).$$

Thus, the lower bound (14) is established.

Proof (Statement 2) Let $\ell \geq 2$ be fixed and $s \to \infty$. Substituting $z = s/(s + \ell)$ in (14)–(16) yields

$$Q = \frac{(1-z)(1-(1-z)^\ell) - (1-\hat{q})z(1-z)^\ell}{1-(1-z)^\ell} = \frac{\ell}{s+\ell} - \frac{\ell^\ell e^{-\ell}}{s^\ell} + O\left(\frac{1}{s^{\ell+1}}\right),$$

$$\hat{q} = 1 - (1-Q)^s = 1 - e^{-\frac{\ell s}{\ell+s} + \Theta\left(\frac{1}{s^{\ell-1}}\right) + \Theta\left(\frac{1}{s}\right)} = 1 - e^{-\ell} + O\left(\frac{1}{s}\right)$$

and

$$\begin{aligned} C(s, \ell) &\geq \frac{1}{\ell} \max_{0 \leq Q \leq 1} \mathscr{D}(\ell, Q, \hat{q}) = \frac{1}{\ell} \max_{0 \leq z \leq 1} \mathscr{D}(\ell, Q(z), \hat{q}(z)) \\ &\geq \frac{1}{\ell} \mathscr{D}(\ell, Q(s/(s+\ell)), \hat{q}(s/(s+\ell))), \end{aligned}$$

where

$$\begin{aligned} \mathscr{D}(\ell, Q, \hat{q}) \triangleq\ & (1-Q)\,\ell \log_2 z - (1-\hat{q}) \log_2 (1-(1-z)^\ell) \\ & + \ell \left(\frac{(1-Q)}{z}(1-z) - \left(\frac{(1-Q)}{z} - \hat{q} \right)(1-z)^\ell \right) \log_2 (1-z) + + \ell h(Q). \end{aligned}$$

Therefore, one can write

$$\begin{aligned} C(s, \ell) \geq\ & \left(\frac{s}{s+\ell} + \frac{\ell^\ell e^{-\ell}}{s^\ell} + O\left(\frac{1}{s^{\ell+1}}\right) \right) \log_2 \left(\frac{s}{s+\ell} \right) \\ & - \left(\frac{e^{-\ell}}{\ell} + O\left(\frac{1}{s}\right) \right) \log_2 \left(1 - \left(\frac{\ell}{s+\ell} \right)^\ell \right) \\ & + \left(1 + O\left(\frac{1}{s^\ell}\right) \right) \frac{\ell}{s+\ell} \log_2 \left(\frac{\ell}{s+\ell} \right) \\ & - \left(e^{-\ell} + O\left(\frac{1}{s}\right) \right) \left(\frac{\ell}{s+\ell} \right)^\ell \log_2 \left(\frac{\ell}{s+\ell} \right) \\ & - \left(\frac{\ell}{s+\ell} - \frac{\ell^\ell e^{-\ell}}{s^\ell} + O\left(\frac{1}{s^{\ell+1}}\right) \right) \log_2 \left(\frac{\ell}{s+\ell} - \frac{\ell^\ell e^{-\ell}}{s^\ell} + O\left(\frac{1}{s^{\ell+1}}\right) \right) \end{aligned}$$

$$-\left(\frac{s}{s+\ell}+\frac{\ell^\ell e^{-\ell}}{s^\ell}+O\left(\frac{1}{s^{\ell+1}}\right)\right)\log_2\left(\frac{s}{s+\ell}+\frac{\ell^\ell e^{-\ell}}{s^\ell}+O\left(\frac{1}{s^{\ell+1}}\right)\right)$$
$$=\frac{\ell^{\ell-1}\log_2 e}{e^\ell\, s^\ell}+O\left(\frac{\log_2 s}{s^{\ell+1}}\right).$$

