Higher-Order MDS Codes

Ron M. Roth, *Fellow, IEEE*

***Abstract*—An improved Singleton-type upper bound is presented for the list decoding radius of linear codes, in terms of the code parameters $[n,k,d]$ and the list size L. L-MDS codes are then defined as codes that attain this bound (under a slightly stronger notion of list decodability), with 1-MDS codes corresponding to ordinary linear MDS codes. Several properties of such codes are presented; in particular, it is shown that the 2-MDS property is preserved under duality. Finally, explicit constructions for 2-MDS codes are presented through generalized Reed–Solomon (GRS) codes.**

***Index Terms*—List decoding, MDS codes, Reed–Solomon codes, Singleton bound.**

I. Introduction

Hereafter, we let F be the finite field $\mathrm{GF}(q)$. Let $\mathcal{C}$ be a code in F^n and let $L \in \mathbb{Z}^+$ and $\tau \in \mathbb{Z}_{\geq 0}$ be given. We say that $\mathcal{C}$ is (τ, L)-*list decodable* if for every $\boldsymbol{y} \in F^n$ there are no $L+1$ distinct codewords of $\mathcal{C}$ at Hamming distance $\leq \tau$ from $\boldsymbol{y}$, i.e.,

$$\left|\mathcal{C} \cap (\boldsymbol{y} + \mathcal{B}(n,\tau))\right| \leq L,$$

where $\mathcal{B}(n,\tau)$ is the set of vectors in F^n with Hamming weight $\leq \tau$. When $\mathcal{C}$ is a linear $[n,k]$ code over F, it is (τ, L)-list decodable if there are no $L+1$ distinct vectors

$$\boldsymbol{e}_0, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_L \in \mathcal{B}(n,\tau)$$

that are in the same coset of $\mathcal{C}$ within F^n. Equivalently, if H denotes a parity-check matrix of $\mathcal{C}$, the syndromes $H\boldsymbol{e}_m^\top$ for $m \in \{0,1,\ldots,L\}$ are not all equal.

The notion of list decoding was introduced by Elias [11] and has been an active area of research in the last 25 years, where the focus has been primarily on constructing list decodable codes with efficient decoding algorithms; among the most notable contributions one can mention [17], [18], [23], and [29]. In addition, several papers presented bounds on the parameters of list decodable codes, mostly in an asymptotic setting [4], [15], [16], [19].

This work will be focusing on non-asymptotic bounds for list decoding and on characterizations of (linear) codes that attain them. We recall next the well-known sphere-packing bound for list decoding, which was proved in [12]. We use $V_q(n,\tau)$ to denote the size (volume) of $\mathcal{B}(n,\tau)$:

$$V_q(n,\tau) = |\mathcal{B}(n,\tau)| = \sum_{i=0}^{\tau} \binom{n}{i} (q-1)^i.$$

This work was supported by Grant 1713/20 from the Israel Science Foundation. Part of this work was presented at the IEEE Int'l Symposium on Information Theory (ISIT), June 2022.

Ron M. Roth is with the Computer Science Department, Technion, Haifa 3200003, Israel. Email: ronny@cs.technion.ac.il

Theorem 1 (List decoding sphere-packing bound)**.** *If $\mathcal{C} \subseteq F^n$ is (τ, L)-list decodable, then*

$$|\mathcal{C}| \leq L \cdot \frac{q^n}{V_q(n,\tau)}.$$

In particular, if $\mathcal{C}$ is a linear $[n,k]$ code over F, we have:

$$q^{n-k} \geq \frac{1}{L} \cdot V_q(n,\tau) \geq \frac{1}{L} \cdot \left(\frac{n(q-1)}{\tau}\right)^{\tau}. \tag{1}$$

The following theorem was proved in [27, Theorem 1.2] (see also [24, Theorem 2.6]).

Theorem 2 (List decoding Singleton bound)**.** *If $\mathcal{C} \subseteq F^n$ is (τ, L)-list decodable, then*

$$|\mathcal{C}| \leq L \cdot q^{n-\tau-\lfloor \tau/L \rfloor}. \tag{2}$$

In particular, if $\mathcal{C}$ is a linear $[n,k]$ code over F and $L < q$, then

$$\tau + \left\lfloor \frac{\tau}{L} \right\rfloor \leq n-k.$$

From $\lfloor \tau/L \rfloor = \lceil (\tau - L + 1)/L \rceil$ we get that the latter inequality is equivalent to

$$\tau \leq \frac{L(n-k) + L - 1}{L+1}. \tag{3}$$

As part of this work, we present conditions under which (3) can be improved to

$$\tau \leq \frac{L(n-k)}{L+1}, \tag{4}$$

and obtain constructions that attain either (3) or (4).

Remark 1. Looking at the floor values of the right-hand sides of (3) and (4), it is fairly easy to see that the latter is smaller by 1 than the former, except when $n-k \equiv 0$ or L $(\mathrm{mod}\ (L+1))$. □

Remark 2. Theorem 1 is sometimes stronger than (4). E.g., from (1) it follows that the bound (4) can be attained for a given L only if

$$q = \Omega\left(\left(\frac{n}{\tau}\right)^L\right) = \Omega\left(\left(\frac{n}{n-k}\right)^L\right)$$

(where the hidden constants in the $\Omega(\cdot)$ terms depend on L); see Appendix A for a proof. □

Improving the bounds (2) and (3) is also the subject of the very recent work [13]. In particular, it is shown there that when $L \leq q$ and $L^2 \leq \tau < Ln/(L+1)$, the multiplier L in (2) can be replaced by $1 + O(L/\tau)$ (or even by 1 when $\tau \equiv L-1$ $(\mathrm{mod}\ L)$). In Remark 4 below, we will say more about the results of [13] and their relationship with our work.

The recent paper [27] shows that there exist generalized Reed-Solomon (GRS) codes that attain the bound (4) for $L =$

$2, 3$, provided that the field size is sufficiently large (relative to n). In addition, that paper presents an explicit construction of GRS codes that attains (4) for $L = 2$ (although the size of the underlying field grows like 2^{k^n}). Some of our results herein build upon [27]. The existence result of [27] was generalized to every L in the very recent paper [9], where, *inter alia,* it is shown that for given L, n, and k and for sufficiently large q, there exist (τ, L)-list decodable $[n, k]$ GRS codes over F for any τ that satisfies (4). Hence, whenever the bound (4) holds, then it is in fact *tight* for sufficiently large q. (We also mention the earlier paper [20] where a heuristic argument is made that could suggest that when τ satisfies (4) with a strict inequality then, for sufficiently large q, there exist linear $[n, k]$ codes over F that are (τ, L)-list decodable.)

One simple construction that attains the bound (3) is the repetition code of certain lengths.

Example 1. For any given $L, u \in \mathbb{Z}^+$, the bound (3) is attained by the $[n, 1]$ repetition code $\mathcal{C}$ of length $n = (L+1)u-1$, which is (τ, L)-list decodable for $\tau = (Ln-1)/(L+1) = Lu - 1$. Specifically, for any $\boldsymbol{y} \in F^n$, the intersection $\mathcal{C} \cap (\boldsymbol{y} + \mathcal{B}(n, \tau))$ contains all the codewords $(c\, c\, \ldots\, c) \in \mathcal{C}$ that agree with $\boldsymbol{y}$ on at least $n - \tau = u$ coordinates; clearly, there can be no more than $\lfloor n/u \rfloor = L$ such codewords (see also [25, p. 175]). □

In Sections I-A and I-B we describe the results of this work.

A. Summary of improved bounds

We first prove in Section II the next theorem, which states that when $L < q$, the bound (4) can be attained (or surpassed) only by MDS codes. Hereafter, $[a : b]$ stands for the integer subset $\{i \in \mathbb{Z} : a \leq i \leq b\}$, with $[b]$ being a shorthand notation for $[1 : b]$.

Theorem 3. *Given $L \in [q-1]$, let $\mathcal{C}$ be a linear $[n, k]$ code over F and let $\tau \in \mathbb{Z}_{\geq 0}$ be such that*

$$\tau \geq \frac{L(n-k)}{L+1}. \tag{5}$$

Then $\mathcal{C}$ is (τ, L)-list decodable only if $\mathcal{C}$ is MDS.

Remark 3. When $L \in [q-1]$, a linear $[n, k, d]$ code is (τ, L)-list decodable only if $\tau < d$. Indeed, when $\tau \geq d$, all q scalar multiples of any nonzero codeword belong to (the trivial coset) $\mathcal{C}$ thereby implying that $L \geq q$. □

It turns out that the condition $L \in [q-1]$ is not too limiting: we state (in Lemma 8 in Section II) that when the code rate is bounded away from 0 and from 1, the inequality (5) can hold only if $L < q-1$ (and, by Theorem 3, the code is then necessarily MDS), unless L is exponentially large in n. That may justify our focus on (linear) MDS codes in this work.

The next theorem, which we also prove in Section II, establishes the improvement (4) on (3) for a wide range of parameters of MDS codes.

Theorem 4. *Given $L \in \mathbb{Z}^+$, let $\mathcal{C}$ be a linear $[n, k{<}n]$ MDS code over F and write*

$$n - k = (L+1)u + r, \tag{6}$$

where $u \in \mathbb{Z}_{\geq 0}$ and $r \in [L+1]$.[1] *Suppose in addition that*

$$L \leq \binom{k-1+u+r}{k-1}. \tag{7}$$

Then $\mathcal{C}$ is (τ, L)-list decodable only if (4) holds.

The following corollary presents concrete ranges of parameters that satisfy the inequality (7) (and, thus, (4) must hold in these ranges).

Corollary 5. *Using the notation of Theorem 4, if the code $\mathcal{C}$ therein is (τ, L)-list decodable then (4) holds in any one of the following cases:*

(a) $k \geq L$,

(b) $k \geq 2$ *and* $n \geq (L+1)h + k$, *where h is the smallest nonnegative integer that satisfies*[2]

$$\binom{k+h}{k-1} \geq L,$$

(c) $k \geq 2$ *and* $n - k \equiv 0, L-1,$ *or* $L \pmod{(L+1)}$,

(d) $n - k - 1 \leq L \leq \binom{n-1}{k-1}$.

Remark 4. The original posted version of our work contained a weaker statement of Theorem 4 + Corollary 5, which only applied to high-rate MDS codes. Shortly after that version was posted, Goldberg *et al.* posted their (independent) work [13], where they show (in Proposition 3.6 therein) that when $k > (q/(q-1))(L-1)$, a linear $[n, k]$ code over F can be (τ, L)-list decodable only if (4) holds. Their result applies to non-MDS codes as well; for MDS codes, their proof can be modified so that it covers the same range, $k \geq L$, as in Corollary 5(a). □

The $[n, 1, n]$ repetition code over F is excluded from Theorem 4 when $L > 1$ (see Example 1), and so is the $[n, n{-}1, 2]$ parity code when $L \geq n$: obviously, this code is $(\tau{=}1, L{=}n)$-list decodable and, thus, (4) does not hold. More generally, any linear $[n, k]$ MDS code is $(n{-}k, L)$-list decodable for

$$L = \binom{n}{k},$$

since every subset of $[n]$ of size $n - k$ contains the support of exactly one vector in each coset of the code; hence, (4) does not hold for these parameters as well (even when $L < q$). In contrast, we also prove in Section II the following result.

Theorem 6. *Let $\mathcal{C}$ be a linear $[n, k{<}n]$ MDS code over F and let $L \in \mathbb{Z}^+$ be smaller than $\binom{n}{k}$. Then $\mathcal{C}$ is not $(n{-}k, L)$-list decodable whenever*

$$q \geq \binom{n}{k+1}.$$

We note that when $L \geq n - k - 1$,

$$\frac{L(n-k)}{L+1} \geq n - k - 1.$$

[1] Namely, r is the ordinary remainder of $n - k$ when divided by $L + 1$, except when $L + 1$ divides $n - k$, in which case $r = L + 1$.

[2] When $k \geq L$, cases (a) and (b) coincide. Otherwise, as k increases from 2 to $L - 1$, the value of h decreases from $L - 2$ to 1 and, respectively, the lower threshold, $(L + 1)h + k$, on n decreases from $L^2 - L$ to $2L$.

Thus, (4) holds when

$$n-k-1 \le L \le \binom{n}{k} - 1$$

for any linear $[n,k]$ MDS code over a sufficiently large field (and—by Corollary 5(d)—under no conditioning on the field size when $n-k-1 \le L \le \binom{n-1}{k-1}$).

B. Higher-order MDS codes

In Sections III and IV, we turn to studying properties of linear codes that attain the bound (4); we will refer to such codes as L-MDS codes. To let the case $L=1$ coincide with ordinary linear MDS codes, we will need a somewhat stronger notion of list decodability, to be defined next. For $a \in \mathbb{Z}^+$, denote by $\mathbb{Z}/a$ the set of rationals of the form b/a where $b \in \mathbb{Z}$.

Let $\mathcal{C}$ be a code in F^n. Given $L \in \mathbb{Z}^+$ and a nonnegative $\tau \in \mathbb{Z}/(L+1)$, we say that $\mathcal{C}$ is *average-radius* (τ, L)*-list decodable*, or *strongly-*(τ, L)*-list decodable*, if for every $\boldsymbol{y} \in F^n$ there are no $L+1$ distinct codewords $\boldsymbol{c}_0, \boldsymbol{c}_1, \ldots, \boldsymbol{c}_L \in \mathcal{C}$ such that

$$\sum_{m \in [0:L]} \mathsf{w}(\boldsymbol{y} - \boldsymbol{c}_m) \le (L+1)\tau,$$

where $\mathsf{w}(\cdot)$ denotes Hamming weight. When $\mathcal{C}$ is a linear $[n,k]$ code over F, it is strongly-(τ, L)-list decodable if there are no $L+1$ distinct vectors

$$\boldsymbol{e}_0, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_L \in F^n$$

that are in the same coset of $\mathcal{C}$ within F^n and satisfy

$$\sum_{m \in [0:L]} \mathsf{w}(\boldsymbol{e}_m) \le (L+1)\tau. \tag{8}$$

In other words, the *average* (rather than the *maximum*) weight of any $L+1$ distinct vectors in the same coset of $\mathcal{C}$ must exceed τ. The term average-radius list decodability was used in [16] and in several more recent papers (although this concept had appeared earlier in [4] and [19]); in this work we prefer using the (shorter) name strong list decodability, as it will appear many times throughout.

Clearly, any strongly-(τ, L)-list decodable code[3] is also (ordinarily) $(\lfloor \tau \rfloor, L)$-list decodable. Theorem 2 holds also for strong list decodability (see the proof in [27]) and, as we will show, so does Theorem 3, except that in both theorems τ is now in $\mathbb{Z}/(L+1)$. Moreover, we have the following (stronger) counterpart of Theorem 4.[4]

Theorem 7. *Let $\mathcal{C}$ be a linear $[n,k]$ MDS code over F and let $L \in \mathbb{Z}^+$ be such that*

$$L \le \binom{n-1}{k-1}.$$

Then $\mathcal{C}$ is strongly-(τ, L)-list decodable only if τ satisfies (4).

A (linear) L*-MDS code* is a linear $[n,k]$ code over F which is strongly-(τ, L)-list decodable for $\tau = L(n-k)/(L+1)$. In such codes, the sum of weights of the $L+1$ lightest vectors in each coset must exceed $L(n-k)$. Note that the 1-MDS property coincides with the ordinary MDS property.

Section III is devoted to proving some basic properties of L-MDS codes. We state the counterpart of Theorem 3 for strong list decodability, namely, that when $L < q$, every L-MDS code over F is MDS (Theorem 10). We then establish the following closure property for a wide range of values of L: if a linear $[n,k]$ MDS code is L-MDS, then it is ℓ-MDS for smaller values of ℓ as well (Theorem 11). We also show that when $L \ge \binom{n}{k} - k(n-k)$, every linear $[n,k]$ MDS code is L-MDS (Theorem 15).

Remark 5. Generalizing the notion of MDS codes was also the subject of the recent independent work [8]. The notion therein, denoted $\mathrm{MDS}(\ell)$, is algebraic in nature and was given as part of the characterization of certain types of maximally recoverable tensor codes. Interestingly, the authors of [8] have established in their very recent work [9] that the two generalizations—in [8] and herein—are intimately related. Specifically, when $L < n-k$ and under the aforementioned closure condition (Theorem 11), the notions of an L-MDS code and of the dual of an $\mathrm{MDS}(L+1)$ code coincide. By combining the results of [9] with Theorem 7 it follows that whenever L, n, and k are such that these two notions coincide then, for sufficiently large q, the bound (4) is tight also for the strongly list-decodable case. □

In Section IV, we concentrate on the case $L=2$ and prove a necessary and sufficient condition for a linear $[n,k]$ MDS code to be 2-MDS (and, thus, $(\lfloor 2(n-k)/3 \rfloor, 2)$-list decodable), in terms of certain properties of its punctured codes (Theorem 18). We also show that the 2-MDS property is preserved under duality (Theorem 19): interestingly, while this property is known to hold for (ordinary) 1-MDS codes, it does not generalize to 3-MDS codes.

Finally, we present in Section V explicit constructions of GRS codes that are 2-MDS. These constructions improve on the results of [27] in that they apply to smaller fields: the field size can be polynomial in n for any fixed k or $n-k$.

C. Notation

We introduce the following notation. For a vector $\boldsymbol{e} = (e_j)_{j \in [n]} \in F^n$, we denote by $\mathsf{Supp}(\boldsymbol{e})$ the support of $\boldsymbol{e}$. For a vector $\boldsymbol{e} \in F^n$ and a subset $J \subseteq [n]$, we let $(\boldsymbol{e})_J$ be the subvector of $\boldsymbol{e}$ which consists of the coordinates that are indexed by J. This notation extends to $\rho \times n$ matrices H over F, with $(H)_J$ being the $\rho \times |J|$ submatrix of H consisting of the columns that are indexed by J.

