The Generalized Covering Radii of Linear Codes

Dor Elimelech, *Graduate Student Member, IEEE*, Marcelo Firer, *Member, IEEE*, and Moshe Schwartz, *Senior Member, IEEE*

***Abstract*— Motivated by an application to database linear querying, such as private information-retrieval protocols, we suggest a fundamental property of linear codes – the generalized covering radius. The generalized covering-radius hierarchy of a linear code characterizes the trade-off between storage amount, latency, and access complexity, in such database systems. Several equivalent definitions are provided, showing this as a combinatorial, geometric, and algebraic notion. We derive bounds on the code parameters in relation with the generalized covering radii, study the effect of simple code operations, and describe a connection with generalized Hamming weights.**

***Index Terms*— Linear codes, covering radius, generalized Hamming weights, block metric.**

I. Introduction

A COMMON query type in database systems involves a linear combination of the database items with coefficients supplied by the user. As examples we mention partial-sum queries [4], and private information retrieval (PIR) protocols [5]. In essence, one can think of the database server as storing m items, $x_1, \dots, x_m \in \mathbb{F}_{q^\ell}$. A user may query the contents of the database by providing $s_1, \dots, s_m \in \mathbb{F}_q$, and getting in response the linear combination $\sum_{i=1}^m s_i x_i$.

Various aspects of these systems are of interest and in need of optimization, such as the amount of storage at the server, and the required bandwidth for the querying protocol. One important such aspect is that of *access complexity*, paralleling a similar concern studied in distributed storage systems [15], [25]. In a straightforward implementation, the time required to access the elements of the database needed to compute the answer to a user query is directly proportional to the number of non-zero coefficients among $s_1, \dots, s_m$. This may prove to be a bottleneck, in particular since in schemes like PIR, the coefficients are random, and therefore a typical query would require the database server to access a fraction of $1 - \frac{1}{q}$ of the items.

Manuscript received December 11, 2020; revised May 14, 2021; accepted September 4, 2021. Date of publication September 27, 2021; date of current version November 22, 2021. The work of Dor Elimelech was supported in part by Israel Science Foundation (ISF) under Grant 1052/18. The work of Marcelo Firer was supported in part by Fapesp under Grant 13/25977-7 and in part by CNPq under Grant 304046/2017-5. The work of Moshe Schwartz was supported in part by the German Israeli Project Cooperation (DIP) under Grant PE2398/1-1. An earlier version of this paper was presented in part at the IEEE International Symposium on Information Theory (ISIT) 2021. *(Corresponding author: Dor Elimelech.)*

Dor Elimelech and Moshe Schwartz are with the School of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Beer Sheva 8410501, Israel (e-mail: doreli@post.bgu.ac.il; schwartz@ee.bgu.ac.il).

Marcelo Firer is with the Institute of Mathematics, Statistics and Scientific Computing, University of Campinas, Campinas 13083-859, Brazil (e-mail: mfirer@ime.unicamp.br).

Communicated by S. Ghorpade, Associate Editor for Coding Theory.

Digital Object Identifier 10.1109/TIT.2021.3115433

A trade-off between access complexity and storage amount was suggested for PIR in [28], echoing a similar suggestion for databases made in [18]. The suggestion calls for a carefully designed set of linear combinations to be pre-computed and stored by the server. Instead of storing $\overline{x} = (x_1, \dots, x_m)$ as is, the server stores $\overline{h}_1 \cdot \overline{x}, \dots, \overline{h}_n \cdot \overline{x}$, where each $\overline{h}_i \in \mathbb{F}_q^m$ describes a linear combination. Assume that the matrix H, whose columns are $\overline{h}_1, \dots, \overline{h}_n$, is a parity-check matrix for a code with covering radius r. Thus, when the user queries the database using $\overline{s} = (s_1, \dots, s_m)$, by the properties of the covering code, $\overline{s}$ may be computed using a linear combination of at most r columns of H. Hence, at most r pre-computed combinations that are stored in the database need to be accessed in order to provide the user with the requested linear combination. The trade-off between access complexity and storage amount follows, since instead of storing m elements, the server now stores $n \geqslant m$ linear combinations, and so n is lower bounded by the smallest possible length for a code with covering radius r and redundancy m over $\mathbb{F}_q$. These code parameters have been thoroughly studied and are well understood [6].

We now take access-complexity optimization one step further. The database server naturally receives a stream of queries, say $\overline{s}_1, \overline{s}_2, \dots$. Those may arrive from the same user, or from multiple distinct users. Instead of handling each of the queries separately, accessing r pre-computed linear combinations for each query, the server may group together t queries, $\overline{s}_1, \dots, \overline{s}_t$ and, hopefully, access fewer than $r \cdot t$ pre-computed linear combinations as it would in a naive implementation. Thus, both storage amount and latency are traded-off for a reduced access complexity.

The motivation mentioned above leads us to the following combinatorial problem: Design a set of vectors, $\overline{h}_1, \dots, \overline{h}_n \in \mathbb{F}_q^m$ (describing linear combinations to pre-compute), such that every t vectors, $\overline{s}_1, \dots, \overline{s}_t \in \mathbb{F}_q^m$ (describing user queries), may be obtained by accessing at most r of the elements $\overline{h}_1, \dots, \overline{h}_n$. When viewed as columns of a parity-check matrix for a code, this becomes a *generalized covering radius* definition. It bears a resemblance to the generalized Hamming weight of codes, introduced by Wei [27] to characterize the performance of linear codes over a wire-tap channel.

The goal of this paper is to study the generalized covering radius as a fundamental property of linear codes. Our main contributions are the following:

1) We discuss three definitions for the generalized covering radius of a code, highlighting the combinatorial, geometric, and algebraic properties of this concept, and showing them to be equivalent.

2) We derive bounds that tie the various parameters of codes to the generalized covering radii. In particular, we prove an asymptotic upper bound on the minimum rate of binary codes with a prescribed second generalized covering radius, thus showing an improvement over the naive approach. The bound on the minimal rate is attained by almost all codes.
3) We determine the effect simple code operations have on the generalized covering radii: code extension, puncturing, the $(u, u+v)$ construction, and direct sum.
4) We discuss a connection between the generalized covering radii and the generalized Hamming weights of codes by showing that the latter is in fact a packing problem with some rank relaxation.

The paper is organized as follows: Preliminaries and notations are presented in Section II. We study various definitions of the generalized covering radius, and show them to be equivalent, in Section III. Section IV is devoted to the derivation of bounds on the generalized covering radii. Basic operations on codes are studied in Section V, and a relation with the generalized Hamming weights in Section VI. We conclude with a discussion of the results and some open questions in Section VII.

II. Preliminaries

For all $n \in \mathbb{N}$, we define $[n] \triangleq \{1, 2, \dots, n\}$. If A is a finite set and $t \in \mathbb{N}$, we denote by $\binom{A}{t}$ the set of all subsets of A of size exactly t. We use $\mathbb{F}_q$ to denote the finite field of size q, and denote $\mathbb{F}_q^* \triangleq \mathbb{F}_q \setminus \{0\}$. Given a vector space V over $\mathbb{F}_q$, we denote by $\begin{bmatrix} V \\ t \end{bmatrix}$ the set of all vector subspaces of V of dimension $t \in \mathbb{N}$. We use lower-letters, v, to denote scalars, overlined lower-case letters, $\overline{v}$, to denote vectors, and either bold lower-case letters, $\mathbf{v}$, or upper-case letter, V, to denote matrices. Whether vectors are row vectors or column vectors is deduced from context.

If H is a matrix with n columns, we denote by $\overline{h}_i$ its i-th column. For $I = \{i_1, i_2, \dots, i_t\} \in \binom{[n]}{t}$, we denote by H_I the restriction of H to the columns whose indices are in I, i.e., $H_I \triangleq (\overline{h}_{i_1}, \dots, \overline{h}_{i_t})$. We shall also use $\langle H_I \rangle$ to denote the linear space spanned by the columns of H_I, i.e.,

$$\langle H_I \rangle \triangleq \left\langle \overline{h}_{i_1}, \overline{h}_{i_2}, \dots, \overline{h}_{i_t} \right\rangle .$$

Given $\overline{v} = (v_1, \dots, v_n) \in \mathbb{F}_q^n$, the support of $\overline{v}$ is defined by

$$\operatorname{supp}(\overline{v}) \triangleq \{i \in [n] | v_i \neq 0\} .$$

Whenever required, for a subset $V \subseteq \mathbb{F}_q^n$ we define

$$\operatorname{supp}(V) \triangleq \bigcup_{\overline{v} \in V} \operatorname{supp}(\overline{v}).$$

The Hamming weight of $\overline{v}$ is then defined as $\operatorname{wt}(\overline{v}) \triangleq |\operatorname{supp}(\overline{v})|$. If $\overline{v}' \in \mathbb{F}_q^n$, then the Hamming distance between $\overline{v}$ and $\overline{v}'$ is given by $d(\overline{v}, \overline{v}') \triangleq \operatorname{wt}(\overline{v} - \overline{v}')$. We also extend the definition to the distance between a vector and a set, namely, for a set $C \subseteq \mathbb{F}_q^n$,

$$d(\overline{v}, C) \triangleq \min \left\{ d(\overline{v}, \overline{c}) \middle| \overline{c} \in C \right\} .$$

Two shapes that will be useful to us are the ball and the cube. For a non-negative integer r, the Hamming ball of radius r centered at $\overline{v} \in \mathbb{F}_q^n$ is defined as

$$B_{r,n,q}(\overline{v}) \triangleq \left\{ \overline{v}' \in \mathbb{F}_q^n \middle| d(\overline{v}, \overline{v}') \leqslant r \right\} .$$

The cube with support $I \in \binom{[n]}{r}$ centered at $\overline{v} \in \mathbb{F}_q^n$ is defined as

$$Q_{I,n,q}(\overline{v}) \triangleq \left\{ \overline{v}' \in \mathbb{F}_q^n \middle| \operatorname{supp}(\overline{v}' - \overline{v}) \subseteq I \right\} .$$

We shall omit the subscripts n and q whenever they may be inferred from the context. We observe that

$$\bigcup_{I \in \binom{[n]}{r}} Q_I(\overline{v}) = B_r(\overline{v}).$$

III. The Generalized Covering Radii

We would now like to introduce the concept of generalized covering radius. We present several definitions, with varying approaches, be they combinatorial, algebraic, or geometric. We then show all of the definitions are in fact equivalent (at least, when linear codes are concerned).

Our first definition stems directly from the application outlined in the introduction – database queries.

Definition 1: Let C be an $[n, k]$ linear code over $\mathbb{F}_q$, given by an $(n-k) \times n$ parity-check matrix $H \in \mathbb{F}_q^{(n-k)\times n}$. For every $t \in \mathbb{N}$ we define the *t-th generalized covering radius*, $R_t(C)$, to be the minimal integer $r \in \mathbb{N}$ such that for every set $S \in \binom{\mathbb{F}_q^{n-k}}{t}$ there exists $I \in \binom{[n]}{r}$ such that $S \subseteq \langle H_I \rangle$. That is,

$$R_t(C) \triangleq \max_{\substack{S \subseteq \mathbb{F}_q^{n-k} \\ |S|=t}} \min_{\substack{I \subseteq [n] \\ S \subseteq \langle H_I \rangle}} |I| .$$

While $R_t(C)$ certainly depends on the code C, for the sake of brevity we sometimes write R_t when we can infer C from the context. At first glance it seems as if R_t does not only depend on C, but also on the choice of parity-check matrix H. However, the following lemma shows this is not the case.

Lemma 2: Let R_t be the given by a full-rank matrix $H \in \mathbb{F}_q^{(n-k)\times n}$ as in Definition 1. For any $A \in \operatorname{GL}(n-k, q)$ (the group of $(n-k) \times (n-k)$ invertible matrices with coefficients in $\mathbb{F}_q$), let R'_t be the generalized covering radius, as in Definition 1, but using the matrix AH. Then $R_t = R'_t$.

Proof: Given $\overline{s} \in \mathbb{F}_q^{n-k}$, if $\overline{s} = \sum_{i \in I} \alpha_i \overline{h}_i$, then, by linearity, we have that

$$A\overline{s} = A \sum_{i \in I} \alpha_i \overline{h}_i = \sum_{i \in I} \alpha_i A\overline{h}_i .$$

It follows that, given $S \subseteq \mathbb{F}_q^{n-k}$, if $S \subseteq \langle H_I \rangle$, then $A(S) \subseteq \langle AH_I \rangle$. Thus,

$$\min_{\substack{I \subseteq [n] \\ S \subseteq \langle H_I \rangle}} |I| \geqslant \min_{\substack{I \subseteq [n] \\ A(S) \subseteq \langle AH_I \rangle}} |I| .$$

Continuing with the same argument but using A^{-1}, we have

$$\min_{\substack{I \subseteq [n] \\ A(S) \subseteq \langle AH_I \rangle}} |I| \geqslant \min_{\substack{I \subseteq [n] \\ A^{-1}A(S) \subseteq \langle A^{-1}AH_I \rangle}} |I| = \min_{\substack{I \subseteq [n] \\ S \subseteq \langle H_I \rangle}} |I| .$$

It then follows that

$$\min_{\substack{I\subseteq[n]\\ S\subseteq\langle H_I\rangle}} |I| = \min_{\substack{I\subseteq[n]\\ A(S)\subseteq\langle AH_I\rangle}} |I|.$$

As a consequence, if S_0 realizes the maximum condition,

$$R_t = \max_{\substack{S\subseteq\mathbb{F}_q^{n-k}\\ |S|=t}} \min_{\substack{I\subseteq[n]\\ S\subseteq\langle H_I\rangle}} |I| = \min_{\substack{I\subseteq[n]\\ S_0\subseteq\langle H_I\rangle}} |I| = \min_{\substack{I\subseteq[n]\\ A(S_0)\subseteq\langle AH_I\rangle}} |I|.$$

It follows that $R_t \leqslant R'_t$. A symmetric argument, gives the reversed inequality, proving the desired claim. ■

We observe, in Definition 1, that requiring $S \subseteq \langle H_I\rangle$ also ensures $\langle S\rangle \subseteq \langle H_I\rangle$. We therefore must have for all $t \in [n-k]$,

$$R_t \geqslant t. \tag{1}$$

We also observe that R_1 is in fact the covering radius of the code C, and that the generalized covering radii are naturally monotone increasing, i.e.,

$$R_1 \leqslant R_2 \leqslant \ldots \leqslant R_{n-k} = n-k, \tag{2}$$

as well as $R_t = n-k$ for all $t \geqslant n-k$. Thus, the values $R_1, \ldots, R_{n-k}$ are called the *generalized covering-radius hierarchy*. While being monotone increasing, we do note however, that the generalized covering radius R_t is not necessarily strictly increasing in t, as the following example shows.

Example 3: Consider the binary Hamming code C, with parameters $[2^m-1, 2^m-1-m, 3]$. An $m\times(2^m-1)$ parity-check matrix H for C comprises of all binary vectors of length m as columns, except for the all-zero column. One can easily check that $R_t(C) = t$ for all $t \in [m]$.

Assume $m \geqslant 2$. Now take C' to be a $[2^m-2, 2^m-2-m, 3]$ code obtained from C by shortening once. Thus, a parity-check matrix H' for C' is obtained by taking H and deleting one of its columns; let us suppose that the shortening was done in the position corresponding to the all-ones column of H. We now obviously have $R_1(C') = 2$ since in order to cover $\{\overline{1}\}$ two columns of H' are required. However, we also have $R_2(C') = 2$ since any 2-dimensional subspace of $\mathbb{F}_2^m$ has at least two nonzero (and hence linearly independent) vectors that appear as columns of H'.

Aiming for a geometric interpretation of the generalized covering radii, we provide two more equivalent definitions that are increasingly geometric in nature.

