Burst List Decoding of Interleaved Reed–Solomon Codes

Tom Kolan and Ron M. Roth, *Fellow, IEEE*

***Abstract*—It is shown that interleaved Reed–Solomon codes can be list-decoded for burst errors while attaining the generalized Reiger bound for list decoding. A respective decoding algorithm is presented which is (significantly) more efficient than a burst list decoder for a non-interleaved Reed–Solomon code with comparable parameters. Finally, it is shown through counterexamples that, unlike the special case of Reed–Solomon codes, interleaving does not always preserve the list decoding properties of the constituent code.**

***Index Terms*—Burst errors, interleaving, list decoding, Reed–Solomon codes, Reiger bound.**

I. Introduction

Coding schemes for list decoding of isolated (random) errors—especially schemes that are based on Reed–Solomon codes (in short, RS codes) and derivatives thereof—have been studied quite extensively [1]–[6]. Less, however, has been published about list decoding of burst errors. In a recent paper [7], several bounds were obtained on the parameters of list decodable codes for single bursts. In particular, it was shown in [7] that if a linear code $\mathcal{C}$ over $\mathrm{GF}(q)$ has a list decoder that corrects any single burst error of length τ or less, then—under certain additional conditions that will be recalled below—the redundancy r of $\mathcal{C}$ is related to τ and the list size ℓ by the following generalization of the Reiger bound:

$$r \geq \tau + \left\lceil \frac{\tau}{\ell} \right\rceil$$

(with the case $\ell = 1$ corresponding to the classical Reiger bound [8], [9]). It was also shown in [7] that this bound is attained by (non-extended and possibly shortened) RS codes.

One drawback of RS codes is that their length is limited to at most $q-1$. In this work, we show (in Section III) that for the case where the list size ℓ divides the maximum burst length τ, a (τ/ℓ)-level interleaving of an RS code of redundancy $\ell+1$ has a burst list decoder with the specified ℓ and τ (for $\ell = 1$, this result is straightforward and well known). Observing that the overall redundancy of the interleaved code is

$$\frac{\tau}{\ell} \cdot (\ell + 1) = \tau + \frac{\tau}{\ell} \,, \tag{1}$$

we thus obtain a construction that attains the above generalization of the Reiger bound and is τ/ℓ times longer than just plainly taking one (non-interleaved) RS code. Furthermore, we demonstrate that interleaved RS codes can be list-decoded for bursts using an algorithm that is (significantly) more efficient than a burst list decoder for a non-interleaved RS code with comparable parameters (the complexity analysis is presented in Section IV). Finally, in Section V, we show through an example that, unlike the special case of RS codes, interleaving does not always preserve the list decoding properties of the constituent code.

Next, we introduce some definitions and notation.

Throughout this work, we let F denote the finite field $\mathrm{GF}(q)$. The order of an element γ in the multiplicative group F^* is denoted by $\mathrm{ord}(\gamma)$, and $F_\delta[x]$ denotes the set of all univariate polynomials of degree less than δ over F in the indeterminate x. For integers $i < j$, the notation $[i, j\rangle$ stands for the integer interval $\{k \in \mathbb{Z} : i \leq k < j\}$.

We say that a word $\boldsymbol{e} \in F^n$ is a *τ-burst* if either $\boldsymbol{e} = \mathbf{0}$ (the all-zero word) or the indexes i and j (in $[0, n\rangle$) of the first and last nonzero entries in $\boldsymbol{e}$ satisfy $j - i < \tau$. For a nonzero word $\boldsymbol{e}$, the notation $\lambda(\boldsymbol{e})$ will stand for the value of i above, and the value $j - i + 1$ will be referred to as the (actual) *length* of the τ-burst and will be denoted by $L(\boldsymbol{e})$ (where $L(\mathbf{0}) = 0$). The set of all τ-bursts in F^n will be denoted by $\mathcal{B}(n, \tau)$.

For a code $\mathcal{C}$ of length n over F and a word $\boldsymbol{y} \in F^n$, we denote by $\boldsymbol{y} - \mathcal{C}$ the set $\{\boldsymbol{y} - \boldsymbol{c} \,:\, \boldsymbol{c} \in \mathcal{C}\}$ (when $\mathcal{C}$ is linear, this set is a coset of $\mathcal{C}$ within F^n). The minimum (Hamming) distance of $\mathcal{C}$ will be denoted by $\mathsf{d}(\mathcal{C})$.

Let $\mathcal{C}$ be a code of length n over F. A *(list) decoder* for $\mathcal{C}$ is a mapping $\mathcal{D}$ from F^n to the set of subsets of F^n, such that for every $\boldsymbol{y} \in F^n$,

$$\mathcal{D}(\boldsymbol{y}) \subseteq \boldsymbol{y} - \mathcal{C}$$

(we knowingly deviate here from the standard definition of a decoder in that $\mathcal{D}$ returns a list of error words rather than codewords). The *list size* of a decoder $\mathcal{D}$ is the largest size of $\mathcal{D}(\boldsymbol{y})$ over all $\boldsymbol{y} \in F^n$.

We say that $\mathcal{D}$ *detects* any single τ-burst error if for every $\boldsymbol{c} \in \mathcal{C}$ and $\boldsymbol{e} \in \mathcal{B}(n, \tau)$,

$$\mathcal{D}(\boldsymbol{c} + \boldsymbol{e}) = \begin{cases} \{\mathbf{0}\} & \text{if } \boldsymbol{e} = \mathbf{0} \\ \emptyset & \text{otherwise} \end{cases} .$$

Such a decoder for $\mathcal{C}$ exists if and only if for every $\boldsymbol{c} \in \mathcal{C}$,

$$(\boldsymbol{c} - \mathcal{C}) \cap \mathcal{B}(n, \tau) = \{\mathbf{0}\} \,.$$

In particular, such a decoder exists if $\mathsf{d}(\mathcal{C}) > \tau$.

T. Kolan was with the Computer Science Department, Technion—Israel Institute of Technology, Haifa 32000, Israel. (e-mail: tomkolan@gmail.com).

R.M. Roth is with the Computer Science Department, Technion—Israel Institute of Technology, Haifa 32000, Israel. (e-mail: ronny@cs.technion.ac.il).

This work was presented in part at the *IEEE International Symposium on Information Theory,* Cambridge, Massachusetts, July 2012.

This work was supported in part by Grants Nos. 1280/08 and 1092/12 from the Israel Science Foundation, Jerusalem, Israel.

We say that $\mathcal{D}$ *corrects* any single τ-burst error if for every $\boldsymbol{c} \in \mathcal{C}$ and $\boldsymbol{e} \in \mathcal{B}(n,\tau)$,

$$\boldsymbol{e} \in \mathcal{D}(\boldsymbol{c}+\boldsymbol{e}) \ .$$

Equivalently, for every $\boldsymbol{y} \in F^n$,

$$(\boldsymbol{y}-\mathcal{C}) \cap \mathcal{B}(n,\tau) \subseteq \mathcal{D}(\boldsymbol{y}) \ . \tag{2}$$

An (ℓ,τ)-*burst list decoder* for $\mathcal{C}$ is a decoder for $\mathcal{C}$ of list size at most ℓ that corrects any single τ-burst error. From (2), it readily follows that such a decoder exists if and only if for every $\boldsymbol{y} \in F^n$,

$$|(\boldsymbol{y}-\mathcal{C}) \cap \mathcal{B}(n,\tau)| \le \ell \ .$$

The next theorem is the generalization of the Reiger bound which was proved in [7], specialized to linear codes.

Theorem 1.1 ([7, Thms. 2.2–2.3]): Let $\mathcal{C}$ be a linear code of length n over F and let τ and ℓ be positive integers that satisfy the following three conditions:

1) Either $(\ell+1)\tau \le n$, or $\ell \,|\, \tau$ and $2\tau \le n$.
2) There is a decoder for $\mathcal{C}$ that detects any single τ-burst error.
3) There is an (ℓ,τ)-burst list decoder for $\mathcal{C}$.

Then the redundancy r of $\mathcal{C}$ satisfies

$$r \ge \tau + \left\lceil \frac{\tau}{\ell} \right\rceil \ .$$

Conversely, it was also shown in [7] that RS codes attain this bound.

Theorem 1.2 ([7, Thm. 4.1]): For $0 \le r < n < q$, let $\mathcal{C}_{\mathrm{RS}}(n,r)$ denote the RS code of length n and redundancy r over F with a parity-check matrix

$$H_{\mathrm{RS}} = H_{\mathrm{RS}}(n,r) = \left(\alpha^{st}\right)_{s=0,\,t=0}^{r-1\ \ n-1} \ , \tag{3}$$

where $\alpha \in F^*$ with $\mathrm{ord}(\alpha) \ge n$. There is an (ℓ,τ)-burst list decoder for $\mathcal{C}_{\mathrm{RS}}(n,r)$, whenever ℓ and τ are positive integers that satisfy

$$r \ge \tau + \left\lceil \frac{\tau}{\ell} \right\rceil \ .$$

(Note that condition 2 in Theorem 1.1 is also satisfied in this case, since $\mathsf{d}(\mathcal{C}_{\mathrm{RS}}) = r+1 > \tau$.)