Given a $\rho \times n$ matrix H over F and $L+1$ subsets $J_0, J_1, \ldots, J_L$ of $[n]$, we define the matrix $M =$

[3] Strong (τ, L)-list decodability is defined (only) when $L \in \mathbb{Z}^+$ and τ is nonnegative in $\mathbb{Z}/(L+1)$; thus, when we say that a code is (or is not) strongly-(τ, L)-list decodable, it will also imply that L and τ are in their valid range. This convention also applies to ordinary (τ, L)-list decodability, in which case $L \in \mathbb{Z}^+$ and $\tau \in \mathbb{Z}_{\ge 0}$.

[4] In the context of strong list decodability, the difference between the bounds (3) and (4) becomes more profound (compared to ordinary list decodability), due to the finer grid of the possible values of τ.

$M_{J_0,J_1,\ldots,J_L}(H)$ by

$$M = \left(\begin{array}{c|c|c|c|c} -(H)_{J_0} & (H)_{J_1} & & & \\ \hline -(H)_{J_0} & & (H)_{J_2} & & \\ \hline \vdots & & & \ddots & \\ \hline -(H)_{J_0} & & & & (H)_{J_L} \end{array}\right), \quad (9)$$

where empty blocks denote all-zero submatrices. The matrix M has $L\rho$ rows and $\sum_{m\in[0:L]} |J_m|$ columns (the matrix H will typically be taken as an $(n-k)\times n$ parity-check matrix of a linear $[n,k]$ code over F).

Remark 6. While the role of J_0 in (9) may seem to differ from that of the rest of the subsets J_m, in the uses of the matrix M in this work, the $L+1$ subsets will enjoy full symmetry. For example, when $L\rho = \sum_{m\in[0:L]} |J_m|$, the matrix M is square and its determinant is the same as that of the $((L+1)\rho)\times((L+1)\rho)$ matrix

$$\left(\begin{array}{c|c|c|c|c} I_\rho & (H)_{J_0} & & & \\ \hline I_\rho & & (H)_{J_1} & & \\ \hline \vdots & & & \ddots & \\ \hline I_\rho & & & & (H)_{J_L} \end{array}\right) \quad (10)$$

(with I_ρ standing for the $\rho\times\rho$ identity matrix): the symmetry among the subsets J_m is apparent in (10). □

II. Proofs of bounds

Proof of Theorem 3. Suppose that $\mathcal{C}$ is a linear $[n,k,d{\le}n{-}k]$ (non-MDS) code over F and let $\boldsymbol{c}$ be a codeword of $\mathcal{C}$ of weight d; without loss of generality we may assume that $\boldsymbol{c}$ takes the form

$$\big(\underbrace{1\,1\,\ldots\,1}_{d \text{ times}}\,0\,0\,\ldots\,0\big) \;\; (\in F^n)\,.$$

Let Y be a set of $L+1$ $(\le q)$ distinct elements of F and consider a vector $\boldsymbol{y}\in F^n$ in which each element of Y appears at least $\lfloor d/(L{+}1)\rfloor$ times among the first d entries of $\boldsymbol{y}$, while the remaining entries of $\boldsymbol{y}$ are all zero. Thus, there exist $L{+}1$ distinct scalar multiples $\boldsymbol{c}_0,\boldsymbol{c}_1,\ldots,\boldsymbol{c}_L$ of $\boldsymbol{c}$, each coinciding with $\boldsymbol{y}$ on at least $\lfloor d/(L+1)\rfloor$ out of the first d coordinates, namely, $\mathsf{w}(\boldsymbol{y}-\boldsymbol{c}_m)\le d-\lfloor d/(L+1)\rfloor$ for each $m\in[0:L]$. Hence, $\mathcal{C}$ is (τ,L)-list decodable only when

$$\tau \le d - \left\lfloor \frac{d}{L+1}\right\rfloor - 1 < \frac{L\cdot d}{L+1} \le \frac{L(n-k)}{L+1}.$$

□

The next lemma (which we prove in Appendix A) shows that the condition $L\in[q-1]$ in Theorem 3 holds in most cases of interest. Specifically, the lemma applies to any (not necessarily linear) code $\mathcal{C}\subseteq F^n$ whose rate, $R=(\log_q|\mathcal{C}|)/n$, is bounded away from 0 and from 1, and to L whose growth rate (with n) is bounded from above by some constant. The statement of the lemma makes use of the function $\eta_q : [0,1/2]\to\mathbb{R}$ which is defined for every $\varepsilon\in[0,1/2]$ by

$$\eta_q(\varepsilon) = \mathsf{h}\left(\frac{q{-}1}{q}(1{-}\varepsilon)\right) - (1{-}\varepsilon)\cdot\mathsf{h}(1/q), \quad (11)$$

where $\mathsf{h}(x) = -x\log_2 x - (1{-}x)\log_2(1{-}x)$. (In the proof of the lemma we show that $\eta_q(\varepsilon)$ is bounded from below by the value $\eta_2(\varepsilon)$, which is independent of q and is positive for every $\varepsilon\in(0,1/2]$.

Lemma 8. *Given a fixed real $\varepsilon\in(0,1/2]$, let $\mathcal{C}\subseteq F^n$ be a code of rate $R\in[\varepsilon,1{-}\varepsilon]$, let $L\in\mathbb{Z}^+$ be such that*

$$L < \frac{1}{\sqrt{2n}}\cdot 2^{\eta_q(\varepsilon)\cdot n}, \quad (12)$$

and let $\tau\in\mathbb{Z}_{\ge 0}$ be such that

$$\tau \ge \frac{Ln(1-R)}{L+1}. \quad (13)$$

Then $\mathcal{C}$ is (τ,L)-list decodable only if $L<q-1$.

Combining Theorem 3 with Lemma 8 yields the following corollary.

Corollary 9. *Given $\varepsilon\in(0,1/2]$, let $\mathcal{C}$ be a linear $[n,k]$ code over F of rate $k/n\in[\varepsilon,1{-}\varepsilon]$, let $L\in\mathbb{Z}^+$ satisfy (12), and let $\tau\in\mathbb{Z}_{\ge0}$ satisfy (5). Then $\mathcal{C}$ is (τ,L)-list decodable only if $\mathcal{C}$ is MDS.*

Remark 7. As $q\to\infty$, the value of $\eta_q(\varepsilon)$ approaches $\mathsf{h}(\varepsilon)$, which is the growth rate of $\binom{n}{\varepsilon n}$ (assuming that εn is an integer). Thus, in terms of growth rates, the requirement (12) in this case is tight, since we have seen that linear $[n,k{=}\varepsilon n]$ MDS codes are (τ,L)-list decodable for $L=\binom{n}{k}$ and $\tau=n-k>L(n-k)/(L+1)$, regardless of q (as long as MDS codes exist over F, e.g., when $q\ge n-1$). □

Proof of Theorem 4. Let

$$\tau = \left\lceil\frac{L(n-k)+1}{L+1}\right\rceil \overset{(6)}{=} \left\lceil\frac{L((L+1)u+r)+1}{L+1}\right\rceil = Lu+r \quad (14)$$

and note that τ is the smallest integer that is greater than the right-hand side of (4). We construct $L+1$ distinct vectors $\boldsymbol{e}_0,\boldsymbol{e}_1,\ldots,\boldsymbol{e}_L\in F^n$ of weight $\le\tau$ which belong to the same coset of $\mathcal{C}$; this will imply that $\mathcal{C}$ is not (τ,L)-list decodable.

We recall that every subset of $[n]$ of size $n-k+1$ is a support of $q-1$ codewords of minimum weight $n-k+1$ of the MDS code $\mathcal{C}$, and those $q-1$ codewords are scalar multiples of each other [22, Ch. 11, §3]. Denoting

$$s = k+u+r \quad (15)$$

(which, by (6), is in $[k+1:n]$), there are

$$\binom{s-1}{k-1} \overset{(7)}{\ge} L$$

codewords of weight $n-k+1$ that have a 1 as their first coordinate and their $k-1$ zero entries are all located within the first s coordinates of the codeword. Let $\boldsymbol{c}_1,\boldsymbol{c}_2,\ldots,\boldsymbol{c}_L$ be L such codewords.

We turn to defining the $L+1$ vectors $\boldsymbol{e}_m$. To this end, we partition $[s+1:n]$, which is of size

$$\begin{aligned} n-s &\stackrel{(15)}{=} n-k-u-r \\ &\stackrel{(6)}{=} ((L+1)u+r)-u-r \\ &= Lu, \end{aligned} \tag{16}$$

into L distinct subsets T_m, $m \in [L]$, each of size u. The vector $\boldsymbol{e}_0 \in F^n$ is defined by

$$\begin{array}{rcl} (\boldsymbol{e}_0)_{[s]} &=& (1\,0\,0\,\ldots\,0) \\ (\boldsymbol{e}_0)_{T_m} &=& (\boldsymbol{c}_m)_{T_m}, \quad m \in [L] \end{array},$$

and the remaining L vectors are defined by

$$\boldsymbol{e}_m = \boldsymbol{e}_0 - \boldsymbol{c}_m, \quad m \in [L]. \tag{17}$$

Clearly, the $L+1$ vectors $\boldsymbol{e}_m$ are distinct and they all belong to the same coset of $\mathcal{C}$.

We next show that $\mathsf{w}(\boldsymbol{e}_m) \le \tau$ for every $m \in [0:L]$. For $m=0$ we have

$$\mathsf{w}(\boldsymbol{e}_0) \le 1+n-s \stackrel{(16)}{\le} Lu+1 \stackrel{(14)}{\le} \tau,$$

and for $m \in [L]$ we have

$$\begin{aligned} \mathsf{w}(\boldsymbol{e}_m) &= \underbrace{\mathsf{w}\left((\boldsymbol{e}_m)_{[s]}\right)}_{s-k} + \underbrace{\mathsf{w}\left((\boldsymbol{e}_m)_{[s+1:n]}\right)}_{\le n-s-|T_m|} \\ &\le (s-k)+(n-s-u) \\ &= n-k-u \stackrel{(6)}{=} Lu+r \stackrel{(14)}{=} \tau. \end{aligned}$$

□

Proof of Corollary 5. We show that the inequality (7) holds in each of the cases (b)–(d) (with case (b) becoming case (a) when $h=0$).

(b) From $n-k \ge (L+1)h$ we get $u+r \ge h+1$ and, so, the right-hand side of (7) satisfies

$$\binom{k-1+u+r}{k-1} \ge \binom{k+h}{k-1} \ge L.$$

(c) Here $r \ge L-1$ and, so, $k-1+u+r \ge L$ and the right-hand side of (7) is at least $\binom{L}{k-1} \ge L$.

(d) When $n-k-1 \le L$ we have $u=0$ and $r=n-k$ and, so, the right-hand side of (7) equals $\binom{n-1}{k-1}$. □

Proof of Theorem 6. Let H be an $(n-k)\times n$ parity-check matrix of $\mathcal{C}$. There are at most $V_q(n, n-k-1)$ column vectors in F^{n-k} that can be written as linear combinations of less than $n-k$ columns of H. By our assumptions,

$$\begin{aligned} V_q(n,n-k-1) &< \binom{n}{n-k-1} \cdot q^{n-k-1} \\ &= \binom{n}{k+1} \cdot q^{n-k-1} \le q^{n-k} \end{aligned}$$

(except when $k=n-1$, where the first inequality is weak and the second is strict; in fact, for this case the theorem holds without any restrictions on q). Hence, there is a column vector $\boldsymbol{s} \in F^{n-k}$ that can be written as a proper linear combination of the columns of $(H)_J$, for any subset $J \subseteq [n]$ of size $n-k$; namely, each column of $(H)_J$ appears with a nonzero coefficient in the linear combination. In particular, for distinct subsets J we get different linear combinations. We conclude that $\mathcal{C}$ cannot be $(n-k, L)$-list decodable when $L < \binom{n}{n-k} = \binom{n}{k}$. □

Remark 8. If we remove any conditioning on the field size, then there do exist MDS codes that are $(n-k, L)$-list decodable for some $L < \binom{n}{k}$ (by Corollary 5(d), such L should also satisfy $L > \binom{n-1}{k-1}$). For example, the linear $[8,4]$ code over $\mathrm{GF}(7)$ in Example 2 below is $(4, 50)$-list decodable, and the linear $[6,2]$ code over $\mathrm{GF}(5)$ in Example 6 below is $(4,11)$-list decodable. □

Proof of Theorem 7. We simplify the proof of Theorem 4, redefining s therein to be n (instead of (15)); respectively, the subsets T_m become empty and $\boldsymbol{e}_0$ becomes $(1\,0\,0\,\ldots\,0)$. Keeping the definition of $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_L$ as in (17), each has weight $n-k$ and, therefore, the left-hand side of (8) equals $L(n-k)+1$. Hence, $\mathcal{C}$ is strongly-(τ, L)-list decodable only if $\tau \le L(n-k)/(L+1)$. □

III. Generalizations of the MDS property

A. L-MDS codes

Recall the definitions of strong list decodability and L-MDS codes from Section I-B. Specifically, an L-MDS code is a linear $[n,k]$ code over F which is strongly-(τ, L)-list decodable for $\tau = L(n-k)/(L+1)$. Clearly, such a code is also $(\lfloor \tau \rfloor, L)$-list decodable (in the ordinary sense), but the converse is generally not true. The next example presents a code which is $(\lceil \tau \rceil, L)$-list decodable yet is not strongly-(τ, L)-list decodable.

Example 2. Let $\mathcal{C}$ be the linear $[8,4]$ code over $F = \mathrm{GF}(7)$ with a parity-check matrix

$$H = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & 3 & 4 & 5 & 6 & 0 \\ 0 & 1 & 4 & 2 & 2 & 4 & 1 & 0 \\ 0 & 1 & 1 & 6 & 1 & 6 & 6 & 1 \end{pmatrix}.$$

This code is a doubly-extended Reed–Solomon code and is therefore MDS [22, p. 323]. By exhaustively checking the cosets of the code we find that this code is (ordinarily) $(\tau{=}4, L{=}50)$-list decodable; note that for these parameters,

$$4 = \tau = \left\lceil \frac{L(n-k)}{L+1} \right\rceil > \frac{L(n-k)}{L+1} = \frac{200}{51}$$

(and, so, the bound (4) is exceeded in this case). On the other hand, the coset that contains the vector $(0\,0\,0\,0\,0\,2\,6\,4)$ has a weight distribution as shown in Table 2, with A_w standing for the number of vectors of weight w in the coset. It follows that

TABLE I
Weight distribution of a coset of the code in Example 2.

A_0	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8
0	0	0	5	45	162	566	921	702

the sum of weights of the lightest $L+1 = 51$ vectors in the coset is

$$3 \cdot A_3 + 4 \cdot A_4 + 5 \cdot 1 = 200 = L(n-k),$$

which means that $\mathcal{C}$ is not strongly-$(200/51, 50)$-list decodable and, therefore, is not 50-MDS (however, this code turns out to be L-MDS for every $L \geq 51$). □

We next present some properties of L-MDS codes, starting with the following theorem, which is the counterpart of Theorem 3 for strong list decodability.

Theorem 10. *Given $L \in [q-1]$, every L-MDS code over F is MDS.*

Proof. If $\mathcal{C}$ is not MDS, then the left-hand side of (8) would not exceed $L(n-k)$ when the $L+1$ $(\leq q)$ vectors $\boldsymbol{e}_m$ are taken as $L+1$ distinct scalar multiples of a codeword of weight $\leq n-k$, with one of the scalars being zero. □

We also have the following closure property.

Theorem 11. *Let $\mathcal{C}$ be a linear $[n, k]$ MDS code over F and let $L \in \mathbb{Z}^+$ be such that*

$$L < \max\left\{ \binom{n-1}{k-1} \Big/ \binom{\lceil (n+k)/2 \rceil - 2}{k-1}, k \right\} + 1. \quad (18)$$

If $\mathcal{C}$ is L-MDS, then it is also ℓ-MDS for every $\ell \in [L]$.

We prove the theorem using the following lemma, which is proved in Appendix A.

Lemma 12. *Given $w, \ell, s, t \in \mathbb{Z}^+$, let $J_1, J_2, \ldots, J_\ell$ be subsets of $[w]$, each of size at least s. If*

$$\ell < \max\left\{ \binom{w}{t} \Big/ \binom{w-s}{t}, t+1 \right\},$$

then there exists a subset $X \subseteq [w]$ of size at most t that intersects with J_m, for each $m \in [\ell]$.

Proof of Theorem 11. The proof is by contradiction. Suppose that ($\mathcal{C}$ is L-MDS and) ℓ is the largest in $[L]$ for which $\mathcal{C}$ is not ℓ-MDS, namely, there exist ℓ distinct vectors $\boldsymbol{e}_0, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_\ell \in F^n$ such that

$$\sum_{m \in [0:\ell]} \mathsf{w}(\boldsymbol{e}_m) \leq \ell(n-k) \quad (19)$$

and, for an $(n-k) \times n$ parity-check matrix $H = (\boldsymbol{h}_j)_{h \in [n]}$ of $\mathcal{C}$:

$$\boldsymbol{s} = H\boldsymbol{e}_0^\top = H\boldsymbol{e}_1^\top = \cdots = H\boldsymbol{e}_\ell^\top. \quad (20)$$

We note that $\boldsymbol{s}$ is nonzero and, therefore, each vector $\boldsymbol{e}_m$ is nonzero: if $\boldsymbol{s}$ were zero then at least ℓ of the vectors $\boldsymbol{e}_m$ would be nonzero codewords of $\mathcal{C}$, in which case the left-hand side of (19) would be at least $\ell(n-k+1)$. We also note that since the difference between any two distinct vectors $\boldsymbol{e}_m$ is a nonzero codeword of $\mathcal{C}$, we have $\mathsf{w}(\boldsymbol{e}_m) \geq d/2 = (n-k+1)/2$ for all $m \in [0:\ell]$ except, possibly, for one index m, say $m = 0$. We also assume without loss of generality that the last entry of $\boldsymbol{e}_0$ is nonzero.