Definition 4: Let C be an $[n,k]$ linear code over $\mathbb{F}_q$. Then for every $t \in \mathbb{N}$ we define the t-th generalized covering radius, $R_t(C)$, to be the minimal integer $r \in \mathbb{N}$ such that for every $\overline{v}_1, \ldots, \overline{v}_t \in \mathbb{F}_q^n$, there exist codewords $\overline{c}_1, \ldots, \overline{c}_t \in C$ and there exists $I \in \binom{[n]}{r}$, such that $\overline{v}_i \in Q_I(\overline{c}_i)$ for all $i \in [t]$.

Lemma 5: Let C be an $[n,k]$ linear code over $\mathbb{F}_q$. Then the values of R_t from Definitions 1 and 4 are the same.

Proof: Fix a parity-check matrix H for C (with full rank). Denote the numbers from Definition 1 and Definition 4 by R_t and R'_t, respectively.

For the first direction, let $\overline{v}_1, \ldots, \overline{v}_t \in \mathbb{F}_q^n$. Consider $\overline{s}_1, \ldots, \overline{s}_t \in \mathbb{F}_q^{n-k}$ given by $\overline{s}_i = H\overline{v}_i$ for all $i \in [t]$. By Definition 1 of R_t, there exists a set $\{i_1, \ldots, i_{R_t}\} = I \in \binom{[n]}{t}$ such that $\overline{s}_1, \ldots, \overline{s}_n \in \langle H_I\rangle$. That is, for each $\ell \in [t]$, there exist scalars $w_{\ell,1}, \ldots, w_{\ell,R_t} \in \mathbb{F}_q$ such that $\overline{s}_\ell = \sum_{j=1}^{R_t} w_{\ell,j}\overline{h}_{i_j}$. We define $\overline{w}_\ell \in \mathbb{F}_q^n$ to be the vector containing $w_{\ell,1}, \ldots, w_{\ell,R_t}$ in the positions of I, and 0 otherwise. Let $\overline{c}_\ell \triangleq \overline{w}_\ell - \overline{v}_\ell$. We note that $\overline{c}_\ell \in C$, as

$$H\overline{c}_\ell = H\overline{w}_\ell - H\overline{v}_\ell = \overline{s}_\ell - \overline{s}_\ell = \overline{0}.$$

On the other hand, $\mathrm{supp}(\overline{c}_\ell - \overline{v}_\ell) = \mathrm{supp}(\overline{w}_\ell) \subseteq I$, and in particular $\overline{v}_\ell \in Q_I(\overline{c}_\ell)$. This shows that $R_t \geqslant R'_t$.

For the second direction of the proof, assume we have vectors $\overline{s}_1 \ldots, \overline{s}_t \in \mathbb{F}_q^{n-k}$. Since H has full rank, there exist $\overline{v}_1, \ldots, \overline{v}_t \in \mathbb{F}_q^n$ such that $H\overline{v}_i = \overline{s}_i$ for all $i \in [t]$. From Definition 4 of R'_t, there exists a set $I \in \binom{[n]}{t}$ such that for all $i \in [t]$, $\mathrm{supp}(\overline{v}_i - \overline{c}_i) \subseteq I$. For each $i \in [t]$, we define $\overline{w}_i \triangleq \overline{v}_i - \overline{c}_i$, and we have

$$H\overline{w}_i = H(\overline{v}_i - \overline{c}_i) = \overline{s}_i.$$

Since $\mathrm{supp}(\overline{w}_i) \subseteq I$, for all $i \in [t]$, it follows that $s_1, \ldots, s_t \in \langle H_I\rangle$. This shows that $R'_t \geqslant R_t$.

Combining the two directions together we obtain that the values of R_t from Definitions 1 and 4 are the same. ■

We now move to a "classical" covering in the geometric sense. It involves a covering of a space with certain shapes. We shall require an extension of the cube to a t-cube. Given a non-negative integer r and support $I \in \binom{[n]}{r}$, the t-cube centered at

$$\mathbf{v} = \begin{bmatrix} \overline{v}_1 \\ \vdots \\ \overline{v}_t \end{bmatrix} \in \mathbb{F}_q^{t\times n},$$

is defined as

$$Q_{I,n,q}^{(t)}(\mathbf{v}) \triangleq \left\{ \mathbf{v}' = \begin{bmatrix} \overline{v}'_1 \\ \vdots \\ \overline{v}'_t \end{bmatrix} \in \mathbb{F}_q^{t\times n} \;\middle|\; \forall i \in [t], \mathrm{supp}(\overline{v}'_i - \overline{v}_i) \in I \right\}.$$

This brings us to the definition of a t-ball centered at $\mathbf{v}$ given by,

$$B_{r,n,q}^{(t)}(\mathbf{v}) \triangleq \bigcup_{I\in\binom{[n]}{r}} Q_I^{(t)}(\mathbf{v}),$$

where we say r is the radius of the t-ball. Again, we shall omit the subscripts n and q whenever they may be inferred from the context. This is indeed a generalization of the ball since

$$B_r^{(1)}(\mathbf{v}) = B_r(\mathbf{v}).$$

Thus, a superscript of $^{(1)}$ will generally be omitted unless a special need for emphasis arises.

In fact, the ball $B_r^{(t)}(\mathbf{v})$ realizes a ball in the natural sense, in the following metric we now define. The space we operate in is $\mathbb{F}_q^{t\times n}$. The *t-weight* of a matrix $\mathbf{v} \in \mathbb{F}_q^{t\times n}$, with row vectors denoted $\overline{v}_i$, is defined as

$$\mathrm{wt}^{(t)}(\mathbf{v}) \triangleq \left| \bigcup_{i\in[t]} \mathrm{supp}(\overline{v}_i) \right|.$$

We now define the t-distance between $\mathbf{v}, \mathbf{v}' \in \mathbb{F}_q^{t\times n}$ as

$$d^{(t)}(\mathbf{v}, \mathbf{v}') \triangleq \mathrm{wt}^{(t)}(\mathbf{v} - \mathbf{v}').$$

In particular, this also shows that $d^{(t)}$ is translation invariant, i.e., for all $\mathbf{v},\mathbf{v}',\mathbf{v}''\in\mathbb{F}_q^{t\times n}$,

$$d^{(t)}(\mathbf{v}+\mathbf{v}'',\mathbf{v}'+\mathbf{v}'')=d^{(t)}(\mathbf{v},\mathbf{v}').$$

It is easily seen now that the t-ball is in fact a ball in the metric induced by the t-distance, i.e.,

$$B_r^{(t)}(\mathbf{v})=\left\{\mathbf{v}'\in\mathbb{F}_q^{t\times n}\middle| d^{(t)}(\mathbf{v},\mathbf{v}')\leqslant r\right\}.$$

We also note that $d^{(1)}$ is simply the Hamming distance function, hence our previous observation of a 1-ball being a ball in the Hamming metric.

Definition 6: Let C be an $[n,k]$ linear code over $\mathbb{F}_q$. Then for every $t\in\mathbb{N}$, we define the t-th generalized covering radius, R_t, to be the minimal integer r such that t-balls centered at

$$C^t\triangleq\left\{\begin{bmatrix}\overline{c}_1\\ \vdots\\ \overline{c}_t\end{bmatrix}\middle|\forall i\in[t],\overline{c}_i\in C\right\},$$

cover $\mathbb{F}_q^{t\times n}$, i.e.,

$$\bigcup_{\mathbf{c}\in C^t}B_r^{(t)}(\mathbf{c})=\mathbb{F}_q^{t\times n}.$$

Lemma 7: Let C be an $[n,k]$ linear code over $\mathbb{F}_q$. Then the values of R_t from Definitions 1, 4, and 6, are the same.

Proof: The proof is straightforward. We observe that for every $\overline{v}_1,\dots,\overline{v}_t\in\mathbb{F}_q^n$ there are $\overline{c}_1,\dots,\overline{c}_t\in C$ and a support $I\in\binom{[n]}{r}$ such that $\overline{v}_i\in Q_I(\overline{c}_i)$ for all $i\in[t]$ if and only if

$$\begin{bmatrix}\overline{v}_1\\ \vdots\\ \overline{v}_t\end{bmatrix}\in B_r^{(t)}\left(\begin{bmatrix}\overline{c}_1\\ \vdots\\ \overline{c}_t\end{bmatrix}\right).$$

Thus, the minimal integer r which defines R_t is the same in Definitions 4 and 6. By Lemma 5, it is also the same as in Definition 1. ■

We would like to comment that if we denote the *columns* of $\mathbf{v}\in\mathbb{F}_q^{t\times n}$ by $\widehat{v}_1,\dots,\widehat{v}_n\in\mathbb{F}_q^t$, then

$$\mathrm{wt}^{(t)}(\mathbf{v})=\left|\left\{j\in[n]\middle|\widehat{v}_j\neq\overline{0}\right\}\right|.$$

This metric is known in the literature as the *block metric* and it was introduced, independently, by Gabidulin [14] and Feng [19].

For our last approach, we make the obvious next step, resulting in an algebraic definition of the generalized covering radii. Assume $\mathbf{v}\in\mathbb{F}_q^{t\times n}$ has rows $\overline{v}_1,\dots,\overline{v}_t\in\mathbb{F}_q^n$. Using the well-known isomorphism $\mathbb{F}_q^t\cong\mathbb{F}_{q^t}$, we can then read each column of $\mathbf{v}$ as a single element from $\mathbb{F}_{q^t}$. More precisely, fix a basis for $\mathbb{F}_{q^t}$ as a vector space over $\mathbb{F}_q$, say, $\beta_1,\dots,\beta_t\in\mathbb{F}_{q^t}$, and associate with $\mathbf{v}$ above the vector

$$\mathbf{v}=\begin{bmatrix}\overline{v}_1\\ \vdots\\ \overline{v}_t\end{bmatrix}\in\mathbb{F}_q^{t\times n}\quad\mapsto\quad\Xi(\mathbf{v})\triangleq\sum_{i=1}^t\beta_i\overline{v}_i\in\mathbb{F}_{q^t}^n.\tag{3}$$

Note that this mapping is in fact a bijection. Under this mapping, a t-ball is mapped to a ball, namely,

$$\Xi(B_{r,n,q}^{(t)}(\mathbf{v}))=B_{r,n,q^t}^{(1)}(\Xi(\mathbf{v})),\tag{4}$$

where we emphasize that the two balls are over different alphabets.

Definition 8: Let C be an $[n,k]$ linear code over $\mathbb{F}_q$. Assume $G\in\mathbb{F}_q^{k\times n}$ is a generator matrix for C, namely,

$$C=\left\{\overline{u}G\middle|\overline{u}\in\mathbb{F}_q^k\right\}.$$

Let $t\in\mathbb{N}$, and let C' be the linear code over $\mathbb{F}_{q^t}$ generated by the same matrix G, namely,

$$C'=\left\{\overline{u}G\middle|\overline{u}\in\mathbb{F}_{q^t}^k\right\}.$$

Then we define the t-th generalized covering radius R_t of C as the covering radius of C', namely,

$$R_t(C)\triangleq R_1(C').$$

Lemma 9: Let C be an $[n,k]$ linear code over $\mathbb{F}_q$. Then the values of R_t from Definitions 1, 4, 6, and 8, are the same.

Proof: Assume the notation of Definition 8. Let $\mathbf{c}\in C^t$, with rows $\overline{c}_1,\dots,\overline{c}_t\in C$, and let $\overline{u}_i\in\mathbb{F}_q^k$ be such that $\overline{c}_i=\overline{u}_iG$. As in (3), assume $\beta_1,\dots,\beta_t\in\mathbb{F}_{q^t}$ is a basis for $\mathbb{F}_{q^t}$ over $\mathbb{F}_q$. Then

$$\Xi(\mathbf{c})=\sum_{i=1}^t\beta_i\overline{c}_i=\left(\sum_{i=1}^t\beta_i\overline{u}_i\right)G.$$

Hence, $\Xi(\mathbf{c})\in C'$, where C' is the code generated by G over $\mathbb{F}_{q^t}$. A symmetric argument gives that Ξ is in fact a bijection between C^t and C'. The claim now follows from Definition 6 and Lemma 7. ■

As a final comment to this section, our original approach to generalize the covering radii of a code, introduced in Definition 1, arises from the interest in querying databases by linear combinations (as, for example, used in PIR), and it uses the parity-check matrix of a code, hence it makes sense only for linear codes. This is not the case for the approach in Definition 6, where R_t is defined intrinsically as a metric invariant. This means that we can use this definition to generalize the covering radii for general (non-linear) codes.

IV. Bounds

A crucial part in our understanding of any figure of merit, is the limits of values it can take. Thus, we devote this section to the derivation of bounds on the generalized covering radii of codes. We put an emphasis on asymptotic bounds, that, given the normalized t-th covering radius, bound the best possible rate. We present a straightforward ball-covering argument for a lower bound. We then also present a trivial upper bound. Our main result is an asymptotic upper bound that improves upon the trivial one, and thus showing there is merit to the usage of generalized covering radii to improve database querying, as described in Section I. Our upper bound is non-constructive, and uses a probabilistic method. It shall be made constructive (albeit, not useful) in Section V.

As is standard, we will require the size of a t-ball. Since the metrics involved are all translation invariant, the size of the ball does not depend on the choice of center. We therefore use

$$V_{r,n,q}^{(t)}\triangleq\left|B_{r,n,q}^{(t)}(\mathbf{0})\right|.$$

Thus, (4) gives the following immediate corollary.

Corollary 10: For all integers n, t, r and a prime power q,

$$V^{(t)}_{r,n,q} = V_{r,n,q^t} = \sum_{i=0}^{r} \binom{n}{i} (q^t - 1)^i.$$

We also recall the definition of the q-ary entropy function,

$$H_q(x) = x \log_q(q-1) - x \log_q(x) \; - (1-x) \log_q(1-x).$$

Using Stirling's approximation, it is well known that

$$V_{r,n,q} = \begin{cases} q^{nH_q(r/n) - o(n)} & 0 \leqslant \frac{r}{n} \leqslant 1 - \frac{1}{q}, \\ q^{n-o(n)} & 1 - \frac{1}{q} < \frac{r}{n} \leqslant 1, \end{cases}$$

and thus,

$$V^{(t)}_{r,n,q} = V_{r,n,q^t} = \begin{cases} q^{tnH_{q^t}(r/n) - o(n)} & 0 \leqslant \frac{r}{n} \leqslant 1 - \frac{1}{q^t}, \\ q^{tn-o(n)} & 1 - \frac{1}{q^t} < \frac{r}{n} \leqslant 1. \end{cases} \tag{5}$$

Let $k_t(n, r, q)$ denote the smallest dimension of a linear code C over $\mathbb{F}_q$ with length n and t-covering radius $R_t(C) \leqslant r$. The following theorem was proved in [7].

Theorem 11 [7]: For all $n, r \in \mathbb{N}$, and a prime power q,

$$\begin{aligned} n - \log_q V_{r,n,q} &\leqslant k_1(n, r, q) \\ &\leqslant n - \log_q V_{r,n,q} + 2\log_2 \; n - \log_q n + O(1). \end{aligned}$$

It is convenient to study normalized parameters with respect to the length of the code. If C is an $[n, k]$ linear code, we define its normalized parameters,

$$\kappa \triangleq \frac{k}{n}, \qquad \rho_t \triangleq \frac{R_t}{n}.$$

Note that we use κ for the rate of the code, and not R, to avoid confusion with the covering radius. For $t \in \mathbb{N}$ and a normalized covering radius $0 \leqslant \rho \leqslant 1$, the minimal rate achieving ρ is defined to be

$$\kappa_t(\rho, q) \triangleq \liminf_{n\to\infty} \frac{k_t(n, \rho n, q)}{n}.$$

In this notation, Theorem 11 gives an asymptotically tight expression,

$$\kappa_1(n, \rho) = \begin{cases} 1 - H_q(\rho) & 0 \leqslant \rho < 1 - \frac{1}{q}, \\ 0 & 1 - \frac{1}{q} \leqslant \rho \leqslant 1. \end{cases} \tag{6}$$

A. General Bounds

For a simple lower bound we use the ball-covering argument.