This completes the proof of Statement 2. □

4 Proofs of Lemmas

Proof (Lemma 1) Let $s \geq 2$, $0 < Q < 1$, $Q < q < \min\{1, sQ\}$ be fixed parameters. Assume also $k \triangleq \lfloor qN \rfloor$ and $N \to \infty$. With the help of the *type* (see [2], [5]) terminology:

$$\{n(\mathbf{a})\}, \quad \mathbf{a} \triangleq (a_1, a_2, \ldots, a_s) \in \{0,1\}^s, \quad 0 \leq n(\mathbf{a}) \leq N, \quad \sum_{\mathbf{a}} n(\mathbf{a}) = N,$$

the probability of event (24) in the ensemble $\{N, t, Q\}$ can be written as follows:

$$\mathscr{P}_2^{(N)}(s, Q, k) = \binom{N}{\lfloor QN \rfloor}^{-s} \cdot \sum_{(32)} \frac{N!}{\prod_{\mathbf{a}} n(\mathbf{a})!}, \tag{31}$$

and the sum in the right-hand side of (31) is taken over all types $\{n(\mathbf{a})\}$ provided that

$$n(\boldsymbol{0}) = N - k, \qquad \sum_{\mathbf{a}:\, a_i = 1} n(\mathbf{a}) = \lfloor QN \rfloor \quad \text{for any } i \in [s]. \tag{32}$$

For every type $\{n(\mathbf{a})\}$ we will consider the corresponding distribution $\tau : \tau(\mathbf{a}) = \frac{n(\mathbf{a})}{N}, \quad \forall\, \mathbf{a} \in \{0,1\}^s$. Applying the Stirling approximation, we obtain the following logarithmic asymptotic behavior of a term in the sum (31):

$$-\log_2 \frac{N!}{\prod_{\mathbf{a}} n(\mathbf{a})!}\binom{N}{\lfloor QN \rfloor}^{-s} = NF(\tau, Q, q)(1 + o(1)),$$

where

$$F(\tau, Q, q) = \sum_{\mathbf{a}} \tau(\mathbf{a}) \log_2 \tau(\mathbf{a}) + sH(Q). \tag{33}$$

Thus, one can reduce the calculation of $\mathscr{A}(s, Q, q)$ defined by (26) to the search of the minimum:

$$\mathscr{A}(s, Q, q) = \min_{\tau \in (48):(49)} F(\tau, Q, q) \triangleq F(\hat{\tau}, Q, q), \tag{34}$$

$$\{\tau : \forall\, \mathbf{a} \quad 0 < \tau(\mathbf{a}) < 1\}, \tag{35}$$

$$\sum_{\mathbf{a}} \tau(\mathbf{a}) = 1, \quad \tau(\boldsymbol{0}) = 1 - q, \quad \sum_{\mathbf{a}:\, a_i = 1} \tau(\mathbf{a}) = Q \quad \forall\, i \in [s], \tag{36}$$

where the restrictions (49) are induced by the definition of type and the properties (32).

To find the minimum (47) and the extremal distribution $\{\hat{\tau}\}$ we use the method of Lagrange multipliers. The Lagrangian is

$$\Lambda \triangleq \sum_{\tau(\mathbf{a})} \tau(\mathbf{a}) \log_2 \tau(\mathbf{a}) + sh(Q) + \lambda_0 (\tau(\boldsymbol{0}) + q - 1) + \sum_{i=1}^{s} \lambda_i \left(\sum_{\mathbf{a}:a_i=1} \tau(\mathbf{a}) - Q \right) + \lambda_{s+1} \left(\sum_{\mathbf{a}} \tau(\mathbf{a}) - 1 \right).$$

Therefore, the necessary conditions for the extremal distribution $\{\hat{\tau}\}$ look similar for any $\mathbf{a} \neq \boldsymbol{0}$, and the condition has the specific form for the case $\mathbf{a} = \boldsymbol{0}$:

$$\begin{cases} \frac{\partial \Lambda}{\partial \tau(\boldsymbol{0})} = \log_2 \hat{\tau}(\boldsymbol{0}) + \log_2 e + \lambda_0 + \lambda_{s+1} = 0, \\ \frac{\partial \Lambda}{\partial \tau(\mathbf{a})} = \log_2 \hat{\tau}(\mathbf{a}) + \log_2 e + \lambda_{s+1} + \sum\limits_{i=1}^{s} a_i \lambda_i = 0 \quad \text{for any } \mathbf{a} \neq \boldsymbol{0}. \end{cases} \tag{37}$$

It turns out that the matrix of second derivatives of the Lagrangian is diagonal and positive definite in the region (48), and the function $F(\tau, Q)$ defined by (33) is strictly $\cup$-convex in the region (48). The Karush–Kuhn–Tacker theorem [10] states that each solution $\tau \in$ (48) satisfying system (37) and constraints (49) gives a local minimum of $F(\tau, Q)$. Thus, if there exists a solution of the system (37) and (49) in the region (48), then it is unique and gives a minimum in the minimization problem (47)–(49).