Next, we construct a subset $J \subseteq [n]$ which intersects with $\mathsf{Supp}(\boldsymbol{e}_m)$, for each $m \in [0:\ell]$. Specifically, we let $J = \{n\} \cup X$, where X is the subset guaranteed by Lemma 12 when applied with $w = n-1$, ℓ $(\leq L-1)$, $s = \lceil d/2 \rceil = \lfloor (n-k)/2 \rfloor + 1$, $t = k-1$, and $J_m = \mathsf{Supp}(\boldsymbol{e}_m)$, $m \in [\ell]$. By the lemma we then get that $|J| \leq k$. Writing $J' = [n] \setminus J$, it follows that the $(n-k) \times |J'|$ submatrix $(H)_{J'} = (\boldsymbol{h}_j)_{j \in J'}$ has full rank $n-k$, which means that there exists a subset $J_{\ell+1} \subseteq J'$ of size $|J_{\ell+1}| \leq n-k$ such that

$$\boldsymbol{s} = \sum_{j \in J_{\ell+1}} a_j \boldsymbol{h}_j,$$

for some $\boldsymbol{a} = (a_j)_{j \in J_{\ell+1}}$ over F. Define the vector $\boldsymbol{e}_{\ell+1} \in F^n$ by $(\boldsymbol{e}_{\ell+1})_{J_{\ell+1}} = \boldsymbol{a}$ and $(\boldsymbol{e}_{\ell+1})_{[n] \setminus J_{\ell+1}} = \boldsymbol{0}$. We have

$$\mathsf{w}(\boldsymbol{e}_{\ell+1}) \leq |J_{\ell+1}| \leq n-k \quad \text{and} \quad \boldsymbol{s} = H\boldsymbol{e}_{\ell+1}^\top. \quad (21)$$

Moreover, for each $m \in [0:\ell]$ there exists a coordinate on which $\boldsymbol{e}_{\ell+1}$ is zero while $\boldsymbol{e}_m$ is not, namely, $\boldsymbol{e}_{\ell+1} \neq \boldsymbol{e}_m$. Combining (19)–(20) with (21) yields that $\mathcal{C}$ is not $(\ell+1)$-MDS, thereby contradicting our assumption that ℓ is the largest in $[L]$ for which $\mathcal{C}$ is not ℓ-MDS. □

Remark 9. Fixing the rate $R = k/n$ to be bounded away from 0 and from 1, the expression in the right-hand side of (18) can be shown to grow exponentially with n. □

There are few cases where the MDS property implies the L-MDS property for $L > 1$, as shown in the following examples and in Theorem 15 below.

Example 3. We verify that when $k \geq n-2$, any linear $[n, k]$ MDS code $\mathcal{C}$ over F is L-MDS for every $L \in \mathbb{Z}^+$. Let $\boldsymbol{e}_0, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_L$ be distinct vectors in the same coset of $\mathcal{C}$. If all of them have weight 2 or more, then their sum of weights is at least $2(L+1) \geq L(n-k)+2$. Otherwise, if, say, $\mathsf{w}(\boldsymbol{e}_0) \leq 1$ then $\mathsf{w}(\boldsymbol{e}_m) \geq n-k+1-\mathsf{w}(\boldsymbol{e}_0)$ for all $m \in [L]$ and, so,

$$\sum_{m \in [0:L]} \mathsf{w}(\boldsymbol{e}_m) \geq L(n-k+1) - (L-1) \cdot \underbrace{\mathsf{w}(\boldsymbol{e}_0)}_{\leq 1} \geq L(n-k)+1$$

(where the last inequality is strict when $\boldsymbol{e}_0 = \boldsymbol{0}$); i.e., $\mathcal{C}$ is L-MDS. Considering the case $\mathsf{w}(\boldsymbol{e}_0) = 1$, it follows from the proof of Theorem 7 (and of Theorem 4) that in this case, there are exactly $\binom{n-1}{k-1}$ vectors of weight $n-k$ in the coset that contains $\boldsymbol{e}_0$, and all the remaining vectors in that coset have larger weight. Hence, when $L > \binom{n-1}{k-1}$ (namely, when we are outside the range of L to which Theorem 7 applies), we have

$$\sum_{m \in [0:L]} \mathsf{w}(\boldsymbol{e}_m) = \underbrace{\mathsf{w}(\boldsymbol{e}_0)}_{1} + \sum_{m \in [L]} \mathsf{w}(\boldsymbol{e}_m) > 1 + L(n-k).$$

We conclude that for linear $[n, k{\geq}n-2]$ MDS codes, the bound (4) is exceeded when $L > \binom{n-1}{k-1}$. □

Example 4. The $[n, k{=}1]$ repetition code over F is L-MDS for every $L \in \mathbb{Z}^+$: given $L+1$ distinct vectors $\boldsymbol{e}_0, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_L$ in the same coset of the code, write $J_m = \mathsf{Supp}(\boldsymbol{e}_m)$ and let

$J'_m = [n] \setminus J_m$. Since nonzero codewords have weight n, the sets J'_m must be disjoint and, so,

$$\begin{aligned}\sum_{m\in[0:L]} \mathsf{w}(\boldsymbol{e}_m) &= \sum_{m\in[0:L]} (n - |J'_m|) \\ &= (L+1)n - \sum_{m\in[0:L]} |J'_m| \\ &\geq (L+1)n - n \geq L(n-k)+1,\end{aligned}$$

with the last inequality being strict when $L > 1$. Hence, $\mathcal{C}$ is L-MDS and, in addition, (4) is exceeded when we are outside the range of Theorem 7. □

Example 5. We show that every linear $[n, k{=}2]$ MDS code $\mathcal{C}$ is 2-MDS. Let $\boldsymbol{e}_0, \boldsymbol{e}_1, \boldsymbol{e}_2$ be distinct vectors in the same coset of $\mathcal{C}$ and write $J_m = \mathsf{Supp}(\boldsymbol{e}_m)$ and $J'_m = [n] \setminus J_m$. Since nonzero codewords have weight $\geq n-k-1 = n-1$, we have $|J_m \cup J_\ell| \geq n-1$ for $0 \leq m < \ell \leq 2$ and, so,

$$|J'_m \cap J'_\ell| = |(J_m \cup J_\ell)'| \leq 1.$$

From $\left|\bigcup_{m\in[0:2]} J'_m\right| \leq n$ we get, by the inclusion–exclusion principle:

$$\begin{aligned}&\sum_{m\in[0:2]} |J'_m| \\ &= \Big|\bigcup_{m\in[0:2]} J'_m\Big| + \sum_{0\leq m<\ell\leq 2} |J'_m \cap J'_\ell| - \Big|\bigcap_{m\in[0:2]} J'_m\Big| \\ &\leq \; n+3.\end{aligned}$$

Hence, for $L = k = 2$:

$$\sum_{m\in[0:2]} \mathsf{w}(\boldsymbol{e}_m) = \sum_{m\in[0:2]} (n - |J'_m|) \geq 2n-3 > L(n-k). \quad □$$

Contrary to Examples 3 and 4, however, we cannot claim that the codes in the last example are L-MDS for $L > 2$. We demonstrate this in the next example.

Example 6. Let $\mathcal{C}$ be the linear $[6,2]$ MDS code over $F = \mathrm{GF}(5)$ with the generator matrix

$$G = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 2 & 3 & 4 \end{pmatrix}.$$

The following four vectors can be verified to be in the same coset of $\mathcal{C}$ within F^6:

$$(1\,1\,3\,0\,0\,0),\ (0\,4\,0\,0\,3\,1),\ (4\,0\,0\,1\,0\,4),\ (0\,0\,1\,2\,1\,0).$$

Therefore, $\mathcal{C}$ is not $(3,3)$-list decodable and, hence, is not 3-MDS. In fact, $\mathcal{C}$ is not L-MDS also when $L = 4,5,6$, yet it is L-MDS for $L \geq 7$. □

With every MDS code $\mathcal{C}$ we can associate a threshold, $L_0(\mathcal{C})$, which is the smallest positive integer such that $\mathcal{C}$ is L-MDS when $L \geq L_0(\mathcal{C})$ (such a threshold always exists since every code is L-MDS for $L \geq |\mathcal{C}|$). We provide an upper bound on $L_0(\mathcal{C})$ in the next theorem, which we prove using the next two lemmas: the first presents a property of the weight distribution of a coset of an MDS code and follows from MacWilliams' identities, and the second is proved in Appendix A.

Lemma 13 ([5]). *Let $\mathcal{C}$ be a linear $[n,k]$ MDS code over F and let $(A_w)_{w\in[0:n]}$ be the weight distribution of some coset X of $\mathcal{C}$ within F^n (where A_w is the number of vectors of weight w in X). Then*

$$\sum_{w=0}^{n-k} \binom{n-w}{k} A_w = \binom{n}{k}.$$

Lemma 14. *Given positive integers $k < n$, define the rational sequence $(\vartheta_w)_{w\in[0:n-k-1]}$ by*

$$\vartheta_w = \frac{1}{n{-}k{-}w} \left(\binom{n-w}{k} - (n{-}k{-}w{+}1) \right). \tag{22}$$

This sequence is all-zero when $k=1$ and is strictly decreasing when $k > 1$.

Theorem 15. *Let $\mathcal{C}$ be a linear $[n,k]$ MDS code over F. Then $\mathcal{C}$ is L-MDS for every $L \in \mathbb{Z}^+$ such that*

$$L \geq \binom{n}{k} - k(n-k). \tag{23}$$

Proof. Fix X to be any coset of $\mathcal{C}$ within F^n and let $(A_w)_{w\in[0:n]}$ be the weight distribution of X. We distinguish between two (disjoint) cases.

Case 1: the weight distribution satisfies the inequality

$$\sum_{w=0}^{n-k-1} (n{-}k{-}w) A_w \leq n-k-1. \tag{24}$$

For a vector $\boldsymbol{e} \in X$, define its *deficiency* by

$$\delta(\boldsymbol{e}) = \max\{n-k-\mathsf{w}(\boldsymbol{e}), 0\}.$$

Observing that the left-hand side of (24) is the sum of the deficiencies of all the vectors in X, we then get that for any $L+1$ distinct vectors $\boldsymbol{e}_0, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_L \in X$:

$$\sum_{m\in[0:L]} \delta(\boldsymbol{e}_m) \leq \sum_{\boldsymbol{e}\in X} \delta(\boldsymbol{e}) \leq n-k-1. \tag{25}$$

Hence,

$$\begin{aligned}\sum_{m\in[0:L]} \mathsf{w}(\boldsymbol{e}_m) &\geq \sum_{m\in[0:L]} (n-k-\delta(\boldsymbol{e}_m)) \\ &= (L+1)(n-k) - \sum_{m\in[0:L]} \delta(\boldsymbol{e}_m) \\ &\overset{(25)}{\geq} (L+1)(n-k) - (n-k-1) \\ &= L(n-k)+1\end{aligned}$$

(regardless of whether L satisfies (23)).

Case 2: the weight distribution satisfies the inequality

$$\sum_{w=0}^{n-k-1} (n{-}k{-}w) A_w \geq n-k. \tag{26}$$

Let $\boldsymbol{e}_0, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_L$ be distinct vectors in X. Then

$$\begin{aligned}&\sum_{m\in[0:L]} \mathsf{w}(\boldsymbol{e}_m) \\ &\quad\geq \sum_{w=0}^{n-k} w A_w + \Big(L+1-\sum_{w=0}^{n-k} A_w\Big)(n-k+1) \\ &\quad= (L+1)(n-k+1) - \sum_{w=0}^{n-k} (n{-}k{-}w{+}1) A_w.\end{aligned}$$

We seek to show that the last expression is at least $L(n-k)+1$, namely, that

$$L + n - k - \sum_{w=0}^{n-k} (n-k-w+1) A_w \geq 0,$$

and, to this end, it suffices (by (23)) to show that

$$\binom{n}{k} - \sum_{w=0}^{n-k} (n-k-w+1) A_w \geq (k-1)(n-k)$$

which, by Lemma 13, is equivalent to

$$\sum_{w=0}^{n-k-1} \left(\binom{n-w}{k} - (n-k-w+1) \right) A_w \geq (k-1)(n-k). \tag{27}$$

For $w \in [0 : n-k-1]$, define

$$B_w = (n-k-w) A_w$$

and let the sequence $(\vartheta_w)_{[0:n-k-1]}$ be as in (22). Then (26) and (27) become, respectively,

$$\sum_{w=0}^{n-k-1} B_w \geq n-k \tag{28}$$

and

$$\sum_{w=0}^{n-k-1} \vartheta_w B_w \geq (k-1)(n-k). \tag{29}$$

By Lemma 14, the left-hand side of (29), when seen as a function of nonnegative integer $(n-k)$-tuples $(B_w)_{w \in [0:n-k-1]}$ that satisfy (28), is minimized when $B_w = 0$ for $w \in [0 : n-k-2]$ and $B_{n-k-1} = n-k$. For this choice, we get equality in (29). □

The range (23) of L in Theorem 15 is tight for $k \in \{1, n-1, n\}$ (see Examples 3 and 4) and for the code in Example 6, yet is loose for $k = n-2$ when $n \geq 4$ (Example 3) and for the code in Example 2, and also for the (rather peculiar) codes in the next example.

Example 7. Let $F = \mathrm{GF}(11)$ and $n = 6$ and, for $k \in [n]$, let $\mathcal{C}_k$ be the linear $[n, k]$ code over F with the $(n-k) \times n$ parity-check matrix $H = (\alpha_j^i)_{i \in [0:n-k-1], j \in [n]}$, where

$$(\alpha_j)_{j \in [n]} = (0\ 1\ 4\ 9\ 5\ 3).$$

This code is a GRS code over F and is therefore MDS. An exhaustive check reveals that $\mathcal{C}_k$ is L-MDS for *every* $k \in [n]$ and *every* $L \in \mathbb{Z}^+$.[5] The same holds for a similar construction of length $n = 7$ over $F = \mathrm{GF}(73)$, where

$$(\alpha_j)_{j \in [n]} = (0\ 1\ 9\ 8\ 3\ 16\ 34).$$

□

[5]The check is required only for $k \in \{2, 3\}$, as the remaining values of k are covered by Examples 3 and 4. The fact that the elements α_j range over the quadratic residues of GF(11)—including zero—may or may not be coincidental.

B. Lightly-L-MDS codes

The next definition introduces a weaker notion of strong list decodability; the notion is quite artificial, yet it will be useful in our analysis.

Given $L \in \mathbb{Z}^+$ and a nonnegative $\tau \in \mathbb{Z}/(L+1)$, a linear $[n, k, d]$ code $\mathcal{C}$ over F is called *lightly-(τ, L)-list decodable* if there are no $L+1$ *nonzero* vectors

$$\boldsymbol{e}_0, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_L \in \mathcal{B}(n, d-1)$$

in the same coset of $\mathcal{C}$ within F^n that have *disjoint supports* and[6] satisfy (8).

Remark 10. Generally, light list decodability is incomparable with ordinary list decodability. Specifically, among any three vectors within the same coset of a linear $[n, k, d]$ code $\mathcal{C}$ there is at least one vector of weight $\geq d/2$ and the other two vectors have total weight $\geq d$. Hence, when $d > 2n/3$, these three vectors cannot have disjoint supports, implying that $\mathcal{C}$ is lightly-$(\tau, 3)$-list decodable for any τ (but it is not $(\tau, 3)$-list decodable when, say, $\tau = n$ and $|\mathcal{C}| > 3$). Conversely, the linear $[15, 5, 4]$ code over GF(2) which is generated by

$$G = \begin{pmatrix} 1 & 1 & 1 & 1 & & & & & & & & & & & \\ 1 & 1 & & & 1 & 1 & & & & & & & & & \\ 1 & 1 & & & & & 1 & 1 & & & & & & & \\ 1 & 1 & & & & & & & 1 & 1 & & & & & \\ 1 & 1 & & & & & & & & & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

can be verified exhaustively to be $(\tau{=}3, L{=}5)$-list decodable, yet it is not lightly-$(3, 5)$-list decodable: denoting by $\boldsymbol{g}_i$ the ith row of G and letting

$$\boldsymbol{e} = (1\,1\,0\,0\,\ldots\,0),$$

the vectors $\boldsymbol{e}, \boldsymbol{e}+\boldsymbol{g}_1, \boldsymbol{e}+\boldsymbol{g}_2, \ldots, \boldsymbol{e}+\boldsymbol{g}_5$ have disjoint supports and total weight $15 = L \cdot \tau$. □

In analogy to the previous definition of L-MDS codes, we define a (linear) *lightly-L-MDS code* to be a linear $[n, k]$ code over F which is lightly-(τ, L)-list decodable for $\tau = L(n-k)/(L+1)$. Here, too, the lightly-1-MDS property coincides with the ordinary MDS property.

The following lemma provides a necessary and sufficient condition for a linear MDS code $\mathcal{C}$ to be lightly-L-MDS. The condition is expressed in terms of the non-singularity of matrices $M = M_{J_0, J_1, \ldots, J_L}(H)$ (as in (9)) which are computed for a parity-check matrix H of $\mathcal{C}$.

Lemma 16. *Given $L \in \mathbb{Z}^+$, let $\mathcal{C}$ be a linear $[n, k{=}Rn]$ MDS code $\mathcal{C}$ over F where*

$$R \geq 1 - \frac{1}{L}, \tag{30}$$

and let H be an $(n-k) \times n$ parity-check matrix of $\mathcal{C}$. Then $\mathcal{C}$ is lightly-L-MDS, if and only if

$$\det(M_{J_0, J_1, \ldots, J_L}(H)) \neq 0$$

for every $L+1$ disjoint subsets

$$J_0, J_1, \ldots, J_L \subseteq [n]$$

[6]Nonzero vectors that have disjoint supports are necessarily distinct. In fact, requiring the vectors to be nonzero is equivalent to requiring them to be distinct, since distinct vectors in $\mathcal{B}(n, d-1)$ that belong to the same coset of $\mathcal{C}$ cannot be codewords of $\mathcal{C}$ and therefore must all be nonzero.

that satisfy the following conditions:

S1) $|J_m| \le n-k$*, for* $m \in [0:L]$*, and—*

S2) $\sum_{m \in [0:L]} |J_m| = L(n-k)$.

