

List Decoding of Burst Errors

Ron M. Roth, *Fellow, IEEE*, and Pascal O. Vontobel, *Member, IEEE*

Abstract—A generalization of the Reiger bound is presented for the list decoding of burst errors. It is then shown that Reed–Solomon codes attain this bound.

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

I. INTRODUCTION

Many interesting data transmission and storage systems can be modeled as channels that introduce burst errors. Assuming a list decoder at the receiver side, we study requirements that a code must satisfy in order to be suitable for data transmission over such channels, in particular, we investigate lower bounds on the code redundancy. As we will see, the resulting bounds depend on the structure of the code, i.e., we obtain different lower bounds for linear codes and group codes on the one hand, and for unstructured codes on the other hand. These bounds can be seen as generalizations of the classical Reiger bound [1], [2]. Finally, we show that Reed–Solomon codes achieve the above-mentioned redundancy lower bound for linear codes. For proving this latter result, we will derive a generalization of the known formula for the resultant of two polynomials, to a larger number of polynomials that have a certain structure.

We start by presenting several definitions that will be used throughout this work. Let F be an alphabet of size $q \geq 2$ and assume hereafter without loss of generality that F is a finite Abelian group. The set of words of length n over F is denoted by F^n (which is a group under the operation of component-by-component addition of elements of F).

We say that a word $e \in F^n$ is a τ -burst if either $e = \mathbf{0}$ (the all-zero word) or the indexes i and j of the first and last nonzero entries in e satisfy $j - i < \tau$.

Let \mathcal{C} be a code of length n over F . A (list) decoder for \mathcal{C} is a mapping $\mathcal{D} : F^n \rightarrow 2^{\mathcal{C}}$, where $2^{\mathcal{C}}$ denotes the power set of \mathcal{C} . The list size of a decoder \mathcal{D} is the largest size of $\mathcal{D}(\mathbf{y})$ over all $\mathbf{y} \in F^n$.

We say that \mathcal{D} detects any single τ -burst error if for every codeword $c \in \mathcal{C}$ and every τ -burst $e \in F^n$,

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

Manuscript received August 19, 2008. Current version published XXXXX XXXXX, 2009. Some of the material in this paper was presented at the 2008 IEEE International Symposium on Information Theory, Toronto, Canada, July 2008.

R. M. Roth is with the Computer Science Department, Technion, Haifa 32000, Israel. This work was done in part while visiting Hewlett–Packard Laboratories, 1501 Page Mill Road, Palo Alto, CA 94304, USA. This work was supported in part by Grant No. 1280/08 from the Israel Science Foundation. (e-mail: ronny@cs.technion.ac.il).

P. O. Vontobel is with Hewlett–Packard Laboratories, 1501 Page Mill Road, Palo Alto, CA 94304, USA. (e-mail: pascal.vontobel@ieee.org).

Communicated by XXXXX

Digital Object Identifier XXXXX

Such a decoder for \mathcal{C} exists if and only if for any two distinct codewords $c_1, c_2 \in \mathcal{C}$, the difference $c_1 - c_2$ is not a τ -burst.

We say that \mathcal{D} corrects any single τ -burst error if for every codeword $c \in \mathcal{C}$ and every τ -burst $e \in F^n$,

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

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. Such a decoder exists if and only if there are no $\ell+1$ distinct pairs

$$(c_0, e_0), (c_1, e_1), \dots, (c_\ell, e_\ell) ,$$

where each c_i is a codeword, each e_i is a τ -burst, and

$$c_0 + e_0 = c_1 + e_1 = \dots = c_\ell + e_\ell .$$

For the case $\ell = 1$ (conventional single τ -burst decoding), we have the well-known Reiger bound, which states that if a code \mathcal{C} has a $(1, \tau)$ -burst list decoder then the redundancy of \mathcal{C} ,

$$r = n - \log_q |\mathcal{C}| ,$$

is at least 2τ (the bound is usually stated for linear codes—see for example [1, p. 258] or [2, p. 110]—although it holds for nonlinear codes as well).

The Reiger bound holds even under the restriction that the burst errors are *phased* [1, p. 272], namely, the support of the τ -burst error is contained in one of the following sets J_i (assuming that entry indexes start at 0):

$$J_i = \{j : i\tau \leq j < (i+1)\tau\} , \quad 0 \leq i < n/\tau . \quad (1)$$

When non-overlapping τ -blocks over F are regarded as symbols of the alphabet F^τ , a phased τ -burst error becomes a single symbol error over F^τ .

When F is a field, then Reed–Solomon codes over F attain the Reiger bound and, in fact, they are optimal also for the deterministic correction of multiple burst errors (for probabilistic correction, see [3]).

Building upon a result by Parvaresh and Vardy [4], Guruswami and Rudra presented in [5] a construction of codes that have a polynomial-time list decoder that corrects any pattern of up to $r(1-\varepsilon)$ errors, where r is the code redundancy and ε is any fixed small positive real. The Guruswami–Rudra scheme is, in fact, a list decoder for Reed–Solomon codes that corrects multiple *phased* burst errors.

In this work, we consider the problem of list decoding of single burst errors that are not necessarily phased. In Section II, we present lower bounds on the redundancy of codes that have (ℓ, τ) -burst list decoders. In most cases, we will assume that the code also has a decoder that detects any single τ -burst error. In Sections III–IV, we show that Reed–Solomon codes attain the respective lower bound for linear codes.

Remark 1.1: In practice, the code \mathcal{C} serves as the set of images of an *encoding mapping* $\mathcal{E} : \mathcal{M} \rightarrow \mathcal{C}$, where \mathcal{M} is the set of messages to be transmitted through the (noisy) channel. In the context of list decoding, the mapping \mathcal{E} does not have to be lossless (i.e., one-to-one), but then, in determining the list size of a decoder \mathcal{D} , we need to count each codeword c in $\mathcal{D}(\mathbf{y})$ a number of times which equals the number of pre-images of c in \mathcal{M} (equivalently, the list size is the largest number of distinct *messages* that are returned by the decoder). However, when using a many-to-one encoder, the decoding can be ambiguous even when no errors have occurred. Such a feature is undesirable in virtually all practical applications: if ambiguity is to be allowed (through the decoding into a list of size greater than 1), then it should be limited only to cases where errors have occurred—as the probability of that to happen is presumed to be small (yet not negligible). Therefore, our definition of the list size of \mathcal{D} assumes that the encoding is lossless, thereby allowing us to regard codewords as messages. And, as said earlier, we will also want the decoder to be able to tell whether a burst error has occurred. \square

Remark 1.2: Since we focus in this paper on the case of a single burst error, any (ℓ, τ) -burst list decoder can be implemented by enumerating over the location of the first nonzero entry in the burst error, thereby effectively transforming the burst error into a burst *erasure*. Now, in the case of linear codes, erasure decoding amounts to computing a syndrome and solving linear equations and, so, erasures can be decoded in polynomial time. Hence, (ℓ, τ) -burst list decoders for linear codes always have a polynomial-time implementation (although for some linear codes we may get faster implementations by taking advantage of the specific structure of the code). \square

II. GENERALIZED REIGER BOUND

Most of this section will be devoted to generalizing the classical Reiger bound to our list-decoding setup. Interestingly, as we have already mentioned, the resulting lower bounds depend on the structure of the code. We emphasize that these differences in lower bounds are not spurious: we will show (with the help of examples) that there are indeed unstructured codes whose redundancy is lower than the redundancy that is required for group or linear codes.

For completeness reasons, we start this section by presenting a generalization of the classical sphere-packing bound to our list-decoding setup. However, unless the codes are long, namely have a block length of at least $\ell \cdot q^{\tau/\ell}$, this generalized sphere-packing bound will not be better than the generalized Reiger bound.

A. Sphere-Packing Type Bound

Given an alphabet F of size q , denote by $\mathcal{V}_q(n, \tau)$ the number of τ -bursts in F^n ; for $0 \leq \tau \leq n$, this number is given by

$$\mathcal{V}_q(n, \tau) = 1 + (q-1)n + (q-1)^2 \sum_{i=0}^{\tau-2} (n-i-1)q^i.$$

We can formulate the following sphere-packing type bound for burst list decoding; its proof is very similar to its symbol-error counterpart in [6].

Theorem 2.1: Let \mathcal{C} be a code of length n over an alphabet of size $q \geq 2$ and let τ and ℓ be positive integers. Then \mathcal{C} has an (ℓ, τ) -burst list decoder only if the redundancy r of \mathcal{C} satisfies

$$r \geq \log_q \left(\frac{\mathcal{V}_q(n, \tau)}{\ell} \right).$$

Proof: Consider all $|\mathcal{C}| \cdot \mathcal{V}_q(n, \tau)$ pairs (c, e) , where c ranges over all codewords in \mathcal{C} and e ranges over all τ -bursts in F^n . A necessary requirement for an (ℓ, τ) -burst list decoder to exist for \mathcal{C} is that any element of F^n can be written as a sum $c + e$ for at most ℓ such pairs. Therefore, we obtain the inequality

$$q^n \cdot \ell \geq |\mathcal{C}| \cdot \mathcal{V}_q(n, \tau),$$

which implies the inequality in the theorem statement. \square

For $n > 1$, the lower bound in Theorem 2.1 is smaller than $\tau + \log_q(n/\ell)$. In this section, we obtain Reiger-type bounds, which turn out to be better for lengths n that are smaller than $\ell \cdot q^{\tau/\ell}$.

