# Repairing Multiple Failures for Scalar MDS Codes

Jay Mardia, Burak Bartan, and Mary Wootters

*Abstract*—In distributed storage, erasure codes (like Reed–Solomon Codes) are often employed to provide reliability. In this setting, it is desirable to be able to repair one or more failed nodes while minimizing the *repair bandwidth*. In this paper, motivated by Reed-Solomon codes, we study the problem of repairing multiple failed nodes in a scalar MDS code. We extend the framework of (Guruswami and Wootters, 2017) to give a framework for constructing repair schemes for multiple failures in general scalar MDS codes in the centralized repair model. We then specialize our framework to Reed–Solomon codes, and also extend and improve upon recent results of (Dau *et al.*, 2017).

*Index Terms*—Regenerating codes, Reed–Solomon codes, multiple failures.

## I. INTRODUCTION

IN CODING for distributed storage, one wishes to store some data $x \in \Sigma^k$ across $n$ nodes. These nodes will occasionally fail, and *erasure coding* is used to allow for the recovery of $x$ given only a subset of the $n$ nodes. A common solution is to use a *Maximum-Distance Separable* (MDS) code; for example, a Reed-Solomon code. An MDS code encodes a message $x \in \Sigma^k$ into $n$ symbols $c \in \Sigma^n$, in such a way that any $k$ symbols of $c$ determine $x$. By putting the symbols $c_i$ of $c$ on different nodes, this gives a distributed storage scheme which can tolerate $n - k$ node failures.

While this level of worst-case robustness is desirable, in practice it is much more common for only a few nodes to fail, rather than $n - k$ of them. To that end, it is desirable to design codes which are simultaneously MDS and which also admit cheap repair of a few failures. One important notion of "cheap" is *network bandwidth:* the amount of data downloaded from the surviving nodes. The naive MDS repair scheme would involve downloading $k$ complete symbols of $n$. Minimum storage regenerating (MSR) codes [7] improve the situation; these are codes which maintain the MDS property, while substantially reducing repair bandwidth for a single failure.

Most of the work in regenerating codes has focused on this case of a single failure, as in many systems this is the most common case, as documented by Rashmi *et al.* [19]. However, even in [19] it is not uncommon to have multiple failures at once, and some systems employ *lazy repair* to encourage this [13]. Motivated by this, many recent works have considered this case of *multiple failures.* In this work, we focus on the question of multiple failures for scalar MDS codes. Our work is inspired by Reed-Solomon codes—arguably the most commonly-used code for distributed storage—but our framework works more broadly for any scalar MDS code.

### A. Previous Work and Our Contributions

There has been a huge amount of work on regenerating codes, and we refer the reader to the survey [8] for an excellent introduction. Most of the work has focused on a single failure, but recently there has been a great deal of work on multiple failures. Two commonly studied models are the *centralized model* (which we study here), and the *cooperative model.* In the centralized model, a single repair center is responsible for the repair of all failed nodes, while in the cooperative model the replacement nodes may cooperate but are distinct [12], [14], [22].

We focus on the centralized model. Most of the work in this model has focused on achieving the *cut-set bound* for multiple failures [1], [15], [20], [24], [26]–[30]. This extends with well-known cut-set bound for the single-failure case [7], and is only achievable when the *sub-packetization* (that is, the number of sub-symbols that each node stores) is reasonably large; in particular, we (at least) require the subpacketization $t$ to be larger than $n - k$, otherwise the trivial lower bound of $k + t - 1$ is larger than the cut-set bound. Most of the works mentioned above focus on *array codes,* that is, codes where the alphabet $\Sigma = B^t$ is naturally thought of as a vector space over a finite field $B$, and the codes are generally $B$-linear.

Other recent works [4], [5], [17], [26] focused on Reed-Solomon codes, and studied multiple failures for *scalar codes,* where the alphabet $\Sigma = F$ is a finite field, and the codes are required to be linear over $F$. In [26], the goal is again the cut-set bound, and the underlying subpacketization is necessarily exponentially large. In [4] and [5], the sub-packetization is taken to be smaller, on the order of $\log(n)$. This is the natural parameter regime for Reed-Solomon codes, and in this regime the cut-set bound is not achievable for high-rate codes. Our work falls into this second parameter regime. We note that recent work [16] has studied an intermediate regime, establishing a trade-off between the sub-packetization and bandwidth when repairing multiple failures for Reed-Solomon codes.

Beginning with the work of Shanmugam *et al.* [21], the repair properties of scalar MDS codes has been increasingly studied [2]–[6], [9], [11], [16], [21], [23], [25], [26].

J. Mardia and B. Bartan are with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: jmardia@stanford.edu; bbartan@stanford.edu).

M. Wootters is with the Departments of Computer Science and Electrical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: marykw@stanford.edu).

Guruswami and Wootters [11] gave a framework for studying single-failure repair schemes for scalar MDS codes, and the works of Dau *et al.* [4], [5] mentioned above adapt the single-failure scheme from [11] to handle two or three failures for Reed-Solomon codes, in several models, including the centralized model.

In this work, we extend and improve the results of [4], [5]. More precisely, we make the following contributions.

1) Following the setup of [11], we give a general framework for constructing repair schemes of scalar MDS codes for multiple failures. Theorem 1 shows that collections of dual codewords with certain properties naturally give rise to repair schemes for multiple failures. This framework is applicable to any scalar MDS code, and for any number of failures $r \leq n - k$.