Proof. Starting with the "if" part, suppose that $\mathcal{C}$ is not lightly-L-MDS. Then there exist nonzero vectors $\boldsymbol{e}_0, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_L \in \mathcal{B}(n, n-k)$ of disjoint supports such that

$$H\boldsymbol{e}_0^\top = H\boldsymbol{e}_1^\top = \cdots = H\boldsymbol{e}_L^\top \tag{31}$$

and (8) holds for $\tau = L(n-k)/(L+1)$, namely

$$\sum_{m \in [0:L]} \mathsf{w}(\boldsymbol{e}_m) \le L(n-k). \tag{32}$$

By (30) we have $L(n-k) \le n$, so we can extend each support $\mathsf{Supp}(\boldsymbol{e}_m)$ to a subset $J_m \subseteq [n]$ so that the subsets $J_0, J_1, \ldots, J_L$ are disjoint and satisfy conditions (S1)–(S2). From (31) it follows that the column vector

$$\left((\boldsymbol{e}_0)_{J_0} \,|\, (\boldsymbol{e}_1)_{J_1} \,|\, \ldots \,|\, (\boldsymbol{e}_L)_{J_L}\right)^\top ,$$

which is nonzero and of length $L(n-k)$, is in the right kernel of $M_{J_0,J_1,\ldots,J_L}(H)$, namely, this matrix, which is of order $(L(n-k)) \times (L(n-k))$, is singular.

Turning to the "only if" part, suppose that $M_{J_0,J_1,\ldots,J_L}(H)$ is singular for disjoint nonempty subsets $J_0, J_1, \ldots, J_L$ that satisfy conditions (S1)–(S2). Then its right kernel contains a nonzero column vector,

$$\left(\hat{\boldsymbol{e}}_0 \,|\, \hat{\boldsymbol{e}}_1 \,|\, \ldots \,|\, \hat{\boldsymbol{e}}_L\right)^\top , \tag{33}$$

where $\hat{\boldsymbol{e}}_m \in F^{|J_m|}$.

For $m \in [0:L]$, let $\boldsymbol{e}_m \in F^n$ be defined by

$$(\boldsymbol{e}_m)_{J_m} = \hat{\boldsymbol{e}}_m \quad \text{and} \quad (\boldsymbol{e}_m)_{[n]\setminus J_m} = \boldsymbol{0}.$$

It is easy to see that the $L+1$ vectors $\boldsymbol{e}_0, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_L$ satisfy (31), namely, they all belong to the same coset of $\mathcal{C}$ within F^n. These vectors have disjoint supports and, by condition (S1), they are all in $\mathcal{B}(n, n-k)$. Now, if the coset they belong to were the trivial coset $\mathcal{C}$ (namely, if they were all codewords of $\mathcal{C}$), then each $\boldsymbol{e}_m$ ($\in \mathcal{B}(n, n-k)$) would be zero, but then the vector in (33) would be all-zero. We therefore conclude that each $\boldsymbol{e}_m$ is nonzero. Finally, condition (S2) implies (32), i.e., $\mathcal{C}$ is not lightly-L-MDS. □

Remark 11. The lemma holds also when the following condition is added:

S3) $|J_m| + |J_\ell| \ge n-k+1$, for all $0 \le m < \ell \le L$.

Referring to the "if" part of the proof, for any two distinct vectors $\boldsymbol{e}_m$ and $\boldsymbol{e}_\ell$ therein we have $|J_m| + |J_\ell| \ge \mathsf{w}(\boldsymbol{e}_m) + \mathsf{w}(\boldsymbol{e}_\ell) \ge \mathsf{w}(\boldsymbol{e}_m - \boldsymbol{e}_\ell) \ge n-k+1$. □

Remark 12. For the special case $L = 2$, condition (S1) can be tightened to

S1) $2 \le |J_m| \le n-k-1$, for $m \in [0:2]$.

Referring again to the "if" part of the proof, the inequality (32) becomes in this case

$$\sum_{m \in [0:2]} \mathsf{w}(\boldsymbol{e}_m) \le 2(n-k).$$

Combining this with $\mathsf{w}(\boldsymbol{e}_1) + \mathsf{w}(\boldsymbol{e}_2) \ge n-k+1$ then implies $\mathsf{w}(\boldsymbol{e}_0) \le n-k-1$. By symmetry, the same applies to the weights of $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$. Hence, $\mathsf{w}(\boldsymbol{e}_1) \ge n-k+1-\mathsf{w}(\boldsymbol{e}_2) \ge 2$ (and the same holds for $\mathsf{w}(\boldsymbol{e}_0)$ and $\mathsf{w}(\boldsymbol{e}_2)$). □

IV. 2-MDS CODES

In this section, we consider the case of 2-MDS codes over F. By Theorem 10 it follows that when $q > 2$, such codes are necessarily MDS. The case $L = q = 2$ is covered by the next lemma, which is proved in Appendix A.

Lemma 17. *A linear* $[n,k,d]$ *code over* GF(2) *is* 2*-MDS, if and only if it is MDS, namely,* $(k,d) \in \{(1,n), (n-1,2), (n,1)\}$.

The next theorem provides a necessary and sufficient condition for a linear $[n,k]$ MDS code to be 2-MDS (and, thus, $(\lfloor (2(n-k)/3 \rfloor, 2)$-list decodable), in terms of the light list decodability of its punctured codes.[7] For the special case of GRS codes (to be discussed in Section V), the sufficiency part of the theorem is the dual-code version of Lemma 4.4 in [27]; we will make a general statement about duality in Theorem 19 below.[8]

Theorem 18. *A linear* $[n,k]$ *MDS code* $\mathcal{C}$ *over* F *is* 2*-MDS, if and only if for every integer* w *in the range*

$$\max\{0, n-2k\} \le w \le n-k-3, \tag{34}$$

every linear $[n^*{=}n{-}w, k]$ *code* $\mathcal{C}^*$ *that is obtained by puncturing* $\mathcal{C}$ *on any* w *coordinates is lightly-*2*-MDS.*

Proof. The range (34) is empty when $\min\{k, n-k\} \le 2$, yet we showed in Examples 3–5 that in this case, every linear $[n,k]$ MDS code is (unconditionally) 2-MDS. Therefore, we assume from now on in the proof that $3 \le k \le n-3$.

We start with the "if" part. Suppose that $\mathcal{C}$ is MDS but not 2-MDS. Then there exist three distinct vectors $\boldsymbol{e}_0, \boldsymbol{e}_1, \boldsymbol{e}_2 \in F^n$ that belong to the same coset of $\mathcal{C}$ such that

$$\sum_{m \in [0:2]} \mathsf{w}(\boldsymbol{e}_m) \le 2(n-k). \tag{35}$$

By possibly translating these vectors by the same vector, we can assume without loss of generality that

$$\bigcap_{m \in [0:2]} \mathsf{Supp}(\boldsymbol{e}_m) = \emptyset. \tag{36}$$

We will show that when we puncture $\mathcal{C}$ on the coordinates on which any *two* of these supports intersect, the resulting code is not lightly-2-MDS.

For $m \in [0:2]$, let $J_m \subseteq [n]$ be such that $\mathsf{Supp}(\boldsymbol{e}_m) \subseteq J_m$ and the property (36) extends to J_0, J_1, and J_2, namely:

$$\bigcap_{m \in [0:2]} J_m = \emptyset \tag{37}$$

[7]The puncturing of a code $\mathcal{C} \subseteq F^n$ on a subset $X \subseteq [n]$ is the code $\{(\boldsymbol{c})_{[n]\setminus X} \,:\, \boldsymbol{c} \in \mathcal{C}\}$.

[8]The proof of Lemma 4.4 in [27] makes use of (the transpose of) the matrix (10), where H is taken as the *generator matrix* of the GRS code. In fact, the proof of Lemma 4.4 in [27] inspired our upcoming proof of Theorem 19 below.

(clearly, these conditions on J_0, J_1, and J_2 hold if each J_m is taken to be $\mathsf{Supp}(\boldsymbol{e}_m)$). For $0 \le m < \ell \le 2$, let

$$w_{m,\ell} = |J_m \cap J_\ell|.$$

We have

$$\begin{aligned} w_{m,\ell} &= |J_m| + J_\ell| - |J_m \cup J_\ell| \\ &\le |J_m| + |J_\ell| - |\mathsf{Supp}(\boldsymbol{e}_m) \cup \mathsf{Supp}(\boldsymbol{e}_\ell)| \\ &\le |J_m| + |J_\ell| - (n-k+1), \end{aligned} \tag{38}$$

where the last step follows from the minimum distance of $\mathcal{C}$. Hereafter, we will further assume that

$$\sum_{m \in [0:2]} |J_m| = 2(n-k) \tag{39}$$

by taking $J_m = \mathsf{Supp}(\boldsymbol{e}_m)$ for $m \in \{0,1\}$ and selecting J_2 to be of size $2(n-k) - |J_0| - |J_1|$ such that

$$\mathsf{Supp}(\boldsymbol{e}_2) \subseteq J_2 \subseteq [n] \setminus (J_0 \cap J_1).$$

Indeed, by (36), the set on the left is fully contained in the set on the right which, in turn, has size

$$n - w_{0,1} \overset{(38)}{\ge} 2n-k+1-|J_0|-|J_1| > 2(n-k)-|J_0|-|J_1|.$$

Write

$$J = (J_0 \cap J_1) \cup (J_0 \cap J_2) \cup (J_1 \cap J_2),$$

which is of size

$$\begin{aligned} w = |J| &\overset{(37)}{=} w_{0,1} + w_{0,2} + w_{1,2} \\ &\overset{(38)}{\le} 2 \sum_{m \in [0:2]} |J_m| - 3(n-k+1) \\ &\overset{(39)}{=} n-k-3. \end{aligned} \tag{40}$$

We find the allowable range of the size w of J. On the one hand, by (40) we have $w \le n-k-3$. On the other hand, by (37) and the inclusion–exclusion principle we also have

$$2(n-k) - w \overset{(39)}{=} \sum_{m \in [0:2]} |J_m| - |J| = |J_0 \cup J_1 \cup J_2| \le n,$$

namely, $w \ge n-2k$. Thus, $\max\{0, n-2k\} \le w \le n-k-3$ (as in (34)).

For $m \in [0:2]$, let $J_m^* = J_m \setminus J$. The sets J_0^*, J_1^*, and J_2^* are disjoint and

$$\sum_{m \in [0:2]} |J_m^*| = \sum_{m \in [0:2]} |J_m| - 2|J| \overset{(39)}{=} 2(n-k-w). \tag{41}$$

Moreover,

$$\begin{aligned} |J_0^*| &= |J_0| - w_{0,1} - w_{0,2} \\ &= |J_0| + w_{1,2} - w \\ &\overset{(38)}{\le} \sum_{m \in [0:2]} |J_m| - (n-k+1) - w \\ &\overset{(39)}{=} n-k-w-1, \end{aligned} \tag{42}$$

and the same upper bound applies to $|J_1^*|$ and $|J_2^*|$ as well. Also note that (41) and (42) imply for each $m \in [0:2]$ that

$$|J_m^*| \ge 2. \tag{43}$$

Let H be an $(n-k) \times n$ parity-check matrix of $\mathcal{C}$ and let P be an $(n-k-w) \times (n-k)$ matrix whose rows form a basis of the left kernel of $(H)_J$. Write $J' = [n] \setminus J$ and let $H^* = (PH)_{J'}$ and $n^* = n - w$. It can be readily verified that H^* is an $(n^*-k) \times n^*$ parity-check matrix of the linear $[n^*, k]$ MDS code $\mathcal{C}^*$ which is obtained by puncturing $\mathcal{C}$ on the coordinate set J.[9]

For $m \in [0:2]$, define the vector $\boldsymbol{e}_m^* \in F^{n^*}$ by $\boldsymbol{e}_m^* = (\boldsymbol{e}_m)_{J'}$. From

$$H(\boldsymbol{e}_2 - \boldsymbol{e}_0)^\top = H(\boldsymbol{e}_1 - \boldsymbol{e}_0)^\top = \mathbf{0}$$

we get that the vectors $(H)_{J'}(\boldsymbol{e}_m^* - \boldsymbol{e}_0^*)^\top$ for $m \in [2]$ are in the linear span of the columns of $(H)_J$; as such, these vectors are in the right kernel of P, namely,

$$H^*(\boldsymbol{e}_0^*)^\top = H^*(\boldsymbol{e}_1^*)^\top = H^*(\boldsymbol{e}_2^*)^\top.$$

Noting that $\mathsf{Supp}(\boldsymbol{e}_m^*) \subseteq J_m^*$ (with equality when $m \in \{0,1\}$), we conclude that

$$\mathsf{w}(\boldsymbol{e}_m^*) \le |J_m^*| \overset{(42)}{\le} n^* - k - 1,$$

$$\sum_{m \in [0:2]} \mathsf{w}(\boldsymbol{e}_m^*) \le \sum_{m \in [0:2]} |J_m^*| \overset{(41)}{=} 2(n^* - k),$$

and, for $0 \le m < \ell \le 2$,

$$\mathsf{Supp}(\boldsymbol{e}_m^*) \cap \mathsf{Supp}(\boldsymbol{e}_\ell^*) \subseteq J_m^* \cap J_\ell^* = \emptyset.$$

Moreover, by (43) we have $\mathsf{w}(\boldsymbol{e}_m^*) = |J_m^*| \ge 2$ for $m \in \{0,1\}$, namely, $\boldsymbol{e}_0^*$ and $\boldsymbol{e}_1^*$ are nonzero in $\mathcal{B}(n^*, n^*-k-1)$ and, therefore, are in a nontrivial coset of $\mathcal{C}^*$ within F^{n^*}. This, in turn, implies that $\boldsymbol{e}_2^*$ (which is in the same coset) is nonzero too. Thus, starting off with the assumption that $\mathcal{C}$ is not 2-MDS, we have shown that $\mathcal{C}^*$ is not lightly-2-MDS.

Turning to the "only if" part, the proof is carried out by essentially retracing our steps for the "if" part. Let H be an $(n-k) \times n$ parity-check matrix of $\mathcal{C}$ and let a code $\mathcal{C}^*$ be given that is obtained by puncturing $\mathcal{C}$ on the coordinates that are indexed by some subset $J \subseteq [n]$ of size $|J| = w$. Suppose that $\mathcal{C}^*$ is not lightly-2-MDS, namely, there exist nonzero vectors $\boldsymbol{e}_0^*, \boldsymbol{e}_1^*, \boldsymbol{e}_2^* \in \mathcal{B}(n-w, n-w-k)$ of disjoint supports such that

$$\sum_{m \in [0:2]} \mathsf{w}(\boldsymbol{e}_m^*) \le 2(n-k-w) \tag{44}$$

and

$$H^*(\boldsymbol{e}_1^* - \boldsymbol{e}_0^*)^\top = H^*(\boldsymbol{e}_2^* - \boldsymbol{e}_0^*)^\top = \mathbf{0}, \tag{45}$$

where $H^* = (PH)_{J'}$ is an $(n-k-w) \times (n-w)$ parity-check matrix of $\mathcal{C}^*$, with P being an $(n-k-w) \times (n-k)$ matrix whose rows form a basis of the left kernel of $(H)_J$. From (45) it follows that the vectors $(H)_{J'}(\boldsymbol{e}_m^* - \boldsymbol{e}_0^*)^\top$ for $m \in [2]$ are in the right kernel of P, which means that they

[9] This is easily seen if we first apply elementary linear operations to the rows of H so that the w columns of $(H)_J$ become (distinct) standard unit vectors in F^{n-k}. The rows of P can be taken as the remaining elements of the (transposed) standard basis of F^{n-k}, which means that $(PH)_{J'}$ is obtained from H simply by removing the rows and columns that contain the 1's in $(H)_J$.

are in the linear span of the columns of $(H)_J$; namely, there exist $\boldsymbol{x}_1, \boldsymbol{x}_2 \in F^w$ such that, for $m \in [2]$:

$$(H)_{J'}(\boldsymbol{e}_m^* - \boldsymbol{e}_0^*)^\top + (H)_J \boldsymbol{x}_m^\top = \boldsymbol{0}. \tag{46}$$

For $m \in [0:2]$, define the vector $\boldsymbol{e}_m \in F^n$ by $(\boldsymbol{e}_m)_{J'} = \boldsymbol{e}_m^*$ and $(\boldsymbol{e}_m)_J = \boldsymbol{x}_m$, where $\boldsymbol{x}_0 = \boldsymbol{0}$. By (44) and (46) we get, respectively, that

$$\begin{aligned}\sum_{m\in[0:2]} \mathsf{w}(\boldsymbol{e}_m) &= \sum_{m\in[0:2]} \mathsf{w}(\boldsymbol{e}_m^*) + \sum_{m\in[0:2]} \mathsf{w}(\boldsymbol{x}_m) \\ &\le \sum_{m\in[0:2]} \mathsf{w}(\boldsymbol{e}_m^*) + 2w \le 2(n-k)\end{aligned}$$

and

$$H\boldsymbol{e}_0^\top = H\boldsymbol{e}_1^\top = H\boldsymbol{e}_2^\top .$$

Moreover, the vectors $\boldsymbol{e}_m$ are distinct since the vectors $\boldsymbol{e}_m^*$ are nonzero with disjoint supports. Hence, $\mathcal{C}$ is not 2-MDS. ☐

Remark 13. When applying Theorem 18 to test whether a given linear $[n,k]$ MDS code $\mathcal{C}$ is 2-MDS, we can use Lemma 16 to check if each punctured $[n-w,k]$ code $\mathcal{C}^*$ is lightly-2-MDS. Specifically, since $\mathcal{C}$ is MDS then so is $\mathcal{C}^*$; moreover, for the range (34), the rate of $\mathcal{C}^*$ is $k/(n-w) \ge 1/2$, namely, the inequality (30) holds as well. ☐

It is well known that the MDS property is preserved under duality [22, p. 318]. We next apply Theorem 18 to show that the same can be said about the 2-MDS property.

Theorem 19. *A linear $[n,k]$ code over F is 2-MDS, if and only if its dual code is.*

Proof. Suppose that $\mathcal{C}$ is MDS but not 2-MDS, namely, there exist three distinct vectors $\boldsymbol{e}_0, \boldsymbol{e}_1, \boldsymbol{e}_2 \in F^n$ that belong to the same coset of $\mathcal{C}$ such that (35) holds. We proceed by applying the proof of Theorem 18—verbatim—up to Eq. (43). Henceforth, we show that when we puncture the dual code $\mathcal{C}^\perp$ on the coordinates that belong to *none* of the supports of these three vectors, the resulting code is not lightly-2-MDS. This, in turn, will imply (by Theorem 18) that $\mathcal{C}^\perp$ is not 2-MDS.