II. Tools

In this section, we consider the case $\tau = \ell$ and prove a refinement of Theorem 1.2 for this case: we show that the (ℓ,ℓ)-burst list decoder $\mathcal{D}$ guaranteed in Theorem 1.2 can be assumed to satisfy a certain relationship between the size of each list $\mathcal{D}(\boldsymbol{y})$ and the lengths of the bursts in $\mathcal{D}(\boldsymbol{y})$. In particular, it will follow from our result that if one or more of the decoded bursts turns out to have actual length that is strictly smaller than ℓ, then such "deficiency" in the burst length implies that the size of the decoded list size must, in fact, be strictly smaller than ℓ. The stronger properties of $\mathcal{D}$ that we show will then be used in Section III to prove that interleaved RS codes have an (ℓ,τ)-burst list decoder and, thus, they attain the bound of Theorem 1.1.

We recall the next theorem from [7], which will be used in the sequel.

Theorem 2.1 ([7, Thm. 3.1]): For integers $1 < m \le r < q$, let $\beta_0, \beta_1, \ldots, \beta_{m-1}$ be elements in F^* and let $\gamma \in F^*$ be such that $\mathrm{ord}(\gamma) \ge r$. Also, let $\mu_0, \mu_1, \ldots, \mu_{m-1}$ be positive integers such that

$$\sum_{i=0}^{m-1} \mu_i = r \ ,$$

and, for each $i \in [0,m\rangle$, define the polynomial

$$M_i(x;\beta_i,\gamma) = \prod_{t=0}^{r-1-\mu_i} (x - \beta_i\gamma^t) \tag{4}$$

(which is regarded as a univariate polynomial in the indeterminate x, with β_i and γ serving as parameters). The following two conditions are equivalent:

(i) There exist polynomials

$$u_i(x) \in F_{\mu_i}[x], \quad i \in [0,m\rangle \ ,$$

not all zero, such that

$$\sum_{i=0}^{m-1} u_i(x) M_i(x;\beta_i,\gamma) = 0 \ .$$

(ii) For some distinct $h,k \in [0,m\rangle$ and some integer b in the range $-\mu_h < b < \mu_k$,

$$\frac{\beta_k}{\beta_h} = \gamma^b \ .$$

Remark 2.1: Theorem 2.1 holds vacuously also when $m = 1$, provided that we formally define $M_0(x,\cdot,\cdot)$ to be the constant 1 in this case. ∎

The next theorem is our refinement of Theorem 1.2 for the case $\tau = \ell$.

Theorem 2.2: For integers $0 < \ell < r < n < q$, let $\mathcal{C}_{\mathrm{RS}}(n,r)$ be as in Theorem 1.2, and let the decoder $\mathcal{D}$ for $\mathcal{C}_{\mathrm{RS}}(n,r)$ be defined for every $\boldsymbol{y} \in F^n$ by

$$\mathcal{D}(\boldsymbol{y}) = \big(\boldsymbol{y} - \mathcal{C}_{\mathrm{RS}}(n,r)\big) \cap \mathcal{B}(n,\ell) \ .$$

Then $\mathcal{D}$ is an (ℓ,ℓ)-burst list decoder for $\mathcal{C}_{\mathrm{RS}}(n,r)$ and it satisfies the following two properties:

(i) For every $\boldsymbol{y} \in \mathcal{C}_{\mathrm{RS}}(n,\ell{+}1)$,

$$\mathcal{D}(\boldsymbol{y}) = \{\boldsymbol{0}\} \ .$$

(ii) For every $\boldsymbol{y} \in F^n \setminus \mathcal{C}_{\mathrm{RS}}(n,\ell{+}1)$,

$$\sum_{\boldsymbol{e} \in \mathcal{D}(\boldsymbol{y})} (\ell + 1 - L(\boldsymbol{e})) \le \ell \ . \tag{5}$$

Proof: First, $\mathcal{D}$ is a decoder that corrects any single ℓ-burst, since it satisfies the containment (2) (with equality) for $\tau = \ell$. To verify property (i), note that $\boldsymbol{y} - \mathcal{C}_{\mathrm{RS}}(n,r) \subseteq \mathcal{C}_{\mathrm{RS}}(n,\ell{+}1)$ and, therefore, the Hamming weight of every nonzero word in $\boldsymbol{y} - \mathcal{C}_{\mathrm{RS}}(n,r)$ exceeds $\ell{+}1$. Property (ii) follows from Lemma 2.3 below. Finally, since each term $\ell{+}1{-}L(\boldsymbol{e})$ in (5) is a positive integer, we get that $|\mathcal{D}(\boldsymbol{y})| \le \ell$ for every $\boldsymbol{y} \in F^n$ and, so, $\mathcal{D}$ is indeed an (ℓ,ℓ)-burst list decoder for $\mathcal{C}_{\mathrm{RS}}(n,r)$. ∎

Remark 2.2: The inequality (5) can be rewritten as

$$|\mathcal{D}(\boldsymbol{y})| \le \ell - \sum_{\boldsymbol{e} \in \mathcal{D}(\boldsymbol{y})} (\ell - L(\boldsymbol{e})) \ .$$

Thus, referring to the discussion at the beginning of this section, we indeed see that if there are nonzero bursts $\boldsymbol{e} \in \mathcal{D}(\boldsymbol{y})$ with actual length $L(\boldsymbol{e}) < \ell$, then they force the size of the list $\mathcal{D}(\boldsymbol{y})$ to be strictly smaller than ℓ. ∎

The next lemma establishes property (ii) in Theorem 2.2.

Lemma 2.3: For integers $0 < \ell < n < q$, let $\mathcal{C}_{\rm RS} = \mathcal{C}_{\rm RS}(n, \ell+1)$ be as defined in Theorem 1.2 and let

$$\boldsymbol{e}_0, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_{m-1}$$

be distinct words that belong to the same nontrivial coset $\boldsymbol{y} - \mathcal{C}_{\rm RS}$ (where $\boldsymbol{y} \notin \mathcal{C}_{\rm RS}$). Then

$$\sum_{i=0}^{m-1} (\ell + 1 - L(\boldsymbol{e}_i)) \leq \ell \; . \tag{6}$$

Proof: First, note that words $\boldsymbol{e}_i$ with $L(\boldsymbol{e}_i) > \ell$ contribute non-positive terms to the left-hand size of (6). Hence, it suffices to prove the lemma under the assumption that $\boldsymbol{e}_i \in \mathcal{B}(n, \ell)$ for all $i \in [0, m\rangle$. Furthermore, we will assume that $n = \mathrm{ord}(\alpha)$ (otherwise, append $\mathrm{ord}(\alpha) - n$ zeroes to each $\boldsymbol{e}_i$ and apply the proof to $\mathcal{C}_{\rm RS}(\mathrm{ord}(\alpha), \ell+1)$).

Our proof builds upon the one given in [7] for Theorem 1.2 but requires additional arguments in order to get to the finer result.

For $i \in [0, m\rangle$, write

$$\lambda_i = \lambda(\boldsymbol{e}_i) \; , \quad \tau_i = L(\boldsymbol{e}_i) \; , \quad \text{and} \quad \mu_i = \ell + 1 - \tau_i \; ; \tag{7}$$

note that the support of $\boldsymbol{e}_i$ is contained in the set

$$J_i = [\lambda_i, \lambda_i + \tau_i\rangle \; . \tag{8}$$

We need to show that $\sum_{i=0}^{m-1} \mu_i \leq \ell$.