2) We instantiate Theorem 1 with two different schemes in Theorems 2 and 3 that both give non-trivial repair schemes for multiple failures. While the scheme in Theorem 2 is asympototically better (as $n$, the length of the code, tends to infinity), the scheme in Theorem 3 is better for small $n$.

   Our schemes are the first in this parameter regime to work for $r > 3$, and additionally they improve over previous work [4], [5] when $r = 2, 3$. More precisely, we obtain the following bounds:

   - Theorem 3 improves and generalizes the scheme of Dau *et al.* for Reed-Solomon codes in the centralized model [5]. More precisely, in Theorem 3, for any $r \leq r_0$, for some $r_0 = O(\sqrt{\log(n)})$, we give schemes for high-rate (say, $1-\varepsilon$), length $n$ Reed-Solomon codes which have repair bandwidth (measured in bits)

$$\left( (n-r) \cdot r - \frac{r(r-1)}{2} \left( \frac{1}{\varepsilon} - 1 \right) \right) \cdot \log_2(1/\varepsilon).$$

   For comparison, the scheme of Dau *et al.* worked for $r = 2, 3$, and had bandwidth $(n-r) \cdot r \cdot \log_2(1/\varepsilon)$. Thus, for $r = 2, 3$ Theorem 3 improves the bandwidth by $\frac{r(r-1)}{2}(1/\varepsilon - 1) \log_2(1/\varepsilon)$ bits, and for larger $r$ we give the first non-trivial schemes for Reed-Solomon Codes with small subpacketization.

   When $r = 1$, this collapses to the scheme of [11], which is optimal.

   - Theorem 2 improves over Theorem 3 asymptotically, but it is not as good for small values of $n$. For this theorem we again generalize the constructions of [6], [11], but we do so in a different way and do not go the same route as [5]. We obtain repair schemes for length $n$ Reed-Solomon codes of rate $1 - \varepsilon$, which have repair bandwidth (in bits) at most

$$\min_{r' \geq r} \left( (n - r') \left( \log_2(n) - \left\lfloor \log_2 \left( \frac{n\varepsilon + r' - 1}{2r' - 1} \right) \right\rfloor \right) \right)$$

   When $r = 1$, this too collapses to the scheme of [11], which is optimal. When $1 < r < r_0$ for some $r_0 = \Omega(n^{1-\varepsilon})$, this bound is approximately $(n - r) \log_2(2r/\varepsilon)$, which is on the order of $\log_2(2r/\varepsilon)$ bits per surviving node. The bound becomes trivial (that is, equal to $kt$) when $r \geq r_0$.
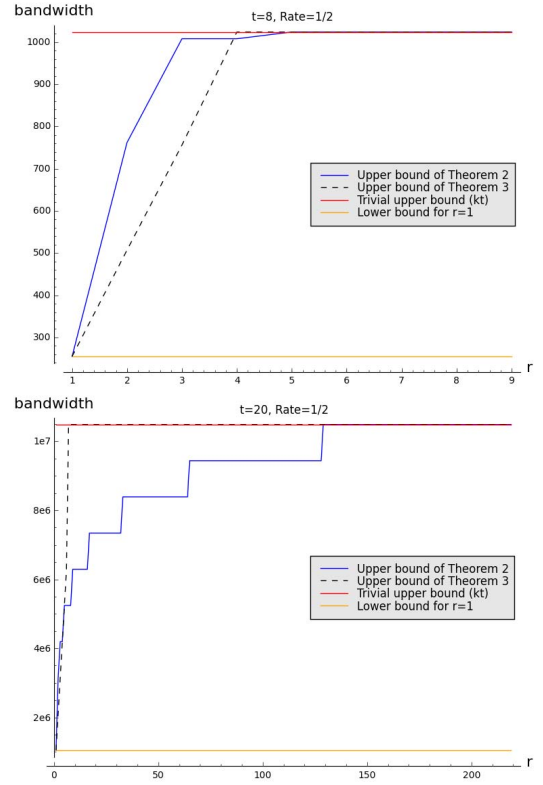


Fig. 1. The bound of Theorems 2 and 3, for rate $1/2$ Reed-Solomon codes with $n = 2^t$, for $t = 8$ (top) and $t = 20$ (bottom). When $t$ is small, Theorem 3 is better (requires less bandwidth); when $t$ is large, Theorem 2 is better. Other works [4], [5] on multiple failures in this parameter regime corrected up to two or three failures only.

We compare our two bounds, along with the trivial bound of bandwidth $kt$ and the optimal bound for $r = 1$, in Figure 1.

We emphasize that the codes in Theorems 2 and 3 are simply Reed-Solomon codes that use all their evaluation points; that is, our results imply that this one classical code can be repaired from a growing number of failures with non-trivial bandwidth, and the repair behavior degrades gracefully as the number of failures increases to $n - k$. However, we do not have a matching lower bound for larger $r$, and we suspect that further improvements are possible.

### B. Organization

In Section II we set up notation and give formal definitions for the problems we consider. In Section III, we give Theorem 1, which provides a framework for constructing repair schemes for multiple failures for general scalar MDS codes. In Section IV, we give Theorems 2 and 3, which specialize Theorem 1 to Reed-Solomon codes, and gives the results advertised above.