Let G be a $k \times n$ generator matrix of $\mathcal{C}$ and let $\boldsymbol{u}_0$, $\boldsymbol{u}_1$, and $\boldsymbol{u}_2$ be distinct vectors in F^k such that

$$\boldsymbol{e}_0 + \boldsymbol{u}_0 G = \boldsymbol{e}_1 + \boldsymbol{u}_1 G = \boldsymbol{e}_2 + \boldsymbol{u}_2 G.$$

Letting

$$K = \bigcup_{m\in[0:2]} J_m = J \cup \bigcup_{m\in[0:2]} J_m^* \tag{47}$$

and denoting $K' = [n] \setminus K$, we have

$$(\boldsymbol{e}_1)_{J_0^* \cup K'} = (\boldsymbol{e}_2)_{J_0^* \cup K'} = \boldsymbol{0}$$

and, so,

$$(\boldsymbol{u}_1 - \boldsymbol{u}_2)(G)_{J_0^* \cup K'} = \boldsymbol{0}.$$

Similarly,

$$(\boldsymbol{u}_2 - \boldsymbol{u}_0)(G)_{J_1^* \cup K'} = \boldsymbol{0} \quad \text{and} \quad (\boldsymbol{u}_0 - \boldsymbol{u}_1)(G)_{J_2^* \cup K'} = \boldsymbol{0}.$$

Defining $\boldsymbol{a}_m = \boldsymbol{u}_{m+1} - \boldsymbol{u}_{m+2}$ (with indexes taken modulo 3) we thus get for every $m \in [0:2]$ that $\boldsymbol{a}_m \ne \boldsymbol{0}$ and

$$\boldsymbol{a}_m (G)_{J_m^* \cup K'} = \boldsymbol{0}. \tag{48}$$

Moreover,

$$\sum_{m\in[0:2]} \boldsymbol{a}_m = \boldsymbol{0}. \tag{49}$$

Write

$$w^* = |K'| = n - |K| \overset{(41)+(47)}{=} 2k - n + w, \tag{50}$$

let P be a $(k-w^*) \times k$ matrix whose rows form a basis of the left kernel of $(G)_{K'}$, and let $H^* = (PG)_K$. We observe that H^* is a $(k-w^*) \times (n-w^*)$ parity-check matrix of the linear $[n^*{=}n{-}w^*, k^*{=}n{-}k]$ code, $\mathcal{C}^*$, which is obtained by puncturing the dual code $\mathcal{C}^\perp$ on the coordinate set K'. It follows from (48) that each vector $\boldsymbol{a}_m$ belongs to the row span of P, namely, we can write $\boldsymbol{a}_m = \boldsymbol{b}_m P$ for a unique nonzero $\boldsymbol{b}_m \in F^{k-w^*}$. By (48)–(49) we conclude that for $m \in [0:2]$,

$$\boldsymbol{b}_m (H^*)_{J_m^*} = \boldsymbol{0},$$

and, in addition,

$$\sum_{m\in[0:2]} \boldsymbol{b}_m = \boldsymbol{0},$$

namely, the (nonzero) vector $(\boldsymbol{b}_1 \,|\, \boldsymbol{b}_2)$ is in the left kernel of the matrix $M_{J_0^*, J_1^*, J_2^*}(H^*)$ (as in (9)). By (41) and (50), this matrix has $2(k-w^*) = 2(n-k-w)$ rows and the same number of columns and, therefore, it is singular.

Finally, we show that $\mathcal{C}^*$ and the subsets J_0^*, J_1^*, and J_2^* satisfy the conditions of Lemma 16 for list size $L = 2$ and code parameters $[n^*, k^*]$. These subsets are obviously disjoint; by (43) they are nonempty; the code rate is given by

$$\frac{k^*}{n^*} = \frac{n-k}{n-w^*} \overset{(50)}{=} \frac{n-k}{2(n-k)-w} \ge \frac{1}{2}$$

(as in (30)); $\mathcal{C}^*$ is MDS; and (42) and (41) (along with the equality $n^* - k^* = n - k - w$) imply, respectively, conditions (S1) and (S2).

Thus, starting off with the assumption that $\mathcal{C}$ is not 2-MDS, we obtain from Lemma 16 that $\mathcal{C}^*$ is not lightly-2-MDS. Hence, by Theorem 18 we conclude that $\mathcal{C}^\perp$ is not 2-MDS. ☐

We note that Theorem 19 does not generalize to 3-MDS codes: the code in Example 6 is not 3-MDS while, by Example 3, its dual code is.

Another application of Theorem 18 is the next result, which states that over sufficiently large fields (namely, exponential in the code length), almost all linear codes are 2-MDS. The proof can be found in Appendix A.

Theorem 20. *Given $n \in \mathbb{Z}^+$ and $k \in [n]$, all but a fraction $O(5^n/q)$ of the linear $[n,k]$ codes over F are 2-MDS.*

V. The 2-MDS GRS case

In this section, we describe an explicit construction of 2-MDS $[n, n{-}\rho]$ GRS codes over extension fields F of $\mathrm{GF}(2)$ of size q which is polynomial n, provided that the redundancy ρ is regarded as a constant (by Theorem 19, the dual code will be a 2-MDS $[n,\rho]$ code over F). The degree of the polynomial in n, however, grows rapidly with ρ, so this result may have a limited practical value; still, the field size herein is generally much smaller than that in the explicit construction of [27].

The construction will be presented in Section V-B; then, in Section V-C, we fine-tune the construction to yield more favorable parameters for the special case $\rho = 3$. Section V-A introduces some notation that will be used in the analysis of the construction.

A. Notation and preliminary analysis

Recall that an $[n, k]$ GRS code over F is a linear $[n, k]$ code with an $(n-k) \times n$ parity-check matrix $H_{\mathrm{GRS}} = (H_{i,j})$ of the form

$$H_{i,j} = v_j \alpha_j^i, \quad i \in [0 : n{-}k{-}1],\ j \in [n], \tag{51}$$

where $\alpha_1, \alpha_2, \ldots, \alpha_n$ are the code locators, which are distinct elements of F, and $v_1, v_2, \ldots, v_n$ are the column multipliers, which are nonzero elements of F. GRS codes are MDS and are closed under puncturing and under duality [22, pp. 303–304]. While the freedom of selecting the column multipliers is needed in order to establish these closures, the choice of their values is immaterial for the purpose of this work, so we will assume henceforth that they are all 1. Substituting $v_j = 1$ in (51) and writing $\boldsymbol{\alpha} = (\alpha_j)_{j \in [n]}$, the respective GRS code will be denoted by $\mathcal{C}_k(\boldsymbol{\alpha})$.

Given an integer $\rho \geq 3$, let $\boldsymbol{\rho} = (\rho_0, \rho_1, \rho_2)$ be a partition of 2ρ where $2 \leq \rho_0 \leq \rho_1 \leq \rho_2 < \rho$ are positive integers (that sum to 2ρ) and let $(\Upsilon_0, \Upsilon_1, \Upsilon_2)$ be the following partition of the set $[2\rho]$:

$$\begin{aligned} \Upsilon_0 &= [\rho_0], \\ \Upsilon_1 &= [\rho_0 + 1 : \rho_0 + \rho_1], \\ \Upsilon_2 &= [\rho_0 + \rho_1 + 1 : 2\rho] \end{aligned} \tag{52}$$

(so that $|\Upsilon_m| = \rho_m$, for $m \in [0:2]$). For a vector $\boldsymbol{x} = (x_\ell)_{\ell \in [2\rho]}$ of indeterminates, we define the $(2\rho) \times (2\rho)$ parametrized matrix $M_{\boldsymbol{\rho}}(\boldsymbol{x})$ by

$$M_{\boldsymbol{\rho}}(\boldsymbol{x}) = \left(\begin{array}{c|c|c} (-x_\ell^i)_{i=0,\ell\in\Upsilon_0}^{\rho-1} & (x_\ell^i)_{i=0,\ell\in\Upsilon_1}^{\rho-1} & \\ \hline (-x_\ell^i)_{i=0,\ell\in\Upsilon_0}^{\rho-1} & & (x_\ell^i)_{i=0,\ell\in\Upsilon_2}^{\rho-1} \end{array} \right). \tag{53}$$

Thus, $\det(M_{\boldsymbol{\rho}}(\boldsymbol{x})$ is an element of the ring, $F[\boldsymbol{x}]$, of multivariate polynomials in the entries of $\boldsymbol{x}$ over F.

When applying Lemma 16 to test whether a given $[n, k]$ GRS code $\mathcal{C}_k(\boldsymbol{\alpha})$ is lightly-2-MDS, we need to check whether $\det(M_{J_0,J_1,J_2}(H_{\mathrm{GRS}})) \neq 0$ for all triples (J_0, J_1, J_2) of disjoint nonempty subsets of $[n]$ that satisfy condition (S1) in Remark 12 and condition (S2) in Lemma 16. Equivalently,we need to check whether

$$\det(M_{\boldsymbol{\rho}}(\boldsymbol{x})) \neq 0$$

for all partitions $\boldsymbol{\rho} = (\rho_0, \rho_1, \rho_2)$ of $2(n-k)$ with $2 \leq \rho_0 \leq \rho_1 \leq \rho_2 < n-k$ and for all triples (J_0, J_1, J_2) of disjoint subsets of $[n]$ of sizes $|J_m| = \rho_m$, while substituting $(\boldsymbol{x})_{\Upsilon_m} = (\boldsymbol{\alpha})_{J_m}$, for $m \in [0:2]$.

Example 8. For $\rho = 3$ and the partition $(2, 2, 2)$:

$$M_{2,2,2}(\boldsymbol{x}) = \left(\begin{array}{cc|cc|cc} -1 & -1 & 1 & 1 & & \\ -x_1 & -x_2 & x_3 & x_4 & & \\ -x_1^2 & -x_2^2 & x_3^2 & x_4^2 & & \\ \hline -1 & -1 & & & 1 & 1 \\ -x_1 & -x_2 & & & x_5 & x_6 \\ -x_1^2 & -x_2^2 & & & x_5^2 & x_6^2 \end{array} \right),$$

and we have

$$\begin{aligned} \det(M_{2,2,2}(\boldsymbol{x})) &= -(x_2 - x_1)(x_4 - x_3)(x_6 - x_5) \\ &\quad \cdot \big(x_1 x_2 (x_3 + x_4 - x_5 - x_6) \\ &\quad + x_3 x_4 (x_5 + x_6 - x_1 - x_2) \\ &\quad + x_5 x_6 (x_1 + x_2 - x_3 - x_4) \big) \\ &= -(x_2 - x_1)(x_4 - x_3)(x_6 - x_5) \\ &\quad \cdot \det(S_{2,2,2}(\boldsymbol{x})), \end{aligned}$$

where

$$S_{2,2,2}(\boldsymbol{x}) = \begin{pmatrix} 1 & -x_1 - x_2 & x_1 x_2 \\ 1 & -x_3 - x_4 & x_3 x_4 \\ 1 & -x_5 - x_6 & x_5 x_6 \end{pmatrix}. \tag{54}$$

Namely, the expansion of $\det(S_{2,2,2}(\boldsymbol{x}))$ yields the sum of all 12 monomials of the form

$$\pm x_j \cdot \prod_{\ell \in \Upsilon_m} x_\ell, \tag{55}$$

where $m \in [0:2]$ and $j \in \Upsilon_{m+1} \cup \Upsilon_{m+2}$ (with the indexes $m{+}1$ and $m{+}2$ taken modulo 3), and the minus sign is taken when $j \in \Upsilon_{m+2}$.

For $n - k = 3$, the subsets J_m in the test of Lemma 16 (with condition (S1) taken from Remark 12) are all of size 2, namely, $\boldsymbol{x}$ will range over all (unordered) triples of (unordered) pairs of code locators. $\square$

For general $\rho \geq 3$ and partition $\boldsymbol{\rho}$ of $[2\rho]$, the Leibniz expansion of $\det(M_{\boldsymbol{\rho}}(\boldsymbol{x}))$ results in a sum of monomials of the form

$$\pm \prod_{\ell \in [2\rho]} x_\ell^{r_\ell}, \tag{56}$$

where $\boldsymbol{r} = (r_\ell)_{\ell \in [2\rho]}$ ranges over the elements in $[0 : \rho - 1]^{2\rho}$ that satisfy the following conditions:

R1) for each element $r \in [0 : \rho - 1]$ there are exactly two distinct indexes ℓ and ℓ' in $[2\rho]$ for which $r_\ell = r_{\ell'} = r$, and—

R2) those two indexes belong to distinct sets Υ_m

(see [27, Lemma 4.2]). The number of such monomials is given by[10]

$$\begin{aligned} N_{\boldsymbol{\rho}} &= \binom{\rho}{\rho-\rho_0 \quad \rho-\rho_1 \quad \rho-\rho_2} \prod_{m\in[0:2]} (\rho_m!) \\ &= \rho! \prod_{m\in[0:2]} \frac{\rho_m!}{(\rho-\rho_m)!} \end{aligned}$$

(which is always an even integer), and each monomial is of total degree

$$\sum_{\ell\in[2\rho]} r_\ell = 2 \sum_{r\in[0:\rho-1]} r = \rho(\rho-1). \tag{57}$$

In particular, $\det(M_{\boldsymbol{\rho}}(\boldsymbol{x}))$ is a homogeneous multivariate polynomial in the entries of $\boldsymbol{x}$ over F.

More properties of $\det(M_{\boldsymbol{\rho}}(\boldsymbol{x}))$ are presented in Appendix B.

Write

$$N(\rho) = \frac{1}{2} \max_{\boldsymbol{\rho}} N_{\boldsymbol{\rho}},$$

where the maximization is over all partitions $\boldsymbol{\rho} = (\rho_0, \rho_1, \rho_2)$ of 2ρ such that $2 \le \rho_0 \le \rho_1 \le \rho_2 < \rho$. It can be readily verified that the maximum is attained when ρ_0, ρ_1, and ρ_2 are (as close as they can get to being) equal, in which case the Stirling approximation for the binomial coefficients yields

$$N(\rho) \approx 2\sqrt{\pi\rho} \cdot \left(\frac{4}{3\,\mathrm{e}^2}\right)^{\rho} \cdot \rho^{2\rho}$$

(where e is the base of natural logarithms) [22, p. 309]; in particular,

$$N(\rho) < \rho^{2\rho}. \tag{58}$$

B. 2-MDS GRS construction for small fixed redundancy

We describe in this section a construction of a 2-MDS $[n, n-\rho]$ GRS code over extension fields F of GF(2) of size (much) smaller than $n^{\rho^{2\rho}}$.

Let ρ and h be integers such that $3 \le \rho < 2^h$; note that by (58) we have $2^{\rho(\rho-1)h} \ge N(\rho)$. Write $n = 2^h$ (which will be the code length) and let K be the field $\mathrm{GF}(2^{\rho(\rho-1)h})$ (which is of size at least $N(\rho)$). Also, let $\{b_j\}_{j\in[n]}$ be the set of elements of $\mathrm{GF}(2^h)$ and let β be an element in K that is not in any proper subfield of K. For $j \in [n]$, define the following n elements of K:

$$\beta_j = \beta + b_j. \tag{59}$$

These elements have the following property: for any two multisets $J, J' \subseteq [n]$, each of size $\le \rho(\rho-1)$, the equality

$$\prod_{j\in J} \beta_j = \prod_{j'\in J'} \beta_{j'}$$

can hold only if $J = J'$ (see [6], [7], [10]). As such, the set $\{\beta_j\}_{j\in[n]}$ forms a (generalized multiplicative) *Sidon set* [3].

Fix arbitrarily $N(\rho)$ distinct elements $\xi_1, \xi_2, \ldots, \xi_{N(\rho)} \in K$ and for each $j \in [n]$, let $\lambda_j(z)$ be the (unique) polynomial of degree $< N(\rho)$ over K that interpolates through the $N(\rho)$ points $\left\{(\xi_i, \beta_j^{2i-1})\right\}_{i\in[N(\rho)]}$:

$$\lambda_j(\xi_i) = \beta_j^{2i-1}, \quad i \in [N(\rho)]. \tag{60}$$

We now define the underlying field of the code to be the extension field $F = \mathrm{GF}(2^{\mu(\rho)\rho(\rho-1)h})$ of K of extension degree

$$\mu(\rho) = \rho(\rho-1)(N(\rho)-1)+1.$$

Finally, letting γ be a root in F of a degree-$\mu(\rho)$ irreducible polynomial over K,[11] our construction is the GRS code $\mathcal{C}_{n-\rho}(\boldsymbol{\alpha})$ over F with the code locators

$$\alpha_j = \lambda_j(\gamma), \quad j \in [n]. \tag{61}$$

We turn to analyzing the construction. Let $\kappa : [2\rho] \to [n]$ be an arbitrary injective mapping and substitute

$$\boldsymbol{x} = (x_\ell)_{\ell\in[2\rho]} \leftarrow \left(\alpha_{\kappa(\ell)}\right)_{\ell\in[2\rho]} \tag{62}$$

into the monomials (56) to obtain the following $N_{\boldsymbol{\rho}}$ elements of F:

$$\prod_{\ell\in[2\rho]} \alpha_{\kappa(\ell)}^{r_\ell}, \tag{63}$$

where $\boldsymbol{r} = (r_\ell)_{\ell\in[2\rho]}$ ranges over all (2ρ)-tuples that satisfy conditions (R1)–(R2). We show that the elements (63) do not sum to zero, namely, $\det(M_{\boldsymbol{\rho}}(\boldsymbol{x})) \ne 0$ under the substitution (62).