Suppose to the contrary that $\sum_{i=0}^{m-1} \mu_i \geq \ell + 1$. Without real loss of generality, we can assume hereafter in the proof that the latter inequality holds with equality; otherwise, we can increase some of the τ_i's (effectively replacing some of the actual burst lengths by upper bounds on these lengths) and decrease the respective values of μ_i $(= \ell + 1 - \tau_i)$ accordingly, to achieve the equality

$$\sum_{i=0}^{m-1} \mu_i = \ell + 1 \; . \tag{9}$$

For $i \in [0, m\rangle$, let

$$M_i(x) = M_i(x; \alpha^{\lambda_i}, \alpha) = \prod_{t=0}^{\tau_i - 1} (x - \alpha^{\lambda_i + t}) \tag{10}$$

be the specialization of (4) to $\gamma = \alpha$, $\beta_i = \alpha^{\lambda_i}$, $r = \ell + 1$, and μ_i as in (7). The next steps in our proof are very similar to the proof of Theorem 1.2 in [7], with r and ℓ therein replaced by $\ell + 1$ and $m - 1$, respectively. We include these steps in Appendix A, for the sake of readability and completeness. It follows from Appendix A that there exist polynomials

$$u_i(x) \in F_{\mu_i}[x] \; , \quad i \in [0, m\rangle \; , \tag{11}$$

not all zero, such that

$$\sum_{i=0}^{m-1} u_i(x) M_i(x) = 0 \; . \tag{12}$$

Combining (9)–(12) with Theorem 2.1, we conclude that there exist distinct $h, k \in [0, m\rangle$ and some integer b in the range $-\mu_h < b < \mu_k$ such that

$$\alpha^{\lambda_k - \lambda_h} = \alpha^b \; ;$$

namely,

$$\lambda_k - \lambda_h \equiv b \pmod{n} \; , \tag{13}$$

where we have used our assumption that $n = \mathrm{ord}(\alpha)$. Without loss of generality we can assume further that $0 \leq b \; (< \mu_k)$, or else simply switch between the roles of h and k. Also, since $\mathcal{C}_{\rm RS}$ is cyclic when $n = \mathrm{ord}(\alpha)$, we can rotate the $\boldsymbol{e}_i$'s, all by the same number ρ of positions, and the resulting new words will all belong to the same coset of $\mathcal{C}_{\rm RS}$. We should then add, modulo n, the integer ρ to each λ_i (and, respectively, to each index in each set J_i in (8)), to obtain the correct index values for the rotated words. Thus, (13) still holds (for the same b) regardless of the value of ρ; in particular, we can select ρ so that $(0 \leq) \; \lambda_h \leq \lambda_k \; (< n)$.

Doing so, we get that

$$|J_h \setminus J_k| \leq \lambda_k - \lambda_h = b \; . \tag{14}$$

On the other hand, since $\boldsymbol{e}_h$ and $\boldsymbol{e}_k$ are distinct yet belong to the same coset of $\mathcal{C}_{\rm RS}$, the difference $\boldsymbol{e}_h - \boldsymbol{e}_k$ is a nonzero codeword of $\mathcal{C}_{\rm RS}$ and, as such, its Hamming weight is at least $\ell + 2$. Hence,

$$|J_h \cup J_k| > \ell + 1 \tag{15}$$

and, so,

$$\begin{aligned} \mu_k &= \ell + 1 - \tau_k \\ &\overset{(15)}{<} |J_h \cup J_k| - |J_k| = |J_h \setminus J_k| \\ &\overset{(14)}{\leq} \lambda_k - \lambda_h = b \; , \end{aligned}$$

thereby contradicting the fact that $b < \mu_k$. ∎

Remark 2.3: We have excluded from Lemma 2.3 the case where $\boldsymbol{y} - \mathcal{C}$ is the trivial coset $\mathcal{C}$, since the inequality (6) does not hold when $m = 1$ and $\boldsymbol{e}_0 = \boldsymbol{0}$. ∎

It can be shown (by counterexamples) that Theorem 2.2(ii) and Lemma 2.3 would no longer hold if we attempted to generalize them to arbitrary linear codes that attain the bound of Theorem 1.1 for $\tau = \ell$, not even when the codes are maximum distance separable (MDS). See Example 5.1 in Section V.

Furthermore, it is rather easy to see that Theorem 2.2(ii) and Lemma 2.3 (or, rather, the discussion at the beginning of this section) would no longer hold if we attempted to generalize them to $\tau \neq \ell$. Specifically, suppose that ℓ is a proper divisor of τ and let $\mathcal{D}$ be the following list decoder for $\mathcal{C}_{\rm RS}(n, r)$, where $r = \tau + (\tau/\ell)$:

$$\mathcal{D}(\boldsymbol{y}) = \big(\boldsymbol{y} - \mathcal{C}_{\rm RS}(n, r)\big) \cap \mathcal{B}(n, \tau) \; .$$

This decoder is a minimal (ℓ, τ)-burst list decoder in the sense that it satisfies the containment (2) with equality. Yet, there may be $\boldsymbol{y} \in F^n$ for which the list $\mathcal{D}(\boldsymbol{y})$ has size (exactly) ℓ, even though it contains a burst of actual length smaller than τ. In fact, there will *always* be such a $\boldsymbol{y}$ when $\ell = 1$ and $\tau > 1$: for any $\boldsymbol{e} \in \mathcal{B}(n, \tau - 1)$ we have $\mathcal{D}(\boldsymbol{e}) = \{\boldsymbol{e}\}$, i.e., $|\mathcal{D}(\boldsymbol{e})| = \ell = 1$.

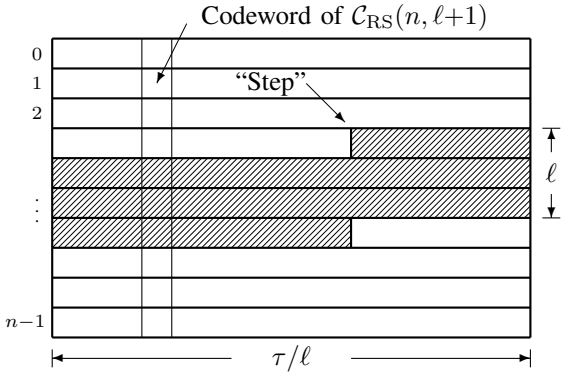

Fig. 1. Interleaved array, with a τ-burst error marked by the shaded area.

III. Burst List Decoding of interleaved RS Codes

In this section, we show that when ℓ divides τ, a (τ/ℓ)-level interleaving of $\mathcal{C}_{\rm RS}(n, \ell{+}1)$ yields a code $\mathbb{C}$ that has an (ℓ, τ)-burst list decoder. Specifically, the code $\mathbb{C}$ is defined by

$$\mathbb{C} = \mathbb{C}(n, \ell, \tau) = \Big\{ (\boldsymbol{c}_0 | \boldsymbol{c}_1 | \ldots | \boldsymbol{c}_{\tau/\ell - 1}) \ : \ \boldsymbol{c}_j \in \mathcal{C}_{\rm RS}(n, \ell{+}1) \text{ for } j \in [0, \tau/\ell\rangle \Big\} \ ;$$

namely, it consists of all $n \times (\tau/\ell)$ arrays over F whose columns are codewords $\boldsymbol{c}_j$ of $\mathcal{C}_{\rm RS} = \mathcal{C}_{\rm RS}(n, \ell{+}1)$. When transmitted over a noisy channel, the array is sent row by row, in which case a τ-burst error is seen as ℓ-burst errors in the columns of the array, as shown in Figure 1. The shaded area in the figure represents a largest possible set of entries that can be affected by a single τ-burst: generally, entries in that area (including leading or trailing entries) can still be error-free, in which case some columns in the array may incur bursts of length less than ℓ. While this observation is straightforward, it is those lightly corrupted—yet still corrupted—columns that could potentially fail the decoding of the array by requiring the list size to be greater than ℓ. Indeed, this could happen if $\mathcal{C}_{\rm RS}$ were replaced by an arbitrary constituent code $\mathcal{C}$, even when $\mathcal{C}$ is (ℓ, ℓ)-burst list decodable and $\mathsf{d}(\mathcal{C}) \geq \ell + 2$: see Example 5.1 in Section V and Example B.1 in Appendix B. However, as we show, Theorem 2.2 will guarantee decoding success for $\mathbb{C}$ when $\mathcal{C} = \mathcal{C}_{\rm RS}$.

The following theorem is the main result of this paper.

Theorem 3.1: Given positive integers $\ell \leq \tau < n < q$ such that $\ell \,|\, \tau$, let $\mathbb{C}$ be the (τ/ℓ)-level interleaving of $\mathcal{C}_{\rm RS} = \mathcal{C}_{\rm RS}(n, \ell{+}1)$. Then $\mathbb{C}$ has an (ℓ, τ)-burst list decoder. In particular, $\mathbb{C}$ attains the bound of Theorem 1.1 when $2\ell \leq n$.

We will prove Theorem 3.1 by introducing an (ℓ, τ)-burst list decoding algorithm for $\mathbb{C}$. The algorithm is presented in Figure 2, and the remaining part of this section is devoted to analyzing that algorithm.

The input to the algorithm is an $n \times (\tau/\ell)$ array Y over F, which is assumed to be a copy of a code array of $\mathbb{C}$, possibly corrupted by a τ-burst error. The algorithm consists of three main loops: (A), (B), and (C). Loop (A) iterates over the columns of Y, as long as the columns are codewords of $\mathcal{C}_{\rm RS}$. If Y is found to be error-free, then Loops (B) and (C) are skipped, and the algorithm returns a list of size 1, containing the all-zero $n \times (\tau/\ell)$ array.