Proposition 12: For any $n, t \in \mathbb{N}$, prime power q, and $0 \leqslant \rho \leqslant 1 - \frac{1}{q^t}$,

$$\kappa_t(\rho, q) \geqslant 1 - H_{q^t}(\rho).$$

Proof: Let C be an $[n, k]$ code over $\mathbb{F}_q$ with $R_t(C) \leqslant \rho n$. For any $\mathbf{c} \in C^t$ consider the t-ball of radius $R_t(C)$ centered at $\mathbf{c}$, $B^{(t)}_{R_t(C)}(\mathbf{c})$. By Definition 6,

$$\bigcup_{\mathbf{c}\in C^t} B^{(t)}_{R_t(C)}(\mathbf{c}) = \mathbb{F}_q^{t\times n}.$$

Thus, using Corollary 10,

$$q^{kt} \cdot V^{(t)}_{R_t(C),n,q} = \left|C^t\right| \cdot V^{(t)}_{R_t(C),n,q} = \sum_{\mathbf{c}\in C^t} \left|B^{(t)}_{R_t(C)}(\mathbf{c})\right| \geqslant q^{nt},$$

and therefore,

$$\frac{k}{n} \geqslant 1 - \frac{\log_{q^t} V^{(t)}_{R_t(C),n,q}}{n}.$$

Using (5) we get,

$$\kappa \geqslant 1 - H_{q^t}\left(\frac{R_t(C)}{n}\right) + o(1) \geqslant 1 - H_{q^t}(\rho) + o(1).$$

This bound holds for an arbitrary $[n, k]$ code with t-covering radius at most ρn. Therefore, we have

$$\frac{k_t(n, \rho n, q)}{n} \geqslant 1 - H_{q^t}(\rho) + o(1),$$

and by taking $\liminf$ we conclude. ■

For an upper bound, we first make the following observation.

Proposition 13: Let C be an $[n, k]$ code over $\mathbb{F}_q$. Then for all $t \in \mathbb{N}$,

$$R_t \leqslant t \cdot R_1.$$

Proof: Let H be a parity-check matrix for C. By Definition 1, given $S = \{\overline{s}_1, \ldots, \overline{s}_t\} \in \binom{\mathbb{F}_q^{n-k}}{t}$, there exist $I_i \in \binom{[n]}{R_1}$ such that $\overline{s}_i \in \langle H_{I_i} \rangle$, for all $i \in [t]$. Define $I \triangleq \bigcup_{i\in[t]} I_i$, then $S \subseteq \langle H_I \rangle$. It follows that

$$R_t \leqslant |I| \leqslant \sum_{i=1}^{t} |I_i| = t \cdot R_1.$$

■

We can now give the following naive upper bound.

Proposition 14: For any $n, t \in \mathbb{N}$, $t \geqslant 2$, prime power q, and $0 \leqslant \rho \leqslant 1$,

$$\kappa_t(\rho, q) \leqslant 1 - H_q\left(\frac{\rho}{t}\right).$$

Proof: By Proposition 13,

$$\kappa_t(\rho, q) \leqslant \kappa_1\left(\frac{\rho}{t}, q\right).$$

We then combine (6) with the fact that $t \geqslant 2$ implies $\frac{\rho}{t} \leqslant 1 - \frac{1}{q}$, to obtain the desired result. ■

Proposition 13 is in fact a consequence of the following, more general, upper bound. This upper bound shows the generalized covering radii are sub-additive.

Proposition 15: Let C be an $[n, k]$ code over $\mathbb{F}_q$. Then for all $t_1, t_2 \in \mathbb{N}$,

$$R_{t_1+t_2} \leqslant R_{t_1} + R_{t_2}.$$

Proof: Let H be a parity-check matrix for C. Given $S \in \binom{\mathbb{F}_q^{n-k}}{t_1+t_2}$, partition it arbitrarily to $S = S_1 \cup S_2$, where $|S_1| = t_1$ and $|S_2| = t_2$. By Definition 1 there exist

$$I_1 \in \binom{[n]}{R_{t_1}} \quad \text{and} \quad I_2 \in \binom{[n]}{R_{t_2}},$$

such that $S_1 \subseteq \langle H_{I_1} \rangle$, and $S_2 \subseteq \langle H_{I_2} \rangle$. Define $I \triangleq I_1 \cup I_2$, then $S = S_1 \cup S_2 \subseteq \langle H_I \rangle$. The claim now follows. ■

B. Upper Bounding the Binary Case With $t = 2$

The upper bound we now present improves upon the trivial one from Proposition 14. Since it is significantly more complex, and has many moving parts, we focus on the binary case with $t = 2$ only. We follow a similar strategy to the one employed by [6, Theorem 12.3.5] for the covering radius, though major adjustments are required due to the more involved nature of this generalized problem. In essence, we show the existence of a covering code using the probabilistic method. The probability is nearly 1, implying almost *all* codes are at least as good as this bound. The main result is Theorem 22.

We outline the proof strategy to facilitate reading this section. We use the probabilistic method by choosing a random generator matrix for a code and bounding the probability that balls centered at the codewords indeed cover the entire space. To do so, we study the random variable that counts how many codewords cover a given point in space. To get a handle on this variable, we bound its expectation and variance.

We first recall the following useful lemma from [7, Lemma 1] concerning the average intersection of a set with its translations. Though originally proved for vectors, it also holds (with exactly the same proof) for matrices.

Lemma 16 [7]: For any $S \subseteq \mathbb{F}_q^{t\times n}$,

$$\frac{1}{q^{tn}} \sum_{\mathbf{v}\in\mathbb{F}_q^{t\times n}} |S \cap (S+\mathbf{v})| = \frac{|S|^2}{q^{tn}}.$$

Let $k, n \in \mathbb{N}$ such that $n \geqslant 2$ and $t \leqslant k < n$. We consider the random matrix $G \in \mathbb{F}_2^{k\times n}$, with rows $\overline{g}_1, \dots, \overline{g}_k$ independently and uniformly drawn from $\mathbb{F}_2^n$. Let C be the random code with generator matrix G.

For a matrix $\mathbf{u} \in \mathbb{F}_2^{t\times k}$, let $\mathbf{c_u} \in C^t$ be defined by $\mathbf{c_u} \triangleq \mathbf{u}G$. Clearly,

$$C^t = \left\{\mathbf{c_u} \middle| \mathbf{u} \in \mathbb{F}_2^{t\times k}\right\}.$$

The next lemma shows a connection between the rank of $\mathbf{u}$ and the statistical independence of the rows of $\mathbf{c_u}$. We remark that the probability of $\mathbf{u}$ being a full rank matrix goes to 1 as $k \to \infty$.

Lemma 17: If $\mathbf{u} \in \mathbb{F}_2^{t\times k}$ has full rank, then $\mathbf{c_u}$ is uniformly distributed on $\mathbb{F}_2^{t\times n}$. In particular, the rows of $\mathbf{c_u}$ are statistically independent.

Proof: Consider the function $f_{\mathbf{u}} : \mathbb{F}_2^{k\times n} \to \mathbb{F}_2^{t\times n}$ given by $f_{\mathbf{u}}(A) = \mathbf{u}A$. Since $\mathbf{u}$ has full rank, $f_{\mathbf{u}}$ is surjective and it is $2^{(k-t)n}$ to one. Hence, for any subset $S \subseteq \mathbb{F}_2^{t\times n}$, the size of the pre-image $f_{\mathbf{u}}^{-1}(S)$ is $2^{(k-t)n}|S|$. We recall that the generator matrix G is uniformly distributed on $\mathbb{F}_2^{k\times n}$. Hence,

$$\begin{aligned}\mathbb{P}[\mathbf{c_u} \in S] &= \mathbb{P}[\mathbf{u}G \in S] = \mathbb{P}[f_{\mathbf{u}}(G) \in S]\\ &= \mathbb{P}\left[G \in f_{\mathbf{u}}^{-1}(S)\right] = \frac{\left|f_{\mathbf{u}}^{-1}(S)\right|}{2^{kn}}\\ &= \frac{2^{(k-t)n}|S|}{2^{kn}} = \frac{|S|}{2^{tn}} = \frac{|S|}{\left|\mathbb{F}_2^{t\times n}\right|}.\end{aligned}$$

This completes the proof. ■

In preparation for bounding the variance of a certain random variable yet to be defined, we shall need to study the probability that pairs of codewords reside in the same ball. For $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{F}_2^{2\times k}$, we consider the matrix $\mathbf{c}_{\mathbf{u}_1,\mathbf{u}_2} \in \mathbb{F}_2^{4\times k}$ defined by

$$\mathbf{c}_{\mathbf{u}_1,\mathbf{u}_2} \triangleq \begin{bmatrix}\mathbf{c}_{\mathbf{u}_1}\\ \mathbf{c}_{\mathbf{u}_2}\end{bmatrix}.$$

We first show that the probability the two codewords are contained in the same ball is maximized by the ball centered at $\mathbf{0}$.

Lemma 18: Let $1 \leqslant r \leqslant n-k$ be an integer and $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{F}_2^{2\times k}$ with full rank, such that $\operatorname{rank}\left[\begin{smallmatrix}\mathbf{u}_1\\ \mathbf{u}_2\end{smallmatrix}\right] = 3$. Then for any $\mathbf{v} \in \mathbb{F}_2^{2\times n}$ we have

$$\mathbb{P}\left[\mathbf{c}_{\mathbf{u}_1,\mathbf{u}_2} \in \left(B_{r,n,2}^{(2)}(\mathbf{v})\right)^2\right] \leqslant \mathbb{P}\left[\mathbf{c}_{\mathbf{u}_1,\mathbf{u}_2} \in \left(B_{r,n,2}^{(2)}(\mathbf{0})\right)^2\right],$$

where

$$\left(B_{r,n,2}^{(2)}(\mathbf{v})\right)^2 = B_{r,n,2}^{(2)}(\mathbf{v}) \times B_{r,n,2}^{(2)}(\mathbf{v}) \subseteq \mathbb{F}_2^{4\times n}$$

is the Cartesian product of the ball $B_{r,n,2}^{(2)}(\mathbf{v})$ with itself.

Proof: Let $\overline{u}_1, \overline{u}_2, \overline{u}_3, \overline{u}_4$ denote the rows of $\left[\begin{smallmatrix}\mathbf{u}_1\\ \mathbf{u}_2\end{smallmatrix}\right]$ and $\mathbf{v} = \left[\begin{smallmatrix}\overline{v}_1\\ \overline{v}_2\end{smallmatrix}\right] \in \mathbb{F}_2^{2\times n}$. Without loss of generality, we assume that $\overline{u}_1$, $\overline{u}_2$, and $\overline{u}_3$, are linearly independent. By this assumption,

$$\overline{u}_4 = a_1 \cdot \overline{u}_1 + a_2 \cdot \overline{u}_2 + a_3 \cdot \overline{u}_3,$$

for some $a_1, a_2, a_3 \in \mathbb{F}_2$. Let $\overline{c}_1, \overline{c}_2$ and $\overline{c}_3, \overline{c}_4$ be the rows of $\mathbf{c}_{\mathbf{u}_1}$ and $\mathbf{c}_{\mathbf{u}_2}$, respectively. We have

$$\begin{bmatrix}\overline{c}_1\\ \overline{c}_2\\ \overline{c}_3\\ \overline{c}_4\end{bmatrix} = \begin{bmatrix}\mathbf{c}_{\mathbf{u}_1}\\ \mathbf{c}_{\mathbf{u}_2}\end{bmatrix} = \begin{bmatrix}\mathbf{u}_1\\ \mathbf{u}_2\end{bmatrix} \cdot G = \begin{bmatrix}\overline{u}_1\\ \overline{u}_2\\ \overline{u}_3\\ a_1\overline{u}_1 + a_2\overline{u}_2 + a_3\overline{u}_3\end{bmatrix} \cdot G,$$

where G is the random generator matrix of the code. Thus,

$$\overline{c}_4 = a_1 \cdot \overline{c}_1 + a_2 \cdot \overline{c}_2 + a_3 \cdot \overline{c}_3,$$

and by Lemma 17, $\left[\begin{smallmatrix}\overline{c}_1\\ \overline{c}_2\\ \overline{c}_3\end{smallmatrix}\right]$ is uniformly distributed on $F_2^{3\times n}$.

We define

$$\begin{bmatrix}\widetilde{c}_1\\ \widetilde{c}_2\\ \widetilde{c}_3\\ \widetilde{c}_4\end{bmatrix} \triangleq \begin{bmatrix}\overline{c}_1\\ \overline{c}_2\\ \overline{c}_3\\ \overline{c}_4\end{bmatrix} - \begin{bmatrix}\overline{v}_1\\ \overline{v}_2\\ \overline{v}_1\\ \overline{v}_2\end{bmatrix} = \mathbf{c}_{\mathbf{u}_1,\mathbf{u}_2} - \begin{bmatrix}\mathbf{v}\\ \mathbf{v}\end{bmatrix}.$$

By the translations invariance of the metric $d^{(2)}$,

$$\mathbf{c}_{\mathbf{u}_1,\mathbf{u}_2} \in \left(B_{r,n,2}^{(2)}(\mathbf{v})\right)^2 \iff \mathbf{c}_{\mathbf{u}_1,\mathbf{u}_2} - \begin{bmatrix}\mathbf{v}\\ \mathbf{v}\end{bmatrix} \in \left(B_{r,n,2}^{(2)}(\mathbf{0})\right)^2.$$

We note that the map $\psi : \mathbb{F}_2^{3\times n} \to \mathbb{F}_2^{3\times n}$ given by

$$\psi(\mathbf{z}) = \mathbf{z} - \begin{bmatrix}\overline{v}_1\\ \overline{v}_2\\ \overline{v}_1\end{bmatrix}$$

is a bijection, and therefore,

$$\begin{bmatrix}\widetilde{c}_1\\ \widetilde{c}_2\\ \widetilde{c}_3\end{bmatrix} = \psi\left(\begin{bmatrix}\overline{c}_1\\ \overline{c}_2\\ \overline{c}_3\end{bmatrix}\right)$$

is uniformly distributed on $\mathbb{F}_2^{3\times n}$ as well.

We divide our analysis into cases, depending on the value of (a_1, a_2, a_3). Since $\mathrm{rank}(\mathbf{u}_1) = \mathrm{rank}(\mathbf{u}_2) = 2$, and rank $\left[\begin{smallmatrix}\mathbf{u}_1\\ \mathbf{u}_2\end{smallmatrix}\right] = 3$, the combinations $(a_1, a_2, a_3) = (0, 0, 0)$ and $(a_1, a_2, a_3) = (0, 0, 1)$ are impossible.