Suppose to the contrary that the elements (63) sum to zero, namely,

$$\sum_{\boldsymbol{r}} \prod_{\ell\in[2\rho]} \alpha_{\kappa(\ell)}^{r_\ell} = 0. \tag{64}$$

By the definition of the code locators in (61):

$$\sum_{\boldsymbol{r}} \prod_{\ell\in[2\rho]} \left(\lambda_{\kappa(\ell)}(\gamma)\right)^{r_\ell} = 0. \tag{65}$$

On the other hand, by the definition of the polynomials $\lambda_j(z)$:

$$\deg \prod_{\ell\in[2\rho]} \left(\lambda_{\kappa(\ell)}(z)\right)^{r_\ell} \overset{(57)}{\le} \rho(\rho-1)(N(\rho)-1). \tag{66}$$

Noting that $\{\gamma^i\}_{i\in[0:\mu(\rho)-1]}$ is a basis of F over K, we get from (65) and (66) the following polynomial identity:

$$\sum_{\boldsymbol{r}} \prod_{\ell\in[2\rho]} \left(\lambda_{\kappa(\ell)}(z)\right)^{r_\ell} = 0. \tag{67}$$

Hence,

$$\sum_{\boldsymbol{r}} \prod_{\ell\in[2\rho]} \left(\lambda_{\kappa(\ell)}(\xi_i)\right)^{r_\ell} = 0, \quad i \in [N(\rho)],$$

[10] We do the enumeration over $\boldsymbol{r}$ by selecting for each $m \in [0:2]$ some ordering on Υ_m and selecting a list $\boldsymbol{m} = (m_0, m_1, \ldots, m_{\rho-1}) \in [0:2]^\rho$ such that each element $m \in [0:2]$ appears in $\boldsymbol{m}$ exactly $\rho - \rho_m$ times. Then, for $r = 0, 1, \ldots, \rho-1$, we set iteratively $r_\ell = r_{\ell'} = r$, where ℓ and ℓ' are the next indexes in line in Υ_{m_r+1} and Υ_{m_r+2}, respectively, according to the ordering that was selected on these subsets. The formula for $N_{\boldsymbol{\rho}}$ is the product of the number of possible lists $\boldsymbol{m}$ and the number of different orderings on each subset Υ_m.

[11] Since $\gcd(\mu(\rho), \rho(\rho-1)) = 1$, we can take the polynomial to be irreducible over $\mathrm{GF}(2^h)$ [21, p. 107]. Moreover, when $\gcd(\mu(\rho), h) = 1$ (which happens when, say, all the prime factors of h are prime factors of $\rho(\rho-1)(N(\rho)-1)$, e.g., when h is a power of 2), we can take the polynomial to be irreducible over GF(2).

which, by (60), becomes

$$\sum_{\boldsymbol{r}}\Big(\prod_{\ell\in[2\rho]}\beta_{\kappa(\ell)}^{r_\ell}\Big)^{2i-1}=0,\quad i\in[N(\rho)],$$

or

$$\sum_{\boldsymbol{r}}\theta_{\boldsymbol{r}}^{2i-1}=0,\quad i\in[N(\rho)],\tag{68}$$

where

$$\theta_{\boldsymbol{r}}=\prod_{\ell\in[2\rho]}\beta_{\kappa(\ell)}^{r_\ell}.$$

Now, the elements $\theta_{\boldsymbol{r}}$ are distinct for distinct $\boldsymbol{r}$ due to the Sidon property of the elements β_j, which means that the respective column vectors,

$$\Big(\theta_{\boldsymbol{r}}\ \theta_{\boldsymbol{r}}^3\ \ldots\ \theta_{\boldsymbol{r}}^{2N(\rho)-1}\Big)^\top\ \Big(\in K^{N(\rho)}\Big),$$

are linearly independent over $\mathrm{GF}(2)$, as they are (at most $2N(\rho)$) distinct columns of a parity-check matrix of a binary BCH code with minimum distance $> 2N(\rho)$ [22, Ch. 7, §6]. On the other hand, (68) implies that these vectors sum to zero, thereby reaching a contradiction.

Thus, we have shown that $\det(M_{\boldsymbol{\rho}}(\boldsymbol{x}))\neq 0$ under the substitution (62) and, so, we conclude from Lemma 16 that the code $\mathcal{C}_{n-\rho}(\boldsymbol{\alpha})$ is lightly-2-MDS. Furthermore, going through the analysis it is fairly easy to see that it implies that any puncturing of this code is lightly-2-MDS too. Hence, by Theorem 18, the code $\mathcal{C}_{n-\rho}(\boldsymbol{\alpha})$ is 2-MDS.

The field size q of F is related to the code length n and the redundancy ρ by:

$$q=2^{\mu(\rho)\rho(\rho-1)h}=n^{\mu(\rho)\rho(\rho-1)}<n^{N(\rho)\rho^2(\rho-1)^2}\overset{(58)}{\ll}n^{\rho^{2\rho+4}}.$$

In comparison, (the dual code of) the explicit construction in Theorem 1.7 in [27] requires a field size of $q=2^{\rho^n}$.

C. Fine tuning for the case of redundancy 3

We consider in this section the special case of $\rho=3$ (as in Example 8). We recall that by Theorem 1, the field size in this case must be at least $\Omega(n^2)$. In comparison, Theorem 1.6 in [27] implies the existence of (the dual code of) a 2-MDS $[n,n-3]$ GRS code over a field of size $O(n^6)$. The argument therein in fact suggests that such a code can be constructed by a randomized algorithm requiring $O(n^6)$ operations in F.[12]

Using the following (deterministic) iterative procedure, the field F can be taken to be of size $O(n^5)$. We select the first five code locators $(\alpha_j)_{j\in[5]}$ to be arbitrary distinct elements of F. Assuming now that we have selected the code locators $(\alpha_j)_{j\in[t-1]}$ for some $t>5$, we select α_t to be a new element in F such that $\det(S_{2,2,2}(\boldsymbol{x}))\neq 0$ (see (54)), where we substitute $(\boldsymbol{x})_{\Upsilon_m}\leftarrow(\alpha_1\,\alpha_2\,\ldots\,\alpha_t)_{J_m}$, $m\in[0:2]$, with $\{J_0,J_1,J_2\}$ ranging over all the (unordered) triples of disjoint subsets of $[t]$ such that $t\in J_0$. In other words, α_t should *not* solve the following linear equation in x_1,

$$\begin{aligned}&\big(x_2(x_3+x_4-x_5-x_6)-x_3x_4+x_5x_6\big)\cdot x_1\\&\quad= x_3x_4(x_2-x_5-x_6)-x_5x_6(x_2-x_3-x_4),\end{aligned}$$

for any assignment of (already selected) distinct code locators to $x_2,x_3,\ldots,x_6$. For any such assignment, this equation has at most one solution, unless the coefficients on both sides vanish, namely:

$$\begin{aligned}x_2(x_3+x_4-x_5-x_6)-x_3x_4+x_5x_6&=0\\x_3x_4(x_2-x_5-x_6)-x_5x_6(x_2-x_3-x_4)&=0.\end{aligned}$$

Yet these two equations cannot hold simultaneously. Indeed, regarding them as linear equations in x_2, they could both hold only if the 2×2 matrix

$$\begin{pmatrix}x_3+x_4-x_5-x_6 & -x_3x_4+x_5x_6\\ x_3x_4-x_5x_6 & -x_3x_4(x_5+x_6)+x_5x_6(x_3+x_4)\end{pmatrix}$$

were singular; but the determinant of this matrix equals $(x_3-x_5)(x_3-x_6)(x_4-x_5)(x_4-x_6)$, which is nonzero for all the assignments to x_3,x_4,x_5,x_6. It readily follows that when

$$15\cdot\binom{n-1}{5}<q,$$

we will be able to find a qualifying value for α_t, as long as $t\le n$.

In what follows, we modify the construction of Section V-B to yield an explicit construction of a 2-MDS $[n,n-3]$ GRS code over $\mathrm{GF}(n^{32})$, where n is an odd power of 2. Here, by "explicit" we mean that for any $j\in[n]$, the complexity of computing the jth code locator is polylogarithmic in n.[13]

Let $h\ge 3$ be an odd positive integer, write $n=2^h$, and let β be an element in $K=\mathrm{GF}(2^{2h})$ that is not in $\mathrm{GF}(2^h)$.[14] For $j\in[n]$, define the elements $\beta_j\in K$ as in (59): here the Sidon property means that $\beta_i\beta_j=\beta_{i'}\beta_{j'}$ implies $\{i,j\}=\{i',j'\}$. Fix six distinct elements $\xi_1,\xi_2,\ldots,\xi_6\in K$ and for each $j\in[n]$, let $\lambda_j(z)$ be of degree <6 over K that interpolates through $\big\{(\xi_i,\beta_j^{2i-1})\big\}_{i\in[6]}$ as in (60).

Let $F=\mathrm{GF}(2^{32h})=\mathrm{GF}(n^{32})$ and let γ be a root in F of a degree-16 irreducible polynomial over $\mathrm{GF}(4)$ (e.g., the polynomial $x^{16}+x^3+x^2+\omega$, where ω is a root of x^2+x+1); since h is odd, this polynomial is also irreducible over K. Our construction is the GRS code $\mathcal{C}_{n-3}(\boldsymbol{\alpha})$ with the code locators as in (61).

Let $\kappa:[6]\to[n]$ be an injective mapping and let $\Upsilon_0=\{1,2\}$, $\Upsilon_1=\{3,4\}$, and $\Upsilon_2=\{5,6\}$. Define the set

$$\mathcal{D}=\big\{(m,j)\,:\,m\in[0:2],\ \ j\in[6]\setminus\Upsilon_m\big\}$$

and the following 12 elements of F:

$$\theta_{m,j}=\beta_{\kappa(j)}\cdot\prod_{\ell\in\Upsilon_m}\beta_{\kappa(\ell)},\quad (m,j)\in\mathcal{D}.$$

[12] In fact, it is fairly easy to show that when $|F|$ grows with n at least as n^6 then, with high probability, any $3\times n$ matrix over F is a parity-check matrix of a linear MDS code over F which is also lightly-2-MDS and, thus, by Theorem 18, it is 2-MDS.

[13] Complexity is measured here by binary operations, as the arithmetic of the fields involved—including finding their representations—can be carried out in time complexity which is polylogarithmic in the field size [28].

[14] In the special case where h is a power of 3, the field K can be constructed as the polynomial ring modulo the binary polynomial $x^{2h}+x^h+1$, which is irreducible over $\mathrm{GF}(2)$ [14, p. 96].

We show that out of these 12 elements, there is at least one that differs from all the rest. Suppose to the contrary that for any given $(m,j) \in \mathcal{D}$ there exists $(m',j') \neq (m,j)$ in $\mathcal{D}$ such that $\theta_{m,j} = \theta_{m',j'}$. Taking $(m,j) = (0,3)$, we obtain:

$$\beta_{\kappa(3)}\beta_{\kappa(1)}\beta_{\kappa(2)} = \beta_{\kappa(j')} \prod_{\ell \in \Upsilon_{m'}} \beta_{\kappa(\ell)}.$$

By the Sidon property we have $1,2,3 \notin \Upsilon_{m'} \cup \{j'\}$, implying that $(m',j') = (2,4)$:

$$\beta_{\kappa(3)}\beta_{\kappa(1)}\beta_{\kappa(2)} = \beta_{\kappa(4)}\beta_{\kappa(5)}\beta_{\kappa(6)}.$$

Taking now $(m,j) = (0,4)$ we get, respectively:

$$\beta_{\kappa(4)}\beta_{\kappa(1)}\beta_{\kappa(2)} = \beta_{\kappa(3)}\beta_{\kappa(5)}\beta_{\kappa(6)}.$$

The last two equations, in turn, lead to the contradiction:

$$\left(\beta_{\kappa(4)}/\beta_{\kappa(3)}\right)^2 = 1.$$

Our analysis now continues as in Section V-B. We assume to the contrary that under the substitution (62), the elements (55) sum to zero:

$$\sum_{(m,j)\in\mathcal{D}} \alpha_{\kappa(j)} \cdot \prod_{\ell \in \Upsilon_m} \alpha_{\kappa(\ell)} = 0$$

(compare with (64)). This, in turn, implies the polynomial identity

$$\sum_{(m,j)\in\mathcal{D}} \lambda_{\kappa(j)}(z) \cdot \prod_{\ell \in \Upsilon_m} \lambda_{\kappa(\ell)}(z) = 0$$

(similarly to (67)), and by substituting $x = \xi_i$ we obtain:

$$\sum_{(m,j)\in\mathcal{D}} \theta_{m,j}^{2i-1} = 0, \quad i \in [6] \tag{69}$$

(compare with (68)). Now, when $\theta_{m,j} = \theta_{m',j'}$ for two index pairs $(m,j) \neq (m',j')$ in $\mathcal{D}$, the respective two terms in (69) cancel each other; yet we have shown that at least one of the elements $\theta_{m,j}$ differs from all the rest. Denoting by $\mathcal{D}'$ the (nonempty) subset of index pairs in $\mathcal{D}$ that remain after such a cancellation, we have:

$$\sum_{(m,j)\in\mathcal{D}'} \theta_{m,j}^{2i-1} = 0, \quad i \in [6].$$

However, this is absurd: there can be no $|\mathcal{D}'| \leq 12$ linearly dependent columns in the parity-check matrix of a binary BCH code whose minimum distance exceeds 12.

APPENDIX A
SKIPPED PROOFS

Proof of Remark 2. Suppose that $\mathcal{C} \subseteq F^n$ is (τ, L)-list decodable where

$$\tau = \left\lceil \frac{L(n-k)}{L+1} \right\rceil, \tag{70}$$

i.e.,

$$q^{(L+1)\tau/L} \geq q^{n-k}.$$

Combining with (1) yields

$$q^{(L+1)\tau/L} \geq \frac{1}{L} \cdot \left(\frac{n(q-1)}{\tau}\right)^\tau.$$

Taking the τth root of both sides and then dividing by q we obtain

$$q^{1/L} \geq L^{-1/\tau} \cdot \frac{q-1}{q} \cdot \frac{n}{\tau} \geq \frac{1}{2L} \cdot \frac{n}{\tau}.$$

Hence,

$$q = \Omega\left(\left(\frac{n}{\tau}\right)^L\right) \stackrel{(70)}{=} \Omega\left(\left(\frac{n}{n-k}\right)^L\right),$$

where the hidden constants depend on L. □

Proof of Lemma 8. We will use the sphere-packing bound along with the inequality

$$V_q(n,\tau) \geq \frac{1}{\sqrt{2n}} \cdot q^{n\mathsf{H}_q(\tau/n)}, \tag{71}$$

where

$$\mathsf{H}_q(x) = \begin{cases} \dfrac{\mathsf{h}(x) + x\,\log_2(q-1)}{\log_2 q} & \text{if } x \in [0,(q-1)/q] \\ 1 & \text{if } x \in [(q-1)/q, 1] \end{cases}$$

(see [22, p. 309]). The function $x \mapsto \mathsf{H}_q(x)$ is continuous, non-decreasing, and concave on $[0,1]$, and is strictly increasing and strictly concave on $[0,(q-1)/q]$, with values ranging from $\mathsf{H}_q(0) = 0$ to $\mathsf{H}_q((q-1)/q) = 1$.

Define the function $x \mapsto f_q(x)$ on $[0,1]$ by

$$f_q(x) = \mathsf{H}_q\left(\frac{q-1}{q}\,x\right) - x \tag{72}$$

$$= \frac{1}{\log_2 q}\left(\mathsf{h}\left(\frac{q-1}{q}\,x\right) - x \cdot \mathsf{h}(1/q)\right); \tag{73}$$

this function will be of interest since

$$\eta_q(\varepsilon) \stackrel{(11)}{=} (\log_2 q) \cdot f_q(1-\varepsilon). \tag{74}$$

We have $f_q(0) = f_q(1) = 0$ which, when combined with (strict) concavity and continuity, implies that the minimum of f_q on the interval $[\varepsilon, 1-\varepsilon]$ is attained at one of the boundaries:

$$\min_{x\in[\varepsilon,1-\varepsilon]} f_q(x) = \min\left\{f_q(\varepsilon), f_q(1-\varepsilon)\right\} > 0.$$

In fact, for any $\varepsilon \in [0,1/2]$ we always have $f_q(\varepsilon) \geq f_q(1-\varepsilon)$, making $f_q(1-\varepsilon)$ the minimum. To verify this, we consider the following function $\varepsilon \mapsto \varphi_q(\varepsilon)$ on $[0,1/2]$:

$$\varphi_q(\varepsilon) = \mathsf{h}\left(\frac{q-1}{q}\,(1-\varepsilon)\right) - \mathsf{h}\left(\frac{q-1}{q}\,\varepsilon\right).$$

Differentiating twice with respect to ε yields

$$\frac{d^2\varphi_q(\varepsilon)}{d\varepsilon^2} = \frac{q-1}{\ln 2} \cdot \frac{1-2\varepsilon}{\varepsilon(1-\varepsilon)\,(1+(q-1)\varepsilon)\,(q-(q-1)\varepsilon)},$$

which is positive for every $\varepsilon \in (0,1/2)$. Hence φ_q is convex on $[0,1/2]$ and, since $\varphi_q(0) = \mathsf{h}(1/q)$ and $\varphi_q(1/2) = 0$, we get that it is bounded from above by the line $\varepsilon \mapsto (1-2\varepsilon)\cdot\mathsf{h}(1/q)$:

$$\varphi_q(\varepsilon) \leq (1-2\varepsilon) \cdot \mathsf{h}(1/q).$$

Therefore,

$$f_q(1-\varepsilon) - f_q(\varepsilon) \stackrel{(73)}{=} \frac{1}{\log_2 q}\left(\varphi_q(\varepsilon) - (1-2\varepsilon)\cdot\mathsf{h}(1/q)\right) \leq 0.$$

Now, let $\mathcal{C}$ be (τ, L)-list decodable where L and τ satisfy (12) and (13), and suppose in addition that $L \geq q-1$. By Theorem 1, the inequality (71), and the monotonicity of $\mathsf{H}_q(\cdot)$ we must have

$$\begin{aligned} 1-R \quad &\geq \quad \mathsf{H}_q\left(\frac{L}{L+1}(1-R)\right) - \frac{\log_q\left(L\sqrt{2n}\right)}{n} \\ &\overset{L\geq q-1}{\geq} \mathsf{H}_q\left(\frac{q-1}{q}(1-R)\right) - \frac{\log_q\left(L\sqrt{2n}\right)}{n}. \quad (75) \end{aligned}$$

Thus, for $R \in [\varepsilon, 1-\varepsilon]$,

$$\begin{aligned} f_q(1-\varepsilon) \quad &= \quad \min_{x\in[\varepsilon,1-\varepsilon]} f_q(x) \;\leq\; f_q(1-R) \\ &\overset{(72)+(75)}{\leq} \quad \frac{\log_q\left(L\sqrt{2n}\right)}{n} \quad \overset{(12)}{<} \quad \frac{\eta_q(\varepsilon)}{\log_2 q}, \end{aligned}$$

thereby contradicting (74). Hence, we must have $L < q-1$.