B. Generalized Reiger Bound for Group Codes

A code \mathcal{C} of length n over (a finite Abelian group) F is called a *group code* over F if it is a subgroup of the group F^n . In particular, if F is a field, then every linear code over F is a group code over F .

For group codes, the conditions for the existence of decoders that detect or correct any single τ -burst are simplified. Specifically, a group code \mathcal{C} has a decoder that detects any single τ -burst if and only if the all-zero codeword is the only τ -burst in \mathcal{C} . And such a code has an (ℓ, τ) -burst list decoder if and only if no $\ell+1$ distinct τ -bursts belong to the same coset of \mathcal{C} within F^n . In particular, if \mathcal{C} is a linear code over a field F , then these τ -bursts cannot have the same syndrome (with respect to any parity-check matrix of \mathcal{C}).

The following theorem is a generalization of the Reiger bound to burst list decoders for group codes.

Theorem 2.2: Let \mathcal{C} be a group code of length n over F and let τ and ℓ be positive integers that satisfy the following three conditions:

- 1) $(\ell+1)\tau \leq n$.
- 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

$$r \geq \left(1 + \frac{1}{\ell}\right)\tau.$$

Proof: Our proof strategy will be to show that if r is not large enough, then we can exhibit $\ell+1$ distinct pairs (c_i, e_i) of codewords c_i and τ -bursts e_i that add up to the same word.

Writing $q = |F|$, we therefore assume that $r < (\ell+1)\tau/\ell$, or, equivalently,

$$\left(\frac{q^n}{|\mathcal{C}|}\right)^\ell < q^{(\ell+1)\tau}. \quad (2)$$

Let J_0, J_1, \dots, J_ℓ be disjoint subsets of integers where each J_i consists of τ consecutive elements from $\{0, 1, \dots, n-1\}$; condition 1 indeed guarantees that such subsets exist. For $i = 0, 1, \dots, \ell$, denote by \mathcal{S}_i the set of all words in F^n whose support is contained in J_i , and define the set \mathcal{S} by

$$\mathcal{S} = \{(\mathbf{v}_1 - \mathbf{v}_0 \mid \mathbf{v}_2 - \mathbf{v}_1 \mid \dots \mid \mathbf{v}_\ell - \mathbf{v}_{\ell-1}) : \mathbf{v}_i \in \mathcal{S}_i \text{ for } i = 0, 1, \dots, \ell\},$$

where the vertical bar “ \mid ” denotes string concatenation. Note that \mathcal{S} is a subset of

$$(F^n)^\ell = \underbrace{F^n \times F^n \times \dots \times F^n}_{\ell \text{ times}}$$

and that for any two elements of \mathcal{S} ,

$$\mathbf{v} = (\mathbf{v}_1 - \mathbf{v}_0 \mid \mathbf{v}_2 - \mathbf{v}_1 \mid \dots \mid \mathbf{v}_\ell - \mathbf{v}_{\ell-1}) \quad (3)$$

and

$$\mathbf{v}' = (\mathbf{v}'_1 - \mathbf{v}'_0 \mid \mathbf{v}'_2 - \mathbf{v}'_1 \mid \dots \mid \mathbf{v}'_\ell - \mathbf{v}'_{\ell-1}), \quad (4)$$

one has $\mathbf{v} = \mathbf{v}'$ if and only if $\mathbf{v}_i = \mathbf{v}'_i$ for all $i = 0, 1, \dots, \ell$. Therefore,

$$|\mathcal{S}| = \prod_{i=0}^{\ell} |\mathcal{S}_i| = q^{(\ell+1)\tau} > \left(\frac{q^n}{|\mathcal{C}|}\right)^\ell,$$

where the inequality follows from (2). This means that $|\mathcal{S}|$ is greater than the number of cosets of the subgroup $\mathcal{C}^\ell = \mathcal{C} \times \mathcal{C} \times \dots \times \mathcal{C}$ of $(F^n)^\ell$ under the component-by-component addition of elements of F^n . By the pigeon-hole principle, there must be two distinct elements in \mathcal{S} , say \mathbf{v} and \mathbf{v}' as in (3)–(4), which are in the same coset of \mathcal{C}^ℓ . Write $\mathbf{e}_i = \mathbf{v}_i - \mathbf{v}'_i$ for $i = 0, 1, \dots, \ell$; then $\mathbf{e}_i \in \mathcal{S}_i$ for all i and

$$(\mathbf{e}_1 - \mathbf{e}_0 \mid \mathbf{e}_2 - \mathbf{e}_1 \mid \dots \mid \mathbf{e}_\ell - \mathbf{e}_{\ell-1}) = \mathbf{v} - \mathbf{v}' \in \mathcal{C}^\ell. \quad (5)$$

Next, we claim that $\mathbf{e}_i \neq \mathbf{0}$ for all $i \in \{0, 1, 2, \dots, \ell\}$. Otherwise, since $\mathbf{v} \neq \mathbf{v}'$, there has to be an index i for which $\mathbf{e}_i = \mathbf{0}$ yet either $\mathbf{e}_{i-1} \neq \mathbf{0}$ (provided that $i > 0$) or $\mathbf{e}_{i+1} \neq \mathbf{0}$ (provided that $i < \ell$). But then,

$$\mathbf{e}_{i\pm 1} - \mathbf{e}_i = \mathbf{e}_{i\pm 1} \in \mathcal{C} \cap \mathcal{S}_{i\pm 1},$$

thereby contradicting condition 2, as \mathcal{C} would have a codeword that is a nonzero τ -burst.

As our next step, we claim that $\mathbf{e}_i \neq \mathbf{e}_j$ for all $0 \leq i < j \leq \ell$: indeed, since $\mathcal{S}_i \cap \mathcal{S}_j = \{\mathbf{0}\}$, then $\mathbf{e}_i = \mathbf{e}_j$ implies that both \mathbf{e}_i and \mathbf{e}_j are zero, which is impossible.

For $i = 0, 1, \dots, \ell$, define the words $\mathbf{c}_0, \mathbf{c}_1, \dots, \mathbf{c}_\ell \in F^n$ iteratively by $\mathbf{c}_0 = \mathbf{0}$ and

$$\mathbf{c}_{i+1} = \mathbf{c}_i + \mathbf{e}_i - \mathbf{e}_{i+1}, \quad 0 \leq i < \ell.$$

Since \mathcal{C} is a group code, it follows from (5) that each \mathbf{c}_i is in fact a codeword of \mathcal{C} . Thus, we have found $\ell+1$ distinct pairs

$$(\mathbf{c}_0, \mathbf{e}_0), (\mathbf{c}_1, \mathbf{e}_1), \dots, (\mathbf{c}_\ell, \mathbf{e}_\ell),$$

where each \mathbf{c}_i is a codeword of \mathcal{C} , each \mathbf{e}_i is a τ -burst, and

$$\mathbf{c}_0 + \mathbf{e}_0 = \mathbf{c}_1 + \mathbf{e}_1 = \dots = \mathbf{c}_\ell + \mathbf{e}_\ell.$$

This, in turn, contradicts condition 3. \square

Observe that in the proof of Theorem 2.2, we did not make any assumptions on the sets J_0, J_1, \dots, J_ℓ , other than

satisfying the following two properties: (i) these sets are disjoint, and (ii) each J_i consists of τ consecutive elements from $\{0, 1, \dots, n-1\}$. If we now select any particular $\ell+1$ sets J_0, J_1, \dots, J_ℓ that satisfy these two properties, then Theorem 2.2 still holds even if the burst error is restricted *a priori* to have support that is contained in one of the sets J_i . In particular, if the subsets J_i are taken as in (1), then we get that Theorem 2.2 holds also for the restricted case of phased burst errors. (When there is no such *a priori* restriction on the location of the burst errors, then, as we show in Theorem 2.3 below, condition 1 in Theorem 2.2 can be relaxed to include more pairs (ℓ, τ) .)

Remark 2.1: As pointed out earlier, when we regard non-overlapping τ -blocks over F as symbols of the alphabet F^τ , a phased τ -burst error becomes a single symbol error. Assuming that τ divides n , we can view \mathcal{C} as a code $\tilde{\mathcal{C}}$ of length n/τ over F^τ , and the redundancy of $\tilde{\mathcal{C}}$ therefore equals $(n/\tau) - \log_{q^\tau} |\tilde{\mathcal{C}}| = (1/\tau)(n - \log_q |\mathcal{C}|)$ (i.e., $(1/\tau)$ times the redundancy of \mathcal{C}). The proof of Theorem 2.2 then implies that the code $\tilde{\mathcal{C}}$ has a decoder that detects a single error and a list decoder of size ℓ that corrects a single error, only if the redundancy of $\tilde{\mathcal{C}}$ is at least $(1/\tau) \cdot (1 + (1/\ell))\tau = 1 + (1/\ell)$. In fact, this is precisely the statement we get when we plug in $\tau = 1$ in Theorem 2.2. \square

Remark 2.2: As we demonstrate in Example 2.3 below, condition 2 is generally *not* implied by condition 3. However, there are exceptions, for example when \mathcal{C} is a linear code over the field $F = \text{GF}(q)$ and $\ell < q$, then condition 2 is actually implied by condition 3. To see this, recall from Section I that the violation of condition 2 implies that there are two codewords $\mathbf{c}_1, \mathbf{c}_2 \in \mathcal{C}$ such that there is a τ -burst \mathbf{e} such that $\mathbf{e} = \mathbf{c}_1 - \mathbf{c}_2 \in \mathcal{C}$. With this, the equation $(\mathbf{c}_1 - \alpha\mathbf{e}) + (\alpha - 1)\mathbf{e} = \mathbf{c}_2$, $\alpha \in F$, shows that there are $|F| = q > \ell$ distinct ways of writing a hypothetical channel output vector \mathbf{c}_2 as a sum of a codeword and a τ -burst, thereby violating condition 3.