### II. PRELIMINARIES

In this section, we set up notation, and formally introduce the definitions that we will work with throughout the paper.

### A. Notation

We use the notation $[n]$ to mean the set of integers $\{1, \ldots, n\}$, and for vectors $v = (v_1, v_2, \ldots, v_n)$,

$w = (w_1, w_2, \ldots, w_n) \in F^n$, we use $\langle v, w \rangle = \sum_{i \in [n]} v_i w_i$ to denote the standard inner product.

*1) Matrix and Vector Notation:* Unless otherwise noted, vectors $v$ are treated as column vectors; the $i$'th entry of a vector $v$ is denoted $v_i$. For a vector $v \in F^n$ and a set $I \subseteq [n]$, with $I = \{i_1, \ldots, i_r\}$ and $i_1 < i_2 < \cdots < i_r$, $v_I$ denotes the (column) vector $(v_{i_1}, v_{i_2}, \ldots, v_{i_r})$. For a vector $v \in F^m$, we will use **set** $(v)$ to denote the set **set** $(v) = \{v_i : i \in [m]\}$.

For a matrix $M$, we use $M[:, i]$ to refer to the $i$'th column and $M[i, :]$ to refer to the $i$'th row of $M$. For sets $I, J$, we will use $M[I, J]$ to refer to the submatrix of $M$ containing the rows indexed by $I$ and the columns indexed by $J$; and we will extend this to $M[I, :]$ and $M[:, J]$ to mean the submatrix formed by the rows in $I$ or columns in $J$, respectively. Our notation is 1-indexed.

*2) Finite Field Notation:* Throughout this paper, $F$ denotes a finite field, and $B \subset F$ denotes a subfield of $F$. We use $F^*$ and $B^*$ to denote the group of units in $F$ and $B$ respectively, and $F^*/B^*$ to denote the quotient group. For a set of elements $S \subseteq F$, we will use $\mathrm{span}_B(S)$ to denote the linear span over $B$ of $S$:

$$\mathrm{span}_B(S) = \left\{ \sum_{x \in S} a_x \cdot x : a_x \in B \right\}.$$

We will similarly use $\dim_B$ to refer to the dimension over $B$. Finally, for a field $F$ with a subfield $B$, so that $F$ has degree $t$ over $B$, the field trace $\mathrm{tr}_{F/B} : F \to B$ is defined by

$$\mathrm{tr}_{F/B}(x) := \sum_{i=0}^{t-1} x^{|B|^i}.$$

The function $\mathrm{tr}_{F/B}$ is a $B$-linear function from $F$ to $B$. We refer the reader to, for example, [10] for a primer/refresher on finite fields.

### B. Definitions

Let $\mathcal{C} \subset \Sigma^n$ be a code of block length $n$ over an alphabet $\Sigma$. As described in the introduction, we imagine the $n$ symbols of a codeword $c = (c_1, c_2, \ldots, c_n) \in \mathcal{C}$ are distributed between $n$ different nodes, so that node $i$ stores the symbol $c_i$.

*1) The Exact Repair Problem:* In the *exact repair problem,* one node, Node $i$, is unavailable, and the goal is to *repair* it (that is, recover $c_i$) using only information from the remaining nodes. Of course, any MDS code can achieve this: by definition, all of $c$ is determined by any $k$ symbols, and so any $k$ surviving nodes determine all of $c$ and in particular the missing information $c_i$. But, as described in the introduction, we hope to do better than this, in terms of the amount of data downloaded.

Formally, suppose that $\Sigma \simeq B^t$ is a vector space over some *base field* $B$. Thus, the contents of a node (a symbol $c_i \in \Sigma$) are $t$ sub-symbols from $B$. When a node fails, a replacement node or repair center can contact a surviving node, which may do some computation and return some number—possibly fewer than $t$—sub-symbols from $B$. The parameter $t$ is called the *sub-packetization.* Formally, we define an exact repair scheme as follows.

*Definition 1:* An exact repair scheme *for a code $\mathcal{C} \subset \Sigma^n$ is defined as follows. For each $i \in [n]$, there is a collection of functions*

$$\left\{ g_{i,j} : j \in [n] \setminus \{i\} \right\},$$

*so that*

$$g_{i,j} : \Sigma \to B^{b_{i,j}}$$

*for some non-negative integer $b_{i,j}$. and so that for all $c \in \mathcal{C}$, $c_i$ is determined from $\{g_{i,j}(c_j) : j \in [n] \setminus \{i\}\}$. The* bandwidth *of this scheme (measured relative to $B$) is the total number of elements of $B$ required to repair any node:*

$$\text{bandwidth} = \max_{i \in [n]} \sum_{j \in [n] \setminus \{i\}} b_{i,j}.$$

*Remark 1 (Variants):* The definition above is not the only definition of regenerating codes, and is missing several parameters often considered. For example, we may also limit number of nodes contacted, requiring the repair scheme to only contact $d$ out of the surviving nodes. We may also allow for the nodes to store more elements of $B$ than the original data blocks to (in the lingo of regenerating codes, to move away from the MSR setting and toward the MBR setting). However, the goal of the current work is to study multiple failures in scalar MDS codes.

*2) Multiple Failures:* In this work, we will focus on the centralized model of multiple repair [1]. In this model, a repair center is in charge of the repair for all the nodes. We count as bandwidth the information downloaded by this repair center, but not between the center and any of the replacement nodes. Formally, we have the following definition.