Finally, we show that $\eta_q(\varepsilon)$ is bounded from below by $\eta_2(\varepsilon)$ (which is positive and independent of q). We consider the bivariate function $(\varepsilon, q) \mapsto \eta_q(\varepsilon)$ as if q were real and differentiate with respect to q to obtain

$$\frac{\partial \eta_q(\varepsilon)}{\partial q} = \frac{1-\varepsilon}{q^2} \cdot \log_2\left(\frac{q}{1-\varepsilon} - (q-1)\right).$$

This derivative is positive for any $\varepsilon \in (0,1)$ and, so, $\eta_q(\varepsilon) \geq \eta_2(\varepsilon)$ for any field size q. □

Proof of Lemma 12. The lemma is immediate when $\ell \leq t$, so we assume hereafter in the proof that $\ell < \binom{w}{t}/\binom{w-s}{t}$. We construct the set X using the following iterative procedure. We start with $X \leftarrow \emptyset$, $Y \leftarrow [w]$, and $\mathcal{L} \leftarrow \{J_m\}_{m\in[\ell]}$, and for $i \leftarrow 1, 2, \ldots, t$, we select to X an element $y \in Y$ for which the size of

$$\left\{J_m \in \mathcal{L} \,:\, y \in J_m\right\} \quad (76)$$

is the largest (namely, there is no element in Y that is contained in more subsets $J_m \in \mathcal{L}$ than y). We then remove from $\mathcal{L}$ the subsets J_m that contain y and remove y from Y. We will show that after no more than t iterations, the set $\mathcal{L}$ becomes empty.

For $i \in [0:t]$, let ℓ_i denote the size of $\mathcal{L}$ right after the ith iteration, with $\ell_0 = \ell$. Right after that iteration we have

$$\sum_{J_m\in\mathcal{L}} |J_m| \geq \ell_i \cdot s,$$

which means that during the $(i+1)$st iteration, the size of (76) for a maximizing y is bounded from below by:

$$\left\lceil \frac{1}{w-i} \sum_{J_m\in\mathcal{L}} |J_m| \right\rceil \geq \left\lceil \frac{\ell_i \cdot s}{w-i} \right\rceil.$$

This leads to the inequality

$$\ell_{i+1} \leq \ell_i - \left\lceil \frac{\ell_i \cdot s}{w-i} \right\rceil = \left\lfloor \ell_i \cdot \frac{w-i-s}{w-i} \right\rfloor$$

and, ignoring the integer truncation, we get the upper bound:

$$\ell_t \leq \ell_0 \cdot \prod_{i=0}^{t-1} \frac{w-i-s}{w-i} = \ell \cdot \binom{w-s}{t} \Big/ \binom{w}{t}.$$

But $\ell < \binom{w}{t}/\binom{w-s}{t}$ and, so $\ell_t < 1$, which means that $\ell_t = 0$. □

Proof of Lemma 14. The case $k=1$ is easily verified, so we assume hereafter that $k > 1$.

Let $w \in [n-k-2]$; then

$$n - w \geq k + 2 \geq k + 1 + \frac{1}{k-1} = \frac{k^2}{k-1}$$

and, so,

$$n - w \leq k(n-k-w).$$

The latter inequality, in turn, is equivalent to

$$(n-k-w)^2 \leq (n-k-w-1)(n-w)$$

which, in turn, is equivalent to

$$(n-k-w)\binom{n-w-1}{k} \leq (n-k-w-1)\binom{n-w}{k}.$$

Making the inequality strict by adding 1 to the right-hand side yields

$$(n-k-w)\binom{n-w-1}{k} < (n-k-w-1)\binom{n-w}{k} + 1,$$

which can be equivalently written as

$$\begin{aligned} &(n-k-w)\left(\binom{n-w-1}{k} - (n-k-w)\right) \\ &\quad < \quad (n-k-w-1)\left(\binom{n-w}{k} - (n-k-w+1)\right). \end{aligned}$$

Finally, dividing both sides by $(n-k-w)(n-k-w+1)$ yields

$$\vartheta_{w+1} < \vartheta_w.$$

□

Proof of Lemma 17. The "if" part follows from Examples 3 and 4. Turning to the "only if" part, let $\mathcal{C}$ be a linear $[n,k]$ code over F where $2 \leq k \leq n-2$ and suppose, without loss of generality, that $\mathcal{C}$ has a $k \times n$ generator matrix of the form $G = (I \,|\, A)$. We show that $\mathcal{C}$ cannot be 2-MDS by distinguishing between several cases, according to the number of zero entries in A. Denote by $\boldsymbol{g}_i$ the ith row of G.

Case 1: A contains at least two zero entries (say, within the first two rows). Take $\boldsymbol{e}_0 = \boldsymbol{0}$, $\boldsymbol{e}_1 = \boldsymbol{g}_1$, and $\boldsymbol{e}_2 = \boldsymbol{g}_2$. These three vectors are in the same (trivial) coset $\mathcal{C}$ and $\sum_{m\in[0:2]} \mathsf{w}(\boldsymbol{e}_m) \leq 2(n-k)$.

Case 2: A contains one zero entry (say, the last entry in the first row). Take $\boldsymbol{e}_0 = (\boldsymbol{0}_k \,|\, \boldsymbol{1}_{n-k-1} \,|\, 0)$ (with k leading zeros), $\boldsymbol{e}_1 = \boldsymbol{e}_0 + \boldsymbol{g}_1$, and $\boldsymbol{e}_2 = \boldsymbol{e}_0 + \boldsymbol{g}_2$. These three vectors are in the same coset of $\mathcal{C}$ and $\sum_{m\in[0:2]} \mathsf{w}(\boldsymbol{e}_m) = n-k+2 \leq 2(n-k)$.

Case 3: A contains no zero entries (i.e., it is all-1s). Take $\boldsymbol{e}_0 = (\boldsymbol{0}_k \,|\, \boldsymbol{1}_{n-k})$, $\boldsymbol{e}_1 = \boldsymbol{e}_0 + \boldsymbol{g}_1$, and $\boldsymbol{e}_2 = \boldsymbol{e}_0 + \boldsymbol{g}_2$. Again, these three vectors are in the same coset of $\mathcal{C}$ and $\sum_{m\in[0:2]} \mathsf{w}(\boldsymbol{e}_m) = n-k+2 \leq 2(n-k)$. □

Proof of Theorem 20. We first recall that for a uniformly distributed random $(n-k) \times n$ matrix $\boldsymbol{H}$ over F we have[15] (see [22, pp. 444–445]):

$$\mathsf{Prob}\{\operatorname{rank}(\boldsymbol{H}) < n-k\} = 1 - \prod_{i=0}^{n-k-1} \left(1 - q^{i-n}\right) \leq \sum_{i=0}^{n-k-1} q^{i-n} < \frac{1}{q^k(q-1)}. \quad (77)$$

In what follows, we consider uniformly distributed random linear $[n, k]$ codes over F; yet, we will find it more convenient to replace this probability space by that of uniformly distributed random $(n-k) \times n$ matrices over F, which will stand for the parity-check matrices of the codes. These two probability spaces would be the same if we conditioned the matrices to be of full rank $n-k$; however, since we seek only an upper bound on the probability of a code being non-2-MDS, the effect of this conditioning amounts to a constant factor (which approaches 1 as q increases). Specifically, for any $(n-k) \times n$ matrix H over F:

$$\begin{aligned}\mathsf{Prob}\{\boldsymbol{H} = H \,|\, \operatorname{rank}(\boldsymbol{H}) = n-k\} &\leq \frac{\mathsf{Prob}\{\boldsymbol{H} = H\}}{\mathsf{Prob}\{\operatorname{rank}(\boldsymbol{H}) = n-k\}} \\ &\overset{(77)}{<} \frac{1}{1 - q^{-k}/(q-1)} \cdot \mathsf{Prob}\{\boldsymbol{H} = H\}.\end{aligned}$$

For a random $(n-k) \times n$ matrix $\boldsymbol{H}$, we let $\mathcal{C}(\boldsymbol{H}) = \ker(\boldsymbol{H})$ be the code over F which is defined by the parity-check matrix $\boldsymbol{H}$.

Denote by $\mathcal{A}_{\mathrm{MDS}}$ the event that every set of $n-k$ columns in a random $(n-k) \times n$ matrix $\boldsymbol{H}$ is linearly independent (namely, that $\boldsymbol{H}$ is a parity-check matrix of an MDS code). A union bound yields the following upper bound on the probability of the complement event:

$$\mathsf{Prob}\{\overline{\mathcal{A}_{\mathrm{MDS}}}\} \overset{(77)}{<} \binom{n}{n-k} \cdot \frac{q^{-k}}{q-1} \leq \frac{2^n}{q}. \quad (78)$$

Thus, by Examples 3–5, the theorem holds when $\min\{k, n-k\} \leq 2$. We assume hereafter in the proof that $3 \leq k \leq n-3$.

Fix a subset $J \subseteq [n]$ of size $|J| = w$ in the (nonempty) range (34) and write $J' = [n] \setminus J$ and $r = n-k-w$. Also, fix three nonzero row vectors $\boldsymbol{e}_0, \boldsymbol{e}_1, \boldsymbol{e}_2 \in F^{n-w}$, with entries indexed by J', such that their supports are disjoint, $\sum_{m \in [0:2]} \mathsf{w}(\boldsymbol{e}_m) \leq 2r$, and the first nonzero entry in $\boldsymbol{e}_0$ equals 1.

Denote by $\mathcal{A}_J$ the event that $\operatorname{rank}((\boldsymbol{H})_J) = w$; note that $\mathcal{A}_{\mathrm{MDS}} \subseteq \mathcal{A}_J$. Also, let $\boldsymbol{H}^* = (\boldsymbol{P}\boldsymbol{H})_{J'}$, where $\boldsymbol{P}$ is an $r \times (n-k)$ matrix whose rows form a basis of the left kernel[16] of $(\boldsymbol{H})_J$; we recall from the proof of Theorem 18 that $\boldsymbol{H}^*$ is the parity-check matrix of the code $(\mathcal{C}(\boldsymbol{H}))_{J'}$, which is obtained by puncturing $\mathcal{C}(\boldsymbol{H})$ on the coordinate set J.

[15] We use boldface letters for random matrices and normal font for their realizations.

[16] The matrix $\boldsymbol{P}$ will be used only under conditioning on $\mathcal{A}_J$, in which case it indeed has r rows.

Define

$$\mathsf{P}(\boldsymbol{e}_0, \boldsymbol{e}_1, \boldsymbol{e}_2; J) = \mathsf{Prob}\{(\boldsymbol{H}^*\boldsymbol{e}_0^\top = \boldsymbol{H}^*\boldsymbol{e}_1^\top = \boldsymbol{H}^*\boldsymbol{e}_2^\top) \cap \mathcal{A}_{\mathrm{MDS}}\},$$

which is the probability that $\mathcal{C}(\boldsymbol{H})$ is MDS yet the three vectors $\boldsymbol{e}_0$, $\boldsymbol{e}_1$, and $\boldsymbol{e}_2$ are in the same coset of the punctured code $(\mathcal{C}(\boldsymbol{H}))_{J'}$. We have:

$$\begin{aligned}&\mathsf{P}(\boldsymbol{e}_0, \boldsymbol{e}_1, \boldsymbol{e}_2; J) \\ &\leq \mathsf{Prob}\{\boldsymbol{H}^*\boldsymbol{e}_0^\top = \boldsymbol{H}^*\boldsymbol{e}_1^\top = \boldsymbol{H}^*\boldsymbol{e}_2^\top \,|\, \mathcal{A}_J\} \cdot \mathsf{Prob}\{\mathcal{A}_J\} \\ &= q^{-2r} \cdot \mathsf{Prob}\{\mathcal{A}_J\} \leq q^{-2r}, \end{aligned} \quad (79)$$

where the equality follows from the following two facts:

(a) the random $r \times (n-w)$ matrix $\boldsymbol{H}^*$ is uniformly distributed (even under the conditioning on $\mathcal{A}_J$, since none of the elements in J indexes any column of $\boldsymbol{H}^*$), and—

(b) the vectors $\boldsymbol{e}_m$ have disjoint supports, and, so, the respective syndromes $\boldsymbol{H}^*\boldsymbol{e}_m^\top$ are statistically independent.

Let now $\mathcal{A}_{\mathrm{light}}$ be the event that every puncturing of $\mathcal{C}(\boldsymbol{H})$ on any w coordinates in the range (34) results in a lightly-2-MDS code. By a union bound we get:

$$\begin{aligned}&\mathsf{Prob}\{\overline{\mathcal{A}_{\mathrm{light}}} \cap \mathcal{A}_{\mathrm{MDS}}\} \\ &\leq \sum_J \sum_{(\boldsymbol{e}_0, \boldsymbol{e}_1, \boldsymbol{e}_2)} \mathsf{P}(\boldsymbol{e}_0, \boldsymbol{e}_1, \boldsymbol{e}_2; J) \\ &= \sum_J \sum_{(J_0, J_1, J_2)} \sum_{\substack{(\boldsymbol{e}_0, \boldsymbol{e}_1, \boldsymbol{e}_2): \\ \mathsf{Supp}(\boldsymbol{e}_m) = J_m}} \mathsf{P}(\boldsymbol{e}_0, \boldsymbol{e}_1, \boldsymbol{e}_2; J),\end{aligned}$$

where J ranges over all subsets of $[n]$ of size w in the range (34) and (J_0, J_1, J_2) ranges over all triples of nonempty disjoint subsets of J' such that $\sum_{m \in [0:2]} |J_m| \leq 2r$. Hence,

$$\begin{aligned}&\mathsf{Prob}\{\overline{\mathcal{A}_{\mathrm{light}}} \cap \mathcal{A}_{\mathrm{MDS}}\} \\ &\overset{(79)}{\leq} \frac{q^{-2r}}{q-1} \sum_J \sum_{(J_0, J_1, J_2)} \prod_{m \in [0:2]} (q-1)^{|J_m|} \\ &\leq \frac{1}{q-1} \left(\frac{q-1}{q}\right)^{2r} \sum_J \sum_{(J_0, J_1, J_2)} 1 \\ &\leq \frac{5^n}{q-1} \left(\frac{q-1}{q}\right)^{2r} = O(5^n/q).\end{aligned} \quad (80)$$

We conclude that

$$\begin{aligned}&\mathsf{Prob}\{\overline{\mathcal{A}_{\mathrm{light}} \cap \mathcal{A}_{\mathrm{MDS}}}\} \\ &= \mathsf{Prob}\{\overline{\mathcal{A}_{\mathrm{light}}} \cup \overline{\mathcal{A}_{\mathrm{MDS}}}\} \\ &= \mathsf{Prob}\{\overline{\mathcal{A}_{\mathrm{light}}} \cap \mathcal{A}_{\mathrm{MDS}}\} + \mathsf{Prob}\{\overline{\mathcal{A}_{\mathrm{MDS}}}\} \\ &\overset{(78)+(80)}{=} O(5^n/q),\end{aligned}$$

and the sought result follows from Theorem 18. □

Remark 14. The upper bound, $O(5^n/q)$, on the fraction of non-2-MDS codes in Theorem 20 is not the best possible; e.g., it can be improved when $k \neq n/2$, yet we omit the details. □

APPENDIX B
PROPERTIES OF THE DETERMINANT OF $M_{\boldsymbol{\rho}}(\boldsymbol{x})$

Given an integer $\rho \geq 3$, let $\boldsymbol{\rho} = (\rho_0, \rho_1, \rho_2)$ be a partition of 2ρ where $2 \leq \rho_0 \leq \rho_1 \leq \rho_2 < \rho$ and let the subsets Υ_0, Υ_1, and Υ_2 be defined as in (52). Let $\boldsymbol{x} = (x_\ell)_{\ell \in [2\rho]}$ be a vector of indeterminates and define the matrix $M_{\boldsymbol{\rho}}(\boldsymbol{x})$ as in (53). For each $m \in [0:2]$, write $\boldsymbol{x}_m = (\boldsymbol{x})_{\Upsilon_m}$ and define the following degree-ρ_m polynomial in z over $F[\boldsymbol{x}_m]$:

$$\sigma_m(z) = \sigma_m(z; \boldsymbol{x}_m) = \prod_{\ell \in \Upsilon_m} (z - x_\ell) = \sum_{j=0}^{\rho_m} \sigma_{m,j} z^{\rho_m - j},$$

where $\sigma_{m,j} = \sigma_{m,j}(\boldsymbol{x}_m)$ is a degree-j homogeneous polynomial in $F[\boldsymbol{x}_m]$ (in particular, $\sigma_{m,0} = 1$ and $\sigma_{m,\rho_m} = \prod_{\ell \in \Upsilon_m}(-x_\ell)$). We associate with $\sigma_m(z)$ the following $(\rho - \rho_m) \times \rho$ matrix over $F[\boldsymbol{x}_m]$,

$$E_\rho(\boldsymbol{x}_m) = \begin{pmatrix} \sigma_{m,0} & \sigma_{m,1} & \cdots & \sigma_{m,\rho_m} & 0 & \cdots & 0 \\ 0 & \sigma_{m,0} & \sigma_{m,1} & \cdots & \sigma_{m,\rho_m} & 0 & \vdots \\ \vdots & \ddots & \ddots & \ddots & \cdots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_{m,0} & \sigma_{m,1} & \cdots & \sigma_{m,\rho_m} \end{pmatrix},$$

and define the $\rho \times \rho$ matrix $S_{\boldsymbol{\rho}}(\boldsymbol{x})$ by stacking the matrices $E_\rho(\boldsymbol{x}_m)$, $m \in [0:2]$:

$$S_{\boldsymbol{\rho}}(\boldsymbol{x}) = \left(\begin{array}{c} E_\rho(\boldsymbol{x}_0) \\ \hline E_\rho(\boldsymbol{x}_1) \\ \hline E_\rho(\boldsymbol{x}_2) \end{array} \right) \tag{81}$$

(note that the number of rows in $S_{\boldsymbol{\rho}}(\boldsymbol{x})$ is indeed $\sum_{m \in [0:2]}(\rho - \rho_m) = \rho$). This matrix can be seen as the generalization to three polynomials—namely, $\sigma_0(z)$, $\sigma_1(z)$, and $\sigma_2(z)$—of the Sylvester matrix of two polynomials; the determinant of the Sylvester matrix equals the resultant of the two polynomials [31, §9.2].