Clearly, if \mathcal{C} is linear then its redundancy r is always an integer. In this case, the lower bound of Theorem 2.2 can be written as

$$r \geq \tau + \left\lceil \frac{\tau}{\ell} \right\rceil. \quad (6)$$

\square

The next theorem is a modification of Theorem 2.2 where condition 1 is relaxed from $(\ell+1)\tau \leq n$ to $2\tau \leq n$ for pairs (ℓ, τ) in which ℓ divides τ .

Theorem 2.3: Theorem 2.2 holds also when condition 1 therein is relaxed to include pairs (ℓ, τ) such that $\ell \mid \tau$ and $2\tau \leq n$.

Proof: Again, the proof strategy will be to show that if r is not large enough, then we can exhibit $\ell+1$ distinct pairs of codewords and τ -bursts that add up to the same word.

Writing $q = |F|$ and $b = \tau/\ell$, we therefore assume that $r < (\ell+1)b$, or, equivalently,

$$|\mathcal{C}| = q^{n-r} > q^{n-(\ell+1)b}. \quad (7)$$

Next, we partition \mathcal{C} into $q^{n-2\tau}$ subsets $\mathcal{C}(\mathbf{v})$, where \mathbf{v} ranges over $F^{n-2\tau}$: each subset $\mathcal{C}(\mathbf{v})$ consists of all codewords of \mathcal{C}

| | | | | | | | | | | | | |
|--------------|----------------|----------------|----------------|-----------------------|-----------------------|-----------------------|-----------------------|----------|------------------------|------------------------|----------------------|--|
| c_0 | | | | | | $\mathbf{y}_{\ell+1}$ | $\mathbf{y}_{\ell+2}$ | \cdots | $\mathbf{y}_{2\ell-2}$ | $\mathbf{y}_{2\ell-1}$ | $\mathbf{y}_{2\ell}$ | |
| c_1 | \mathbf{y}_1 | | | | | $\mathbf{y}_{\ell+2}$ | $\mathbf{y}_{\ell+3}$ | \cdots | $\mathbf{y}_{2\ell-1}$ | $\mathbf{y}_{2\ell}$ | | |
| c_2 | \mathbf{y}_1 | \mathbf{y}_2 | | | | $\mathbf{y}_{\ell+3}$ | $\mathbf{y}_{\ell+4}$ | \cdots | $\mathbf{y}_{2\ell}$ | | | |
| \vdots | \vdots | \vdots | \ddots | | \ddots | | \ddots | | \vdots | | | |
| $c_{\ell-1}$ | \mathbf{y}_1 | \mathbf{y}_2 | \cdots | $\mathbf{y}_{\ell-2}$ | $\mathbf{y}_{\ell-1}$ | | | | | $\mathbf{y}_{2\ell}$ | | |
| c_ℓ | \mathbf{y}_1 | \mathbf{y}_2 | \mathbf{y}_3 | \cdots | $\mathbf{y}_{\ell-1}$ | \mathbf{y}_ℓ | | | | | | |

Fig. 1. Configuration of the codewords c_0, c_1, \dots, c_ℓ (blank elements represent “don’t cares”).

whose $(n-2\tau)$ -suffix equals \mathbf{v} . Clearly, there is at least one word \mathbf{v}' for which

$$|\mathcal{C}(\mathbf{v}')| \geq \frac{|\mathcal{C}|}{q^{n-2\tau}} = \frac{|\mathcal{C}|}{q^{n-2\ell b}} > q^{(\ell-1)b},$$

where the strict inequality follows from (7). We let \mathcal{C}' denote the set of all (2τ) -prefixes of the codewords in $\mathcal{C}(\mathbf{v}')$; note that \mathcal{C}' is a code of length 2τ over F , and since \mathcal{C} satisfies the conditions of the theorem, then so does \mathcal{C}' (except that \mathcal{C}' is not necessarily a group code, yet it is still a coset of such a code within $F^{2\tau}$).

Let J_0, J_1, \dots, J_ℓ be defined by

$$J_i = \{j : ib \leq j < ib + \tau\}, \quad 0 \leq i \leq \ell.$$

For every $0 \leq i < \ell$ we have $|J_i \cup J_{i+1}| = \tau + b = (\ell+1)b$. Since the length of \mathcal{C}' is $2\tau = 2\ell b$ and its size is greater than $q^{(\ell-1)b}$, we conclude by the pigeon-hole principle that \mathcal{C}' must contain two distinct codewords, say \mathbf{u}_i and \mathbf{u}'_i , which agree on all positions except possibly those that are indexed by $J_i \cup J_{i+1}$.

For $i = 0, 1, \dots, \ell$, define the codewords $c_0, c_1, \dots, c_\ell \in \mathcal{C}'$ iteratively by selecting $c_0 \in \mathcal{C}'$ arbitrarily and letting

$$c_{i+1} = c_i + \mathbf{u}_i - \mathbf{u}'_i, \quad 0 \leq i < \ell.$$

Thus, for every $0 \leq i < \ell$, the codewords c_i and c_{i+1} agree on all positions except possibly those that are indexed by $J_i \cup J_{i+1}$.

Let $\mathbf{y} \in F^{2\tau}$ be such that it agrees with c_0 on its last τ ($= \ell b$) positions and with c_ℓ on its first τ positions. Write

$$\mathbf{y} = (\mathbf{y}_1 \mid \mathbf{y}_2 \mid \cdots \mid \mathbf{y}_{2\ell}),$$

where each \mathbf{y}_j is a b -block over F . From the construction of the codewords c_i we get by a simple backward induction on i that the (ib) -prefix of c_i is given by

$$(\mathbf{y}_1 \mid \mathbf{y}_2 \mid \cdots \mid \mathbf{y}_i).$$

Similarly, by a forward induction on i it follows that the $((\ell-i)b)$ -suffix of c_i is given by

$$(\mathbf{y}_{\ell+i+1} \mid \mathbf{y}_{\ell+i+2} \mid \cdots \mid \mathbf{y}_{2\ell}).$$

Thus, the configuration of the codewords c_0, c_1, \dots, c_ℓ is as shown in Figure 1.

For $0 \leq i < \ell$, define $e_i = \mathbf{y} - c_i$. From Figure 1 we readily see that the support of e_i is contained in J_i and, so, e_i is a τ -burst. Obviously,

$$c_0 + e_0 = c_1 + e_1 = \cdots = c_\ell + e_\ell (= \mathbf{y}),$$

which means that we will establish the contradiction once we show that the codewords c_0, c_1, \dots, c_ℓ are all distinct. Indeed, suppose that c_0, c_1, \dots, c_i are distinct yet $c_{i+1} = c_m$ for some $m \leq i$. Since $c_{i+1} - c_i = \mathbf{u}_i - \mathbf{u}'_i \neq \mathbf{0}$, we must actually have $m < i$. But then it follows from Figure 1 that the two (distinct) codewords c_i and c_m would share the ℓ blocks

$$\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_i, \quad \text{and} \quad \mathbf{y}_{\ell+i+1}, \mathbf{y}_{\ell+i+2}, \dots, \mathbf{y}_{2\ell}$$

and, as such, they would differ on at most τ positions, thereby contradicting condition 2. \square

Remark 2.3: One may ask if condition 1 in Theorems 2.2 and 2.3 can be further relaxed to requiring only that $2\tau \leq n$ (without restricting τ to be an integer multiple of ℓ). The code we present in the next example shows that, in general, Theorems 2.2 and 2.3 no longer hold under such a relaxation. \square

Example 2.1: We present a linear code \mathcal{C} of length $n = 8$ and redundancy $r = 4$ over $F = \text{GF}(q)$ which satisfies conditions 2 and 3 in the theorem for $\ell = 2$ and $\tau = 3$. For these parameters, condition 1 in Theorems 2.2 and 2.3 is violated, and the redundancy lower bounds in these theorems indeed do not hold. In particular, the specialized redundancy lower bound (6) for linear codes in Remark 2.2 does not hold either.

The code \mathcal{C} is generated by the matrix

$$G = \begin{pmatrix} 1 * * 0 1 0 0 0 \\ 0 1 * 0 1 1 0 0 \\ 0 0 1 1 0 * 1 0 \\ 0 0 0 1 0 * * 1 \end{pmatrix},$$

where the stars stand hereafter for arbitrary elements of F . Since the rows of G form a diagonal band of 5-bursts, it follows that none of the nonzero codewords of \mathcal{C} is a 4-burst and, so, \mathcal{C} satisfies condition 2 of Theorem 2.2. Furthermore, if $c_0 + e_0 = c_1 + e_1$ for distinct codewords $c_0, c_1 \in \mathcal{C}$ and nonzero 3-burst errors e_0 and e_1 , then the leftmost entries in e_0 and e_1 have to be at least two positions apart. (Similarly, the rightmost entries in e_0 and e_1 have to be at least two positions apart.) We next show that a violating configuration

$$c_0 + e_0 = c_1 + e_1 = c_2 + e_2$$

cannot exist (for distinct $c_0, c_1, c_2 \in \mathcal{C}$) by distinguishing between several cases.