Case 1: If $(a_1, a_2, a_3) = (1, 1, 1)$, a simple calculation shows that we have,

$$\widetilde{c}_4 = \widetilde{c}_1 + \widetilde{c}_2 + \widetilde{c}_3.$$

Thus, $\mathbf{c}_{\mathbf{u}_1,\mathbf{u}_2}$ and $\mathbf{c}_{\mathbf{u}_1,\mathbf{u}_2} - \left[\begin{smallmatrix}\mathbf{v}\\ \mathbf{v}\end{smallmatrix}\right]$ have the same distribution, so

$$\begin{aligned}\mathbb{P}&\left[\mathbf{c}_{\mathbf{u}_1,\mathbf{u}_2} \in \left(B^{(2)}_{r,n,2}(\mathbf{v})\right)^2\right]\\ &= \mathbb{P}\left[\mathbf{c}_{\mathbf{u}_1,\mathbf{u}_2} - \begin{bmatrix}\mathbf{v}\\ \mathbf{v}\end{bmatrix} \in \left(B^{(2)}_{r,n,2}(\mathbf{0})\right)^2\right]\\ &= \mathbb{P}\left[\mathbf{c}_{\mathbf{u}_1,\mathbf{u}_2} \in \left(B^{(2)}_{r,n,2}(\mathbf{0})\right)^2\right].\end{aligned}$$

Case 2: If $(a_1, a_2, a_3) = (1, 1, 0)$, a similar calculation as in the previous case shows that

$$\widetilde{c}_4 = \widetilde{c}_1 + \widetilde{c}_2 + \overline{v}_1.$$

For any $\overline{z} \in \mathbb{F}_2^n$ we consider the set

$$S_{\overline{z}} \triangleq \left\{ \begin{bmatrix}\overline{w}_1\\ \overline{w}_2\\ \overline{w}_3\end{bmatrix} \in \mathbb{F}_2^{3\times n} \;\middle|\; \begin{bmatrix}\overline{w}_1\\ \overline{w}_2\\ \overline{w}_3\\ \overline{w}_1 + \overline{w}_2 + \overline{z}\end{bmatrix} \in \left(B^{(2)}_r(\mathbf{0})\right)^2 \right\}.$$

Since $\left[\begin{smallmatrix}\widetilde{c}_1\\ \widetilde{c}_2\\ \widetilde{c}_3\end{smallmatrix}\right]$ and $\left[\begin{smallmatrix}\overline{c}_1\\ \overline{c}_2\\ \overline{c}_3\end{smallmatrix}\right]$ are uniformly distributed, to prove the theorem's claim is equivalent to showing that $\left|S_{\overline{0}}\right| \geqslant \left|S_{\overline{v}_1}\right|$, which is also equivalent to $\left|S_{\overline{0}} \setminus S_{\overline{v}_1}\right| \geqslant \left|S_{\overline{v}_1} \setminus S_{\overline{0}}\right|$.

If $\overline{v}_1 = \overline{0}$, this condition is automatically satisfied. Otherwise, we will prove our claim by showing that for $\overline{v}'_1$ obtained by zeroing one of the bits of $\overline{v}_1$ we have $\left|S_{\overline{v}'_1} \setminus S_{\overline{v}_1}\right| \geqslant \left|S_{\overline{v}_1} \setminus S_{\overline{v}'_1}\right|$. Then, repeating this arguments and zeroing all the non-zero bits of $\overline{v}_1$ we conclude the desired inequality.

Indeed, we find an injection $S_{\overline{v}_1} \setminus S_{\overline{v}'_1} \to S_{\overline{v}'_1} \setminus S_{\overline{v}_1}$. Let $i \in [n]$ be an index such that the i-th bit of $\overline{v}_1$ is 1. Denote by $\overline{e}_i$ the i-th standard unit vector, and set $\overline{v}'_1 = \overline{v}_1 + \overline{e}_i$. Let $\left[\begin{smallmatrix}\overline{w}_1\\ \overline{w}_2\\ \overline{w}_3\end{smallmatrix}\right] \in S_{\overline{v}_1} \setminus S_{\overline{v}'_1}$. We have

$$\begin{bmatrix}\overline{w}_1\\ \overline{w}_2\end{bmatrix} \in B^{(2)}_r(\mathbf{0}), \quad \begin{bmatrix}\overline{w}_3\\ \overline{w}_1 + \overline{w}_2 + \overline{v}_1\end{bmatrix} \in B^{(2)}_r(\mathbf{0})$$

and

$$\begin{bmatrix}\overline{w}_3\\ \overline{w}_1 + \overline{w}_2 + \overline{v}_1 + \overline{e}_i\end{bmatrix} \notin B^{(2)}_r(\mathbf{0}).$$

Thus,

$$\mathrm{supp}(\overline{w}_1 + \overline{w}_2 + \overline{v}_1) \subsetneqq \mathrm{supp}(\overline{w}_1 + \overline{w}_2 + \overline{v}_1 + \overline{e}_i).$$

Since the i-th bit of $\overline{v}_1 + \overline{e}_i$ is 0, it is only possible if the i-th bit of $\overline{w}_1 + \overline{w}_2$ is 1. Hence,

$$i \in \mathrm{supp}\left(\begin{bmatrix}\overline{w}_1\\ \overline{w}_2\end{bmatrix}\right). \tag{7}$$

We define

$$\phi\begin{bmatrix}\overline{w}_1\\ \overline{w}_2\\ \overline{w}_3\end{bmatrix} \triangleq \begin{bmatrix}\overline{w}_1 + \overline{e}_i\\ \overline{w}_2\\ \overline{w}_3\end{bmatrix}.$$

By (7) we have,

$$\mathrm{supp}\left(\phi\begin{bmatrix}\overline{w}_1\\ \overline{w}_2\\ \overline{w}_3\end{bmatrix}\right) \subseteq \mathrm{supp}\left(\begin{bmatrix}\overline{w}_1\\ \overline{w}_2\\ \overline{w}_3\end{bmatrix}\right).$$

Hence,

$$\begin{bmatrix}\overline{w}_1 + \overline{e}_i\\ \overline{w}_2\end{bmatrix} \in B^{(2)}_r(\mathbf{0}).$$

Furthermore, $(\overline{w}_1 + \overline{w}_2 + \overline{e}_i) + \overline{v}'_1 = \overline{w}_1 + \overline{w}_2 + \overline{v}_1$, and so

$$\begin{bmatrix}\overline{w}_3\\ (\overline{w}_1 + \overline{w}_2 + \overline{e}_i) + \overline{v}'_1\end{bmatrix} = \begin{bmatrix}\overline{w}_3\\ \overline{w}_1 + \overline{w}_2 + \overline{v}_1\end{bmatrix} \in B^{(2)}_r(\mathbf{0}).$$

That is, $\phi\left[\begin{smallmatrix}\overline{w}_1\\ \overline{w}_2\\ \overline{w}_3\end{smallmatrix}\right] \in S_{\overline{v}'_1}$. On the other hand,

$$\begin{bmatrix}\overline{w}_3\\ (\overline{w}_1 + \overline{w}_2 + \overline{e}_i) + \overline{v}_1\end{bmatrix} = \begin{bmatrix}\overline{w}_3\\ \overline{w}_1 + \overline{w}_2 + \overline{v}'_1\end{bmatrix} \notin B^{(2)}_r(\mathbf{0}).$$

Hence, $\phi\left[\begin{smallmatrix}\overline{w}_1\\ \overline{w}_2\\ \overline{w}_3\end{smallmatrix}\right] \notin S_{\overline{v}_1}$. This shows that ϕ maps $S_{\overline{v}_1} \setminus S_{\overline{v}'_1}$ to $S_{\overline{v}'_1} \setminus S_{\overline{v}_1}$. Clearly ϕ is injective and it is the desired map.

Case 3: If $(a_1, a_2, a_3) = (1, 0, 1)$ we have $\overline{c}_4 = \overline{c}_3 + \overline{c}_1$, or equivalently, $\overline{c}_1 = \overline{c}_3 + \overline{c}_4$. This case is equivalent to the case where $(a_1, a_2, a_3) = (1, 1, 0)$ with $\overline{c}_1, \overline{c}_2$ and $\overline{c}_3, \overline{c}_4$ switching rolls.

Case 4: If $(a_1, a_2, a_3) = (0, 1, 1)$ this is equivalent to the case where $(a_1, a_2, a_3) = (1, 1, 0)$.

Case 5: If $(a_1, a_2, a_3) = (0, 1, 0)$ we have

$$\widetilde{c}_4 = \widetilde{c}_2.$$

Thus, $\mathbf{c}_{\mathbf{u}_1,\mathbf{u}_2}$ and $\mathbf{c}_{\mathbf{u}_1,\mathbf{u}_2} - \left[\begin{smallmatrix}\mathbf{v}\\ \mathbf{v}\end{smallmatrix}\right]$ have the same distribution, and the case is completed as Case 1.

Case 6: If $(a_1, a_2, a_3) = (1, 0, 0)$ then we have

$$\widetilde{c}_4 = \widetilde{c}_1 + (\overline{v}_1 + \overline{v}_2).$$

Similarly to Case 2, where $(a_1, a_2, a_3) = (1, 1, 0)$, we show that $\left|S_{\overline{0}}\right| \geqslant \left|S_{\overline{v}_1+\overline{v}_2}\right|$. We use the same technique in order to show that we increase $\left|S_{\overline{v}_1+\overline{v}_2}\right|$ when we flip a bit of $\overline{v}_1 + \overline{v}_2$ from 1 to 0, and the same mapping ϕ. ∎

For any $\mathbf{v} \in \mathbb{F}_2^{2\times n}$ we define $X_{\mathbf{v}}$ to be the number of codewords in C^2 that are generated from full-rank coefficients matrices, and that r-cover $\mathbf{v}$. Formally,

$$X_{\mathbf{v}} \triangleq \sum_{\substack{\mathbf{u}\in\mathbb{F}_2^{2\times k}\\ \mathrm{rank}(\mathbf{u})=2}} \mathbb{I}_{\left\{\mathbf{v}\in B^{(2)}_r(\mathbf{c}_{\mathbf{u}})\right\}},$$

where $\mathbb{I}_A$ is the indicator function of the event A. Clearly, $X_{\mathbf{v}}$ depends on n, k, and r, although we omit them from our notation. The random variable $X_{\mathbf{v}}$ plays an important role in our main result, and we study its properties in preparation for the main theorem. We first bound the expectation of $X_{\mathbf{v}}$.

Lemma 19: For $0 \leqslant \rho < \frac{3}{4}$, $n, k, r \in \mathbb{N}$, $3 \leqslant k \leqslant n$, $r = \rho n$, and $\mathbf{v} \in \mathbb{F}_2^{2\times n}$,

$$V^{(2)}_{r,n,2} \cdot 2^{2k-1-2n} < \mathrm{E}[X_{\mathbf{v}}] < V^{(2)}_{r,n,2} \cdot 2^{2k-2n}.$$

Proof: By Lemma 17 for $\mathbf{u} \in \mathbb{F}_2^{2\times k}$ with full rank, $\mathbf{c_u}$ is uniformly distributed on $\mathbb{F}_2^{2\times n}$. Therefore,

$$\begin{aligned}
\mathrm{E}[X_{\mathbf{v}}] &= \sum_{\substack{\mathbf{u}\in\mathbb{F}_2^{2\times k}\\ \mathrm{rank}(\mathbf{u})=2}} \mathrm{E}\left[\mathbb{I}_{\left\{\mathbf{v}\in B_r^{(2)}(\mathbf{c_u})\right\}}\right] \\
&= \sum_{\substack{\mathbf{u}\in\mathbb{F}_2^{2\times k}\\ \mathrm{rank}(\mathbf{u})=2}} \mathbb{P}\left[\mathbf{v}\in B_r^{(2)}(\mathbf{c_u})\right] \\
&= \sum_{\substack{\mathbf{u}\in\mathbb{F}_2^{2\times k}\\ \mathrm{rank}(\mathbf{u})=2}} \frac{\left|B_r^{(2)}(\mathbf{c_u})\right|}{2^{2n}} = (2^k-1)(2^k-2)\frac{V_{r,n,2}^{(2)}}{2^{2n}}.
\end{aligned}$$

For $k \geqslant 3$ we have

$$2^{2k-1} < (2^k-1)(2^k-2) < 2^{2k},$$

which gives us the desired result. ■

We consider the functions $f_1, f_2 : [0,1] \to \mathbb{R}$ defined by

$$f_1(\rho) \triangleq \max_{\substack{0\leqslant\alpha\leqslant\rho\\ 0\leqslant\beta\leqslant\alpha\\ 0\leqslant\gamma\leqslant\rho-\alpha+\beta}} \left(H_2(\alpha) + \alpha H_2\left(\frac{\beta}{\alpha}\right) + 2(\alpha-\beta) + (1-\alpha+\beta)H_2\left(\frac{\gamma}{1-\alpha+\beta}\right)\right),$$

and

$$f_2(\rho) \triangleq \max_{\substack{0\leqslant\alpha\leqslant\rho\\ 0\leqslant\beta\leqslant\rho-\alpha}} \left(H_2(\alpha) + 2(1-\alpha)H_2\left(\frac{\beta}{1-\alpha}\right) + 2\alpha\right).$$

We then define

$$f(\rho) \triangleq \max\left(f_1(\rho), f_2(\rho)\right),$$

which we will use in order to bound $\mathrm{Var}(X_{\mathbf{v}})$.

Lemma 20: For $0 \leqslant \rho < \frac{3}{4}$, $n, k, r \in \mathbb{N}$, $k \leqslant n$, $r = \rho n$, and $\mathbf{v} \in \mathbb{F}_2^{2\times n}$,

$$\mathrm{Var}(X_{\mathbf{v}}) \leqslant 7\,\mathrm{E}[X_{\mathbf{v}}] + 2^{3(k-n)+n(f(\rho)+o(1))}$$

Proof: To simplify notation, we denote

$$\eta_{\mathbf{u}} \triangleq \mathbb{I}_{\left\{\mathbf{v}\in B_r^{(2)}(\mathbf{c_u})\right\}}.$$

We then calculate,

$$\begin{aligned}
\mathrm{Var}(X_{\mathbf{v}}) &= \mathrm{Var}\left(\sum_{\substack{\mathbf{u}\in\mathbb{F}_2^{2\times k}\\ \mathrm{rank}(\mathbf{u})=2}} \eta_{\mathbf{u}}\right) \\
&= \sum_{\substack{\mathbf{u}\in\mathbb{F}_2^{2\times k}\\ \mathrm{rank}(\mathbf{u})=2}} \mathrm{Var}\left(\eta_{\mathbf{u}}\right) \\
&\quad + \sum_{\substack{\mathbf{u}_1\in\mathbb{F}_2^{2\times k}\\ \mathrm{rank}(\mathbf{u}_1)=2}} \sum_{\substack{\mathbf{u}_2\in\mathbb{F}_2^{2\times k}\\ \mathrm{rank}(\mathbf{u}_2)=2}} \mathrm{Cov}\left(\eta_{\mathbf{u}_1}, \eta_{\mathbf{u}_2}\right). \qquad (8)
\end{aligned}$$

We separate the sums in (8) into four parts, and bound each one of them individually.

For a Bernoulli random variable $Z \sim \mathrm{Ber}(p)$ we have

$$\mathrm{Var}(Z) = \mathrm{E}[Z^2] - \mathrm{E}[Z]^2 = p - p^2 \leqslant p = \mathrm{E}[Z]. \qquad (9)$$

Applying this bound to the first sum of (8), we have

$$\sum_{\substack{\mathbf{u}\in\mathbb{F}_2^{2\times k}\\ \mathrm{rank}(\mathbf{u})=2}} \mathrm{Var}\left(\eta_{\mathbf{u}}\right) \leqslant \sum_{\substack{\mathbf{u}\in\mathbb{F}_2^{2\times k}\\ \mathrm{rank}(\mathbf{u})=2}} \mathrm{E}\left[\eta_{\mathbf{u}}\right] = \mathrm{E}[X_{\mathbf{v}}]. \qquad (10)$$

We now consider the double sum in (8), which we separate into three parts, according to $\mathrm{rank}\left(\left[\begin{smallmatrix}\mathbf{u}_1\\ \mathbf{u}_2\end{smallmatrix}\right]\right)$.

If $\mathrm{rank}\left(\left[\begin{smallmatrix}\mathbf{u}_1\\ \mathbf{u}_2\end{smallmatrix}\right]\right) = 4$, by Lemma 17, $\mathbf{c}_{\mathbf{u}_1}$ and $\mathbf{c}_{\mathbf{u}_2}$ are statistically independent and therefore so are $\eta_{\mathbf{u}_1}$ and $\eta_{\mathbf{u}_2}$. Thus, the covariance in that case is 0.