*Definition 2:* An exact centralized repair scheme *for $r$ failures for a code $\mathcal{C} \subset \Sigma^n$ is defined as follows. For each set $I \subseteq [n]$ of size at most $r$, there is a collection of functions*

$$\left\{ g_{I,j} : j \in [n] \setminus I \right\},$$

*so that*

$$g_{I,j} : \Sigma \to B^{b_{I,j}}$$

*for some non-negative integer $b_{I,j}$, and so that for all $c \in \mathcal{C}$, and all $i \in I$, $c_i$ is determined from $\{g_{I,j}(c_j) : j \in [n] \setminus I\}$. The* bandwidth *of this scheme (measured relative to $B$) is the total number of elements of $B$ required to repair the nodes in any set $I$:*

$$\text{bandwidth} = \max_{I \subset [n], |I| \leq r} \sum_{j \in [n] \setminus \{i\}} b_{I,j}.$$

Definition 2 is perhaps the simplest possible definition of the exact repair problem for multiple failures. As per Remark 1, we could spice up the definition of the exact repair problem in many ways; and beyond that following the work of [5] for Reed-Solomon codes in other models, we could include in our measure of bandwidth some way to capture the cost of communication between the multiple replacement nodes. However, addressing even this simplest case is interesting and much is unknown, so we will focus on this case for the current work, and we hope that the insights of this work may be extended to more complicated models.

*3) Linear Repair Schemes and Scalar MDS Codes:* As mentioned in the introduction, most of the work on regenerating codes explicitly views the alphabet $\Sigma$ as a vector space over some field $B$. However, for many codes commonly used in distributed storage—notably Reed-Solomon codes—it is more common to view the alphabet $\Sigma$ as a finite field $F$. Such codes are termed "scalar" MDS codes [21]. However, if $B \subseteq F$ is a subfield so that the degree of $F$ over $B$ is $t$, then $F$ is in fact a vector space of dimension $t$ over $B$, and so the set-up above makes sense. We focus on this setting for the rest of the paper: that is, $\mathcal{C} \subset F^n$ is a linear subspace which has the property that any $k$ symbols of a codeword $c \in \mathcal{C}$ determine $c$.

In this setting, while more restrictive[1] than that of Definition 1, there is additional algebraic structure which, it turns out, very nicely characterizes exact repair schemes for a scalar MDS code $\mathcal{C} \subset F^n$ (for a single failure) in terms of the *dual code* $\mathcal{C}^\perp := \{v \in F^n : \langle c, v \rangle = 0 \forall c \in \mathcal{C}\}$. More formally, we define a *repair matrix* for a symbol $i \in [n]$ as follows.

*Definition 3:* Let $\mathcal{C} \subseteq F^n$ be an MDS code over $F$, and suppose that $B$ is a subfield of $F$, so that $F$ has degree $t$ over $B$. Let $i \in [n]$. A repair matrix *with bandwidth $b$ for an index $i$ is a matrix $M \in F^{n \times t}$ with the following properties:*

1) *The columns of $M$ are codewords in the dual code $\mathcal{C}^\perp$.*
2) *The elements of the $i$'th row $M[i, :]$ of $M$ have full rank over $B$.*
3) *We have*

$$\sum_{j \in [n] \setminus \{i\}} \dim_B \left\{ \mathrm{span}_B \left\{ \mathit{set}\left( M[j, :] \right) \right\} \right\} = b.$$

One of the main results of [11] was that repair matrices precisely characterize linear repair schemes. We say that a repair scheme as in Definition 1 is *linear* if the functions $g_{i,j}$, along with the function that determines $c_i$, are all $B$-linear. The work of [11] showed that a (scalar) MDS code $\mathcal{C}$ admits a linear repair scheme with bandwidth $b$ if and only if, for all $i \in [n]$, there is a repair matrix with bandwidth at most $b$ for $i$.

## III. FRAMEWORK

In this section, we extend the framework of [11] to the case of multiple repairs. We first define an analog of repair matrices for multiple repair.

*Definition 4:* Let $\mathcal{C} \subseteq F^n$ be a MDS code over $F$, and suppose that $B$ is a subfield of $F$, so that $F$ has degree $t$ over $B$. Let $I \subset [n]$ have size $r$. A multiple-repair matrix *with bandwidth $b$ for $I$ is a matrix $M \in F^{n \times rt}$ with the following properties:*

1) *The columns of $M$ are codewords in the dual code $\mathcal{C}^\perp$.*
2) *The submatrix $M[I, :]$ has full rank over $B$, in the sense that for all nonzero $y \in B^{rt}$, $M[I, :] \cdot y \neq 0$.*
3) *We have*

$$\sum_{j \in [n] \setminus I} \dim_B \left\{ \mathrm{span}_B \left\{ \mathit{set}\left( M[j, :] \right) \right\} \right\} = b.$$

Our main theorem is that an MDS code $\mathcal{C}$ admits a (linear) repair scheme for a set $I$ of failed nodes with bandwidth $b$ if there exists a multiple-repair matrix with bandwidth $b$ for $I$.

*Theorem 1:* Let $\mathcal{C} \subset F^n$ be an MDS code, and let $B \subset F$ be a subfield so that $F$ has degree $t$ over $B$. Suppose that for all $I \subseteq [n]$ of size $r$, there is a multiple-repair matrix $M \in F^{n \times rt}$ with bandwidth at most $b$ for $I$. Then $\mathcal{C}$ admits an exact centralized repair scheme for $r$ failures with bandwidth $b$.