Note that the symmetry of problem yields the equality $v \triangleq \lambda_1 = \lambda_2 = \cdots = \lambda_s$. Let $u \triangleq \log_2 e + \lambda_{s+1}$ and $w \triangleq \lambda_0$. One can rewrite (49) and (37) as follows:

$$\begin{cases} (1)\ \log_2 \hat{\tau}(\mathbf{a}) + u + v \sum_{i=1}^{s} a_i = 0 \quad \text{for any } \mathbf{a} \neq \boldsymbol{0}, \\ (2)\ \log_2 \hat{\tau}(\boldsymbol{0}) + u + w = 0, \\ (3)\ \hat{\tau}(\boldsymbol{0}) = 1 - q, \\ (4)\ \sum_{\mathbf{a}} \hat{\tau}(\mathbf{a}) = 1, \\ (5)\ \sum_{\mathbf{a}:a_i=1} \hat{\tau}(\mathbf{a}) = Q \quad \text{for any } i \in [s]. \end{cases} \tag{38}$$

Let $y \triangleq \frac{1}{1+2^{-v}}$. The first equation of the system (38) means that

$$\hat{\tau}(\mathbf{a}) = \frac{1}{2^u y^s} (1-y)^{\sum a_j} y^{s - \sum a_j} \quad \text{for any } \mathbf{a} \neq \boldsymbol{0}. \tag{39}$$

Substituting (39) into the Eq. (5) allows us to obtain

$$\sum_{\mathbf{a}:a_i=1} \frac{1}{2^u y^s} (1-y)^{\sum a_j} y^{s - \sum a_j} = \frac{1-y}{2^u y^s},$$

and therefore the solution u is determined by the equality

$$u = \log_2 \left(\frac{1-y}{Q y^s} \right). \tag{40}$$

Substituting (39), (40) and the third equation of (38) into the Eq. (4) of the system (38) we have

$$q = \sum_{\mathbf{a} \neq 0} \hat{\tau}(\mathbf{a}) = \frac{Q(1-y^s)}{1-y},$$

i.e. the Eq. (28). Thus, the conditions (49) and (37) have the unique solution τ in the region (48):

$$\hat{\tau}(\boldsymbol{0}) = 1 - q,$$
$$\hat{\tau}(\mathbf{a}) = \frac{Q}{1-y}(1-y)^{\sum a_j} y^{s-\sum a_j} \quad \text{for any } \mathbf{a} \neq \boldsymbol{0}, \tag{41}$$

where the parameters q and y are related by the Eq. (28). To get the exact formula (27), the substitution of (41) into (33) is sufficient.

Let us prove the properties of the function (27). Note that the function $q(y) = Q\frac{1-y^s}{1-y}$ (28) monotonically increases in the interval $y \in (0, 1)$ and takes the values Q and sQ at the ends of the interval, respectively. That is why one can consider the function (27) as the function $\mathscr{F}(s, Q, y) \triangleq \mathscr{A}(s, Q, q(y))$ of the parameter y in the interval $y \in (0, y_1)$, where $q(y_1) = \min\{1, sQ\}$. The derivative of the function $\mathscr{F}(s, Q, y)$ equals

$$\mathscr{F}'(s, Q, y) = q'(y) \log_2 \left(\frac{Qy^s}{1 - Q - y + Qy^s} \right). \tag{42}$$

Thus, $\mathscr{F}(s, Q, y)$ decreases in the interval $y \in (0, 1 - Q)$, increases in the interval $y \in (1-Q, y_1)$, is $\cup$-convex, attains the minimal value 0 at $y_0 = 1-Q$ and $q(y_0) = 1-(1-Q)^s$. □