If $\mathrm{rank}(\mathbf{u}_1) = \mathrm{rank}(\mathbf{u}_2) = \mathrm{rank}\left(\left[\begin{smallmatrix}\mathbf{u}_1\\ \mathbf{u}_2\end{smallmatrix}\right]\right) = 2$, by the Cauchy-Schwarz inequality, Lemma 17 and (9) we have

$$\begin{aligned}
\mathrm{Cov}\left(\eta_{\mathbf{u}_1}, \eta_{\mathbf{u}_2}\right) &\leqslant \sqrt{\mathrm{Var}\left(\eta_{\mathbf{u}_1}\right)\mathrm{Var}\left(\eta_{\mathbf{u}_2}\right)} \\
&\leqslant \sqrt{\mathrm{E}\left[\eta_{\mathbf{u}_1}\right]\mathrm{E}\left[\eta_{\mathbf{u}_2}\right]} \\
&= \sqrt{\mathbb{P}\left[\mathbf{v}\in B_r^{(2)}(\mathbf{c}_{\mathbf{u}_1})\right]\mathbb{P}\left[\mathbf{v}\in B_r^{(2)}(\mathbf{c}_{\mathbf{u}_2})\right]} \\
&= \sqrt{\mathbb{P}\left[\mathbf{c}_{\mathbf{u}_1}\in B_r^{(2)}(\mathbf{v})\right]\mathbb{P}\left[\mathbf{c}_{\mathbf{u}_2}\in B_r^{(2)}(\mathbf{v})\right]} \\
&= \frac{\left|B_r^{(2)}(\mathbf{v})\right|}{2^{2n}} = \mathrm{E}\left[\eta_{\mathbf{u}_1}\right].
\end{aligned}$$

Since $\mathrm{rank}(\mathbf{u}_1) = \mathrm{rank}(\mathbf{u}_2)$, we have that $\mathbf{u}_2 = A\cdot\mathbf{u}_1$ where $A \in \mathrm{GL}(2,2)$ is a 2×2 invertible matrix over $\mathbb{F}_2$. Summation over all pairs $\mathbf{u}_1, \mathbf{u}_2$ such that $\mathrm{rank}(\mathbf{u}_1) = \mathrm{rank}(\mathbf{u}_2) = \mathrm{rank}\left(\left[\begin{smallmatrix}\mathbf{u}_1\\ \mathbf{u}_2\end{smallmatrix}\right]\right) = 2$ gives

$$\begin{aligned}
&\sum_{\substack{\mathbf{u}_1\in\mathbb{F}_2^{2\times k}\\ \mathrm{rank}(\mathbf{u}_1)=2}} \sum_{A\in\mathrm{GL}(2,2)} \mathrm{Cov}\left(\eta_{\mathbf{u}_1}, \eta_{A\mathbf{u}_1}\right) \\
&\leqslant \sum_{\substack{\mathbf{u}_1\in\mathbb{F}_2^{2\times k}\\ \mathrm{rank}(\mathbf{u}_1)=2}} \sum_{A\in\mathrm{GL}(2,2)} \mathrm{E}\left[\eta_{\mathbf{u}_1}\right] \\
&= \sum_{A\in\mathrm{GL}(2,2)} \sum_{\substack{\mathbf{u}_1\in\mathbb{F}_2^{2\times k}\\ \mathrm{rank}(\mathbf{u}_1)=2}} \mathrm{E}\left[\eta_{\mathbf{u}_1}\right] \\
&= \sum_{A\in\mathrm{GL}(2,2)} \mathrm{E}[X_{\mathbf{v}}] = 6\,\mathrm{E}[X_{\mathbf{v}}],
\end{aligned}$$

where the last equality follows since there are exactly six 2×2 invertible matrices over $\mathbb{F}_2$.

We are left with the case of $\mathrm{rank}\left(\left[\begin{smallmatrix}\mathbf{u}_1\\ \mathbf{u}_2\end{smallmatrix}\right]\right) = 3$. We start by bounding $\mathrm{Cov}(\eta_{\mathbf{u}_1}, \eta_{\mathbf{u}_2})$, and then evaluate this bound, dividing our analysis into three cases according to the linear dependence structure of $\mathbf{u}_1$ and $\mathbf{u}_2$.

By Lemma 18,

$$\begin{aligned}
\mathrm{Cov}(\eta_{\mathbf{u}_1}, \eta_{\mathbf{u}_2}) &= \mathrm{E}[\eta_{\mathbf{u}_1}\eta_{\mathbf{u}_2}] - \mathrm{E}[\eta_{\mathbf{u}_1}]\,\mathrm{E}[\eta_{\mathbf{u}_2}] \\
&\leqslant \mathrm{E}[\eta_{\mathbf{u}_1}\eta_{\mathbf{u}_2}] = \mathbb{P}\left[\mathbf{c}_{\mathbf{u}_1}, \mathbf{c}_{\mathbf{u}_1} \in B_r^{(2)}(\mathbf{v})\right] \\
&\leqslant \mathbb{P}\left[\mathbf{c}_{\mathbf{u}_1}, \mathbf{c}_{\mathbf{u}_1} \in B_r^{(2)}(\mathbf{0})\right].
\end{aligned}$$

As before, let $\overline{u}_1, \overline{u}_2, \overline{u}_3, \overline{u}_4$ denote the rows of $\left[\begin{smallmatrix}\mathbf{u}_1\\ \mathbf{u}_2\end{smallmatrix}\right]$. Without loss of generality, assume $\overline{u}_1, \overline{u}_2, \overline{u}_3$ are linearly independent and $\overline{u}_4 \in \langle \overline{u}_1, \overline{u}_2, \overline{u}_3 \rangle$, that is,

$$\overline{u}_4 = a_1 \cdot \overline{u}_1 + a_2 \cdot \overline{u}_2 + a_3 \cdot \overline{u}_3,$$

for some $a_1, a_2, a_3 \in \mathbb{F}_2$. Thus, $\left[\begin{smallmatrix}\overline{u}_1\\ \overline{u}_2\\ \overline{u}_3\end{smallmatrix}\right] G$ is uniformly distributed on $\mathbb{F}_2^{3\times n}$ and therefore $\mathbb{P}\left[\mathbf{c}_{\mathbf{u}_1}, \mathbf{c}_{\mathbf{u}_2} \in B_r^{(2)}(\mathbf{0})\right]$ is proportional to the number of triples $\overline{w}_1, \overline{w}_2, \overline{w}_3 \in \mathbb{F}_2^n$ such that

$$\begin{bmatrix}\overline{w}_1\\ \overline{w}_2\end{bmatrix}, \begin{bmatrix}\overline{w}_3\\ a_1 \cdot \overline{w}_1 + a_2 \cdot \overline{w}_2 + a_3 \cdot \overline{w}_3\end{bmatrix} \in B_r^{(2)}(\mathbf{0}).$$

We enumerate the number of such triples in every dependence structure, which we denote by $N(a_1, a_2, a_3)$. Since $\mathrm{rank}(\mathbf{u}_2) = 2$ the combinations $a_1 = a_2 = a_3 = 0$ and $a_1 = a_2 = 0$, $a_3 = 1$ are impossible. Thus we have six cases, and we will show that they can be reduced to three cases.

Case 1 - $(a_1, a_2, a_3) = (1, 1, 0)$: If $(a_1, a_2, a_3) = (1, 1, 0)$, then $\overline{w}_4 = \overline{w}_1 + \overline{w}_2$. Hence, the number of triples in this case is given by

$$N(1,1,0) = \sum_{i=0}^{r} \binom{n}{i} \sum_{j=0}^{i} \binom{i}{j} 2^{i-j} \sum_{\ell=0}^{r-i+j} \binom{n-i+j}{\ell} 2^{i-j}.$$

Here, the integer i runs over all possible values for the size of $\mathrm{supp}\left(\left[\begin{smallmatrix}\overline{w}_1\\ \overline{w}_2\end{smallmatrix}\right]\right)$, namely, between 0 and r, and $\binom{n}{i}$ counts the number of ways to choose this support. The integer j runs over all possible values for the number of overlapping 1's between $\overline{w}_1$ and $\overline{w}_2$, $\binom{i}{j}$ counts the number of ways to choose these overlapping positions, and 2^{i-j} counts the number of ways to distribute the remaining 1's between $\overline{w}_1$ and $\overline{w}_2$. The integer ℓ runs over all possible values of the number of non-overlapping 1's between $\overline{w}_3$ and $\overline{w}_1 + \overline{w}_2$, and $\binom{n-i+j}{\ell}$ counts the number of ways to choose those non-overlapping 1's in the remaining $n - (i - j)$ coordinates. Finally, 2^{i-j} counts the number of ways to choose overlapping 1's from $\mathrm{supp}(\overline{w}_1 + \overline{w}_2)$ to $\overline{w}_3$.

Case 2 - $(a_1, a_2, a_3) = (1, 1, 1)$: By similar calculations as in the first case we obtain,

$$N(1,1,1) = N(1,1,0).$$

Case 3 - $(a_1, a_2, a_3) = (1, 0, 0)$: In this case

$$N(1,0,0) = \sum_{i=0}^{r} \binom{n}{i} \left(\sum_{j=0}^{r-i} \binom{n-i}{j} 2^{i} \right)^2.$$

The other three cases can be reduced to one of the previous one. The case where $(a_1, a_2, a_3) = (0, 1, 0)$ is trivially equivalent to the case of $(1, 0, 0)$. If $(a_1, a_2, a_3) = (1, 0, 1)$ we have $\overline{w}_4 = \overline{w}_1 + \overline{w}_3$, and therefore $\overline{w}_1 = \overline{w}_3 + \overline{w}_4$. Thus, this case is equivalent to the case of $(1, 1, 0)$ with $\overline{w}_1, \overline{w}_2$ and $\overline{w}_3, \overline{w}_4$ switching parts. Similarly the case where $(a_1, a_2, a_3) = (0, 1, 1)$ is equivalent to the case of $(1, 1, 0)$.

We recall that $r = \rho n$. Fix some $0 \leqslant i \leqslant r$, $0 \leqslant j \leqslant i$ and $0 \leqslant \ell \leqslant \rho - i + j$. We denote

$$i = \alpha n \quad j = \beta n \quad \ell = \gamma n.$$

The constraints on i, j and ℓ impose

$$0 \leqslant \alpha \leqslant \rho, \quad 0 \leqslant \beta \leqslant \alpha, \quad 0 \leqslant \gamma \leqslant \rho - \alpha + \beta.$$

Using the well-known identity $\binom{n}{\alpha n} = 2^{n(H_2(\alpha)+o(1))}$ we have

$$\begin{aligned}\binom{n}{i}\binom{i}{j} 2^{i-j} \binom{n-i+j}{\ell} 2^{i-j} &= 2^{n\left(H_2(\alpha) + \alpha H_2\left(\frac{\beta}{\alpha}\right) + 2(\alpha-\beta) + (1-\alpha+\beta) H_2\left(\frac{\gamma}{1-\alpha+\beta}\right) + o(1)\right)} \\ &\leqslant 2^{n(f_1(\rho)+o(1))}.\end{aligned}$$

and therefore

$$\begin{aligned}\sum_{i=0}^{r} \binom{n}{i} \sum_{j=0}^{i} \binom{i}{j} 2^{i-j} \sum_{\ell=0}^{r-i+j} \binom{n-i+j}{\ell} 2^{i-j} \\ \leqslant n^3 2^{n(f_1(\rho)+o(1))} = 2^{n(f_1(\rho)+o(1))}.\end{aligned}$$

In a similar fashion we obtain

$$\sum_{i=0}^{r} \binom{n}{i} \left(\sum_{j=0}^{r-i} \binom{n-i}{j} 2^{i} \right)^2 \leqslant 2^{n(f_2(\rho)+o(1))}.$$

Combining the bounds we obtain

$$\begin{aligned}\mathrm{Cov}(\eta_{\mathbf{u}_1}, \eta_{\mathbf{u}_2}) &\leqslant \mathbb{P}\left[\mathbf{c}_{\mathbf{u}_1}, \mathbf{c}_{\mathbf{u}_1} \in B_r^{(2)}(\mathbf{0})\right] \\ &\leqslant \frac{2^{n(\max(f_1(\rho), f_2(\rho))+o(1))}}{2^{3n}} \\ &= 2^{n(f(\rho)-3+o(1))}.\end{aligned}$$

Summing overall $\mathbf{u}_1, \mathbf{u}_2$ such that $\mathrm{rank}\left(\left[\begin{smallmatrix}\mathbf{u}_1\\ \mathbf{u}_2\end{smallmatrix}\right]\right) = 3$ gives

$$\begin{aligned}&\sum_{\mathrm{rank}\left(\left[\begin{smallmatrix}\mathbf{u}_1\\ \mathbf{u}_2\end{smallmatrix}\right]\right)=3} \mathrm{Cov}(\eta_{\mathbf{u}_1}, \eta_{\mathbf{u}_2}) \\ &\leqslant \sum_{\mathrm{rank}\left(\left[\begin{smallmatrix}\mathbf{u}_1\\ \mathbf{u}_2\end{smallmatrix}\right]\right)=3} 2^{n(f(\rho)-3+o(1))} \\ &\leqslant \left(2^k - 1\right)\left(2^k - 2\right)\left(2^k - 4\right) 2^3 2^{n(f(\rho)-3+o(1))} \\ &\leqslant 2^{n(f(\rho)-3+o(1))+3k}.\end{aligned}$$

Summing the upper bounds on all of the parts in the sum (8) we obtain the desired bound and complete the proof. ■

The functions $f_1(\rho)$ and $f_2(\rho)$ are given by maximizing the two functions,

$$\begin{aligned}f_1'(\alpha, \beta, \gamma) &\triangleq \Bigg(H_2(\alpha) + \alpha H_2\left(\frac{\beta}{\alpha}\right) + 2(\alpha - \beta) \\ &\qquad + (1 - \alpha + \beta) H_2\left(\frac{\gamma}{1-\alpha+\beta}\right)\Bigg), \\ f_2'(\alpha, \beta) &\triangleq \left(H_2(\alpha) + 2(1-\alpha) H_2\left(\frac{\beta}{1-\alpha}\right) + 2\alpha\right).\end{aligned}$$

We observe that the parameter ρ only controls the maximization domain. Using standard analysis techniques we can find the exact expression for $f(\rho)$. Let us denote

$$s(\rho) \triangleq \frac{1}{10}\left(1 + 8\rho - \sqrt{1 + 16\rho - 16\rho^2}\right).$$

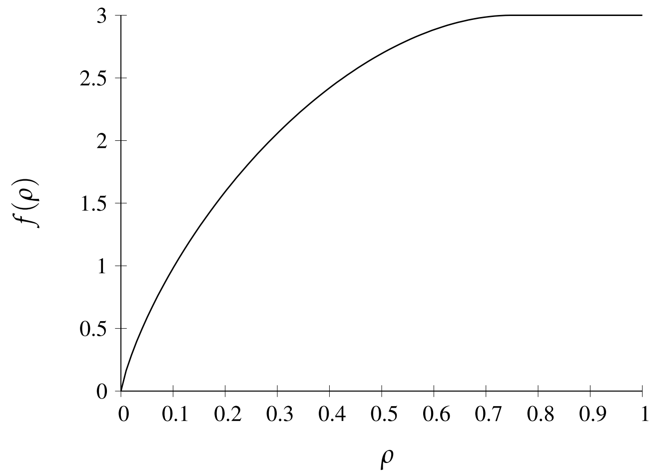

Fig. 1. The function $f(\rho)$ from Lemma 20.