Input: received array $Y = (\boldsymbol{y}_0 | \boldsymbol{y}_1 | \ldots | \boldsymbol{y}_{\tau/\ell - 1})$ over F.
Data structures:
  Integer intervals $V \subseteq [0, n\rangle$;
  Sets $\mathcal{S}$, $\mathcal{S}'$ of intervals $V \subseteq [0, n\rangle$;
  Arrays E_V over F indexed by intervals V.
{
  // Locate the first erroneous column (if any): (A)
  for $\big(j \leftarrow 0;\ j < \tau/\ell;\ j{+}{+}\big)$ {
    if $\big(\boldsymbol{y}_j \notin \mathcal{C}_{\rm RS}\big)$
      break;
  }
  if $\big(j \geq \tau/\ell\big)$
    Output $\{0_{n \times (\tau/\ell)}\}$;
  else {
    $\mathcal{S} \leftarrow \emptyset$;
    // Decode the first erroneous column and generate (B)
    // the (initial contents of the) interval set $\mathcal{S}$:
    for every $\boldsymbol{e} \in \mathcal{D}(\boldsymbol{y}_j)$
      for $\big(b = \max(\lambda(\boldsymbol{e}){+}L(\boldsymbol{e}){-}\ell, 0){-}1;$
                $b < \min(\lambda(\boldsymbol{e}), n{-}\ell);\ b{+}{+}\big)$ {
        $V \leftarrow [\max(b, 0), b{+}\ell{+}1\rangle$;
        $\mathcal{S} \leftarrow \mathcal{S} \cup \{V\}$;
        $E_V \leftarrow (0_{n \times j} | \boldsymbol{e})$;
      }
    // Decode the remaining columns using the intervals (C)
    // in $\mathcal{S}$ as erasure locators:
    for $\big(k \leftarrow j + 1;\ k < \tau/\ell;\ k{+}{+}\big)$ {
      $\mathcal{S}' \leftarrow \emptyset$;
      for every $V \in \mathcal{S}$
        if $\big(\mathcal{D}_V(\boldsymbol{y}_k)$ contains a word $\boldsymbol{e}$ and $L(\boldsymbol{e}) \leq \ell\big)$ {
          if $\big(\boldsymbol{e} \neq \boldsymbol{0}$ and $V = [\lambda(\boldsymbol{e}), \lambda(\boldsymbol{e}){+}\ell{+}1\rangle\big)$
            $V' \leftarrow V \setminus \{\lambda(\boldsymbol{e}){+}\ell\}$;
          else
            $V' \leftarrow V$;
          // V' qualifies to survive in $\mathcal{S}$:
          $\mathcal{S}' \leftarrow \mathcal{S}' \cup \{V'\}$;
          $E_{V'} \leftarrow (E_V | \boldsymbol{e})$;
        }
      $\mathcal{S} \leftarrow \mathcal{S}'$;
    }
    Output $\{E_V \ : \ V \in \mathcal{S}\}$;
  }
}

Fig. 2. Decoding algorithm for the interleaved code $\mathbb{C}$.

When Y is flagged with errors, Loop (B) is entered with the index j pointing at the first column, $\boldsymbol{y}_j$, in Y that is corrupted. The (ℓ, ℓ)-burst list decoder $\mathcal{D}$ of Theorem 2.2 is applied to that column, resulting in the list

$$\mathcal{D}(\boldsymbol{y}_j) = \big(\boldsymbol{y}_j - \mathcal{C}_{\rm RS}\big) \cap \mathcal{B}(n, \ell) = \{\boldsymbol{e}_0, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_{m-1}\} \ .$$

Writing $\lambda_i = \lambda(\boldsymbol{e}_i)$ and $\tau_i = L(\boldsymbol{e}_i)$, for each $i \in [0, m\rangle$, the following collection of up to $\ell + 1 - \tau_i$ integer intervals $V_{i,b}$ is added to the interval set $\mathcal{S}$ that is computed in Loop (B):

$$\mathcal{S}_i = \Big\{ V_{i,b} = [\max(b, 0), b{+}\ell{+}1\rangle \ : \ \max(\lambda_i{+}\tau_i{-}\ell, 0) - 1 \leq b < \min(\lambda_i, n{-}\ell) \Big\} \ . \quad (16)$$

Next, we state two lemmas regarding the interval set $\mathcal{S}$.

Lemma 3.2: Suppose that E is an $n \times (\tau/\ell)$ array over F that forms a τ-burst when unfolded (as in Figure 1). Assume

in addition that E is in the coset $Y-\mathbb{C}$ and that column j in E equals one of the words in $\mathcal{D}(\boldsymbol{y}_j)$. Then the nonzero entries in E are all confined to rows that are indexed by one of the intervals $V \in \mathcal{S}$.

Proof: Suppose that column j in E equals $\boldsymbol{e}_i \in \mathcal{D}(\boldsymbol{y}_j)$. We show that the nonzero entries in E are confined to rows that are indexed by one of the intervals $V_{i,b} \in \mathcal{S}_i$ in (16).

The projection of the τ-burst to column j forms an ℓ-burst that is indexed by an interval $[b', b'+\ell\rangle$, where

$$\max(\lambda_i+\tau_i-\ell, 0) \le b' \le \min(\lambda_i, n-\ell)$$

(the $\max(\cdot, 0)$ and $\min(\cdot, n-\ell)$ just truncate the lower and upper limits so that the interval $[b', b'+\ell\rangle$ does not extend beyond the array boundaries). Yet, we need to take into account that row $b'-1$ might also be part of the τ-burst at columns $k > j$, due to the "step" shown in Figure 1 (which may occur if the τ-burst does not start right at the beginning of a row). Letting $b = b'-1$, the result follows. ∎

Lemma 3.3:

$$|\mathcal{S}| \le \ell \ .$$

Proof: Using the notation (16), we have

$$|\mathcal{S}| \le \sum_{i=0}^{m-1} |\mathcal{S}_i| \le \sum_{i=0}^{m-1} (\ell+1-\tau_i) \le \ell \ ,$$

where the last inequality follows from Theorem 2.2(ii). ∎

Turning to Loop (C), this loop iterates over the remaining columns in the array, and to each column, an *erasure decoder* $\boldsymbol{y} \mapsto \mathcal{D}_J(\boldsymbol{y})$ for $\mathcal{C}_{\rm RS}$ is applied: given any subset $J \subseteq [0, n\rangle$ and word $\boldsymbol{y} \in F^n$, the set $\mathcal{D}_J(\boldsymbol{y})$ consists of all words in the coset $\boldsymbol{y}-\mathcal{C}_{\rm RS}$ whose support is contained in J. The subsets J are taken as the intervals $V \in \mathcal{S}$, and Lemma 3.2 guarantees that the arrays E_V that are formed in Loop (C) range over all the τ-bursts in the coset $Y-\mathbb{C}$. Furthermore, since $|V| \le \ell+1$ for every $V \in \mathcal{S}$, we get from the distance properties of $\mathcal{C}_{\rm RS}$ that $|\mathcal{D}_V(\boldsymbol{y}_k)| \le 1$ for every k and V; i.e., there is at most one possible column $\boldsymbol{e} \in \mathcal{D}_V(\boldsymbol{y}_k)$ that can be appended to each E_V while still forming (the first $k+1$ columns of) a τ-burst. Hence, by Lemma 3.3, the output list, $\{E_V \ : \ V \in \mathcal{S}\}$, has size at most ℓ. This completes the proof that the algorithm in Figure 1 is an (ℓ, τ)-burst list decoder for $\mathbb{C}$; namely, $\mathbb{C}$ satisfies condition 3 in Theorem 1.1.

It is easy to see that $\mathbb{C}$ also satisfies condition 2; in fact, Loop (A) can serve as a single $(\tau+(\tau/\ell))$-burst error detector for $\mathbb{C}$, where an early "break" from the loop means that at least one of the columns in Y has been subject to an $(\ell+1)$-burst error. Our requirement that $2\ell \le n$ implies that $\mathbb{C}$ satisfies condition 1 in Theorem 1.1, and, since the redundancy of $\mathbb{C}$ is $\tau+(\tau/\ell)$ (see (1)), this code attains the bound of that theorem. This completes the proof of Theorem 3.1.