Case 1: Suppose to the contrary that there exists a violating

configuration with error words of the form

$$\begin{aligned} e_0 &= (**00000) \\ e_1 &= (00***00) , \\ e_2 &= (0000***0) \end{aligned}$$

and assume without loss of generality that $c_1 = \mathbf{0}$. Then, from $c_0 - c_1 = e_1 - e_0$ we deduce that c_0 takes the form

$$c_0 = (*****000) ,$$

which means that c_0 has to be a nonzero scalar multiple of the first row of G . Also, from $c_2 - c_1 = e_1 - e_2$ we get that

$$c_2 = (00*****0) ,$$

which means that c_2 is a nonzero scalar multiple of the third row in G . Therefore, the fourth position in c_0 is zero while it is nonzero in c_2 , and this, in turn, implies that the fourth position in $c_0 - c_2$ is nonzero also. Yet, the latter contradicts the fact that $c_0 - c_2 = e_2 - e_0$.

Case 2: Suppose now that the violating configuration takes the form

$$\begin{aligned} e_0 &= (**00000) \\ e_1 &= (00***00) \\ e_2 &= (00000***0) \end{aligned}$$

(e_0 and e_1 are as in Case 1, yet the support of e_2 is shifted one position to the right). Assuming again that $c_1 = \mathbf{0}$, we get that c_0 has to be a nonzero scalar multiple of the first row of G while c_2 has to be a nonzero linear combination of the last two rows of G . Hence, the fifth position in $c_0 - c_2$ cannot be zero, yet this contradicts the fact that $c_0 - c_2 = e_2 - e_0$.

There are two other violating configurations to consider, which are obtained by reversing the order of coordinates in the error patterns covered by Cases 1 and 2. The proof of contradiction remains the same due to the symmetries of G . \square

C. Generalized Reiger Bound for General Codes

The lower bound on the redundancy in Theorems 2.2 and 2.3 applies to group codes. As the next example shows, this bound does not apply to general codes.

Example 2.2: Let F be an alphabet of size $q \geq 2$ and consider the code \mathcal{C} of length 4 and size $2q-2$ over F which is defined as the union of the following two sets:

$$\mathcal{C}_1 = \{(aaa0) : a \in F \setminus \{0\}\}$$

and

$$\mathcal{C}_2 = \{(0aaa) : a \in F \setminus \{0\}\} .$$

We claim that \mathcal{C} satisfies conditions 2–3 of Theorem 2.2, for $\tau = \ell = 2$. Indeed, every two distinct codewords $c_1, c_2 \in \mathcal{C}$ either differ on each of their first three positions (if $c_1, c_2 \in \mathcal{C}_1$), or on each of their last three positions (if $c_1, c_2 \in \mathcal{C}_2$), or on both their first and last positions (if $c_1 \in \mathcal{C}_1$ and $c_2 \in \mathcal{C}_2$); in either case, the difference $c_1 - c_2$ is not a 2-burst and therefore condition 2 is satisfied.

As for condition 3, suppose to the contrary that there exist three distinct codewords $c_0, c_1, c_2 \in \mathcal{C}$ and respective three 2-bursts $e_0, e_1, e_2 \in F^4$ such that

$$c_0 + e_0 = c_1 + e_1 = c_2 + e_2 .$$

Since \mathcal{C} has been shown to satisfy condition 2, the supports of e_0, e_1 , and e_2 have to be distinct, which means that c_0, c_1 , and c_2 can be assumed to take the form shown in Figure 1, with y_0, y_1, y_2 , and y_3 now being elements of F . In particular, c_0 and c_1 agree on their last position, which is possible only if both belong to \mathcal{C}_1 . Similarly, c_1 and c_2 agree on their first position, implying that both belong to \mathcal{C}_2 . Thus, c_1 belongs to both \mathcal{C}_1 and \mathcal{C}_2 , which is a contradiction since these sets are disjoint.

Now, the redundancy of \mathcal{C} equals $4 - \log_q(2q-2)$ and, for $q > 2$, this number is smaller than 3, which is the lower bound we get for $\tau = \ell = 2$ in Theorem 2.3. \square

In fact, Example 2.2 attains the lower bound in the next result (which applies to list size 2; we will generalize this bound to larger ℓ in Theorem 2.6 below).

Proposition 2.4: Let \mathcal{C} be a code of length n over an alphabet of size $q \geq 2$ and let τ be a positive integer that satisfies the following three conditions:

- 1) τ is even and $2\tau \leq n$.
- 2) There is a decoder for \mathcal{C} that detects any single τ -burst error.
- 3) There is a $(2, \tau)$ -burst list decoder for \mathcal{C} .

Then the redundancy r of \mathcal{C} satisfies

$$\begin{aligned} r &\geq 2\tau - \log_q(2q^{\tau/2} - 2) \\ &= \left(1 + \frac{1}{2}\right)\tau - \log_q 2 + \log_q\left(\frac{1}{1 - q^{-\tau/2}}\right) . \end{aligned}$$

In particular, $r > \left(1 + \frac{1}{2}\right)\tau - \log_q 2$.

Proof: Write $b = \tau/2$, and suppose to the contrary that $r < 2\tau - \log_q(2q^b - 2)$; namely,

$$|\mathcal{C}| = q^{n-r} > q^{n-2\tau} \cdot (2q^b - 2) . \quad (8)$$

Let \mathcal{C}' be the code of length 2τ as defined in the proof of Theorem 2.3; recall that since \mathcal{C} satisfies the three conditions of the theorem, then so does \mathcal{C}' . From (8) we get that

$$|\mathcal{C}'| \geq \frac{|\mathcal{C}|}{q^{n-2\tau}} > 2q^b - 2 ,$$

that is,

$$|\mathcal{C}'| \geq 2q^b - 1 . \quad (9)$$

Let c be a codeword of \mathcal{C} . We say that a codeword $c' \neq c$ in \mathcal{C} is a *right* (respectively, *left*) *neighbor* of c if c and c' share the same suffix (respectively, prefix) of length b . Let \mathcal{C}'_1 (respectively, \mathcal{C}'_2) be the set of all codewords of \mathcal{C}' that have no right (respectively, left) neighbors. Since the b -suffixes of the elements of \mathcal{C}'_1 must all be distinct, we must have $|\mathcal{C}'_1| \leq q^b$. From (9) it follows that the set $\mathcal{C}' \setminus \mathcal{C}'_1$ is nonempty; hence, there is at least one b -block that does not appear as a b -suffix in any element in \mathcal{C}'_1 . Thus, $|\mathcal{C}'_1| \leq q^b - 1$ and, since the same upper bound applies to $|\mathcal{C}'_2|$, we get

$$|\mathcal{C}' \setminus (\mathcal{C}'_1 \cup \mathcal{C}'_2)| \geq (2q^b - 1) - 2(q^b - 1) \geq 1 .$$

We conclude that \mathcal{C}' contains a codeword c_1 that has both a right neighbor c_0 and a left neighbor c_2 , and by condition 2 these two neighbors must be distinct. Yet the codewords c_0, c_1 , and c_2 form the violating configuration of Figure 1, thereby reaching a contradiction. \square

The next lemma will be used to generalize Proposition 2.4 or to larger ℓ .

Lemma 2.5: Let ℓ be an integer greater than 1 and let \mathcal{C} be a code of length 2ℓ over an alphabet of size q . Suppose that \mathcal{C} satisfies conditions 2–3 in Theorem 2.2 for $\tau = \ell$. Then

$$|\mathcal{C}| < \ell \cdot q^{\ell-1}.$$

Proof: We prove the lemma by induction on ℓ . For any integer $\ell > 1$, we denote by $M(\ell)$ the size of the largest code \mathcal{C} of length 2ℓ that satisfies the conditions of the lemma.

The induction base ($\ell = 2$) follows by substituting $\tau = 2$ and $n = 4$ in Proposition 2.4: we get $M(2) \leq 2q - 2$.

Turning to the induction step, given an integer $\ell > 2$, let \mathcal{C} be a code of length 2ℓ and size $M(\ell)$ that satisfies the conditions of the lemma. Let the set \mathcal{C}_1 consist of all codewords c in \mathcal{C} with the property that no codeword in $\mathcal{C} \setminus \{c\}$ agrees with c on its first $\ell-1$ positions. Denote by \mathcal{C}_2 the complement set $\mathcal{C} \setminus \mathcal{C}_1$.