Proof (Lemma 2) Now, compute the conditional probability

$$\mathscr{P}_1^{(N)}(\ell, Q, k) \triangleq \Pr\left\{ \bigvee_{i \in \mathscr{S}} \boldsymbol{x}(i) \succeq \bigwedge_{j \in \mathscr{L}} \boldsymbol{x}(j) \Big/ \left| \bigvee_{i \in \mathscr{S}} \boldsymbol{x}(i) \right| = k \right\}.$$

Let q, $Q \le q \le \min\{1, sQ\}$, be fixed and $k \triangleq \lfloor qN \rfloor$, $\lfloor QN \rfloor \le k \le s\lfloor QN \rfloor$. In terms of *types* (see [2], [5]):

$$\{n(\mathbf{a})\}, \quad \mathbf{a} \triangleq (a_1, a_2, \ldots, a_s) \in \{0,1\}^\ell, \quad 0 \le n(\mathbf{a}) \le N, \quad \sum_{\mathbf{a} \in \{0,1\}^\ell} n(\mathbf{a}) = N, \tag{43}$$

one can rewrite the probability in the following form

$$\mathscr{P}_1^{(N)}(\ell, Q, k) = \sum_{(45)} \frac{N!}{\prod\limits_{\mathbf{a} \in \{0,1\}^\ell} n(\mathbf{a})!} \frac{\binom{k}{n(\boldsymbol{I})}}{\binom{N}{n(\boldsymbol{I})}} \binom{N}{\lfloor QN \rfloor}^{-\ell}, \tag{44}$$

where the summation is taken over all choices of types $\{n(\mathbf{a})\}$ provided that

$$\sum_{\mathbf{a}:\, a_i = 1} n(\mathbf{a}) = \lfloor QN \rfloor \quad \text{for any } i \in [\ell]. \tag{45}$$

Applying the Stirling formula calculate the logarithmic behaviour of a term in (44)

$$\log_2 \left(\frac{N!}{\prod\limits_{\mathbf{a} \in \{0,1\}^\ell} n(\mathbf{a})!} \frac{\binom{k}{n(\boldsymbol{I})}}{\binom{N}{n(\boldsymbol{I})}} \binom{N}{\lfloor QN \rfloor}^{-\ell} \right) = 2^{-NF(\tau, Q, q)(1+o(1))},$$

where

$$F(\tau, Q, q) \triangleq \sum_{\mathbf{a} \in \{0,1\}^\ell} \tau(\mathbf{a}) \log_2 \tau(\mathbf{a}) - q \cdot h\left(\frac{\tau(\boldsymbol{I})}{q} \right) + h(\tau(\boldsymbol{I})) + \ell \cdot h(Q). \tag{46}$$

Here the *probability distribution* $\{\tau(\mathbf{a})\}$ is determined as

$$\tau(\mathbf{a}) \triangleq \frac{n(\mathbf{a})}{N} \quad \text{for any } \mathbf{a} \in \{0,1\}^{\ell}.$$

Since we are interested in

$$\mathscr{D}(\ell, Q, q) = \lim_{N\to\infty} -\frac{\log_2\left(P_1^{(N)}(\ell, Q, k)\right)}{N},$$

we might estimate the following minimum

$$\mathscr{D}(\ell, Q, q) = \min_{\tau\in(48):(49)} F(\tau, Q, q) \triangleq F(\hat{\tau}, Q, q), \tag{47}$$

$$\left\{\tau : \forall\, \mathbf{a} = (a_1, \ldots, a_\ell) \in \{0,1\}^{\ell} \quad 0 < \tau(\mathbf{a}) < 1\right\}, \tag{48}$$

$$\sum_{\mathbf{a}} \tau(\mathbf{a}) = 1, \qquad \sum_{\mathbf{a}:\, a_i=1} \tau(\mathbf{a}) = Q \quad \text{for any } i \in [\ell], \tag{49}$$

where the restrictions (49) are induced by properties (43) and (43).