Remark 3.1: The two "if"s in Loop (C) guarantee that the returned list in Figure 2 contains *only* τ-bursts. Specifically, the outer "if" will disqualify an interval V if the (unique) burst error that is returned by the erasure decoder (at any column) for that V has length $\ell+1$. And the inner "if" guarantees (through shortening of the interval V) that once we incur the "step" in Figure 1, there will be no "step back" in subsequent columns. The surviving (possibly shortened) intervals are recorded into a temporary interval set $\mathcal{S}'$ which, in turn, is copied into $\mathcal{S}$ at the end of each iteration of Loop (C). ∎

IV. Decoding complexity

In this section, we present some implementation details regarding the decoding algorithm in Figure 2 and compute the time complexity of that algorithm. We then compare this complexity with that of an (ℓ, τ)-burst list decoder for a non-interleaved RS code with the same code length.

Loop (A) can be carried out by computing the syndrome, with respect to the parity-check matrix $H_{\rm RS} = H_{\rm RS}(n, \ell+1)$ in (3), of each column in the array Y (these syndromes will be used also for columns j through $\tau/\ell-1$ in Loops (B) and (C)). This computation requires less than $2(\tau/\ell)\cdot(\ell+1)n = O(\tau n)$ arithmetic operations in F.

Loop (B) applies an (ℓ, ℓ)-burst list decoder for $\mathcal{C}_{\rm RS}(n, \ell+1)$ to column j in Y. Such a decoder, in turn, can be implemented by applying iteratively an erasure decoder $\mathcal{D}_J(\cdot)$ to that column, where J ranges over the intervals

$$[b, b+\ell\rangle \ , \quad b \in [0, n-\ell\rangle \ .$$

Given $\boldsymbol{y}_j \in F^n$, let $S(x)$ be the syndrome polynomial in $F_{\ell+1}[x]$ that is associated with $\boldsymbol{y}_j$, namely, the coefficients of $S(x)$ are given by the syndrome $H_{\rm RS}\boldsymbol{y}_j$. Also, for $b \in [0, n-\ell\rangle$, let $\Lambda_b(x)$ be the erasure-locator polynomial

$$\Lambda_b(x) = \prod_{t=0}^{\ell-1} (1-\alpha^{b+t}x) \ .$$

It follows from the known properties of RS decoding (see, for example, [10, Problem 6.11]) that $\mathcal{D}_{[b,b+\ell\rangle}(\boldsymbol{y}_j) \neq \emptyset$ if and only if the coefficient of x^ℓ in the erasure evaluator polynomial,

$$\Gamma_b(x) = \Lambda_b(x)S(x) \bmod x^{\ell+1} \ , \tag{17}$$

is zero (where MOD denotes remaindering). And when that happens, the erasure values can be found by Forney's algorithm, namely, by evaluating the ratio $-\Gamma_b(x)/(x\Lambda_b'(x))$ at $x = \alpha^{-b}, \alpha^{-b-1}, \ldots, \alpha^{-b-\ell+1}$ [10, Section 6.5]. From (17) we see that the polynomials $\Gamma_b(x)$ satisfy the relationship

$$\Gamma_{b+1}(x) = \Gamma_b(x)(1-\alpha^{b+\ell}x)/(1-\alpha^b x) \bmod x^{\ell+1} \ ;$$

as such, they can be computed for all $b \in [0, n-\ell\rangle$ using a total of $O(\ell n)$ arithmetic operations in F. Recalling that $|(\boldsymbol{y}_j-\mathcal{C}_{\rm RS}) \cap \mathcal{B}(n, \ell)| \le \ell$, we will need to apply Forney's algorithm at most ℓ times; therefore, the overall time complexity of Loop (B) is $O(\ell n + \ell^3)$ operations in F.

Loop (C) amounts to applying an RS erasure decoder $\mathcal{D}_V(\cdot)$, for each $V \in \mathcal{S}$, to each of the remaining columns of Y, where the complexity of each application of $\mathcal{D}_V(\cdot)$ is $O\big(|V|(\ell+1)\big) = O(\ell^2)$; namely, Loop (C) can be carried out using $O\big((\tau/\ell)|\mathcal{S}|\ell^2\big) = O(\tau\ell^2)$ operations in F. We conclude that the overall time complexity of the algorithm in Figure 2 is $O(\tau n + \tau\ell^2) = O\big(\ell(N+\ell\tau)\big)$ operations in F, where $N = (\tau/\ell)n$ stands for the effective length of $\mathbb{C}$ as a linear code over F.

Our strategy for implementing an (ℓ, ℓ)-burst list decoder for $\mathcal{C}_{\rm RS}(n, \ell+1)$ (in the analysis of Loop (B)) can be applied more generally to obtain an (ℓ, τ)-burst list decoder for the

(non-interleaved) code $\mathcal{C}_{\mathrm{RS}}(N, \tau + (\tau/\ell))$. The resulting time complexity turns out to be $O\big(\tau(N + \ell\tau)\big)$ operations in F (including the complexity of computing the syndrome), which is τ/ℓ times larger than the decoding complexity of an interleaved RS code of the same length.

V. Counterexample

In Example 5.1 below, we present a construction of a linear MDS code that attains the bound of Theorem 1.1 for $\tau = \ell = 3$, yet violates Theorem 3.1 and—*a fortiori*—also violates Theorem 2.2(ii) and Lemma 2.3. The purpose of this example is to demonstrate that we do need to make essential use of the particular structure of RS codes in order to obtain those results.

We will make use of the following lemma.

Lemma 5.1: Let $\mathcal{C}$ be a linear code of length n over F, and assume that $\mathcal{C}$ has a decoder which detects any θ-burst, for some positive integer θ (in particular, θ can be taken as $\mathsf{d}(\mathcal{C}) - 1$). Then $\mathcal{C}$ has an (ℓ, τ)-burst list decoder, whenever $\tau \le \theta$ and

$$\ell \ge \left\lceil \frac{n - \tau + 1}{\theta - \tau + 1} \right\rceil \ . \tag{18}$$

Proof: Let $\{\boldsymbol{e}_0, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_{m-1}\}$ be the intersection $(\boldsymbol{y} - \mathcal{C}) \cap \mathcal{B}(n, \tau)$ for some $\boldsymbol{y} \in F^n$. We will show that m is bounded from above by the right-hand side of (18). The result obviously holds when $m \le 1$ (and, in particular, when $\boldsymbol{y} \in \mathcal{C}$), so we can assume hereafter in the proof that $m \ge 2$. Without loss of generality we further assume that $\lambda(\boldsymbol{e}_{i-1}) \le \lambda(\boldsymbol{e}_i)$ for all $i \in [1, m\rangle$.

Since $\boldsymbol{e}_i - \boldsymbol{e}_{i-1} \in \mathcal{C} \setminus \{\boldsymbol{0}\}$ yet $\mathcal{C} \cap \mathcal{B}(n, \theta) = \{\boldsymbol{0}\}$, it follows that $\boldsymbol{e}_i - \boldsymbol{e}_{i-1} \notin \mathcal{B}(n, \theta)$; that is, for every $i \in [1, m\rangle$,

$$\lambda(\boldsymbol{e}_i) + L(\boldsymbol{e}_i) - \lambda(\boldsymbol{e}_{i-1}) \ge L(\boldsymbol{e}_i - \boldsymbol{e}_{i-1}) \ge \theta + 1 \ ,$$

or

$$\lambda(\boldsymbol{e}_i) - \lambda(\boldsymbol{e}_{i-1}) \ge \theta + 1 - L(\boldsymbol{e}_i) \ . \tag{19}$$

Therefore,

$$\begin{aligned} n &\ge \lambda(\boldsymbol{e}_{m-1}) + L(\boldsymbol{e}_{m-1}) - \lambda(\boldsymbol{e}_0) \\ &= L(\boldsymbol{e}_{m-1}) + \sum_{i=1}^{m-1} \big(\lambda(\boldsymbol{e}_i) - \lambda(\boldsymbol{e}_{i-1})\big) \\ &\ge L(\boldsymbol{e}_{m-1}) + \sum_{i=1}^{m-1} \big(\theta + 1 - L(\boldsymbol{e}_i)\big) \ , \end{aligned}$$

where the last step follows from (19). We thus obtain

$$\begin{aligned} n &\ge \theta + 1 + \sum_{i=1}^{m-2} \big(\theta + 1 - L(\boldsymbol{e}_i)\big) \\ &\ge \theta + 1 + (m-2)(\theta + 1 - \tau) \\ &= \tau + (m-1)(\theta - \tau + 1) \ , \end{aligned}$$

where the penultimate step is justified by the inequality $L(\boldsymbol{e}_i) \le \tau$. Hence,

$$m \le \left\lfloor \frac{n - \tau}{\theta - \tau + 1} \right\rfloor + 1 = \left\lceil \frac{n - \tau + 1}{\theta - \tau + 1} \right\rceil \ .$$

■

Example 5.1: Let $F = \mathrm{GF}(q)$ where $q > 7$ and let γ be an element in F^* such that $\mathrm{ord}(\gamma) \ge 7$. Consider the linear code $\mathcal{C}$ of length $n = 8$ over F with a generator matrix

$$G = \left(\begin{array}{ccc|cc|ccc} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ \alpha_1 & \alpha_4 & \alpha_5 & 0 & 1 & \alpha_2 & \alpha_3 & \alpha_6 \\ \alpha_1^2 & \alpha_4^2 & \alpha_5^2 & 0 & 1 & \alpha_2^2 & \alpha_3^2 & \alpha_6^2 \\ \alpha_1^3 & \alpha_4^3 & \alpha_5^3 & 0 & 1 & \alpha_2^3 & \alpha_3^3 & \alpha_6^3 \end{array} \right) \ ,$$

where

$$\alpha_t = \frac{1}{1 - \gamma^t} \ , \quad 1 \le t \le 6 \ .$$

One can verify that the entries in the second row are all distinct elements of F. Hence, $\mathcal{C}$ is MDS and $\mathsf{d}(\mathcal{C}) = 5$. It follows from Lemma 5.1 that $\mathcal{C}$ has a $(3, 3)$-burst list decoder and, as such, it attains the bound of Theorem 1.1 for $\tau = \ell = 3$.