*Remark 2:* In fact, the other direction of Theorem 1 is true as well; that is, any exact linear centralized repair scheme gives rise to a multiple-repair matrix. The proof follows a similar outline as the proof in [11], and is formally proved as Theorem 6 in the independent work [17].

*Proof:* Let $I \subset [n]$ be any set of $r$ failures, and let $M \in F^{n \times rt}$ be a multiple-repair matrix with bandwidth $b$ for $I$. For each $j \in [n] \setminus I$, we will show how to use $M$ to construct the functions $g_{I, j} : F \to B^{b_{I, j}}$.

We will choose $b_{I, j}$ (the number of sub-symbols returned by $g_{I, j}$) to be $b_{I, j} = \dim_B \left\{ \mathrm{span}_B \left\{ \mathbf{set}\left( M[j, :] \right) \right\} \right\}$. Then by Definition 4, $\sum_{j \in [n] \setminus I} b_{I, j} \leq b$. Let $\lambda_1, \ldots, \lambda_{b_{I, j}} \in F$ be a basis for the elements of $M[j, :]$ over $B$. (We note that the $\lambda_i$ depend on the choice of $j$, but we suppress this for notational clarity). For $x \in F$, we choose

$$g_{I, j}(x) = (\mathrm{tr}_{F/B}(\lambda_1 \cdot x), \mathrm{tr}_{F/B}(\lambda_2 \cdot x), \cdots, \mathrm{tr}_{F/B}(\lambda_{b_{I, j}} \cdot x)).$$

We first observe that, by Property 3 in Definition 4, the total bandwidth of this scheme is $b$ symbols of $B$. We next need to show that this repair scheme works; that is, we need to show that for all $c \in \mathcal{C}$, the values $\left\{ g_{I, j}(c_j) : j \in [n] \setminus I \right\}$ determine $\{c_i : i \in I\}$.

By Property 1 in Definition 4, for all $\ell \in [rt]$, we have $M[:, \ell] \in \mathcal{C}^\perp$. This means that for all $c \in \mathcal{C}$, and for all $\ell \in [rt]$,

$$0 = \sum_{i \in [n]} c_i \cdot M[i, \ell]$$

$$\sum_{i \in I} c_i \cdot M[i, \ell] = - \sum_{j \in [n] \setminus I} c_j \cdot M[j, \ell]$$

$$\mathrm{tr}_{F/B}\left( \sum_{i \in I} c_i \cdot M[i, \ell] \right) = \mathrm{tr}_{F/B}\left( - \sum_{j \in [n] \setminus I} c_j \cdot M[j, \ell] \right)$$

$$\sum_{i \in I} \mathrm{tr}_{F/B}(c_i \cdot M[i, \ell]) = - \sum_{j \in [n] \setminus I} \mathrm{tr}_{F/B}(c_j \cdot M[j, \ell]).$$

We claim that the right-hand side above can be constructed from the values $\left\{ g_{I, j}(c_j) : j \in [n] \setminus I \right\}$. Indeed, write $M[j, \ell] = \sum_{i=1}^{b_{I, j}} a_{i, \ell, j} \lambda_i$ for some coefficients $a_{i, \ell, j} \in B$. Then,

$$- \sum_{j \in [n] \setminus I} \mathrm{tr}_{F/B}(c_j \cdot M[j, \ell])$$

$$= - \sum_{j \in [n] \setminus I} \mathrm{tr}_{F/B}\left( c_j \cdot \sum_{i=1}^{b_{I, j}} a_{i, \ell, j} \lambda_i \right)$$

$$= - \sum_{j \in [n] \setminus I} \sum_{i=1}^{b_{I, j}} a_{i, \ell, j} \mathrm{tr}_{F/B}\left( c_j \cdot \lambda_i \right),$$

---

[1]The difference is that an array code with the MDS property need not be itself a linear code over $\Sigma$ (and indeed this may not even make sense if $\Sigma$ is not a field), while a scalar MDS code is by definition linear over $\Sigma$.

and the values $\mathrm{tr}_{F/B}\left(c_j \cdot \lambda_i\right)$ are precisely what is returned by $g_{I,j}(c_j)$. Thus, given the returned information, the repair center can reconstruct the quantities

$$\sum_{i \in I} \mathrm{tr}_{F/B}(c_i \cdot M[i, \ell]) \qquad \forall \ell \in [rt]. \tag{1}$$

Finally, we invoke Property 2 in Definition 4 to show that (1) in fact contain enough information to recover $\{c_i : i \in I\}$. To see this, consider the map $\varphi : F^r \to B^{rt}$ given by

$$\varphi(x) = \begin{pmatrix} \mathrm{tr}_{F/B}(\langle x, M[I, 1]\rangle) \\ \mathrm{tr}_{F/B}(\langle x, M[I, 2]\rangle) \\ \vdots \\ \mathrm{tr}_{F/B}(\langle x, M[I, rt]\rangle) \end{pmatrix}.$$