For $0 \leqslant \rho < \frac{3}{4}$, f_1' is maximized at the point

$$\left(\alpha_1^{\max}(\rho), \beta_1^{\max}(\rho), \gamma_1^{\max}(\rho)\right) = (\rho, \rho - s(\rho), s(\rho)),$$

and for $\frac{3}{4} \leqslant \rho \leqslant 1$ it is maximized at

$$\left(\alpha_1^{\max}(\rho), \beta_1^{\max}(\rho), \gamma_1^{\max}(\rho)\right) = \left(\frac{3}{4}, \frac{1}{4}, \frac{1}{4}\right).$$

Similarly, for $0 \leqslant \rho < \frac{3}{4}$, f_2' is maximized at the point

$$\left(\alpha_2^{\max}(\rho), \beta_2^{\max}(\rho)\right) = (s(\rho), \rho - s(\rho)),$$

and for $\frac{3}{4} \leqslant \rho \leqslant 1$ it is maximized at

$$\left(\alpha_2^{\max}(\rho), \beta_2^{\max}(\rho)\right) = \left(\frac{1}{2}, \frac{1}{4}\right).$$

Curiously, f_1' and f_2' take the same value at their maximum points for any $0 \leqslant \rho \leqslant 1$, and therefore

$$\begin{aligned} f(\rho) &= f_1'\left(\alpha_1^{\max}(\rho), \beta_1^{\max}(\rho), \gamma_1^{\max}(\rho)\right) \\ &= f_2'\left(\alpha_2^{\max}(\rho), \beta_2^{\max}(\rho)\right) \\ &= \begin{cases} H_2(s(\rho)) + 2s(\rho) & \\ \quad +2(1-s(\rho))H_2\left(\frac{\rho - s(\rho)}{1-s(\rho)}\right) & 0 \leqslant \rho < \frac{3}{4}, \\ 3 & \frac{3}{4} \leqslant \rho \leqslant 1. \end{cases} \end{aligned}$$

The function $f(\rho)$ is shown in Figure 1.

We now use the obtained bounds on the expectation and variance of $X_{\mathbf{v}}$ in order to prove (under certain conditions) that with high probability, all $X_{\mathbf{v}}$ are positive. Namely, the entire space is covered.

Proposition 21: Let $0 \leqslant \rho < \frac{3}{4}$, and let $(k_n)_{n=1}^{\infty}$ be a sequence of integers such that

$$k_n > n - \log_4\left(V_{\rho n, n, 2}^{(2)}\right) + \log_2(n)$$

and

$$\limsup_{n\to\infty} \max_{\overline{v}\in\mathbb{F}_2^{2\times n}} \frac{\log_2(\operatorname{Var}(X_{\mathbf{v}}))}{\log_2(\operatorname{E}[X_{\mathbf{v}}])} < s < 2,$$

where $X_{\mathbf{v}}$ is defined with respect to n, k_n, and $r_n = \rho n$.

Let C_n' be the random code with a uniformly distributed $k_n' \times n$ random generator matrix G', where $k_n' = k_n + 2\left\lceil \log_2(n) \right\rceil + 2$. Then

$$R_2(C_n') \leqslant r_n = \rho n$$

with probability that tends to 1 as $n \to \infty$.

Proof: Let G be the matrix obtained by taking the first k_n rows of G', and $\overline{g}_1, \dots, \overline{g}_m$ be the remaining rows, $m = 2\left\lceil \log_2(n) \right\rceil + 2$. We fix some $\varepsilon > 0$ and $\mathbf{v} \in F_2^{2\times n}$. Consider the random code C_n with random generator matrix G. By Chebyshev's inequality

$$\mathbb{P}\left[|X_{\mathbf{v}} - \operatorname{E}[X_{\mathbf{v}}]| \geqslant 2^{\varepsilon} \cdot \operatorname{E}[X_{\mathbf{v}}]^{s/2}\right] \leqslant \frac{\operatorname{Var}(X_{\mathbf{v}})}{2^{2\varepsilon}\operatorname{E}[X_{\mathbf{v}}]^{s}}.$$

By assumption, for sufficiently large n, $\operatorname{Var}(X_{\mathbf{v}}) < \operatorname{E}[X_{\mathbf{v}}]^s$, and therefore

$$\mathbb{P}\left[|X_{\mathbf{v}} - \operatorname{E}[X_{\mathbf{v}}]| \geqslant 2^{\varepsilon} \cdot \operatorname{E}[X_{\mathbf{v}}]^{s/2}\right] < 2^{-2\varepsilon}. \tag{11}$$

We denote $\beta \triangleq \operatorname{E}[X_{\mathbf{v}}]$ and define $\beta(\varepsilon) \triangleq \beta - 2^{\varepsilon}\beta^{s/2}$. From (11) it follows that for sufficiently large n,

$$\mathbb{P}[X_{\mathbf{v}} \leqslant \beta(\varepsilon)] < 2^{-2\varepsilon}.$$

Define the set Q_0 to be the set of elements in $\mathbb{F}_2^{2\times n}$ that are r_n-covered by at most $\beta(\varepsilon)$ codewords from the set $\left\{\mathbf{c}_{\mathbf{u}} \middle| \mathbf{u} \in \mathbb{F}_2^{2\times k}, \operatorname{rank}(\mathbf{u}) = 2\right\} \subseteq C_n^2$. Formally,

$$Q_0 \triangleq \left\{\mathbf{v} \in \mathbb{F}_2^{2\times n} \middle| X_{\mathbf{v}} \leqslant \beta(\varepsilon)\right\}.$$

Let

$$q_0 \triangleq \frac{|Q_0|}{2^{2n}}$$

denote the proportion of Q_0 inside $\mathbb{F}_2^{2\times n}$. We have

$$q_0 = \frac{1}{2^{2n}} \sum_{\mathbf{v}\in\mathbb{F}_2^{2\times n}} \mathbb{I}_{\{X_{\mathbf{v}} \leqslant \beta(\varepsilon)\}},$$

and therefore, for sufficiently large n,

$$\operatorname{E}[q_0] = \mathbb{P}[X_{\mathbf{v}} \leqslant \beta(\varepsilon)] < 2^{-2\varepsilon}. \tag{12}$$

By Markov's inequality and (12) we have

$$\mathbb{P}\left[q_0 \geqslant 2^{-\varepsilon}\right] \leqslant \frac{\operatorname{E}[q_0]}{2^{-\varepsilon}} < 2^{-\varepsilon}. \tag{13}$$

For a set $V \subseteq \mathbb{F}_2^{2\times n}$ we define $Q(V)$ to be the set of r_n-remote points from V. That is,

$$Q(V) \triangleq \mathbb{F}_2^{2\times n} \setminus \bigcup_{\mathbf{v}\in V} B_{r_n}^{(2)}(\mathbf{v}).$$

Clearly, if $\beta(\varepsilon) \geqslant 0$, we have

$$Q(C_n) \subseteq Q_0. \tag{14}$$

We will later choose ε such that $\beta(\varepsilon) \geqslant 0$.

For arbitrary $\overline{x}, \overline{y} \in \mathbb{F}_2^n$ we consider the linear code generated by adding $\overline{x}$ and $\overline{y}$, which is $C_n + \langle \overline{x}, \overline{y} \rangle$. From the definition of Q we have,

$$\begin{aligned} Q\left((C_n + \langle \overline{x}, \overline{y}\rangle)^2\right) &\subseteq Q\left(C_n^2 \cup \left(C_n^2 + \begin{bmatrix}\overline{x}\\ \overline{y}\end{bmatrix}\right)\right) \\ &= Q\left(C_n^2\right) \cap Q\left(C_n^2 + \begin{bmatrix}\overline{x}\\ \overline{y}\end{bmatrix}\right) \\ &= Q\left(C_n^2\right) \cap \left(Q\left(C_n^2\right) + \begin{bmatrix}\overline{x}\\ \overline{y}\end{bmatrix}\right), \end{aligned}$$

where the last equality follows from the invariance of $d^{(2)}$ under translations. Hence, we have

$$\left|Q\left((C_n + \langle \overline{x}, \overline{y}\rangle)^2\right)\right| \leqslant \left|Q\left(C_n^2\right) \cap \left(Q\left(C_n^2\right) + \begin{bmatrix}\overline{x}\\ \overline{y}\end{bmatrix}\right)\right|.$$

Combining this with (14) and Lemma 16 we have

$$\begin{aligned} &\frac{1}{2^{2n}} \sum_{\overline{x},\overline{y}\in\mathbb{F}_2^n} \left|Q\left((C_n + \langle \overline{x}, \overline{y}\rangle)^2\right)\right| \\ &\quad\leqslant \frac{1}{2^{2n}} \sum_{\overline{x},\overline{y}\in\mathbb{F}_2^n} \left|Q\left(C_n^2\right) \cap \left(Q\left(C_n^2\right) + \begin{bmatrix}\overline{x}\\ \overline{y}\end{bmatrix}\right)\right| \\ &\quad= \frac{1}{2^{2n}} \cdot \frac{\left|Q\left(C_n^2\right)\right|^2}{2^{2n}} \leqslant \frac{1}{2^{2n}} \cdot \frac{|Q_0|^2}{2^{2n}} = q_0^2. \end{aligned} \tag{15}$$

Recall the first k_n rows of G' make up G, and the remaining rows are denoted by $\overline{g}_1, \ldots, \overline{g}_m$. For $1 \leqslant \ell \leqslant \frac{m}{2}$ we denote

$$q_\ell \triangleq \frac{\left|Q\left((C_n + \langle \overline{g}_1, \ldots, \overline{g}_{2\ell}\rangle)^2\right)\right|}{2^{2n}}.$$

Since the rows of G and the remaining rows $\overline{g}_1, \ldots, \overline{g}_m$ are independent and uniformly distributed, (15) implies that

$$\mathrm{E}[q_1; q_0] \leqslant q_0^2, \tag{16}$$

and by a similar argument,

$$\mathrm{E}[q_\ell; q_{\ell-1}] \leqslant q_{\ell-1}^2. \tag{17}$$

We fix some $\lambda > 0$ and bound $\mathbb{P}\left[q_1 \leqslant 2^{\lambda-2\varepsilon}\right]$ from below. By Markov's inequality, the law of total probability, and (13), we have

$$\begin{aligned} &\mathbb{P}\left[q_1 \leqslant 2^{\lambda-2\varepsilon}\right] \\ &\quad\geqslant \mathbb{P}[q_1 \leqslant 2^{\lambda-2\varepsilon}; q_0 \leqslant 2^{-\varepsilon}] \cdot \mathbb{P}\left[q_0 \leqslant 2^{-\varepsilon}\right] \\ &\quad\geqslant \left(1 - \mathbb{P}[q_1 > 2^{\lambda-2\varepsilon}; q_0 \leqslant 2^{-\varepsilon}]\right) \cdot \left(1 - 2^{-\varepsilon}\right) \\ &\quad\geqslant \left(1 - \frac{\mathrm{E}[q_1; q_0 \leqslant 2^{-\varepsilon}]}{2^{\lambda-2\varepsilon}}\right) \cdot \left(1 - 2^{-\varepsilon}\right). \end{aligned}$$

By the law of total expectation and (16) we obtain a bound on $\mathrm{E}[q_1; q_0 \leqslant 2^{-\varepsilon}]$,

$$\begin{aligned} &\mathrm{E}[q_1; q_0 \leqslant 2^{-\varepsilon}] \\ &\quad= \mathrm{E}_{q_0}\left[\mathrm{E}[q_1; q_0, q_0 \leqslant 2^{-\varepsilon}]\right] \\ &\quad= \sum_{a\leqslant 2^{-\varepsilon}} \mathrm{E}[q_1; q_0 = a, q_0 \leqslant 2^{-\varepsilon}]\mathbb{P}[q_0 = a; q_0 \leqslant 2^{-\varepsilon}] \\ &\quad= \sum_{a\leqslant 2^{-\varepsilon}} \mathrm{E}[q_1; q_0 = a]\mathbb{P}[q_0 = a; q_0 \leqslant 2^{-\varepsilon}] \\ &\quad\leqslant \sum_{a\leqslant 2^{-\varepsilon}} a^2 \cdot \mathbb{P}[q_0 = a; q_0 \leqslant 2^{-\varepsilon}] \leqslant 2^{-2\varepsilon}, \end{aligned}$$

where summation over $a \leqslant 2^{-\varepsilon}$ is valid since our distribution is discrete with finite support. Altogether,

$$\mathbb{P}\left[q_1 \leqslant 2^{\lambda-2\varepsilon}\right] \geqslant \left(1 - 2^{-\lambda}\right)\left(1 - 2^{-\varepsilon}\right).$$

Repeating the same arguments inductively using (17), we obtain

$$\mathbb{P}\left[q_\ell \leqslant 2^{2^\ell(\lambda-\varepsilon)-\lambda}\right] \geqslant \left(1 - 2^{-\lambda}\right)^\ell \left(1 - 2^{-\varepsilon}\right) \tag{18}$$

for all $1 \leqslant \ell \leqslant \frac{m}{2}$.

We now set

$$\varepsilon = 2\log_2(\log_2 n),$$

and recall the assumption,

$$k_n > n - \log_4\left(V_{\rho n,n,2}^{(2)}\right) + \log_2(n).$$

By Lemma 19 we have

$$\begin{aligned} \beta = \mathrm{E}[X_{\mathbf{v}}] &> V_{\rho n,n,2}^{(2)} \cdot 2^{2k_n-1-2n} \\ &> V_{\rho n,n,2}^{(2)} 2^{2\left(n-\log_4\left(V_{\rho n,n,2}^{(2)}\right)+\log_2(n)\right)-1-2n} = \frac{n^2}{2}. \end{aligned}$$

We also recall that in the beginning of our analysis, we assumed that ε is chosen such that $\beta(\varepsilon) \geqslant 0$ (for sufficiently large n). Indeed, since $\beta > \frac{n^2}{2}$ and $s < 2$,

$$\begin{aligned} \beta(\varepsilon) &= \beta - 2^\varepsilon \beta^{\frac{s}{2}} = \beta - 2^{2\log_2(\log_2 n))}\beta^{\frac{s}{2}} \\ &= \beta - (\log_2 n)^2 \beta^{\frac{s}{2}} \xrightarrow[n\to\infty]{} \infty. \end{aligned}$$

We set $\lambda = \varepsilon - 1$. For sufficiently large n, $\lambda > 0$ and

$$2^{2^{m/2}(\lambda-\varepsilon)-\lambda} < 2^{-2^{m/2}} = 2^{-2^{\lceil \log_2(n)\rceil+1}} < 2^{-2n}.$$

Hence, the event $\left\{q_{\frac{m}{2}} < 2^{2^{m/2}(\lambda-\varepsilon)-\lambda}\right\}$ implies the event that $\left(C_n + \langle \overline{g}_1, \ldots, \overline{g}_m\rangle\right)^2 = C_n'$ has covering radius $R_2(C_n') \leqslant \rho n$, because $q_{\frac{m}{2}} < 2^{-2n}$ implies

$$\left|Q(C_n')\right| = \left|Q\left((C_n + \langle \overline{g}_1, \ldots, \overline{g}_m\rangle)^2\right)\right| < 1,$$

but since this quantity is a non-negative integer, it must be 0. In total,

$$\begin{aligned} \mathbb{P}\left[R_2(C_n') \leqslant \rho n\right] &\geqslant \left(1 - 2^{-\varepsilon}\right)\left(1 - 2^{-\lambda}\right)^{m/2} \\ &= \left(1 - 2^{-\varepsilon}\right)\left(1 - 2^{-\varepsilon+1}\right)^{\lceil \log_2 n\rceil+1} \\ &= \left(1 - (\log_2 n)^{-2}\right)\left(1 - 2(\log_2 n)^{-2}\right)^{\lceil \log_2 n\rceil+1} \\ &= O\left(1 - \frac{1}{(\log_2(n))^2}\right) \xrightarrow[n\to\infty]{} 1. \end{aligned}$$

This completes the proof. ■

Ultimately, we are interested in $\kappa_t(\rho, q)$, which is the asymptotic minimal rate required for covering $(\mathbb{F}_q^n)^t$ (that we identify with $\mathbb{F}_q^{t\times n}$) by t-balls of radius ρn. We now prove our main result for this section.