Let \mathcal{T} be the set of all distinct $(\ell-1)$ -prefixes of the words in \mathcal{C}_2 . No element in \mathcal{T} can appear as an $(\ell-1)$ -prefix in any codeword in \mathcal{C}_1 and, so,

$$|\mathcal{C}_1| \leq q^{\ell-1} - |\mathcal{T}|.$$

Since \mathcal{C} has a decoder that detects any single ℓ -burst error, no two distinct words in \mathcal{C}_2 can have the same ℓ -prefix, which means that at most q words in \mathcal{C}_2 can share the same $(\ell-1)$ -prefix. Hence,

$$|\mathcal{C}_2| \leq q \cdot |\mathcal{T}|$$

and, so,

$$\begin{aligned} M(\ell) &= |\mathcal{C}_1| + |\mathcal{C}_2| \\ &\leq |\mathcal{C}_1| + q \cdot |\mathcal{T}| \\ &\leq (q^{\ell-1} - |\mathcal{T}|) + q \cdot |\mathcal{T}|. \end{aligned} \quad (10)$$

For any element v in the alphabet F of \mathcal{C} , let $\mathcal{C}_2(v)$ denote the set of all codewords in \mathcal{C}_2 that end with v . There exists at least one element $v' \in F$ for which

$$|\mathcal{C}_2(v')| \geq \frac{|\mathcal{C}_2|}{q} = \frac{M(\ell) - |\mathcal{C}_1|}{q}.$$

Let the mapping $\varphi : \mathcal{C}_2(v') \rightarrow F^{2\ell-2}$ be defined by

$$\varphi(x_1 x_2 \dots x_{2\ell-1} v') = x_1 x_2 \dots x_{\ell-1} x_{\ell+1} x_{\ell+2} \dots x_{2\ell-1};$$

namely, $\varphi(\cdot)$ deletes (punctures) the entries of its argument at the ℓ th and (2ℓ) th positions. (Note that, because we assumed that \mathcal{C} satisfies condition 2 in Theorem 2.2 for $\tau = \ell$, the mapping φ turns out to be bijective.) Denote by \mathcal{C}' the set of images of this mapping:

$$\mathcal{C}' = \{\varphi(c) : c \in \mathcal{C}_2(v')\}.$$

Since $\mathcal{C}_2(v')$ satisfies condition 2 for $\tau = \ell$, then \mathcal{C}' has to satisfy that condition for $\tau = \ell-1$; furthermore, $\varphi(\cdot)$ is bijective and, so,

$$|\mathcal{C}'| = |\mathcal{C}_2(v')| \geq \frac{M(\ell) - |\mathcal{C}_1|}{q},$$

or

$$\begin{aligned} M(\ell) &\leq |\mathcal{C}_1| + q \cdot |\mathcal{C}'| \\ &\leq (q^{\ell-1} - |\mathcal{T}|) + q \cdot |\mathcal{C}'|. \end{aligned}$$

Combining the latter inequality with (10) we thus get

$$M(\ell) \leq (q^{\ell-1} - |\mathcal{T}|) + q \cdot \min\{|\mathcal{T}|, |\mathcal{C}'|\}. \quad (11)$$

Next, we show that \mathcal{C}' has an $(\ell-1, \ell-1)$ -burst list decoder. If this were not the case, then there would be a word

$$\mathbf{y}' = y_1 y_2 \dots y_{\ell-1} y_{\ell+1} y_{\ell+2} \dots y_{2\ell-1}$$

in $F^{2\ell-2}$ and respective ℓ words $\mathbf{c}'_0, \mathbf{c}'_1, \dots, \mathbf{c}'_{\ell-1}$ in \mathcal{C}' that would form the violating configuration shown in Figure 2. The respective pre-images $\mathbf{c}_i = \varphi^{-1}(\mathbf{c}'_i)$, all belonging to \mathcal{C}_2 (and hence to \mathcal{C}), would then look like the first ℓ rows in the configuration of Figure 1 (with each block \mathbf{y}_i therein replaced by the element y_i of F). Recall, however, that since \mathcal{C}_2 is the complement set of \mathcal{C}_1 , each codeword in \mathcal{C}_2 agrees on the first $\ell-1$ positions with at least one other codeword in \mathcal{C}_2 . In particular, there is a codeword $\mathbf{c}_\ell \in \mathcal{C}_2$ that agrees with $\mathbf{c}_{\ell-1}$ ($= \varphi^{-1}(\mathbf{c}'_{\ell-1})$) on its first $\ell-1$ positions. The codeword \mathbf{c}_ℓ could therefore serve as the last row in Figure 1, thereby contradicting the fact that \mathcal{C} has an (ℓ, ℓ) -burst list decoder. We conclude that \mathcal{C}' has an $(\ell-1, \ell-1)$ -burst list decoder and, so,

$$|\mathcal{C}'| \leq M(\ell-1).$$

Combining the latter inequality with (11) we get

$$\begin{aligned} M(\ell) &\leq (q^{\ell-1} - |\mathcal{T}|) + q \cdot \min\{|\mathcal{T}|, M(\ell-1)\} \\ &\leq \max_{t \in \mathbb{Z}} \{(q^{\ell-1} - t) + q \cdot \min\{t, M(\ell-1)\}\} \\ &= q^{\ell-1} + (q-1) \cdot M(\ell-1). \end{aligned}$$

The result now follows by the induction hypothesis on $M(\ell-1)$. \square

Theorem 2.6: Let \mathcal{C} be a code of length n over an alphabet of size $q \geq 2$ and let ℓ and τ be positive integers that satisfy the following three conditions:

- 1) $\ell \mid \tau$, $\ell > 1$, and $2\tau \leq n$.
- 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

$$r > \left(1 + \frac{1}{\ell}\right)\tau - \log_q \ell.$$

Proof: Denote by F the alphabet of \mathcal{C} , and let \mathcal{C}' be defined as in the proof of Theorem 2.3. Then \mathcal{C}' is a code of length 2τ over F which satisfies conditions 2–3 and

$$|\mathcal{C}'| \geq \frac{|\mathcal{C}|}{q^{n-2\tau}}. \quad (12)$$

Write $b = \tau/\ell$. By grouping together non-overlapping b -blocks over F , we now regard \mathcal{C}' as a code of length 2ℓ over F^b . As such, \mathcal{C}' satisfies the conditions of Lemma 2.5 for an alphabet of size q^b . Hence,

$$|\mathcal{C}'| < \ell \cdot q^{b(\ell-1)},$$

| | | | | | | | | |
|---------------|----------|--------------|--------------|--------------|---------------|---------------|--------------|---------------|
| c'_0 | | $y_{\ell+1}$ | $y_{\ell+2}$ | \cdots | $y_{2\ell-2}$ | $y_{2\ell-1}$ | | |
| c'_1 | y_1 | | | | $y_{\ell+2}$ | $y_{\ell+3}$ | \cdots | $y_{2\ell-1}$ |
| c'_2 | y_1 | y_2 | | | | $y_{\ell+3}$ | $y_{\ell+4}$ | \cdots |
| \vdots | \vdots | \vdots | \ddots | \ddots | \ddots | \ddots | \vdots | |
| $c'_{\ell-1}$ | y_1 | y_2 | \cdots | $y_{\ell-2}$ | $y_{\ell-1}$ | | | |

Fig. 2. Configuration of the words $c'_0, c'_1, \dots, c'_{\ell-1}$.

which readily implies with (12) that

$$|\mathcal{C}| < \ell \cdot q^{n-2\tau+b(\ell-1)} = \ell \cdot q^{n-b(\ell+1)}.$$

Thus, the redundancy r of \mathcal{C} satisfies

$$\begin{aligned} r &= n - \log_q |\mathcal{C}| \\ &> n - \log_q (\ell \cdot q^{n-b(\ell+1)}) \\ &= b(\ell+1) - \log_q \ell \\ &= \left(1 + \frac{1}{\ell}\right) \tau - \log_q \ell, \end{aligned}$$

as claimed. \square

In all our bounds, we have assumed that the code \mathcal{C} has a decoder that detects any single τ -burst error (condition 2 in the theorems in Sections II-B and II-C). We have also shown in Remark 2.2 that when \mathcal{C} is linear and $\ell < q$, then condition 2 is actually implied by condition 3. One could therefore ask whether condition 2 is at all necessary in order to obtain our bounds. The next example answers this question affirmatively: it exhibits a code that does not satisfy condition 2 and it violates the bound of Proposition 2.4.

Example 2.3: Let F be an alphabet of size $q \geq 2$, select z to be a nonzero element in F , and consider the code \mathcal{C} of length 4 and size $2q$ over F which is defined as the union of the following two sets:

$$\{(a00a) : a \in F\}$$

and

$$\{(azz a) : a \in F\}.$$

We show that \mathcal{C} has a $(2, 2)$ -burst list decoder (while obviously, there is no decoder for \mathcal{C} that can detect any single 2-burst error). Suppose to the contrary that there exist three distinct codewords $c_0, c_1, c_2 \in \mathcal{C}$ and respective three 2-bursts $e_0, e_1, e_2 \in F^4$ such that

$$c_0 + e_0 = c_1 + e_1 = c_2 + e_2.$$

Since no two codewords in \mathcal{C} share the same 2-suffix, there can be at most one 2-burst—say e_0 —whose last two entries are zero. By symmetry, e_2 (say) is the only 2-burst whose first two entries are zero. Thus, e_1 can be zero only in its first and last positions, which brings us to the configuration of Figure 1; namely, c_0 and c_2 are distinct right and left neighbors of c_1 (see the proof of Proposition 2.4). However, this is impossible, since each codeword in \mathcal{C} has exactly one neighbor (which is both a left neighbor and a right neighbor).