We will show that $\varphi$ is invertible. To see this, consider the map $\psi : B^{rt} \to F^r$ given by

$$\psi(y) = M[I, :] \cdot y.$$

This map is clearly $B$-linear and Property 2 says that $\psi$ is injective. By counting dimensions (over $B$), $\psi$ is surjective as well. To conclude, we will observe that $\psi$ is the adjoint of $\varphi$, in the sense that for all $y \in B^{rt}$ and for all $x \in F^r$, we have

$$\langle \varphi(x), y \rangle = \mathrm{tr}_{F/B}\left(\langle x, \psi(y)\rangle\right),$$

and hence since $\psi$ is invertible then $\varphi$ is invertible. Formally, we compute

$$\langle \varphi(x), y \rangle = \sum_{j \in [rt]} y_j \cdot \mathrm{tr}_{F/B}(\langle x, M[I, j]\rangle)$$
$$= \mathrm{tr}_{F/B}\left(\sum_{j \in [rt]} y_j \cdot \langle x, M[I, j]\rangle\right)$$
$$= \mathrm{tr}_{F/B}\left(\langle x, M[I, :] \cdot y\rangle\right)$$
$$= \mathrm{tr}_{F/B}\left(\langle x, \psi(y)\rangle\right).$$

Now, we would like to show that $\varphi$ is injective. Let $x \in F^r$ be nonzero. Then there is some $z \in F^r$ so that $\mathrm{tr}_{F/B}(\langle x, z\rangle) \neq 0$. Because $\psi$ is surjective, there is some $y \in B^{rt}$ so that $\psi(y) = z$. But then

$$\langle \varphi(x), y \rangle = \mathrm{tr}_{F/B}(\langle x, \psi(y)\rangle) = \mathrm{tr}_{F/B}(\langle x, z\rangle) \neq 0,$$

and hence $\varphi(x) \neq 0$ as well. This shows that $\varphi$ is injective; again by dimension counting, we see that $\varphi$ is also surjective and hence invertible.

Thus, given $\varphi(x)$, $x$ is the solution to an invertible system of linear equations over the base field $B$. Thus, we may recover $x$ by solving this system of linear equations. To complete the argument, we observe that the quantities (1) in fact give us $\varphi(c_I)$, where we recall that $c_I$ denotes the restriction of $c$ to $I$. Thus, given (1), we may invert $\varphi$ and recover $\{c_i : c \in I\}$, as desired. □

## IV. CENTRALIZED REPAIR SCHEMES FOR RS CODES WITH MULTIPLE FAILURES

In this section we specialize Theorem 1 to Reed-Solomon codes. The *Reed-Solomon Code* $\mathcal{C}$ of dimension $k$ over $F$ with evaluation points $\alpha_1, \ldots, \alpha_n$ is the set

$$\mathcal{C} = \left\{ \begin{pmatrix} f(\alpha_1) \\ f(\alpha_2) \\ \vdots \\ f(\alpha_n) \end{pmatrix} : f \in F[X], \deg(f) < k \right\}.$$

The dual $\mathcal{C}^{\perp}$ of any Reed-Solomon Code is a *generalized Reed-Solomon Code*

$$\left\{ \begin{pmatrix} \lambda_1 \cdot g(\alpha_1) \\ \lambda_2 \cdot g(\alpha_2) \\ \vdots \\ \lambda_n \cdot g(\alpha_n) \end{pmatrix} : g \in F[X], \deg(g) < n - k \right\}, \tag{2}$$

where $\lambda_1, \ldots, \lambda_n \in F$ are constants which depend on the choice of $\alpha_1, \ldots, \alpha_n$. (We refer the reader to [18] for more details).

Below, we give two constructions instantiating Theorem 1 for Reed-Solomon Codes. Our first scheme, discussed in Section IV-A, is much better as $n \to \infty$. However, for small values of $n$, our second scheme, discussed in Section IV-B, is quantitatively better, and so we include it as it may be of more practical interest.

Both of our schemes generalize the construction of [6] and [11] for a single failure. We briefly review this construction below, as we will need it for our constructions. The scheme of [11] was as follows:

*Proposition 1 [11]: Let $n = |F|$, and let $\mathcal{C} \subseteq F^n$ be the Reed-Solomon code of dimension $k = n - n/|B|$, which uses all evaluation points $F = \{\alpha_1, \ldots, \alpha_n\}$. Let $\delta \in F$, and let $\zeta_1, \ldots, \zeta_t$ be a basis for $F$ over $B$. Then the matrix $M = M(\delta, \zeta_1, \ldots, \zeta_t) \in F^{n \times t}$ with*

$$M[j, w] = \frac{\delta \cdot \mathrm{tr}(\zeta_w(\alpha_j - \alpha_i))}{\alpha_j - \alpha_i} \tag{3}$$

*is a repair matrix for index $i$ with bandwith $n - 1$ symbols of $B$.*

To see that this is indeed a valid repair matrix for $i$, observe that the polynomial

$$h_w(X) = \frac{\delta \cdot \mathrm{tr}_{F/B}(\zeta_w(X - \alpha_i))}{X - \alpha_i}$$
$$= \delta \left( \zeta_w + \zeta_w^{|B|}(X - \alpha_i)^{|B|-1} + \cdots \right.$$
$$\left. \cdots + \zeta_w^{|B|^{t-1}}(X - \alpha_i)^{|B|^{t-1}-1} \right)$$

is indeed a polynomial of degree less than $n - k = n/|B| = |B|^{t-1}$, and so the column $M[:, j]$ is an element of $\mathcal{C}^{\perp}$. Moreover, we have $h_w(\alpha_i) = \delta\zeta_w$, and so $M[i, :] = (\delta\zeta_1, \delta\zeta_2, \ldots, \delta\zeta_t)$ is full rank. Finally, for all $j \neq i$,

$$h_w(\alpha_j) \in \frac{\delta}{\alpha_j - \alpha_i} \cdot B,$$

and hence

$$\mathrm{span}_B\left(\mathrm{set}\left(M[j, :]\right)\right) = \mathrm{span}_B\left\{h_w(\alpha_j) : w \in [t]\right\}$$

has dimension 1 over $B$, and so the bandwidth of the repair matrix is $n - 1$.