Next, we exhibit a coset of $\mathcal{C}$ which contains three nonzero 3-bursts, yet one of them is actually a 2-burst (and, therefore, these words violate the inequality (6) in Lemma 2.3). Let $u(x)$ and $v(x)$ be the following polynomials in $F_4[x]$:

$$u(x) = (1 - \alpha_2^{-1}x)(1 - \alpha_3^{-1}x)(1 - \alpha_6^{-1}x)$$

and

$$v(x) = (1 - \alpha_1^{-1}x)(1 - \alpha_4^{-1}x)(1 - \alpha_6^{-1}x) \ .$$

It is easy to verify that

$$u(0) = v(0) = 1 \quad \text{and} \quad u(1) = v(1) = \gamma^{11} \ .$$

Now, consider the following three (column) words—$\boldsymbol{e}_0$, $\boldsymbol{e}_1$, and $\boldsymbol{e}_2$—in $\mathcal{B}(8, 3)$:

$$\boldsymbol{e}_0 = \begin{pmatrix} u(\alpha_1) \\ u(\alpha_4) \\ u(\alpha_5) - v(\alpha_5) \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{e}_1 = \begin{pmatrix} 0 \\ 0 \\ -v(\alpha_5) \\ -1 \\ -\gamma^{11} \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ v(\alpha_2) \\ v(\alpha_3) \\ 0 \end{pmatrix} .$$

It can be verified that the difference $\boldsymbol{e}_0 - \boldsymbol{e}_1$ equals the following vector, whose entries are the values of $u(x)$ at the elements along the second row of G:

$$\big(u(\alpha_1)\ u(\alpha_4)\ u(\alpha_5)\ u(0)\ u(1)\ u(\alpha_2)\ u(\alpha_3)\ u(\alpha_6)\big)^T \ .$$

Similarly, $\boldsymbol{e}_2 - \boldsymbol{e}_1$ equals the vector whose entries are the values of $v(x)$ at those elements. It follows that $\boldsymbol{e}_0 - \boldsymbol{e}_1$ and $\boldsymbol{e}_2 - \boldsymbol{e}_1$ are codewords of $\mathcal{C}$, which means that $\boldsymbol{e}_0$, $\boldsymbol{e}_1$, and $\boldsymbol{e}_2$ all belong to the same nontrivial coset $\boldsymbol{e}_0 - \mathcal{C}$. Yet, since $L(\boldsymbol{e}_2) < 3$, these three words violate the inequality (6).

A second coset that violates that inequality can be formed by using the polynomials

$$\hat{u}(x) = (1 - \alpha_1^{-1}x)(1 - \alpha_3^{-1}x)(1 - \alpha_6^{-1}x)$$

and

$$\hat{v}(x) = (1 - \alpha_1^{-1}x)(1 - \alpha_4^{-1}x)(1 - \alpha_5^{-1}x) \ ,$$

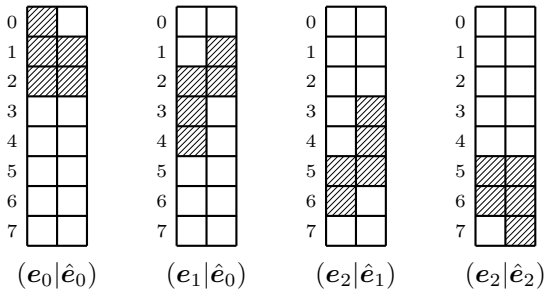

Fig. 3. Four 6-bursts contained in the same coset of $\mathbb{C}$ in Example 5.1.

and defining, respectively,

$$\hat{\boldsymbol{e}}_0=\begin{pmatrix}0\\ \hat{u}(\alpha_4)\\ \hat{u}(\alpha_5)\\ 0\\ 0\\ 0\\ 0\\ 0\end{pmatrix},\quad \hat{\boldsymbol{e}}_1=\begin{pmatrix}0\\ 0\\ 0\\ -1\\ -\gamma^{10}\\ -\hat{u}(\alpha_2)\\ 0\\ 0\end{pmatrix},\quad \hat{\boldsymbol{e}}_2=\begin{pmatrix}0\\ 0\\ 0\\ 0\\ 0\\ \hat{v}(\alpha_2)-\hat{u}(\alpha_2)\\ \hat{v}(\alpha_3)\\ \hat{v}(\alpha_6)\end{pmatrix}.$$

Since $\hat{u}(0)=\hat{v}(0)=1$ and $\hat{u}(1)=\hat{v}(1)=\gamma^{10}$, we again obtain that $\hat{\boldsymbol{e}}_0-\hat{\boldsymbol{e}}_1$ and $\hat{\boldsymbol{e}}_2-\hat{\boldsymbol{e}}_1$ are codewords of $\mathcal{C}$, namely, $\hat{\boldsymbol{e}}_0$, $\hat{\boldsymbol{e}}_1$, and $\hat{\boldsymbol{e}}_2$ are 3-bursts that belong to the same coset $\hat{\boldsymbol{e}}_0-\mathcal{C}$, while $L(\hat{\boldsymbol{e}}_0)<3$.

Finally, we demonstrate that Theorem 3.1 becomes false if we attempt to state it with $\mathcal{C}$ replacing $\mathcal{C}_{\rm RS}(n,\ell{+}1)$ therein. Specifically, we show that the 2-level interleaving $\mathbb{C}$ of $\mathcal{C}$ does not have a $(3,6)$-burst list decoder, by exhibiting an 8×2 array E over F for which the coset $E-\mathbb{C}$ contains more than three 6-bursts. Indeed, taking $E=(\boldsymbol{e}_0|\hat{\boldsymbol{e}}_0)$, the coset $E-\mathbb{C}$ contains the four 8×2 arrays

$$(\boldsymbol{e}_0|\hat{\boldsymbol{e}}_0)\,,\quad (\boldsymbol{e}_1|\hat{\boldsymbol{e}}_0)\,,\quad (\boldsymbol{e}_2|\hat{\boldsymbol{e}}_1)\,,\quad \text{and}\quad (\boldsymbol{e}_2|\hat{\boldsymbol{e}}_2)\,,$$

all of which are 6-bursts (see Figure 3: the bursts in the second and third arrays incur a "step" up when moving from the first column to the second). ∎

In Appendix B, we present another example of a (nonlinear and non-MDS) code $\mathcal{C}$ with $\mathsf{d}(\mathcal{C})=\ell{+}2$ that has an (ℓ,ℓ)-burst list decoder, yet the (τ/ℓ)-level interleaving of $\mathcal{C}$ is (ℓ',τ)-burst list decodable only for ℓ' that grows quadratically with ℓ.

Acknowledgment

The authors wish to thank Pascal Vontobel for his helpful comments.

References

[1] P. Elias, "Error-correcting codes for list decoding," *IEEE Trans. Inf. Theory*, 37 (1991), 5–12.
[2] M. Sudan, "Decoding of Reed–Solomon codes beyond the error-correction bound," *J. Complexity,* 13 (1997), 180–193.
[3] V. Guruswami, M. Sudan, "Improved decoding of Reed–Solomon and algebraic–geometry codes," *IEEE Trans. Inf. Theory,* 45 (1999), 1757–1767.
[4] R. Koetter, A. Vardy, "Algebraic soft-decision decoding of Reed–Solomon codes," *IEEE Trans. Inf. Theory,* 49 (2003), 2809–2825.
[5] F. Parvaresh, A. Vardy, "Correcting errors beyond the Guruswami–Sudan radius in polynomial time," *Proc. 46th Annual IEEE Symp. Foundations of Computer Science (FOCS 2005),* Pittsburgh, PA (2005), 285–294.
[6] V. Guruswami, A. Rudra, "Explicit codes achieving list decoding capacity: error-correcting with optimal redundancy," *IEEE Trans. Inf. Theory,* 54 (2008), 135–150.
[7] R.M. Roth, P.O. Vontobel, "List decoding of burst errors," *IEEE Trans. Inf. Theory,* 55 (2009), 4179–4190.
[8] S. Lin, D.J. Costello, Jr., *Error Control Coding: Fundamentals and Applications,* Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1983.
[9] W.W. Peterson, and E.J. Weldon, Jr., *Error-Correcting Codes*, 2nd ed., MIT Press, Cambridge, Massachusetts, 1972.
[10] R.M. Roth, *Introduction to Coding Theory,* Cambridge University Press, Cambridge, UK, 2006.