Theorem 22: For any $0 < \rho \leqslant 1$,

$$\kappa_2(\rho, 2) \leqslant \begin{cases} 1 - (4H_4(\rho) - f(\rho)) & 0 \leqslant \rho < \frac{3}{4}, \\ 0 & \frac{3}{4} \leqslant \rho \leqslant 1. \end{cases}$$

Proof: Let us first consider $0 \leqslant \rho < \frac{3}{4}$. Fix some $\varepsilon > 0$. For any $n \in \mathbb{N}$ we look for k_n such that, if n is sufficiently large,

$$\frac{\log_2 \mathrm{Var}(X_{\mathbf{v}})}{\log_2 \mathrm{E}[X_{\mathbf{v}}]} < 2 - \frac{\varepsilon}{2}$$

for all $\mathbf{v} \in \mathbb{F}_2^{2\times n}$. By Lemma 20,

$$\mathrm{Var}(X_{\mathbf{v}}) \leqslant 7\,\mathrm{E}[X_{\mathbf{v}}] + 2^{3(k_n-n)+n(f(\rho)+o(1))}.$$

By Lemma 19 we have

$$\mathrm{E}[X_{\mathbf{v}}] \geqslant 2^{2k_n-1-2n} V^{(2)}_{\rho n,n,2} = 2^{2k_n-1-2n+2n(H_4(\rho)+o(1))},$$

and therefore,

$$\frac{\log_2 \mathrm{Var}(X_{\mathbf{v}})}{\log_2 \mathrm{E}[X_{\mathbf{v}}]} \leqslant \frac{3(k_n-n)+n(f(\rho)+o(1))}{2k_n-1-2n+2n(H_4(\rho)+o(1))}.$$

Let k_n^* be the solution of the equation

$$\frac{3(k_n^*-n)+nf(\rho)}{2k_n^*-1-2n+2nH_4(\rho)} = 2-\varepsilon,$$

namely,

$$k_n^* = \frac{2-\varepsilon+n(1+f(\rho)-4H_4(\rho)+2\varepsilon(H_4(\rho)-1))}{1-2\varepsilon}.$$

Define $k_n \triangleq \lfloor k_n^* \rfloor$. Since $\frac{k_n}{k_n^*} \xrightarrow[n\to\infty]{} 1$, for sufficiently large n,

$$\begin{aligned}\frac{\log_2 \mathrm{Var}(X_{\mathbf{v}})}{\log_2 \mathrm{E}[X_{\mathbf{v}}]} &\leqslant \frac{3(k_n-n)+n(f(\rho)+o(1))}{2k_n-1-2n+2n(H_4(\rho)+o(1))}\\ &\leqslant (2-\varepsilon)+\frac{\varepsilon}{2} = 2-\frac{\varepsilon}{2}.\end{aligned}$$

We note that

$$\lim_{\varepsilon\to 0} k_n^* = 2+n(1-4H_4(\rho)+f(\rho)).$$

By standard analysis techniques, for all $0 \leqslant \rho < \frac{3}{4}$,

$$1-4H_4(\rho)+f(\rho) > 1-H_4(\rho).$$

Thus, for a sufficiently small ε we can choose a sufficiently large n, such that

$$k_n > n - \log_4(V^{(2)}_{\rho n,n,2}) + \log_2(n)$$

and

$$\frac{\log_2 \mathrm{Var}(X_{\mathbf{v}})}{\log_2 \mathrm{E}[X_{\mathbf{v}}]} \leqslant 2-\frac{\varepsilon}{2}.$$

By Proposition 21, there exists a sequence of codes $(C_n')_{n=N}^{\infty}$ such that C_n' is an $[n, k_n']$ code where

$$\begin{aligned}k_n' &= k_n + 2\lceil \log_2(n) \rceil + 2,\\ R_2(C_n) &\leqslant \rho n.\end{aligned}$$

Thus, for sufficiently large n,

$$k_2(n,\rho n,2) \leqslant k_n + 2\lceil \log_2(n) \rceil + 2,$$

and therefore,

$$\begin{aligned}\kappa_2(\rho,2) &= \limsup_{n\to\infty} \frac{k_2(n,\rho n,2)}{n}\\ &\leqslant \lim_{n\to\infty} \frac{k_n+2\lceil \log_2(n)\rceil+2}{n}\\ &= \frac{1+f(\rho)-4H_4(\rho)+2\varepsilon(H_4(\rho)-1)}{1-2\varepsilon}.\end{aligned}$$

Taking $\varepsilon \to 0$ we obtain

$$\kappa_2(\rho,2) \leqslant 1+f(\rho)-4H_4(\rho),$$

as desired.

By its definition, $k_t(\rho, q)$ is a decreasing monotonic function in ρ. It is easy to verify that $f(\rho)$ tends to 3 when ρ tends to $\frac{3}{4}$ from the left. Thus, for $\frac{3}{4} \leqslant \rho \leqslant 1$,

$$\begin{aligned}0 \leqslant k_2(\rho,2) &\leqslant \lim_{\rho'\to\left(\frac{3}{4}\right)_-} \kappa_2(\rho',2)\\ &\leqslant \lim_{\rho'\to\left(\frac{3}{4}\right)_-} 1-4H_4(\rho')+f(\rho') = 0.\end{aligned}$$

■

A comparison of the various asymptotic bounds is shown in Figure 2. It is interesting to note that the upper bound of Theorem 22 matches the lower ball-covering bound at $\rho = \frac{3}{4}$, particularly so because the function $f(\rho)$ is defined by the binary entropy function, and not the quaternary entropy function. We also note that the naive upper bound of Proposition 14 is better than the upper bound of Theorem 22 for $\rho \lesssim 0.145$.

V. Simple Code Operations

Some code operations are very common. Among these we can find code extension, code puncturing, the $(u, u+v)$ construction, and direct sum. In this section we show the effect these operations have on the generalized covering radii mirrors their effect on the (regular) covering radius. We use the direct product to turn the non-constructive upper bound of Theorem 22 to an explicit construction, albeit, not a very useful one.

Given a code $C \subseteq \mathbb{F}_q^n$, let

$$C^* \triangleq \left\{(c_1,\dots,c_{n-1}) \middle| (c_1,\dots,c_{n-1},c_n) \in C\right\},$$

be the *punctured code*, and

$$\overline{C} \triangleq \left\{\left(c_1,\dots,c_n,-\sum_{i=1}^n c_i\right) \middle| (c_1,\dots,c_n) \in C\right\},$$

be the *extended code*. Even though puncturing is defined as the removal of the last coordinate, the following results apply to the removal of any single coordinate.

By [6, Theorem 3.1.1, p. 62], $R_1(C^*)$ equals $R_1(C)$ or $R_1(C)-1$ and $R_1(\overline{C})$ equals $R_1(C)$ or $R_1(C)+1$. The same result holds for the generalized covering radii.

Proposition 23: Let C be an $[n,k]$ linear code. Then for any $t \in \mathbb{N}$,

1) $R_t(C^*)$ equals $R_t(C)$ or $R_t(C)-1$;
2) $R_t(\overline{C})$ equals $R_t(C)$ or $R_t(C)+1$.

Proof: Let $G \in \mathbb{F}_q^{k\times n}$ be a generator matrix for C, and let $D \subseteq \mathbb{F}_{q^t}^n$ be the code generated by G over $\mathbb{F}_{q^t}$. The code C^* over $\mathbb{F}_q$ is then generated by G^* which is the matrix obtained from G by removing the last column. Denote by $D^* \subseteq \mathbb{F}_{q^t}^{n-1}$ the code generated by G^* over $\mathbb{F}_{q^t}$. Obviously, D^* is also obtained by puncturing D. By Definition 8, and [6, Theorem 3.1.1, p. 62],

$$\begin{aligned}R_t(C^*) = R_1(D^*) &\in \{R_1(D), R_1(D)-1\}\\ &= \{R_t(C), R_t(C)-1\}.\end{aligned}$$

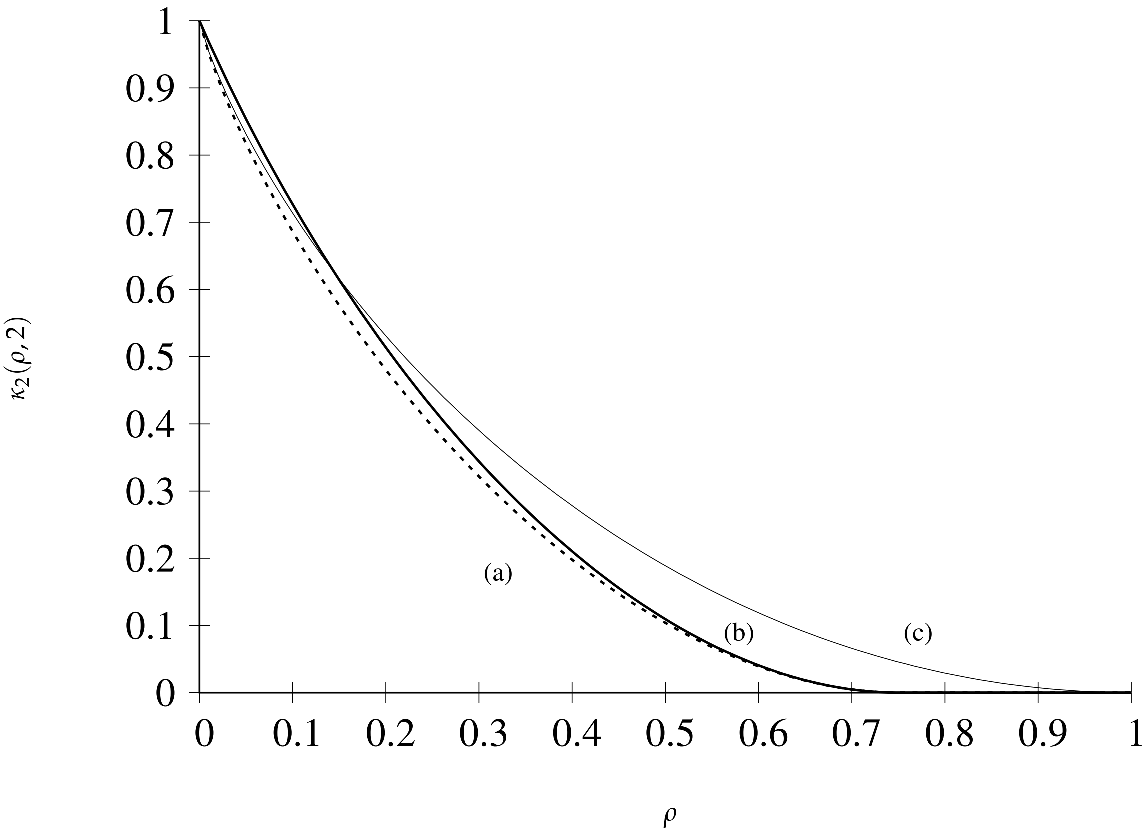

Fig. 2. A comparison of the bounds on $\kappa_2(\rho, 2)$: (a) the ball-covering lower bound, (b) the upper bound of Theorem 22, and (c) the naive upper bound of Proposition 14.

Using a similar approach we prove the case of code extension. We construct the following $k \times (n+1)$ matrix

$$\overline{G} \triangleq [G, -G \cdot \overline{1}].$$

Obviously the code generated over $\mathbb{F}_q$ by $\overline{G}$ is $\overline{C}$. Using G and $\overline{G}$ over $\mathbb{F}_{q^t}$ we get the codes D and $\overline{D}$, respectively. Again, by Definition 8, and [6, Theorem 3.1.1, p. 62],

$$\begin{aligned} R_t(\overline{C}) = R_1(\overline{D}) &\in \{R_1(D), R_1(D)+1\} \\ &= \{R_t(C), R_t(C)+1\}. \end{aligned}$$

■

Assume C_1 and C_2 are $[n, k_1]$ and $[n, k_2]$ codes, respectively. The $(u, u+v)$ construction uses C_1 and C_2 to produce a code

$$C = \left\{(\overline{u}, \overline{u}+\overline{v}) \middle| \overline{u} \in C_1, \overline{v} \in C_2\right\},$$

and by [6, Theorem 3.4.1, p. 66], its covering radius is upper bounded by $R_1(C) \leqslant R_1(C_1) + R_1(C_2)$.

Proposition 24: Let C_i be an $[n, k_i]$ code over $\mathbb{F}_q$, $i = 1, 2$, and let C be the code constructed from C_1 and C_2 using the $(u, u+v)$ construction. Then for any $t \in \mathbb{N}$,

$$R_t(C) \leqslant R_t(C_1) + R_t(C_2).$$

Proof: If $G_i \in \mathbb{F}_q^{k_i \times n}$ is a generator matrix for C_i, $i = 1, 2$, then it is easy to see that

$$G \triangleq \begin{bmatrix} G_1 & G_1 \\ 0 & G_2 \end{bmatrix},$$

is a generator matrix for C. The rest of the proof follows that of Proposition 23 by considering the code generated by G over $\mathbb{F}_{q^t}$. ■

We now look at the direct sum. Given an $[n_1, k_1]$ code C_1, and an $[n_2, k_2]$ code C_2, both over $\mathbb{F}_q$, the direct sum is defined as

$$C_1 \oplus C_2 \triangleq \left\{(\overline{c}_1, \overline{c}_2) \middle| \overline{c}_1 \in C_1, \overline{c}_2 \in C_2\right\},$$

which is an $[n_1+n_2, k_1+k_2]$ code over $\mathbb{F}_q$. It is well known [6, Theorem 3.2.1, p. 63] that

$$R_1(C_1 \oplus C_2) = R_1(C_1) + R_1(C_2).$$

Proposition 25: Let C_i be an $[n_i, k_i]$ code over $\mathbb{F}_q$, for $i = 1, 2$. Then for any $t \in \mathbb{N}$,

$$R_t(C_1 \oplus C_2) = R_t(C_1) + R_t(C_2).$$

Proof: If $G_i \in \mathbb{F}_q^{k_i \times n_i}$ is a generator matrix for C_i, $i = 1, 2$, then it is easy to see that

$$G \triangleq \begin{bmatrix} G_1 & 0 \\ 0 & G_2 \end{bmatrix},$$

is a generator matrix for $C_1 \oplus C_2$. The rest of the proof follows that of Proposition 23 by considering the code generated by G over $\mathbb{F}_{q^t}$. ■

Remark 26: The upper bound presented in Theorem 22 is proved by showing the existence of a sequence of codes in a non-constructive way. We use Proposition 25 in order to find an explicit construction for a code attaining the bound of Theorem 22.

We fix $0 \leqslant \rho \leqslant 1$. In the proof of Theorem 22, we find a sequence of covering codes $(C_n)_n$, where C_n is an $[n, k_n]$ code with covering radius at most ρn and

$$\lim_{n\to\infty} \frac{k_n}{n} = 1 - 4H_4(\rho) + f(\rho).$$

The existence of such a sequence of codes is guaranteed by Proposition 21, where it is proved that, when we randomly choose a $k_n \times n$ generator matrix, the probability to get a code with those properties is lower bounded by $O\left(1 - (\log_2(n))^{-2}\right)$.

We consider the $[n2^{n\cdot k_n}, k_n 2^{n\cdot k_n}]$ code generated by the direct product

$$\widetilde{C}_n \triangleq \bigoplus_{G\in\mathbb{F}_2^{k_n\times n}} C_G,$$

where C_G is the code with generator matrix G. We note that the rate of $\widetilde{C}_n$ is upper bounded by $\frac{k_n}{n}$. By our probabilistic argument, the normalized covering radius of $\widetilde{C}_n$ is upper bounded by $\rho \cdot (1 - O(\log_2(n)^{-2})) + 1 \cdot O(\log_2(n)^{-2})$, which tends to ρ as $n \to \infty$.

This technique of explicitly constructing codes by direct-summing codes is well known and has been used many times in order to make probabilistic proofs constructive, e.g., [3], [24]. The disadvantage of this technique is the enormous block length of the resulting code. In our construction, in order to ensure a normalized covering radius at most $\rho + \varepsilon$ the required block length is $\Omega\left(2^{2^{1/\varepsilon}+1/\sqrt{\varepsilon}}\right)$.

VI. The Generalized Packing Radii

Given an $[n, k]$ linear code C over $\mathbb{F}_q$, the generalized Hamming weight of the code, d_t, $t \in \mathbb{N}$, is defined as the minimal support size containing a linear subcode of C of dimension t, i.e.,

$$d_t \triangleq \min_{C'\in\left[{C \atop t}\right]} \left|\mathrm{supp}(C')\right|.$$

In particular, d_1 is the usual minimum distance of C.