Hence, we conclude that the code \mathcal{C} satisfies condition 3 for $\tau = \ell = 2$ yet violates condition 2 for these parameters. (Note that for $q = 2$, the code \mathcal{C} happens to be linear over $\text{GF}(2)$.) The redundancy of \mathcal{C} equals $4 - \log_q(2q) = 3 - \log_q 2$, which is smaller than the lower bound that we get for $\tau = 2$ in Proposition 2.4. \square

The code in Example 2.3 attains the next bound.

Proposition 2.7: Let \mathcal{C} , q , and τ be as in Proposition 2.4, except that \mathcal{C} is not required to satisfy condition 2. Then the redundancy r of \mathcal{C} satisfies

$$r \geq \left(1 + \frac{1}{2}\right) \tau - \log_q 2.$$

Proof: We follow the steps of the proof of Proposition 2.4, except that (8) is replaced by

$$|\mathcal{C}| = q^{n-r} > 2q^{n-2\tau+b}.$$

and (9) by

$$|\mathcal{C}'| \geq 2q^b + 1.$$

Let \mathcal{C}'_0 be the set of all codewords in \mathcal{C}' that have a right neighbor which is also a left neighbor. By condition 3, each codeword in \mathcal{C}'_0 has exactly one such neighbor (which, obviously, is also an element of \mathcal{C}'_0). Also, no codeword in \mathcal{C}'_0 can have an ordinary neighbor (left or right) in $\mathcal{C}' \setminus \mathcal{C}'_0$, (or else we would get the violating configuration of Figure 1). In particular, no b -suffix (respectively, b -prefix) of a codeword in \mathcal{C}'_0 can appear as such in a codeword that belongs to \mathcal{C}'_1 (respectively, \mathcal{C}'_2), where \mathcal{C}'_1 and \mathcal{C}'_2 are as in the proof of Proposition 2.4. Therefore,

$$|\mathcal{C}'_1|, |\mathcal{C}'_2| \leq q^b - \frac{|\mathcal{C}'_0|}{2}$$

and, so,

$$|\mathcal{C}' \setminus (\mathcal{C}'_0 \cup \mathcal{C}'_1 \cup \mathcal{C}'_2)| \geq (2q^b + 1) - |\mathcal{C}'_0| - 2\left(q^b - \frac{|\mathcal{C}'_0|}{2}\right) \geq 1.$$

We conclude that \mathcal{C}' contains a codeword c_1 that has a right neighbor c_0 and a left neighbor c_2 , and these neighbors are distinct. But this brings us again to the configuration in Figure 1, thereby reaching a contradiction. \square

Example 2.1 demonstrates that for $q \geq 4$, Theorem 2.6 becomes false if we remove from condition 1 therein the assumption that τ is an integer multiple of ℓ (requiring only that $2\tau \leq n$). A similar statement holds for Proposition 2.7 and $q \geq 5$.

III. GENERALIZED RESULTANT OF CERTAIN POLYNOMIALS

This section develops the tools that will be used in Section IV to show that Reed–Solomon codes attain the bound (6). In particular, Theorem 3.2 below presents a generalization of the known formula for the resultant of two polynomials, to a larger number of polynomials that have a certain structure.

For a field F and an integer k , denote by $F_k[x]$ the set of all polynomials over F of degree less than k in the indeterminate x .

Let F be the finite field $\text{GF}(q)$ and let r be a positive integer. Fix α to be a nonzero element in F with multiplicative order at least r , and let $\beta = (\beta_i)_{i=0}^\ell$ be a vector whose $\ell+1$ entries are all nonzero elements of F . Let $\mu_0, \mu_1, \dots, \mu_\ell$ be positive integers such that

$$\sum_{i=0}^{\ell} \mu_i = r. \quad (13)$$

For $i = 0, 1, \dots, \ell$, define

$$\tau_i = r - \mu_i, \quad 0 \leq i \leq \ell,$$

and for an indeterminate x , denote by $M_i(x; \beta_i)$ the expression

$$M_i(x; \beta_i) = \prod_{j=0}^{\tau_i-1} (x - \beta_i \alpha^j).$$

We regard $M_i(x; \beta_i)$ as a univariate polynomial over F in the indeterminate x , with β_i serving as a parameter.

In this section, we prove the following result.

Theorem 3.1: The following two conditions are equivalent:

(i) There exist polynomials

$$u_i(x) \in F_{\mu_i}[x], \quad 0 \leq i \leq \ell, \quad (14)$$

not all zero, such that

$$\sum_{i=0}^{\ell} u_i(x) M_i(x; \beta_i) = 0. \quad (15)$$

(ii) For some distinct i and k in the range $0 \leq i, k \leq \ell$ and some integer t in the range $-\mu_i < t < \mu_k$,

$$\frac{\beta_k}{\beta_i} = \alpha^t.$$

Proof: This theorem is implied by the discussion in the following paragraphs, in particular by Theorem 3.2. \square

For each $i = 0, 1, 2, \dots, \ell$, write

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

(where $M_{i,j} \in F$ is a function of β_i), and define $A_i(\beta_i)$ to be the following $\mu_i \times r$ echelon matrix over F :

$$A_i(\beta_i) = \begin{pmatrix} M_{i,0} & M_{i,1} & \dots & M_{i,\tau_i} & & & \\ & M_{i,0} & M_{i,1} & \dots & M_{i,\tau_i} & & \mathbf{0} \\ \mathbf{0} & & \ddots & \ddots & \dots & \ddots & \\ & & & M_{i,0} & M_{i,1} & \dots & M_{i,\tau_i} \end{pmatrix}. \quad (16)$$

Then, (14)–(15) can be expressed in matrix form as

$$\sum_{i=0}^{\ell} \mathbf{u}_i A_i(\beta_i) = \mathbf{0},$$

where each \mathbf{u}_i is a row vector in F^{τ_i} , and at least one of these vectors is nonzero. Equivalently,

$$\mathbf{u} A = \mathbf{0},$$

where \mathbf{u} is a nonzero vector in F^r and $A = A(\beta)$ is the following $r \times r$ matrix over F :

$$A(\beta) = \begin{pmatrix} A_0(\beta_0) \\ A_1(\beta_1) \\ \vdots \\ A_\ell(\beta_\ell) \end{pmatrix}.$$

Theorem 3.2: (Generalized resultant of $M_i(x; \beta_i)$) For some nonzero constant $\kappa(\alpha) \in F$ (which depends on α but not on β),

$$\det(A(\beta)) = \kappa(\alpha) \cdot \prod_{0 \leq i < k \leq \ell} \prod_{s=0}^{\mu_i-1} \prod_{t=0}^{\mu_k-1} (\beta_k \alpha^s - \beta_i \alpha^t). \quad (17)$$

To prove the latter theorem, we regard β as a vector of indeterminates and

$$\Delta(\beta) = \det(A(\beta))$$

as a multivariate polynomial over F . The properties of this polynomial are summarized in Lemmas 3.3–3.5 below, and Theorem 3.2 will then follow as a direct corollary of these properties.

Given a vector $\xi = (\xi_0 \xi_1 \dots \xi_{m-1})$, we denote by $V(\xi)$ the $m \times m$ Vandermonde matrix

$$V(\xi) = (\xi_t^s)_{s,t=0}^{m-1}.$$

We will use the notation V_m for $V(1 \alpha \alpha^2 \dots \alpha^{m-1})$.

Lemma 3.3: The multivariate polynomial $\Delta(\beta)$ is not identically zero.

Proof: We find an assignment $\beta^* = (\beta_i^*)_{i=0}^\ell$ for β for which $\Delta(\beta^*) \neq 0$. For $i = 0, 1, \dots, \ell$, define the partial sums

$$r_i = \mu_0 + \mu_1 + \dots + \mu_i.$$

Moreover, let

$$\beta_i^* = \alpha^{r_i}.$$

Taking the product of $A_i(\beta_i^*)|_{\beta_i^* = \alpha^{r_i}}$ and V_r , one can check that the nonzero columns of the resulting $\mu_i \times r$ matrix $A_i(\alpha^{r_i}) V_r$ are indexed by integers j in the range $0 \leq j < r_i$. Furthermore, the μ_i columns that are indexed by

$$r_{i-1} \leq j < r_i$$

(with $r_{-1} = 0$) form a $\mu_i \times \mu_i$ non-singular matrix X_i which is obtained by multiplying a Vandermonde matrix to the right by a diagonal matrix; specifically:

$$X_i = \left(\alpha^{(r_{i-1}+t)s} \right)_{s,t=0}^{\mu_i-1} \cdot \text{diag} (M_i(\alpha^{r_{i-1}+t}; \alpha^{r_i}))_{t=0}^{\mu_i-1}. \quad (18)$$

It follows that the respective matrix $A(\beta^*)V_r$ has a block-triangular form and, so,

$$\begin{aligned} \Delta(\beta^*) = \det(A(\beta^*)) &= \frac{\det(A(\beta^*)V_r)}{\det(V_r)} \\ &= \frac{1}{\det(V_r)} \prod_{i=0}^{\ell} \det(X_i) \\ &\neq 0. \end{aligned} \quad (19)$$

□

Lemma 3.4: For each $i = 0, 1, \dots, \ell$, the degree of β_i in $\Delta(\beta)$ is at most $\mu_i \tau_i$.