To find the minimum we apply the Lagrange method, i.e., consider the Lagrangian

$$\begin{aligned}\Lambda \triangleq & \sum_{\mathbf{a}\in\{0,1\}^{\ell}} \tau(\mathbf{a}) \log_2 \tau(\mathbf{a}) - q \cdot h\left(\frac{\tau(\boldsymbol{1})}{q}\right) + h(\tau(\boldsymbol{1})) + \ell \cdot h(Q) \\ & + \mu_0 \cdot \left(\sum_{\mathbf{a}} \tau(\mathbf{a}) - 1\right) + \sum_{i=1}^{\ell} \mu_i \cdot \left(\sum_{\mathbf{a}:\, a_i=1} \tau(\mathbf{a}) - Q\right).\end{aligned} \tag{50}$$

Therefore, the necessary conditions for the extremal distribution $\{\hat{\tau}\}$ look similar for any $\mathbf{a} \neq \boldsymbol{1}$, and for the case $\mathbf{a} = \boldsymbol{1}$, the condition has the specific form :

$$\begin{cases} \frac{\partial \Lambda}{\partial \tau(\mathbf{a})} = \log_2 \hat{\tau}(\mathbf{a}) + \log_2 e + \mu_0 + \sum\limits_{i:\, a_i=1} \mu_i = 0 \quad \text{for any } \mathbf{a} \neq \boldsymbol{1}, \\ \frac{\partial \Lambda}{\partial \tau(\boldsymbol{1})} = \log_2 \hat{\tau}(\boldsymbol{1}) + \log_2 e + \sum\limits_{i=0}^{\ell} \mu_i + \log_2\left(\frac{1-\hat{\tau}(\boldsymbol{1})}{q-\hat{\tau}(\boldsymbol{1})}\right) = 0. \end{cases} \tag{51}$$

The matrix of second derivatives of the Lagrangian is clearly diagonal. Thus, this matrix is positive definite in the region (48) and the function $F(\tau, Q, q)$ defined by (46) is strictly $\cup$-convex in the region (46). The Karush–Kuhn–Tacker theorem [10] states that each solution $\{\hat{\tau}\}$ satisfying system (51) and constraints (49) gives a local minimum of $F(\tau, Q, q)$. Thus, if there exists a solution of the system (51) and (49) in the region (48), then it is unique and gives a minimum in the minimization problem (47)–(49).

Note that the symmetry of problem yields the equality $\mu \triangleq \mu_1 = \mu_2 = \cdots = \mu_\ell$. Let $\hat{\mu} \triangleq \log_2 e + \mu_0$. One can rewrite (51) as

$$\begin{cases} \hat{\mu} + \mu \sum_{i=1}^{\ell} a_i + \log_2(\hat{\tau}(\mathbf{a})) = 0 \quad \text{for any } \mathbf{a} \neq \boldsymbol{1}, \\ \hat{\mu} + \mu\ell + \log_2(\hat{\tau}(\boldsymbol{1})) + \log_2\left(\frac{1-\hat{\tau}(\boldsymbol{1})}{q-\hat{\tau}(\boldsymbol{1})}\right) = 0. \end{cases} \tag{52}$$

The first equations of (52) lead to

$$\hat{\tau}(\mathbf{a}) = \frac{2^{-\hat{\mu}}}{z^{\ell}} \prod P(a_i),$$

where we introduce the Bernoulli distribution

$$P(a) \triangleq \begin{cases} z \triangleq \frac{1}{1+2^{-\mu}} & \text{for } a = 0, \\ 1 - z \triangleq \frac{2^{-\mu}}{1+2^{-\mu}} & \text{for } a = 1. \end{cases}$$

In particular, it follows that

$$\mu = \log_2 \left(\frac{z}{1-z} \right). \tag{53}$$

Since (49) the sum of all probabilities equals 1 we get

$$\hat{\tau}(\boldsymbol{I}) = 1 - \sum_{k=0}^{\ell-1} \binom{\ell}{k} \frac{2^{-\hat{\mu}}}{z^\ell} z^{\ell-k} (1-z)^k = 1 - \frac{2^{-\hat{\mu}}}{z^\ell} \left(1 - (1-z)^\ell \right). \tag{54}$$

The relation (49) of constant weight leads to

$$Q = \frac{2^{-\hat{\mu}}}{z^\ell} \sum_{k=0}^{\ell-2} \binom{\ell-1}{k} z^{\ell-k-1} (1-z)^{k+1} + 1 - \frac{2^{-\hat{\mu}}}{z^\ell} \left(1 - (1-z)^\ell \right) = 1 - \frac{2^{-\hat{\mu}}}{z^{\ell-1}}.$$