Generalized Hamming weights were introduced by Wei in 1991 [27], as a figure of merit to analyze the security performance of a code on a wire-tap channel. Wei proved that the weight hierarchy is strictly increasing and proved the duality theorem, relating the weight hierarchy of a code and its dual. A stronger duality theorem, namely a generalization of MacWilliams identity for the generalized Hamming weight distribution of a code and its dual was provided in [20]. The weight hierarchy of a code was computed for many families of codes [11], [16], [27] and bounds are produced in [1], [8], [17]. Natural generalizations of MDS codes are presented in [10]. The generalized weights were also defined in other metric instances, such as the rank metric [21], [23]. In a very interesting approach to generalized weights, considering a representation of linear codes as a set of points in a projective space, Tsfasman and Vladut [26] transform the generalized weights from a metric problem into a combinatorial-incidence problem. Forney showed in [12] deep connections between the generalized Hamming weight hierarchy of a linear code and the complexity of its minimal trellis diagram and an initial attempt to bound the error probability of a code (in the erasure channel) using the generalized weights was done in [9], [22].

In the following we shall require the size $\lfloor (d_t - 1)/2 \rfloor$. To simplify the presentation we define for all $t \in \mathbb{N}$,

$$\delta_t \triangleq \left\lfloor \frac{d_t - 1}{2} \right\rfloor.$$

We also define the set

$$\mathcal{L}^{(t)}(\mathbb{F}_q^n) \triangleq \left\{ \mathbf{v} \in \mathbb{F}_q^{t\times n} \,\middle|\, \mathrm{rank}(\mathbf{v}) = t \right\}.$$

Lemma 27: Let C be an $[n, k]$ linear code over $\mathbb{F}_q$. Then for every $t \in [k]$, δ_t is the largest integer satisfying that for all $\mathbf{c}, \mathbf{c}' \in C^t$ such that $\mathbf{c} - \mathbf{c}' \in \mathcal{L}^{(t)}(\mathbb{F}_q^n)$,

$$B_{\delta_t}^{(t)}(\mathbf{c}) \cap B_{\delta_t}^{(t)}(\mathbf{c}') = \emptyset.$$

Proof: First, for $0 \leqslant r \leqslant \delta_t$, assume to the contrary that there exist $\mathbf{c}, \mathbf{c}' \in C^t$, $\mathbf{c} - \mathbf{c}' \in \mathcal{L}^{(t)}(\mathbb{F}_q^n)$, and that

$$B_r^{(t)}(\mathbf{c}) \cap B_r^{(t)}(\mathbf{c}') \neq \emptyset.$$

Let $\mathbf{v}$ be in that intersection. Thus, there exist $I, I' \subseteq [n]$, with $|I|, \left|I'\right| \leqslant r \leqslant \delta_t$, such that

$$\mathbf{v} \in Q_I^{(t)}(\mathbf{c}) \qquad \text{and} \qquad \mathbf{v} \in Q_{I'}^{(t)}(\mathbf{c}').$$

But then

$$\mathbf{c} - \mathbf{c}' \in Q_{I\cup I'}^{(t)}(\mathbf{0}).$$

Since $\mathbf{c} - \mathbf{c}' \in \mathcal{L}^{(t)}(\mathbb{F}_q^n)$, the row space of $\mathbf{c} - \mathbf{c}'$ is a t-dimensional subcode of C supported by $\left|I \cup I'\right| \leqslant 2r \leqslant 2\delta_t < d_t$ coordinates, which is a contradiction to the definition of d_t.

For the second direction, assume $r > \delta_t$. By the definition of δ_t, there exists a subcode $C' = \langle \overline{c}_1, \dots, \overline{c}_t \rangle \subseteq C$ of dimension t and support I of size $|I| = \delta_t$. Set $\mathbf{c} \in C^t$ to be the matrix whose rows are $\overline{c}_1, \dots, \overline{c}_t$, and arbitrarily choose $I_1, I_2 \in \binom{[n]}{r}$ such that $I \subseteq I_1 \cup I_2$. We construct, for all $i \in [t]$, a vector $\overline{v}_i$ that agrees with $\overline{c}_i$ in the coordinates I_1, and is 0 elsewhere. Set $\mathbf{v} \in \mathbb{F}_q^{t\times n}$ to be the matrix whose rows are $\overline{v}_1, \dots, \overline{v}_t$, and observe that

$$\mathbf{v} \in Q_{I_1}^{(t)}(\mathbf{0}) \qquad \text{and} \qquad \mathbf{v} \in Q_{I_2}^{(t)}(\mathbf{c}).$$

Hence

$$B_r^{(t)}(\mathbf{0}) \cap B_r^{(t)}(\mathbf{c}) \neq \emptyset,$$

so δ_t is the maximal integer with the desired property. ∎

We observe that for $t = 1$, Lemma 27 becomes the standard packing of Hamming error balls induced by the code C, and δ_1 is the packing radius of the code, and hence, $\delta_1 \leqslant R_1$. It is therefore tempting to conjecture that $\delta_t \leqslant R_t$ for all $t \in [\min\{k, n-k\}]$. However, Lemma 27 does not describe a packing of t-balls, when $t \geqslant 2$, since these may intersect if the difference between their centers is not of full rank.

VII. Conclusion

We proposed a fundamental property of linear codes – the generalized covering-radius hierarchy. It characterizes the trade-off between storage amount, latency, and access complexity in databases queried by linear combinations, as is the case, for example, in PIR schemes. We showed three equivalent definitions for these radii, highlighting their combinatorial, geometric, and algebraic aspects. We derived bounds on the code parameters in relation with the generalized covering radii, and studied the effect simple code operations have on them. Finally, we described a connection between the generalized covering-radius hierarchy and the generalized Hamming weight hierarchy.

While the study of the generalized covering-radius hierarchy has its own independent intellectual merit, let us also place the bound of Theorem 22 back in the context of PIR schemes. Consider the binary case, and assume we allow a latency of $t = 2$, namely, the server waits until two queries arrive and then handles them both. Further assume, that to handle the two queries we allow the server to access at most $\frac{1}{2}$ of its storage. Stated alternatively, the average access per query is a $\frac{1}{4}$ of the storage. By Theorem 22, since $\kappa_2(\frac{1}{2}, 2) \approx 0.11$, there exists a code allowing 89% of the server storage for user information and only 11% for overhead. A naive approach, using $\kappa_1(\frac{1}{4}, 2) \approx 0.19$, implies the storage may contain only 81% user information and 19% overhead.

Many other open problems remain, and we mention but a few. First, extending Theorem 22 to address non-binary generalized covering radii for all t is still an open question, as is closing the gap with the lower bound of Proposition 12.

It would also be interesting to determine the generalized covering-radius hierarchy of known codes. These may be extreme in some cases. As we saw in Example 3, the Hamming code satisfies $R_t = t$, and in particular the covering-radius hierarchy is strictly increasing, that is, $R_t < R_{t+1}$ for all $t \in [n-k-1]$. This property is exclusive to the Hamming code (except other trivial cases).

Proposition 28: Let C be an $[n, k \geqslant 1, d \geqslant 3]$ linear code over $\mathbb{F}_q$. Then

$$R_1 < R_2 < \ldots < R_{n-k} = n - k,$$

if and only if C is the q-ary Hamming code.

Proof: Suppose that the covering-radius hierarchy of a code C is strictly increasing. Since $R_{n-k} = n - k$, we must have $R_1 = 1$. But a linear code with parameters $[n, k, d \geqslant 3]$ with covering radius $R_1 = 1$ is 1-perfect and it must be the q-ary Hamming code with parameters $n = \frac{q^m-1}{q-1}$, $k = n - m$, and $d = 3$. ■

In contrast with the Hamming code, whose generalized covering radii are all distinct, the opposite occurs with MDS codes. As was shown in [2], [13], the (first) covering radius of $[n, k]$ MDS codes is $n - k$, except in rare cases where it is $n - k - 1$. Since the upper limit on the generalized covering radius is $n - k$, the entire hierarchy is either constant, or is a step function.

Finally, we have an algorithmic question: Given a parity-check matrix H for an $[n, k]$ code over $\mathbb{F}_q$, and given vectors $\overline{s}_1, \ldots, \overline{s}_t \in \mathbb{F}_q^{n-k}$, how do we efficiently find R_t columns of H that span the t vectors? These questions, and many others, are left for future research.

References

[1] A. Ashikhmin, A. Barg, and S. Litsyn, "New upper bounds on generalized weights," *IEEE Trans. Inf. Theory*, vol. 45, no. 4, pp. 1258–1263, May 1999.

[2] D. Bartoli, M. Giulietti, and I. Platoni, "On the covering radius of MDS codes," *IEEE Trans. Inf. Theory*, vol. 61, no. 2, pp. 801–811, Feb. 2015.

[3] V. M. Blinovskii, "Covering the Hamming space with sets translated by vectors of a linear code," *Problemy Peredachi Informatsii*, vol. 26, no. 3, pp. 21–26, 1990.

[4] B. Chazelle and B. Rosenberg, "Computing partial sums in multidimensional arrays," in *Proc. 5th Annu. Symp. Comput. Geometry*, 1989, pp. 131–139.

[5] B. Chor, O. Goldreich, E. Kushilevitz, and M. Sudan, "Private information retrieval," in *Proc. IEEE 36th Annu. Found. Comput. Sci.*, Aug. 1995, pp. 41–50.

[6] G. Cohen, I. Honkala, S. Litsyn, and A. Lobstein, *Covering Codes*. Amsterdam, The Netherlands: North-Holland, 1997.

[7] G. Cohen and P. Frankl, "Good coverings of Hamming spaces with spheres," *Discrete Math.*, vol. 56, nos. 2–3, pp. 125–131, 1985.

[8] G. Cohen, S. Litsyn, and G. Zemor, "Upper bounds on generalized distances," *IEEE Trans. Inf. Theory*, vol. 40, no. 6, pp. 2090–2092, Nov. 1994.

[9] F. Didier, "A new upper bound on the block error probability after decoding over the erasure channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 10, pp. 4496–4503, Oct. 2006.

[10] S. T. Dougherty and S.-H. Han, "Higher weights and generalized MDS codes," *J. Korean Math. Soc.*, vol. 47, no. 6, pp. 1167–1182, Nov. 2010.

[11] G. L. Feng, K. K. Tzeng, and V. K. Wei, "On the generalized Hamming weights of several classes of cyclic codes," *IEEE Trans. Inf. Theory*, vol. 38, no. 3, pp. 1125–1130, May 1992.

[12] G. D. Forney, "Density/length profiles and trellis complexity of linear block codes and lattices," in *Proc. IEEE Int. Symp. Inf. Theory*, Trondheim, Norway, Jun. 1994, p. 339.

[13] E. M. Gabidulin and T. Klove, "The Newton radius of MDS codes," in *Proc. Inf. Theory Workshop (ITW)*, Killarney, Ireland, Jun. 1998, pp. 50–51.

[14] E. Gabidulin, "Combinatorial metrics in coding theory," in *Proc. 2nd Int. Symp. Inf. Theory*, Armenia, 1971.

[15] P. Gopalan, C. Huang, H. Simitci, and S. Yekhanin, "On the locality of codeword symbols," *IEEE Trans. Inf. Theory*, vol. 58, no. 11, pp. 6925–6934, Aug. 2012.

[16] P. Heijnen and R. Pellikaan, "Generalized Hamming weights of Q-ary Reed–Muller codes," *IEEE Trans. Inf. Theory*, vol. 44, no. 1, pp. 181–196, Jan. 1998.

[17] T. Helleseth, T. Klove, and O. Ytrehus, "Generalized Hamming weights of linear codes," *IEEE Trans. Inf. Theory*, vol. 38, no. 3, pp. 1133–1140, May 1992.

[18] C.-T. Ho, J. Bruck, and R. Agrawal, "Partial-sum queries in OLAP data cubes using covering codes," *IEEE Trans. Comput.*, vol. 47, no. 12, pp. 1326–1340, Dec. 1998.

[19] K. Feng, L. Xu, and F. J. Hickernell, "Linear error-block codes," *Finite Fields Appl.*, vol. 12, no. 4, pp. 638–652, 2006.

[20] T. Kløve, "Support weight distribution of linear codes," *Discrete Math.*, vols. 106–107, pp. 311–316, Sep. 1992.

[21] J. Kurihara, R. Matsumoto, and T. Uyematsu, "Relative generalized rank weight of linear codes and its applications to network coding," *IEEE Trans. Inf. Theory*, vol. 61, no. 7, pp. 3912–3936, Jul. 2015.

[22] L. C. Lemes and M. Firer, "Generalized weights and bounds for error probability over erasure channels," in *Proc. Inf. Theory Appl. Workshop (ITA)*, San Diego, CA, USA, Feb. 2014, pp. 1–8.

[23] F. Oggier and A. Sboui, "On the existence of generalized rank weights," in *Proc. Int. Symp. Inf. Theory Appl. (ISITA)*, Honolulu, HI, USA, Oct. 2012, pp. 406–410.

[24] A. Potukuchi and Y. Zhang, "Improved efficiency for covering codes matching the sphere-covering bound," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2020, pp. 102–107.

[25] I. Tamo, Z. Wang, and J. Bruck, "Access versus bandwidth in codes for storage," *IEEE Trans. Inf. Theory*, vol. 60, no. 4, pp. 2028–2037, Apr. 2014.

[26] M. A. Tsfasman and S. G. Vladut, “Geometric approach to higher weights,” *IEEE Trans. Inf. Theory*, vol. 41, no. 6, pp. 1564–1588, Nov. 1995.

[27] V. K. Wei, “Generalized Hamming weights for linear codes,” *IEEE Trans. Inf. Theory*, vol. 37, no. 5, pp. 1412–1418, Sep. 1991.

[28] Y. Zhang, E. Yaakobi, T. Etzion, and M. Schwartz, “On the access complexity of PIR schemes,” in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Paris, France, Jul. 2019, pp. 2134–2138.

Dor Elimelech (Graduate Student Member, IEEE) received the B.Sc. degree in mathematics from Ben-Gurion University of the Negev, Israel, and the B.Sc. and M.Sc. degrees in electrical engineering from Ben-Gurion University of the Negev in 2018 and 2020, respectively, where he is currently pursuing the Ph.D. degree in electrical engineering. His research interests include coding theory, ergodic theory, and dynamical systems.

Marcelo Firer (Member, IEEE) received the B.Sc. and M.Sc. degrees from the State University of Campinas, Brazil, in 1989 and 1991, respectively, and the Ph.D. degree from The Hebrew University of Jerusalem in 1997, all in mathematics. He is currently an Associate Professor with the State University of Campinas. His research interests include coding theory, action of groups, and semigroups.

Moshe Schwartz (Senior Member, IEEE) received the B.A. *(summa cum laude)*, M.Sc., and Ph.D. degrees from the Computer Science Department, Technion—Israel Institute of Technology, Haifa, Israel, in 1997, 1998, and 2004, respectively.

He was a Fulbright Post-Doctoral Researcher with the Department of Electrical and Computer Engineering, University of California San Diego, and a Post-Doctoral Researcher with the Department of Electrical Engineering, California Institute of Technology. While on sabbatical from 2012 to 2014, he was a Visiting Scientist at the Massachusetts Institute of Technology (MIT). He is currently a Professor with the School of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Israel. His research interests include algebraic coding, combinatorial structures, and digital sequences.

Prof. Schwartz received the 2009 IEEE Communications Society Best Paper Award in Signal Processing and Coding for Data Storage and the 2020 NVMW Persistent Impact Prize. He served as an Associate Editor of coding techniques and coding theory for IEEE TRANSACTIONS ON INFORMATION THEORY from 2014 to 2021. Since 2021, he has been serving as an Area Editor of coding and decoding for IEEE TRANSACTIONS ON INFORMATION THEORY. Since 2021, he has been an Editorial Board Member of the *Journal of Combinatorial Theory, Series A*.