We have the following theorem, which allows us to use the matrix $S_{\boldsymbol{\rho}}(\boldsymbol{x})$ when applying Lemma 16 to test if a given GRS code is lightly-2-MDS.

Theorem 21. *Using the notation above, the following two conditions are equivalent for every vector $\boldsymbol{\alpha} = (\alpha_\ell)_{\ell \in [2\rho]}$ over any extension field of F.*

(i) $\det(M_{\boldsymbol{\rho}}(\boldsymbol{\alpha})) \neq 0$.

(ii) $\det(S_{\boldsymbol{\rho}}(\boldsymbol{\alpha})) \neq 0$, *and for every $m \in [0:2]$, the entries of $(\boldsymbol{\alpha})_{\Upsilon_m}$ are all distinct.*

Proof. Let Φ be an extension field of F which contains the entries of $\boldsymbol{\alpha}$ and, for $m \in [0:2]$, write $\boldsymbol{\alpha}_m = (\boldsymbol{\alpha})_{\Upsilon_m}$. Clearly, if there are two identical entries in $\boldsymbol{\alpha}_m$ for some $m \in [0:2]$ then $\det(M_{\boldsymbol{\rho}}(\boldsymbol{\alpha})) = 0$. Thus, we assume hereafter in the proof that for each $m \in [0:2]$, all the entries of $\boldsymbol{\alpha}_m$ are distinct. We denote by $\Phi_\rho[z]$ the set of polynomials (in the indeterminate z) of degree less than ρ over Φ.

The condition $\det(M_{\boldsymbol{\rho}}(\boldsymbol{\alpha})) = 0$ is equivalent to having two row vectors, $\boldsymbol{a}_1, \boldsymbol{a}_2 \in \Phi^\rho$, not both zero, such that

$$(\boldsymbol{a}_1 \,|\, \boldsymbol{a}_2)\, M_{\boldsymbol{\rho}}(\boldsymbol{\alpha}) = \boldsymbol{0}.$$

This condition, in turn, is equivalent to having two polynomials, $a_1(z), a_2(z) \in \Phi_\rho[z]$, not both zero, such that

$$a_0(\alpha_\ell) + a_1(\alpha_\ell) = 0, \quad \ell \in \Upsilon_0 \tag{82}$$

and, for $m = 1, 2$:

$$a_m(\alpha_\ell) = 0, \quad \ell \in \Upsilon_m \tag{83}$$

(namely, $\boldsymbol{a}_m$ is the vector of coefficients of $a_m(z)$, with the first entry in $\boldsymbol{a}_m$ being the free coefficient of $a_m(z)$). Conditions (82) and (83) are equivalent to having three polynomials, $a_0(z), a_1(z), a_2(z) \in \Phi_\rho[z]$, not all zero, such that

$$\sum_{m \in [0:2]} a_m(z) = 0$$

and, for $m \in [0:2]$:

$$a_m(\alpha_\ell) = 0, \quad \ell \in \Upsilon_m.$$

The latter equation means that $\sigma_m(z; \boldsymbol{\alpha}_m)$ is a divisor of $a_m(z)$ in $\Phi[z]$, namely, for each $m \in [0:2]$ there exists $b_m(z) \in \Phi_{\rho - \rho_m}(z)$ such that

$$a_m(z) = b_m(z) \cdot \sigma_m(z; \boldsymbol{\alpha}_m).$$

We conclude that the condition $\det(M_{\boldsymbol{\rho}}(\boldsymbol{\alpha})) = 0$ is equivalent to having polynomials $b_m(z) \in F_{\rho - \rho_m}(z)$, $m \in [0:2]$, not all zero, such that

$$\sum_{m \in [0:2]} b_m(z) \cdot \sigma_m(z; \boldsymbol{\alpha}_m) = 0.$$

Denoting by $\boldsymbol{b}_m$ the vector of coefficients (in $\Phi^{\rho - \rho_m}$) of $b_m(z)$, with the free coefficient being the *last* entry in $\boldsymbol{b}_m$, the last equation can be written in vector form as

$$\sum_{m \in [0:2]} \boldsymbol{b}_m \cdot E_\rho(\boldsymbol{\alpha}_m) = \boldsymbol{0},$$

or as

$$(\boldsymbol{b}_0 \,|\, \boldsymbol{b}_1 \,|\, \boldsymbol{b}_2)\, S_{\boldsymbol{\rho}}(\boldsymbol{\alpha}) = \boldsymbol{0},$$

where $\boldsymbol{b}_0, \boldsymbol{b}_1, \boldsymbol{b}_2$ are not all zero. Yet this can hold if and only if $\det(S_{\boldsymbol{\rho}}(\boldsymbol{\alpha})) = 0$. □

It follows from Theorem 21 that $\boldsymbol{\alpha}$ $(\in \Phi^{2\rho})$ is a root of $\det(M_{\boldsymbol{\rho}}(\boldsymbol{x}))$ $(\in F[\boldsymbol{x}])$ if and only if it is a root of

$$\Gamma_{\boldsymbol{\rho}}(\boldsymbol{x}) = \det(S_{\boldsymbol{\rho}}(\boldsymbol{x})) \cdot \prod_{m \in [0:2]} \prod_{\substack{\ell, \ell' \in \Upsilon_m : \\ \ell > \ell'}} (x_\ell - x_{\ell'}).$$

In particular, $\det(M_{\boldsymbol{\rho}}(\boldsymbol{x}))$, as an element of $F[\boldsymbol{x}]$, is divisible by $x_\ell - x_{\ell'}$ for every two indexes $\ell \neq \ell'$ in the same subset Υ_m. Based on numerical results, we conjecture that

$$\det(M_{\boldsymbol{\rho}}(\boldsymbol{x})) = (-1)^{\rho(\rho_1+1)} \cdot \Gamma_{\boldsymbol{\rho}}(\boldsymbol{x}). \tag{84}$$

One evidence that supports this conjecture is that both $\det(M_{\boldsymbol{\rho}}(\boldsymbol{x}))$ and $\Gamma_{\boldsymbol{\rho}}(\boldsymbol{x})$ have the same total degree. Specifically, we pointed out in Section V-A that the Leibniz expansion of $\det(M_{\boldsymbol{\rho}}(\boldsymbol{x}))$ results in a sum of monomials (56) that satisfy conditions (R1) and (R2) and, therefore, each has total degree $\rho(\rho - 1)$. As for the degree of $\Gamma_{\boldsymbol{\rho}}(\boldsymbol{x})$, we have the following result.

Lemma 22. $\det(S_{\boldsymbol{\rho}}(\boldsymbol{x}))$ *is a homogeneous polynomial in* $F[\boldsymbol{x}]$ *of total degree*

$$\rho(\rho-1) - \sum_{m\in[0:2]} \binom{\rho_m}{2} = \rho^2 - \frac{1}{2}\sum_{m\in[0:2]} \rho_m^2.$$

Proof. A typical term in the Leibniz expansion of $\det(S_{\boldsymbol{\rho}}(\boldsymbol{x}))$ takes—up to a sign—the form

$$\prod_{m\in[0:2]} \prod_{i\in[\rho-\rho_m]} \sigma_{m,j_m(i)-i}, \tag{85}$$

where the list

$$\Big((j_0(i))_{i\in[\rho-\rho_0]} \;\Big|\; (j_1(i))_{i\in[\rho-\rho_1]} \;\Big|\; (j_2(i))_{i\in[\rho-\rho_2]} \Big)$$

forms a permutation on $[\rho]$. The term (85), when expressed as a multivariate polynomial in the entries of $\boldsymbol{x}$, has total degree

$$\begin{aligned} &\sum_{m\in[0:2]} \sum_{i\in[\rho-\rho_m]} (j_m(i)-i) \\ &= \Big(\sum_{m\in[0:2]} \sum_{i\in[\rho-\rho_m]} j_m(i)\Big) - \Big(\sum_{m\in[0:2]} \sum_{i\in[\rho-\rho_m]} i\Big) \\ &= \Big(\sum_{j\in[\rho]} j\Big) - \Big(\sum_{m\in[0:2]} \sum_{i\in[\rho-\rho_m]} i\Big) \\ &= \binom{\rho+1}{2} - \sum_{m\in[0:2]} \binom{\rho-\rho_m+1}{2} \\ &= \rho^2 - \frac{1}{2}\sum_{m\in[0:2]} \rho_m^2. \end{aligned}$$

□

It follows from the lemma that the equality (84) holds (up to a multiplying scalar) if $\det(S_{\boldsymbol{\rho}}(\boldsymbol{x}))$ is absolutely irreducible (namely, it is irreducible over any extension field of F) [26, p. 4]; however, this is yet to be shown (or to be disproved).[17]

The definitions of $M_{\boldsymbol{\rho}}(\boldsymbol{x})$ and $S_{\boldsymbol{\rho}}(\boldsymbol{x})$ can be generalized in a straightforward manner to any $L \geq 2$. Given a partition $\boldsymbol{\rho} = (\rho_0, \rho_1, \ldots, \rho_L)$ of $L\rho$ such that $L \leq \rho_0 \leq \rho_1 \leq \cdots \leq \rho_L < \rho$, the matrix $M_{\boldsymbol{\rho}}(\boldsymbol{x})$ will then have order $(L\rho)\times(L\rho)$ and the matrix $S_{\boldsymbol{\rho}}(\boldsymbol{x})$, which is obtained by stacking $L+1$ matrices (instead of three) in (81), will have order $\rho \times \rho$. Theorem 21 generalizes accordingly and, therefore, by Lemma 16, we can use the matrix $S_{\boldsymbol{\rho}}(\boldsymbol{x})$ to test if a given GRS code is lightly-L-MDS. The total degree of $\det(S_{\boldsymbol{\rho}}(\boldsymbol{x}))$ is

$$L\cdot\binom{\rho}{2} - \sum_{m\in[0:L]} \binom{\rho_m}{2} = \frac{L}{2}\cdot\rho^2 - \frac{1}{2}\sum_{m\in[0:L]} \rho_m^2$$

and, based on numerical evidence, we conjecture that

$$\det(M_{\boldsymbol{\rho}}(\boldsymbol{x})) = (-1)^s \cdot \det(S_{\boldsymbol{\rho}}(\boldsymbol{x})) \cdot \prod_{m\in[0:L]} \prod_{\substack{\ell,\ell'\in\Upsilon_m:\\ \ell>\ell'}} (x_\ell - x_{\ell'}),$$

where

$$s = \rho\cdot\Bigg(\sum_{\text{odd } m\in[0:L]} (\rho-\rho_m)\Bigg).$$

For a certain structure of assignments to the vector $\boldsymbol{x}$, the matrix $S_{\boldsymbol{\rho}}(\boldsymbol{x})$ also appears in [24, §§III–IV], as part of the analysis of the burst-error list decodability of (cyclic) Reed–Solomon codes. This matrix was also used in [30] as a tool for settling the GM–MDS conjecture.

[17] In the case of Example 8 it is absolutely irreducible. Note that this is in contrast with the case of the resultant of two polynomials, which is highly reducible: it factors into linear terms over the splitting field of the polynomials [31, p. 142]. Similar behavior of determinants is seen also in other problems (see [1] or [2]).

Acknowledgment

I would like to thank the authors of [13] for making me aware of their work.

References

[1] C. D'Andrea, M. Sombra, "The Cayley–Menger determinant is irreducible for $n \geq 3$," *Siberian Math. J.,* 46 (2005), 71–76.
[2] C. D'Andrea, L.F. Tabera, "Tropicalization and irreducibility of generalized Vandermonde determinants," *Proc. Am. Math. Soc.,* 137 (2009), 3647–3656.
[3] L. Babai, V.T. Sós, "Sidon sets in groups and induced subgraphs of Cayley graphs," *Europ. J. Comb.,* 6 (1985), 101–114.
[4] V. M. Blinovsky, "Bounds for codes in the case of list decoding of finite volume," *Probl. Inf. Transm.,* 22 (1986), 7–19.
[5] P.G. Bonneau, "Weight distribution of translates of MDS codes," *Combinatorica,* 10 (1990), 103–105.
[6] R.C. Bose, "An affine analogue of Singer's theorem," *J. Indian Math. Soc. (N.S.),* 6 (1942), 1–5.
[7] R.C. Bose, S. Chowla, "Theorems in the additive theory of numbers," *Comment. Math. Helv.,* 37 (1962/1963), 141–147.
[8] J. Brakensiek, S. Gopi, V. Makam, "Lower bounds for maximally recoverable tensor codes and higher order MDS codes," *IEEE Trans. Inf. Theory,* to appear.
[9] J. Brakensiek, S. Gopi, V. Makam, "Generic Reed–Solomon codes achieve list-decoding capacity," available online at https://arxiv.org/abs/2206.05256.
[10] S. Chowla, "Solution of a problem of Erdős and Turán in additive-number theory," *Proc. Nat. Acad. Sci. India. Sect. A.,* 14 (1944), 1–2.
[11] P. Elias, "List decoding for noisy channels," TR 335, Research Laboratory of Electronics, MIT, 1957.
[12] P. Elias, "Error-correcting codes for list decoding," *IEEE Trans. Inf. Theory,* 37 (1991), 5–12.
[13] E. Goldberg, C. Shangguan, I. Tamo, "Singleton-type bounds for list-decoding and list-recovery, and related results," available online at http://arxiv.org/abs/2112.05592.
[14] S.W. Golomb, *Shift Register Sequences,* Revised Edition, Aegean Park Press, Laguna Hills, California, 1982.
[15] V. Guruswami, "Limits to list decodability of linear codes," *Proc. 34th ACM Symp. on Theory of Computing (STOC),* Montréal, Québec, 2002, 802–811.
[16] V. Guruswami, S. Narayanan, "Combinatorial limitations of average-radius list decoding," *IEEE Trans. Inf. Theory,* 60 (2014), 5827–5842.
[17] V. Guruswami, A. Rudra, "Explicit codes achieving list decoding capacity: Error-correcting up to the Singleton bound," *IEEE Trans. Inf. Theory,* 54 (2008), 135–150.
[18] V. Guruswami, M. Sudan, "Improved decoding of Reed–Solomon and algebraic-geometric codes," *IEEE Trans. Inf. Theory,* 45 (1999), 1757–1767.
[19] V. Guruswami, S. Vadhan, "A lower bound on list size for list decoding," *IEEE Trans. Inf. Theory,* 56 (2010), 5681–5688.
[20] J. Justesen, Y. Høholdt, "Bounds on list decoding of MDS codes," *IEEE Trans. Inf. Theory,* 47 (2001), 1604–1609.
[21] R. Lidl, H. Niederreiter, *Finite Fields,* Second Edition, Cambridge University Press, Cambridge, 1997.
[22] F.J. MacWilliams, N.J.A. Sloane, *The Theory of Error-Correcting Codes,* North-Holland, Amsterdam, 1977.
[23] F. Parvaresh, A. Vardy, "Correcting errors beyond the Guruswami–Sudan radius in polynomial time," *Proc. 46th Symp. Foundations of Computer Science (FOCS)*, Pittsburgh, PA, 2005, 285–294.
[24] R.M. Roth, P.O. Vontobel, "List decoding of burst errors," *IEEE Trans. Inf. Theory,* 55 (2009), 4179–4190.

[25] G. Ruckenstein, R.M. Roth, "Bounds on the list-decoding radius of Reed–Solomon codes," *SIAM J. Discrete Math.,* 17 (2003), 171–195.

[26] I.R. Shafarevich, *Basic Algebraic Geometry 1: Varieties in Projective Space,* Third Edition, Springer, Heidelberg, 2013.

[27] C. Shangguan, I. Tamo, "Combinatorial list-decoding of Reed–Solomon codes beyond the Johnson radius," *Proc. 52nd Annual ACM Symp. on Theory of Computing (STOC),* Chicago, IL, 2020, 538–551.

[28] V. Shoup, "New algorithms for finding irreducible polynomials over finite fields," *Math. Comput.,* 54 (1990), 435–447.

[29] M. Sudan, "Decoding of Reed–Solomon codes beyond the error-correction bound," *J. Complexity,* 13 (1997), 180–193.

[30] H. Yildiz, B. Hassibi, "Optimum linear codes with support-constraints generator matrices over small fields," *IEEE Trans. Inf. Theory,* 65 (2019), 7868–7875.

[31] R. Zippel, *Effective Polynomial Computation,* Kluwer, Boston, 1993.

Ron M. Roth (Fellow, IEEE) received the B.Sc. degree in computer engineering, the M.Sc. in electrical engineering, and the D.Sc. in computer science from Technion—Israel Institute of Technology, Haifa, Israel, in 1980, 1984, and 1988, respectively. Since 1988 he has been with the Computer Science Department at Technion, where he now holds the General Yaakov Dori Chair in Engineering. During the academic years 1989–91 he was a Visiting Scientist at IBM Research Division, Almaden Research Center, San Jose, California. During 1996–97, 2004–05, and 2011–2012 he was on sabbatical leave at Hewlett–Packard Laboratories, Palo Alto, California, and in 2022 at Hewlett Packard Enterprise, Milpitas, California. He is the author of the book *Introduction to Coding Theory*, published by Cambridge University Press in 2006. Dr. Roth was an associate editor for coding theory in IEEE TRANSACTIONS ON INFORMATION THEORY from 1998 till 2001, and he is now serving as an associate editor in *SIAM Journal on Discrete Mathematics*. His research interests include coding theory, information theory, and their application to storage, computation, and the theory of complexity.