Proof: By inspecting the matrix $A(\beta)$ we see that the largest contribution to the degree of β_i can be made by the leftmost (main) diagonal in $A_i(\beta_i)$: the product of the elements along that diagonal is

$$M_{i,0}^{\mu_i} = (-\alpha^{(\tau_i-1)/2} \beta_i)^{\mu_i \tau_i},$$

and, so, the degree of β_i in $\Delta(\beta)$ can be at most $\mu_i \tau_i$. □

Lemma 3.5: For every distinct $i, k \in \{0, 1, \dots, \ell\}$, the multivariate polynomial $\Delta(\beta)$ is divisible by

$$\prod_{t=0}^{\mu_k-1} (\beta_k - \beta_i \alpha^t)^{\min\{\mu_i, \mu_k-t\}}.$$

Proof: Due to symmetry, it suffices to prove the lemma assuming $i = 0$. Hereafter in this proof, we fix k to be some element in $\{1, 2, \dots, \ell\}$. While it is not too difficult to see that $\beta_k - \beta_0 \alpha^t$ is a factor of $\Delta(\beta)$, we also need to establish the multiplicity of that factor. We do this by introducing μ_0 new indeterminates which are given by the entries of the following vector γ :

$$\gamma = (\gamma_h)_{h=0}^{\mu_0-1}.$$

We define the respective polynomials

$$\sigma_h(x; \gamma_h) = \prod_{j=0}^{\tau_0+h-1} (x - \gamma_h \alpha^j), \quad 0 \leq h < \mu_0,$$

and regard them as univariate polynomials in the indeterminate x over the field

$\Phi = F(\beta_1, \beta_2, \dots, \beta_{k-1}, \beta_{k+1}, \beta_{k+2}, \dots, \beta_\ell, \gamma_0, \gamma_1, \dots, \gamma_{\mu_0-1})$; namely, Φ is the rational function field over F where the indeterminates are all the entries of β and γ , except for β_k . (The analysis in the sequel will involve univariate polynomials in the indeterminate β_k over Φ , as well as the rational function field $\Phi(\beta_k)$.) Notice that when we substitute $\gamma_h = \beta_0$, we get

$$\sigma_h(x; \beta_0) = M_0(x; \beta_0) \cdot \prod_{j=0}^{h-1} (x - \beta_0 \alpha^{\tau_0+j}). \quad (20)$$

Let $S_0(\gamma)$ be the $\mu_0 \times r$ matrix over Φ whose rows are given by the coefficients of $\sigma_h(x; \gamma_h)$, for $0 \leq h < \mu_0$ (i.e., entry (h, j) in $S_0(\gamma)$ is the coefficient of x^j in $\sigma_h(x; \gamma_h)$). It follows from (20) that when we substitute $\gamma = \beta_0 = (\beta_0 \beta_0 \dots \beta_0)$, then $S_0(\beta_0)$ and $A_0(\beta_0)$ are related by

$$S_0(\beta_0) = LA_0(\beta_0), \quad (21)$$

where L is a $\mu_0 \times \mu_0$ lower-triangular matrix having 1's along its main diagonal.

Let $S(\beta_k; \gamma)$ be the following $r \times r$ matrix over the field $\Phi(\beta_k)$:

$$S(\beta_k; \gamma) = S(\beta_1, \beta_2, \dots, \beta_\ell; \gamma) = \begin{pmatrix} S_0(\gamma) \\ A_1(\beta_1) \\ A_2(\beta_2) \\ \vdots \\ A_\ell(\beta_\ell) \end{pmatrix}.$$

From (21) we get that, in $\Phi(\beta_k)$,

$$\Delta(\beta) = \det(S(\beta_k; \beta_0)). \quad (22)$$

Let $f(\beta_k; \gamma)$ be the following univariate polynomial in the indeterminate β_k over Φ :

$$f(\beta_k; \gamma) = \det(S(\beta_k; \gamma)). \quad (23)$$

We verify that for every $0 \leq t < \mu_k$ and every h in the range

$$\mathcal{R}(t) = \left\{ h : \max\{0, t + \mu_0 - \mu_k\} \leq h < \mu_0 \right\},$$

the element $\gamma_h \alpha^t$ is a root of $f(\beta_k; \gamma)$. We do this by demonstrating that for any such t and h , the rows of $S(\gamma_h \alpha^t; \gamma)$ are linearly dependent over Φ . Specifically, we exhibit nonzero $e \in \Phi^{\mu_0}$ and $u \in \Phi^{\mu_k}$ such that

$$eS_0(\gamma) - uA_k(\gamma_h \alpha^t) = \mathbf{0}. \quad (24)$$

Given t and h , let $u(x)$ be the following univariate polynomial over Φ :

$$u(x) = \left(\prod_{j=0}^{t-1} (x - \gamma_h \alpha^j) \right) \left(\prod_{j=t}^{h+\mu_k-\mu_0-1} (x - \gamma_h \alpha^{\tau_k+j}) \right)$$

(where a product over an empty set is defined to be 1). Since $h \in \mathcal{R}(t)$ we have

$$\deg u(x) = h + \mu_k - \mu_0 < \mu_k,$$

so we can take u to be the vector of coefficients of $u(x)$. We readily get that

$$\begin{aligned} u(x)M_k(x; \gamma_h \alpha^t) &= \prod_{j=0}^{\tau_k+h+\mu_k-\mu_0-1} (x - \gamma_h \alpha^j) \\ &= \prod_{j=0}^{\tau_0+h-1} (x - \gamma_h \alpha^j) = \sigma_h(x; \gamma_h). \end{aligned}$$

Hence, (24) holds when e is taken as a unit vector having 1 at position h .

We conclude that, over Φ , the polynomial $f(\beta_k; \gamma)$ is divisible by

$$\prod_{t=0}^{\mu_k-1} \prod_{h \in \mathcal{R}(t)} (\beta_k - \gamma_h \alpha^t).$$

Substituting $\gamma = \beta_0$, it follows that $f(\beta_k; \beta_0)$ is divisible by

$$\prod_{t=0}^{\mu_k-1} (\beta_k - \beta_0 \alpha^t)^{|\mathcal{R}(t)|} = \prod_{t=0}^{\mu_k-1} (\beta_k - \beta_0 \alpha^t)^{\min\{\mu_0, \mu_k-t\}},$$

and, by (22)–(23), so is $\Delta(\beta)$. \square

Proof of Theorem 3.2: The right-hand side of (17) factors over F as follows: for every distinct $i, k \in \{0, 1, \dots, \ell\}$ and every $0 \leq t < \mu_k$, the term

$$\beta_k - \beta_i \alpha^t$$

has multiplicity $\min\{\mu_i, \mu_k - t\}$ in the right-hand side of (17) (for $t = 0$, we regard $\beta_k - \beta_i$ and $\beta_i - \beta_k$ as the same term). By Lemma 3.5 we then get that the right-hand side of (17) divides $\Delta(\beta)$. Furthermore, for each $\{0, 1, \dots, \ell\}$, the degree of β_i in the right-hand side of (17) equals

$$\sum_{\substack{k=0 \\ k \neq i}}^{\ell} \mu_i \mu_k = \mu_i(r - \mu_i) = \mu_i \tau_i.$$

Hence, by Lemmas 3.3 and 3.4 we conclude that the right-hand side of (17) actually equals $\Delta(\beta)$.

Note that the exact expression for $\kappa(\alpha)$ is

$$\kappa(\alpha) = \frac{1}{(\det(V_r))^2} \cdot \prod_{i=0}^{\ell} \left((\det(V_{\mu_i}))^2 \alpha^{\mu_i \tau_i (\tau_i - 1)/2} \prod_{s=0}^{\mu_i - 1} \prod_{t=\mu_i}^{r-1} (\alpha^t - \alpha^s) \right),$$

and its derivation can be found in [7, Appendix II]. \square

Remark 3.1: For $\ell = 1$ (in which case $\tau_0 + \tau_1 = r$), the matrix $A(\beta)$ is the Sylvester matrix [8] of the polynomials $M_0(x; \beta_0)$ and $M_1(x; \beta_1)$ (up to reversal of the order of the rows and columns), and Theorem 3.2 then provides the known formula for the resultant of these polynomials [9, p. 36]. \square

Remark 3.2: For $r = \ell + 1$ (in which case $\mu_i = 1$ for all i), the matrix $A(\beta)$ is related to the $r \times r$ Vandermonde matrix $V(\beta)$ by

$$A(\beta) = V^T(\beta)U(\alpha),$$

where $V^T(\beta)$ is the transpose of $V(\beta)$ and where $U(\alpha)$ does not depend on β and is zero below its main anti-diagonal. Theorem 3.2 then provides the known formula for the determinant of a square Vandermonde matrix. \square

IV. BURST LIST DECODING OF REED–SOLOMON CODES

The goal of this section is to show that the well-known Reed–Solomon codes achieve the generalized Reiger bound for linear codes (see Equation (6) in Remark 2.2).