Appendix A
Part of proof of Lemma 2.3

We include here the analysis which establishes (11)–(12).

For $i\in[0,m\rangle$, denote by H_i the $(\ell{+}1)\times\tau_i$ sub-matrix of $H_{\rm RS}$ in (3) which is formed by the columns of $H_{\rm RS}$ that are indexed by the set J_i defined in (8), namely:

$$H_i=\left(\alpha^{(\lambda_i+t)s}\right)_{s=0,\,t=0}^{\ell\quad \tau_i-1}.$$

Define also the $(\ell{+}1)\times(\ell{+}1)$ matrix T_i by

$$T_i=\left(\begin{array}{c|c} I_i & 0\\ \hline A_i & \end{array}\right),$$

where I_i is a $\tau_i\times\tau_i$ identity matrix and A_i is the $\mu_i\times(\ell{+}1)$ echelon matrix

$$A_i=\begin{pmatrix} M_{i,0} & M_{i,1} & \dots & M_{i,\tau_i} & & \\ & M_{i,0} & M_{i,1} & \dots & M_{i,\tau_i} & \\ \mathbf{0} & & \ddots & \ddots & \dots & \ddots \\ & & & M_{i,0} & M_{i,1} & \dots & M_{i,\tau_i}\end{pmatrix},$$

with $M_{i,j}$ being the coefficients of $M_i(x)$ in (10), namely,

$$M_i(x)=\sum_{j=0}^{\tau_i}M_{i,j}x^j\,.$$

Notice that $A_iH_i=0$ and, so, the product T_iH_i results in an $(\ell{+}1)\times\tau_i$ matrix W_i of the form:

$$W_i=T_iH_i=\begin{pmatrix}\left(\alpha^{(\lambda_i+t)s}\right)_{s,t=0}^{\tau_i-1}\\ \hline 0\end{pmatrix}. \tag{20}$$

Specifically, the first τ_i rows form a non-singular square Vandermonde matrix, whereas the remaining μ_i rows are all zero.

Consider the following $(m{-}1)(\ell{+}1)\times(m{-}1)(\ell{+}1)$ matrix B:

$$B=\begin{pmatrix} H_0 & -H_1 & & & \mathbf{0}\\ H_0 & & -H_2 & & \\ \vdots & & & \ddots & \\ H_0 & \mathbf{0} & & & -H_{m-1}\end{pmatrix}$$

(it is indeed easy to see from (7) and (9) that the sum $\sum_{i=0}^{m-1}\tau_i$ equals $(m{-}1)(\ell{+}1)$). Next, we multiply B to the left by an

$(m-1)(\ell+1) \times (m-1)(\ell+1)$ block-diagonal matrix T which contains the blocks $T_1, T_2, \ldots, T_{m-1}$ along its main diagonal:

$$TB = \begin{pmatrix} \boxed{Z_1} & \boxed{-W_1} & & \mathbf{0} \\ \boxed{Z_2} & & \boxed{-W_2} & \\ \vdots & & & \ddots \\ \boxed{Z_{m-1}} & \mathbf{0} & & \boxed{-W_{m-1}} \end{pmatrix} ,$$

where W_i is as in (20) and

$$Z_i = T_i H_0 = \begin{pmatrix} \big(\alpha^{(\lambda_0+t)s}\big)_{s,t=0}^{\tau_0-1} \\ \hline A_i H_0 \end{pmatrix} .$$

Recalling that $\boldsymbol{e}_0, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_{m-1}$ belong to the same coset of $\mathcal{C}_{\rm RS}$, we get that $H_{\rm RS}\boldsymbol{e}_0 = H_{\rm RS}\boldsymbol{e}_i$ for every $i \in [1, m\rangle$. This implies that B has dependent columns and is therefore singular; hence, so is the $\tau_0 \times \tau_0$ matrix

$$\begin{pmatrix} A_1 H_0 \\ \hline A_2 H_0 \\ \hline \vdots \\ \hline A_{m-1} H_0 \end{pmatrix} ,$$

which is formed by taking the last μ_i rows of each Z_i and stacking them together for all $i \in [1, m\rangle$ (from (9) we get that $\sum_{i=1}^{m-1} \mu_i = \ell + 1 - \mu_0 = \tau_0$). Hence, there exist row vectors $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{m-1}$, not all zero, such that $\boldsymbol{u}_i \in F^{\mu_i}$ and

$$\sum_{i=1}^{m-1} \boldsymbol{u}_i A_i H_0 = \mathbf{0} .$$

Equivalently, due to the structure of A_i and H_0, there exist polynomials

$$u_i(x) \in F_{\mu_i}[x] , \quad i \in [1, m\rangle ,$$

not all zero, such that

$$\sum_{i=1}^{m-1} u_i(\alpha^{\lambda_0+t}) M_i(\alpha^{\lambda_0+t}) = 0 , \quad t \in [0, \tau_0\rangle .$$

However, the latter condition means that the polynomial

$$\sum_{i=1}^{m-1} u_i(x) M_i(x)$$

is divisible by $M_0(x)$; namely, there exists a polynomial $u_0(x) \in F_{\mu_0}[x]$ such that

$$\sum_{i=0}^{m-1} u_i(x) M_i(x) = 0 ,$$

thereby establishing (11)–(12).

APPENDIX B
ANOTHER COUNTEREXAMPLE

In Example B.1 below, we present a (nonlinear) code $\mathcal{C}$ with $\mathsf{d}(\mathcal{C}) = \ell + 2$ that has an (ℓ, ℓ)-burst list decoder, yet, as τ grows, the (τ/ℓ)-level interleaving of $\mathcal{C}$ is (ℓ', τ)-burst list decodable only for ℓ' that grows quadratically with ℓ.

Example B.1: Let ℓ be a positive even integer in the range $4 \le \ell \le 2q$, let $\tau \ge \ell^2/2$ be a positive integer multiple of ℓ, and let n be an integer such that

$$n \ge \ell\Big(\frac{\ell}{2} + 1\Big) + \Big(\frac{\ell}{2} - 1\Big) = \frac{1}{2}\ell(\ell+3) - 1 .$$

For every $j \in [0, \ell/2\rangle$, fix a word set

$$\mathcal{E}_j = \big\{\boldsymbol{e}_{i,j} \ : \ i \in [0, \ell\rangle\big\} \tag{21}$$

(of size ℓ), where each $\boldsymbol{e}_{i,j}$ is a word in F^n with support

$$\mathsf{supp}(\boldsymbol{e}_{i,j}) = [i(\ell/2{+}1) + j, (i{+}1)(\ell/2{+}1) + j\rangle \tag{22}$$

(of size $\ell/2 + 1$; note that by the choice of n, the right-hand side of (22) is indeed contained in $[0, n\rangle$). Figure 4 depicts an $n \times (\tau/\ell)$ array, where in each column $j \in [0, \ell/2\rangle$, the stacked rectangles (labeled "i, j") represent the supports of $\boldsymbol{e}_{i,j} \in \mathcal{E}_j$. (The figure is drawn to scale for $\ell = 8$.)