In [6], it was observed that the trace function above can be replaced with an arbitrary linearized polynomial. More precisely, let $W \subseteq F$ be a subspace of dimension $s$ over $B$. The *subspace polynomial* defined by $W$ is

$$L_W(X) = \prod_{\alpha \in W} (X - \alpha).$$

It is well-known that $L_W$ is a $B$-linear map from $F$ to $F$, of the form

$$L_W(X) = \sum_{j=0}^{s} c_j X^{|B|^j} \tag{4}$$

for coefficients $c_0, \ldots, c_s \in F$. In particular, there is no constant term; $X$ divides $L_W(X)$. Moreover, the coefficient $c_0$ is nonzero, as

$$c_0 = \sum_{\alpha \in W} \prod_{\beta \in W \setminus \{\alpha\}} \beta = \prod_{\beta \in W \setminus \{0\}} \beta \neq 0.$$

Since the kernel of $L_W$ is $W$, the image of $L_W$ is a subspace of dimension $t - s$ over $B$.

With this background in place, we proceed to our constructions.

### A. Main Construction

Our main construction for Reed-Solomon codes generalizes the construction of [6] and [11]. In particular, we will choose $rt$ different low-degree polynomials of a form similar to (3); we follow [6] and replace the trace function with a linearized polynomial. The key is to choose an appropriate modification so that the requirements of Theorem 1 hold. We will see how to do this below, but first we state our main result.

*Theorem 2:* Let $B \subseteq F$ be a subfield, and let $\mathcal{C} \subset F^n$ be a Reed-Solomon code of dimension $k$ over $F$. Choose $1 \leq r \leq n - k$. Then $\mathcal{C}$ admits an exact centralized repair scheme for $r$ failures with bandwidth $b$ at most

$$\min_{r \leq r' \leq n-k} (n - r') \left( t - \left\lfloor \log_{|B|} \left( \frac{n - k + r' - 1}{2r' - 1} \right) \right\rfloor \right). \tag{5}$$

*Remark 3 (Quantitative Interpretation):* Theorem 2 works for any value of $r$ between $1$ and $n - k$, is optimal for $r = 1$, and gives non-trivial bounds for $r \ll n^R$, where $R = k/n$ is the rate of the code. More precisely, using the fact that $t = \log_{|B|}(n)$ and dropping the minimum, we see that the right hand side of (5) is at most

$$(n - r) \cdot \left( \log_{|B|} \left( \frac{2r - 1}{(1 - R) + (r - 1)/n} \right) + 1 \right).$$

When $r \ll n^R$, this is approximately $\log_{|B|} \left( \frac{2r}{1-R} \right)$ subsymbols per surviving node; thus the number of bits that each node has to send increases with both $R$ and $r$, as expected.

*Finally, we see that for some value $r_0 = \Omega(n^R)$, the right hand side of (5) becomes equal to $kt$ and thus trivial. We do not know whether this limit $r = n^R$ is a fundamental limit or an artifact of our approach.*

*The behavior of (5) is shown in Figure 1.*

*Finally, we address the presence of the minimum in (5). The reason for the minimum is that without it, the expression 5*
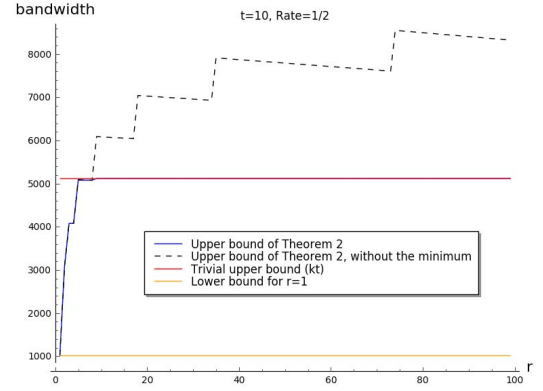


Fig. 2. The bound (5) from Theorem 2, for $t = 10$ and $R = 1/2$, along with the same expression with the minimum ommited. Both bounds become trivial around $r = 10$ when the expression without the minimum crosses the trivial upper bound $kt$. The dashed line (without the minimum) will eventually decrease to meet the trivial bound again at $r = 512$.