This gives the relation between $\hat{\mu}$ and parameters Q and z

$$\hat{\mu} = -\log_2 \left((1-Q) z^{\ell-1} \right). \tag{55}$$

Finally, substituting (53)–(55) to the second equation of (52) yields

$$\log_2 \left(\frac{z - (1-Q)\left(1-(1-z)^\ell\right)}{(q-1)z + (1-Q)\left(1-(1-z)^\ell\right)} \right) + \log_2 \left(\frac{\left(1-(1-z)^\ell\right)}{(1-z)^\ell} \right) = 0$$

This equation determines Q as a function of parameters z, q, s and ℓ

$$Q = \frac{(1-z)(1-(1-z)^\ell) - (1-q)z(1-z)^\ell}{1-(1-z)^\ell}. \tag{56}$$

Notice that for fixed parameters q, s and ℓ there is a bijection between $Q \in [0, 1]$ and $z \in [0, 1]$. From (55) and (56) it follows that

$$\frac{2^{-\hat{\mu}}}{z^\ell} = \frac{1-Q}{z} = \frac{1-q(1-z)^\ell}{1-(1-z)^\ell}. \tag{57}$$

Let us substitute $q = \hat{q} = 1 - (1-Q)^s$. Thus

$$\hat{\tau}(\boldsymbol{I}) = \hat{q}(1-z)^\ell. \tag{58}$$

Remind (47) that

$$F(\hat{\tau}, Q, \hat{q}) = \sum_{\mathbf{a} \in \{0,1\}^\ell} \hat{\tau}(\mathbf{a}) \log_2 \hat{\tau}(\mathbf{a}) - \hat{q} \cdot h\left(\frac{\hat{\tau}(\boldsymbol{I})}{\hat{q}} \right) + h(\hat{\tau}(\boldsymbol{I})) + \ell \cdot h(Q). \tag{59}$$

Let us rewrite the first sum of (59) applying (57):

$$\begin{aligned}\sum_{\mathbf{a}\in\{0,1\}^\ell} \hat{\tau}(\mathbf{a})\log_2\hat{\tau}(\mathbf{a}) &= \hat{\tau}(\boldsymbol{I})\log_2\hat{\tau}(\boldsymbol{I})\\ &\quad+\sum_{i=0}^{\ell-1}\binom{\ell}{i}\frac{2^{-\hat{\mu}}}{z^\ell}(1-z)^i z^{\ell-i}\log_2\left(\frac{2^{-\hat{\mu}}}{z^\ell}(1-z)^i z^{\ell-i}\right)\\ &=\sum_{i=0}^{\ell-1}\binom{\ell}{i}\frac{2^{-\hat{\mu}}}{z^\ell}(1-z)^i z^{\ell-i}\log_2\left(\frac{2^{-\hat{\mu}}}{z^\ell}\right)\\ &\quad+\sum_{i=0}^{\ell-1}\binom{\ell}{i}\frac{2^{-\hat{\mu}}}{z^\ell}(1-z)^i z^{\ell-i}\log_2\left(z^{\ell-i}\right)\\ &\quad+\sum_{i=0}^{\ell-1}\binom{\ell}{i}\frac{2^{-\hat{\mu}}}{z^\ell}(1-z)^i z^{\ell-i}\log_2\left((1-z)^i\right)+\hat{\tau}(\boldsymbol{I})\log_2\hat{\tau}(\boldsymbol{I})\\ &=\left(1-\hat{q}(1-z)^\ell\right)\log_2\left(\frac{1-\hat{q}(1-z)^\ell}{1-(1-z)^\ell}\right)+(1-Q)\,\ell\,\log_2 z\\ &\quad+\frac{(1-Q)}{z}\ell\left((1-z)-(1-z)^\ell\right)\log_2(1-z)+\hat{\tau}(\boldsymbol{I})\log_2\hat{\tau}(\boldsymbol{I}).\end{aligned}$$