Fix $\mathcal{A} = \big\{\boldsymbol{a}_0, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_{\ell/2-1}\big\}$ to be a subset of F^n (of size $\ell/2 \le q$) with $\mathsf{d}(\mathcal{A}) \ge 2\ell + 4$; e.g., take some word $\boldsymbol{a} \in F^n$

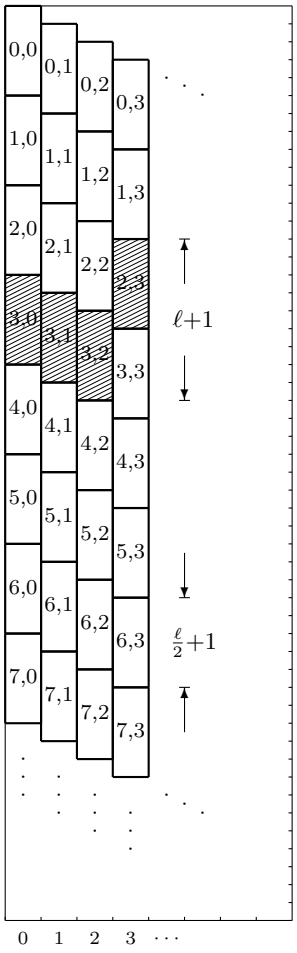

Fig. 4. Figure for Example B.1.

of Hamming weight $2\ell+4$ and let the words in $\mathcal{A}$ be distinct scalar multiples of $\boldsymbol{a}$ (note that $2\ell+4<n$ when $\ell\geq 4$). The code $\mathcal{C}$ is defined as the union of the following $\ell/2$ subsets of F^n (each of size ℓ):

$$\mathcal{C}_j = \boldsymbol{a}_j - \mathcal{E}_j \ , \quad j\in[0,\ell/2\rangle \ .$$

It can be readily verified that $\mathsf{d}(\mathcal{C})=\ell+2$: the Hamming distance between any two distinct words within each subset $\mathcal{C}_j$ is $\ell+2$, while the distance between any two words from distinct subsets is at least $(2\ell+4)-(\ell+2)\geq \ell+2$. In addition, $\mathcal{C}$ has an (ℓ,ℓ)-burst list decoder: indeed, if the intersection $(\boldsymbol{y}-\mathcal{C})\cap\mathcal{B}(n,\ell)$ contained more than ℓ words, then at least two of them, say $\boldsymbol{y}-\boldsymbol{c}_j$ and $\boldsymbol{y}-\boldsymbol{c}_k$, had to correspond to codewords $\boldsymbol{c}_j$ and $\boldsymbol{c}_k$ that belong to distinct subsets $\mathcal{C}_j$ and $\mathcal{C}_k$. However, this is impossible, since $\boldsymbol{c}_j$ and $\boldsymbol{c}_k$, and hence $\boldsymbol{y}-\boldsymbol{c}_j$ and $\boldsymbol{y}-\boldsymbol{c}_k$, differ on more than 2ℓ positions and, therefore, cannot be both in $\mathcal{B}(n,\ell)$.

Let $\mathbb{C}$ be the code over F obtained by a (τ/ℓ)-level interleaving of $\mathcal{C}$, and consider the following $n\times(\tau/\ell)$ array over F:

$$Y = (\boldsymbol{a}_0|\boldsymbol{a}_1|\ldots|\boldsymbol{a}_{\ell/2-1}|\boldsymbol{0}|\boldsymbol{0}|\ldots|\boldsymbol{0}) \ .$$

When we apply an (ℓ,ℓ)-burst list decoder $\mathcal{D}$ for $\mathcal{C}$ to each column of Y, we get, for the first $\ell/2$ columns:

$$\mathcal{D}(\boldsymbol{a}_j) = \mathcal{E}_j \ , \quad j\in[0,\ell/2\rangle$$

(for the remaining columns we get $\mathcal{D}(\boldsymbol{0})=\{\boldsymbol{0}\}$).

Next, for every $(i,k)\in\{(0,\ell/2-1)\}\cup\big([1,\ell\rangle\times[0,\ell/2\rangle\big)$, define the $n\times(\tau/\ell)$ array

$$E_{i,k} = (\boldsymbol{e}_{i,0}|\boldsymbol{e}_{i,1}|\ldots|\boldsymbol{e}_{i,k}|\boldsymbol{e}_{i-1,k+1}|\boldsymbol{e}_{i-1,k+2}|\ldots \\ |\boldsymbol{e}_{i-1,\ell/2-1}|\boldsymbol{0}|\boldsymbol{0}|\ldots|\boldsymbol{0}) \ ;$$

for $k=\ell/2-1$, this definition reduces to:

$$E_{i,\ell/2-1} = (\boldsymbol{e}_{i,0}|\boldsymbol{e}_{i,1}|\boldsymbol{e}_{i,2}|\ldots|\boldsymbol{e}_{i,\ell/2-1}|\boldsymbol{0}|\boldsymbol{0}|\ldots|\boldsymbol{0}) \ .$$

For example, the support of $E_{3,2}$ is marked in Figure 4 by the shaded rectangles. Clearly, $E_{i,k}\in Y-\mathbb{C}$. In addition, it can be verified that each $E_{i,k}$ is a τ-burst. Thus, we have shown that the number of τ-bursts in $Y-\mathbb{C}$ is greater than $\ell(\ell-1)/2$ (which, in turn, is greater than ℓ when $\ell\geq 4$); namely, $\mathbb{C}$ has no $(\ell(\ell-1)/2,\tau)$-burst list decoder. ∎

It still remains open how to make Example B.1 stronger by constructing, for any given even list size ℓ and sufficiently large field F, a *linear* code $\mathcal{C}_{\rm lin}$ over F that satisfies the following three properties:

(L1) $\mathcal{C}_{\rm lin}$ is MDS with $\mathsf{d}(\mathcal{C}_{\rm lin})=\ell+2$.
(L2) $\mathcal{C}_{\rm lin}$ has an (ℓ,ℓ)-burst list decoder (and, so, $\mathcal{C}_{\rm lin}$ attains the bound of Theorem 1.1).
(L3) For every $j\in[0,\ell/2\rangle$, there is a set $\mathcal{E}_j$ as in (21)–(22) whose ℓ elements all belong to the same coset of $\mathcal{C}_{\rm lin}$.

Whenever such a code $\mathcal{C}_{\rm lin}$ exists, it follows from the analysis in Example B.1 that the $(\ell/2)$-level interleaving of $\mathcal{C}_{\rm lin}$ yields a code $\mathbb{C}$ that has no $(\ell(\ell-1)/2,\ell^2/2)$-burst list decoder.

For the special case of $\ell=4$, we have verified by a computer program that properties (L1) and (L2) hold for the linear $[13,8]$ code over $\mathrm{GF}(67)$ that is generated by the following 8×13 matrix:

$$G_{\rm lin} = \left(\begin{array}{cccccccccccccc} 1&1&1&1&1&1&0&0&0&0&0&0&0\\ 1&1&1&0&0&0&1&1&1&0&0&0&0\\ 1&1&1&0&0&0&0&0&0&1&1&1&0\\ \hline 0&1&60&46&14&7&46&0&0&0&0&0&0\\ 0&1&60&46&0&0&0&63&32&46&0&0&0\\ 0&1&60&46&0&0&0&0&0&0&52&30&1\\ \hline 0&0&1&62&36&0&0&0&14&43&25&58&31\\ 0&0&1&62&36&1&0&0&50&59&7&7&50 \end{array}\right) .$$

As for property (L3), it is easy to see from the first three rows of $G_{\rm lin}$ that the following four words belong to the same coset of $\mathcal{C}_{\rm lin}$:

$$\begin{array}{rcl} \boldsymbol{e}_{0,0} &=& (\,66\ 66\ 66\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\,)^T\\ \boldsymbol{e}_{1,0} &=& (\,0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\,)^T\\ \boldsymbol{e}_{2,0} &=& (\,0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 0\ 0\,)^T\\ \boldsymbol{e}_{3,0} &=& (\,0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 0\,)^T \ . \end{array}$$

Similarly, from the next three rows of $G_{\rm lin}$ we get that the following four words belong to the same coset of $\mathcal{C}_{\rm lin}$:

$$\begin{array}{rcl} \boldsymbol{e}_{0,1} &=& (\,0\ 66\ 7\ 21\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\,)^T\\ \boldsymbol{e}_{1,1} &=& (\,0\ 0\ 0\ 0\ 14\ 7\ 46\ 0\ 0\ 0\ 0\ 0\ 0\,)^T\\ \boldsymbol{e}_{2,1} &=& (\,0\ 0\ 0\ 0\ 0\ 0\ 0\ 63\ 32\ 46\ 0\ 0\ 0\,)^T\\ \boldsymbol{e}_{3,1} &=& (\,0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 52\ 30\ 1\,)^T \ . \end{array}$$

Thus, we conclude that the 2-level interleaving of $\mathcal{C}_{\rm lin}$ has no $(6,8)$-burst list decoder.

Tom Kolan was born in Tiberias, Israel, in 1981. He received the B.Sc. degree in mathematics with computer science and the M.Sc. degree in computer science from Technion—Israel Institute of Technology, Haifa, Israel, in 2007 and 2011, respectively. Since 2013 he has been a research scientist at IBM Research Division, Haifa, Israel. His research interests include algebraic coding theory and applications of error-control coding.

Ron M. Roth (M'88–SM'97–F'03) 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, and during 1996–97, 2004–05, and 2011–2012 he was on sabbatical leave at Hewlett–Packard Laboratories, Palo Alto, 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 the theory of complexity.