Let F be the finite field $\text{GF}(q)$ and let α be an element of multiplicative order n in F . For a non-negative integer $r < n$, denote by $\mathcal{C}_{\text{RS}}(n, r)$ the $[n, k=n-r]$ Reed–Solomon code over F with a parity-check matrix

$$H_{\text{RS}} = (\alpha^{sj})_{s=0, j=0}^{r-1, n-1}.$$

The following theorem shows that $\mathcal{C}_{\text{RS}}(n, r)$ attains the bound (6).

Theorem 4.1: There is an (ℓ, τ) -burst list decoder for $\mathcal{C}_{\text{RS}}(n, r)$ whenever ℓ and τ are positive integers that satisfy

$$r \geq \tau + \left\lceil \frac{\tau}{\ell} \right\rceil. \quad (25)$$

Proof: We will assume in the proof that (25) holds with equality; otherwise, just reduce r to the right-hand side of (25).

Recalling the coset characterization of τ -burst errors in Section II-B, we suppose to the contrary that there exist $\ell + 1$ distinct row vectors $e_0, e_1, \dots, e_\ell \in F^n$ such that

$$H_{\text{RS}} e_0^T = H_{\text{RS}} e_1^T = \dots = H_{\text{RS}} e_\ell^T, \quad (26)$$

where the support of each e_i is contained in a subset

$$J_i = \{\lambda_i + t : 0 \leq t < \tau\};$$

here each λ_i is an integer in the range $0 \leq \lambda_i \leq n - \tau$. We observe that since the minimum distance of $\mathcal{C}_{\text{RS}}(n, r)$ is $r + 1$, for every distinct $i, k \in \{0, 1, \dots, \ell\}$ we must have

$$|J_i \cup J_k| > r,$$

which readily implies that for $i \neq k$,

$$|J_i \setminus J_k| > r - \tau = \left\lceil \frac{\tau}{\ell} \right\rceil.$$

Thus, for every distinct $i, k \in \{0, 1, \dots, \ell\}$,

$$\|\lambda_k - \lambda_i\|_n > \left\lceil \frac{\tau}{\ell} \right\rceil, \quad (27)$$

where

$$\|a\|_n = \min(|a|, n - |a|).$$

The sum of the sizes of the sets J_i is $(\ell + 1)\tau$, and this value may be smaller than ℓr in case τ is not divisible by ℓ . For convenience in the sequel, we will now artificially expand some of the sets J_i by one, by adding the element $\lambda_i + \tau$, so that the sum of the sizes becomes exactly ℓr . Letting τ_i be the size of (the possibly expanded) J_i and defining

$$\mu_i = r - \tau_i,$$

we have

$$\sum_{i=0}^{\ell} \mu_i = \sum_{i=0}^{\ell} (r - \tau_i) = (\ell + 1)r - \sum_{i=0}^{\ell} \tau_i = r$$

(see (13)).

Denote by H_i the $r \times \tau_i$ sub-matrix of H_{RS} which is formed by the columns of H that are indexed by J_i , namely:

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

Define the $r \times r$ matrix T_i by

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

where I_i is a $\tau_i \times \tau_i$ identity matrix and $A_i(\cdot)$ is given by (16). Notice that $A_i(\alpha^{\lambda_i})H_i = 0$ and, so, the product $T_i H_i$ results in an $r \times \tau_i$ matrix Y_i which takes the following form:

$$Y_i = T_i H_i = \left(\begin{array}{c} (\alpha^{(\lambda_i + t)s})_{s,t=0}^{\tau_i-1} \\ 0 \end{array} \right). \quad (28)$$

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

Consider the following $\ell r \times \ell r$ matrix B :

$$B = \begin{pmatrix} \begin{array}{cc|c} H_0 & -H_1 & \\ \hline H_0 & & -H_2 \\ \vdots & & \\ \hline H_0 & \mathbf{0} & -H_\ell \end{array} & \mathbf{0} \end{pmatrix}.$$

Next, we multiply B to the left by an $\ell r \times \ell r$ block-diagonal matrix T which contains the blocks T_1, T_2, \dots, T_ℓ along its main diagonal:

$$TB = \begin{pmatrix} \begin{array}{cc|c} Z_1 & -Y_1 & \\ \hline Z_2 & & -Y_2 \\ \vdots & & \\ \hline Z_\ell & \mathbf{0} & -Y_\ell \end{array} & \mathbf{0} \end{pmatrix},$$

where Y_i is given by (28) and

$$Z_i = T_i H_0 = \left(\frac{(\alpha^{\lambda_0+t})_{s,t=0}^{\tau_0-1}}{A_i(\alpha^{\lambda_i}) H_0} \right).$$

Our contradicting assumption (26) implies that B has dependent columns and is therefore singular. This, in turn, implies the singularity of the $\tau_0 \times \tau_0$ matrix

$$\begin{pmatrix} A_1(\alpha^{\lambda_1}) H_0 \\ A_2(\alpha^{\lambda_2}) H_0 \\ \vdots \\ A_\ell(\alpha^{\lambda_\ell}) H_0 \end{pmatrix},$$

which is formed by taking the last μ_i rows of each Z_i and stacking them together for all $i = 1, 2, \dots, \ell$ (notice that $\sum_{i=1}^{\ell} \mu_i = r - \mu_0 = \tau_0$). Hence, there exist row vectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_\ell$, not all zero, such that $\mathbf{u}_i \in F^{\mu_i}$ and

$$\sum_{i=1}^{\ell} \mathbf{u}_i A_i(\alpha^{\lambda_i}) H_0 = \mathbf{0}.$$

Equivalently, there exist polynomials

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

not all zero, such that

$$\sum_{i=1}^{\ell} u_i(\alpha^{\lambda_0+t}) M_i(\alpha^{\lambda_0+t}; \alpha^{\lambda_i}) = 0, \quad 0 \leq t < \tau_0.$$

However, the latter condition means that the polynomial

$$\sum_{i=1}^{\ell} u_i(x) M_i(x; \alpha^{\lambda_i})$$

(which is in $F_r[x]$) is divisible by $M_0(x; \alpha^{\lambda_0})$; namely, there exists a $u_0(x) \in F_{\mu_0}[x]$ such that

$$\sum_{i=0}^{\ell} u_i(x) M_i(x; \alpha^{\lambda_i}) = 0.$$

We then get from Theorem 3.1 that there exist distinct $i, k \in \{0, 1, \dots, \ell\}$ such that

$$\|\lambda_k - \lambda_i\|_n < \max\{\mu_i, \mu_k\} \leq \left\lceil \frac{\tau}{\ell} \right\rceil.$$

This, however, contradicts (27). \square

REFERENCES

- [1] S. Lin and D. J. Costello, Jr., *Error Control Coding: Fundamentals and Applications*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1983.
- [2] W. W. Peterson and E. J. Weldon, Jr., *Error-Correcting Codes*, 2nd ed. Cambridge, MA: MIT Press, 1972.
- [3] V. Y. Krachkovsky, "Reed-Solomon codes for correcting phased error bursts," *IEEE Trans. on Inform. Theory*, vol. 49, no. 11, pp. 2975–2984, Nov. 2003.
- [4] A. Vardy and F. Parvaresh, "Correcting errors beyond the Guruswami-Sudan radius in polynomial time," in *Proc. of the 46th Symp. Foundations of Computer Science (FOCS)*, Pittsburgh, PA, 2005, pp. 285–294.
- [5] V. Guruswami and A. Rudra, "Explicit codes achieving list decoding capacity: error-correcting up to the Singleton bound," in *Proc. of the 38th Annual ACM Symp. Theory of Computing (STOC)*, Seattle, WA, 2006, pp. 1–10.
- [6] P. Elias, "Error-correcting codes for list decoding," *IEEE Trans. on Inform. Theory*, vol. 37, no. 1, pp. 5–12, Jan. 1991.
- [7] R. M. Roth and P. O. Vontobel, "List decoding of burst errors," *preliminary version of the present paper, available on ArXiv under <http://arxiv.org/abs/0808.2837>*, Aug. 2008.
- [8] D. Cox, J. Little, and D. O. Shea, *Ideals, Varieties, and Algorithms*, 2nd ed. New York: Springer, 1997.
- [9] R. Lidl and H. Niederreiter, *Finite Fields*, 2nd ed. Cambridge, UK: Cambridge University Press, 1997.

Ron M. Roth (M'88–SM'97–F'03) was born in Ramat Gan, Israel, in 1958. He 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 and 2004–05 he was on sabbatical leave at Hewlett-Packard Laboratories, Palo Alto, California.

Dr. Roth is the author of the book *Introduction to Coding Theory*, published by Cambridge University Press in 2006. He 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.

Pascal O. Vontobel (S'96–M'97) received the Diploma degree in electrical engineering in 1997, the Post-Diploma degree in information techniques in 2002, and the Ph.D. degree in electrical engineering in 2003, all from ETH Zurich, Zurich, Switzerland.

From 1997 to 2002, he was a Research and Teaching Assistant at the Signal and Information Processing Laboratory at ETH Zurich. After being a Postdoctoral Research Associate at the University of Illinois at Urbana-Champaign, at the University of Wisconsin-Madison (Visiting Assistant Professor), and at the Massachusetts Institute of Technology, he joined the Information Theory Research Group at Hewlett-Packard Laboratories in Palo Alto, CA, in the summer of 2006 as a research scientist. His research interests lie in information theory, communications, and signal processing.

Dr. Vontobel was awarded the ETH medal for his Ph.D. dissertation.