Taking into account (58) the second term of (59) is

$$\begin{aligned}-\hat{q}h\left(\frac{\hat{\tau}(\boldsymbol{I})}{\hat{q}}\right)&=\tau(\boldsymbol{I})\log_2\left(\frac{\hat{\tau}(\boldsymbol{I})}{q}\right)+(q-\hat{\tau}(\boldsymbol{I}))\log_2\left(\frac{q-\hat{\tau}(\boldsymbol{I})}{q}\right)\\ &=\ell\hat{q}(1-z)^\ell\log_2(1-z)+\hat{q}(1-(1-z)^\ell)\log_2(1-(1-z)^\ell).\end{aligned}$$

The third term of (59) is

$$h(\hat{\tau}(\boldsymbol{I}))=-\hat{\tau}(\boldsymbol{I})\log_2\hat{\tau}(\boldsymbol{I})-(1-\hat{\tau}(\boldsymbol{I}))\log_2(1-\hat{\tau}(\boldsymbol{I})).$$

Finally, the last term of (59) is $\ell h(Q)$. Therefore, the value $\mathscr{D}(\ell, Q, \hat{q}) = F(\hat{\tau}, Q, \hat{q})$ can be written as

$$\begin{aligned}\mathscr{D}(\ell,Q,\hat{q})\triangleq{}&\ell h(Q)+(1-Q)\ell\,\log_2 z-(1-\hat{q})\log_2(1-(1-z)^\ell)\\ &+\ell\left(\frac{(1-Q)}{z}(1-z)-\left(\frac{(1-Q)}{z}-\hat{q}\right)(1-z)^\ell\right)\log_2(1-z).\end{aligned}$$

This finishes the proof of Lemma 2. □

Acknowledgements A. D'yachkov, I.V. Vorobyev, N.A. Polyanskii and V.Yu. Shchukin have been supported in part by the Russian Foundation for Basic Research under Grant No. 16-01-00440.

References

1. Bassalygo L.A., Rykov V.V.: Multiple-access hyperchannel. Probl. Inf. Transm. **49**(4), 299–307 (2013).
2. Csiszar I., Korner J.: Information Theory: Coding Theorems for Discrete Memoryless Systems. Cambridge University Press, Cambridge (2011).
3. D'yachkov A.G., Macula A.J., Rykov V.V.: New Applications and Results of Superimposed Code Theory Arising from the Potentialities of Molecular Biology. Numbers, Information and Complexity. Kluwer, Dordrecht (2000).

4. D'yachkov A.G., Vilenkin P., Macula A., Torney D.: Families of finite sets in which no intersection of ℓ sets is covered by the union of s others. J. Comb. Theory A **99**, 195–218 (2002).
5. D'yachkov A.G., Vorobyev I.V., Polyanskii N.A., Shchukin V.Yu.: Bounds on the rate of disjunctive codes. Probl. Inf. Transm. **50**(1), 27–56 (2014).
6. D'yachkov A.G., Vorobyev I.V., Polyanskii N.A., Shchukin V.Yu.: Bounds on the rate of superimposed codes. In: Proceedings of the IEEE International Symposium on Information Theory (ISIT), pp. 2341–2345. IEEE, Honolulu (2014).
7. D'yachkov A.G., Vorobyev I.V., Polyanskii N.A., Shchukin V.Yu.: Almost disjunctive list-decoding codes. Probl. Inf. Transm. **51**(2), 110–131 (2014).
8. D'yachkov A.G., Vorobyev I.V., Polyanskii N.A., Shchukin V.Yu.: Almost cover-free codes and designs. In: Proceedings of the IEEE International Symposium on Information Theory (ISIT), pp. 2899–2903. IEEE, Hong Kong (2015).
9. Erdos P., Frankl P., Furedi Z.: Families of finite sets in which no set is covered by the union of 2 others. J. Comb. Theory A **33**, 158–166 (1982).
10. Galeev E.M., Tikhomirov V.M.: Optimization: Theory, Examples, Problems. Editorial URSS, Moscow (2000) (in Russian).
11. Kautz W.H., Singleton R.C.: Nonrandom binary superimposed codes. IEEE Trans. Inf. Theory. **10**(4), 363–377 (1964).
12. Malyutov M.B.: The separating property of random matrices. Math. Notes **23**(1), 84–91 (1978).
13. Mitchell C.J., Piper F.C.: Key storage in secure networks. Discret. Appl. Math. **21**, 215–228 (1988).
14. Torney D.C.: Sets pooling designs. Ann. Comb. **3**, 95–101 (1999).