*would not be monotonically increasing in $r$, and so the bound given would be obviously weak for some $r$. To illustrate this non-monotonicity, we plot (5) with and without the minimum in Figure 2.*

*Proof of Theorem 2:* Let $s$ be a parameter to be chosen later. Let $W \subseteq F$ be a subspace of dimension $s$ over $B$, and let $L_W$ be the subspace polynomial

$$L_W(X) = \prod_{\alpha \in W} (X - \alpha).$$

As above, let $I \subset [n]$ be the set of failed nodes, so $|I| = r$. Let $Z = \{\zeta_1, \ldots, \zeta_t\}$ be a basis for $F$ over $B$. Define

$$F_I(X) = \prod_{i \in I} (X - \alpha_i),$$

and for $p \in [r]$, define

$$P_{\zeta, p}(X) = \frac{L_W \left( \zeta \cdot F_I(X) \cdot X^{p-1} \right)}{F_I(X)}.$$

This is a generalization of the construction in [6] and [11] to multiple failures: if $r = 1$ and $\alpha_1$ is the only failed node, then the above is

$$P_{\zeta, 1}(X) = \frac{L_W \left( \zeta \cdot (X - \alpha_1) \right)}{(X - \alpha_1)},$$

exactly as in [6].

Now, we construct our repair matrix $M \in F^{n \times rt}$ as follows. We index the rows of $M$ by $j \in [n]$, and the columns by pairs $(\zeta, i)$ for $\zeta \in Z$ and $p \in [r]$. Then we set

$$M[j, (\zeta, p)] := \lambda_j P_{\zeta, p}(\alpha_j),$$

where $\lambda_j$ is as in (2).

We must show that $M$ satisfies the conditions of Theorem 1. In particular, we will show that $M$ has the form shown in Figure 3, which as we will see below implies that Theorem 1 applies. First, we compute the bandwidth of $M$.

Fig. 3. The matrix $M$ constructed in the proof of Theorem 2 for $r = 3$, where we write $I = \{i_1, i_2, i_3\}$, and $[n] \setminus I = \{j_1, \ldots, j_{n-3}\}$. Here, $c_0$ is the coefficient from (4), and $\vec{\zeta} = (\zeta_1, \ldots, \zeta_r)$ is a vector consisting of the elements of the basis $Z$.

*Claim 1: The bandwidth of $M$ is at most $(n - r)(t - s)$.*

*Proof:* For $j \notin I$, the set of symbols $\mathbf{set}(M[j, :])$ that appears in row $j$ of $M$ is precisely

$$\mathbf{set}(M[j, :])$$
$$= \left\{ \lambda_j \cdot P_{\zeta, p}(\alpha_j) : \zeta \in Z, p \in [r] \right\}$$
$$= \left\{ \lambda_j \cdot \frac{L_W(\zeta \cdot F_I(\alpha_j) \cdot \alpha_j^{p-1})}{F_I(\alpha_j)} : \zeta \in Z, p \in [r] \right\}$$
$$\subseteq \frac{\lambda_j}{F_I(\alpha_j)} \mathrm{Im}(L_W),$$

which is a subspace of dimension $t - s$. Since there are $n - r$ such $j$'s by definition the bandwidth of $M$ is bounded by

$$(n - r)(t - s),$$

as desired. □

We next show that $M[I, :]$ has full rank, in the sense that $M[I, :]y \neq 0$ for all nonzero $y \in B^{rt}$.

*Claim 2: $M[I, :]$ has full rank.*

*Proof:* Using (4), we write

$$P_{\zeta, p}(X) = \frac{P_W(\zeta \cdot F_I(X) \cdot X^{p-1})}{F_I(X)}$$
$$= \frac{\sum_{m=0}^{s} c_m (\zeta \cdot F_I(X) \cdot X^{p-1})^{|B|^m}}{F_I(X)}$$
$$= \sum_{m=0}^{s} c_m \cdot \zeta^{|B|^m} \cdot F_I(X)^{|B|^m - 1} \cdot X^{(p-1) \cdot |B|^m},$$

and hence for $i \in I$, all terms vanish except the $m = 0$ term, and we have

$$P_{\zeta, p}(\alpha_i) = c_0 \cdot \zeta \cdot \alpha_i^{p-1}.$$

To show that $M[I, :]$ has full rank, by definition we must show that for all nonzero $y \in B^{rt}$, $M[I, :]y \neq 0$. So let $y \in B^{rt} \setminus \{0\}$, and write $y = (y^{(1)}, y^{(2)}, \ldots, y^{(r)})$, where each $y^{(i)} \in B^t$. By the above characterization, we have

$$\langle M[i, :], y \rangle = \lambda_i \cdot \sum_{p=1}^{r} c_0 \alpha_i^{p-1} \langle \vec{\zeta}, y^{(p)} \rangle,$$

where $\vec{\zeta} = (\zeta_1, \ldots, \zeta_t) \in F^t$. Since the $\zeta_i$ form a basis of $F$ over $B$ and since $y \neq 0$, at least one of the coefficients $\langle \vec{\zeta}, y^{(p)} \rangle$ is nonzero. Thus, $\langle M[i, :], y \rangle = \lambda_i \cdot g(\alpha_i)$, where $g(X) = \sum_{p=1}^{r} c_0 \langle \vec{\zeta}, x^{(p)} \rangle X^{p-1}$ is some nonzero polynomial of degree at most $r - 1$. This means that $M[I, :] \cdot y \neq 0$, or else this polynomial $g(X)$ would have at least $r$ roots, one for each $i \in I$. □

Finally, we must choose $s$. We need for the columns of $M$ to be elements of $\mathcal{C}^\perp$, which by (2) is the same as requiring the polynomials $P_{\zeta, p}(X)$ to have degree strictly less than $n - k$. That is, we require

$$|B|^s \cdot (2r - 1) - r \leq n - k - 1,$$

which is satisfied by the choice of

$$s = \left\lfloor \log_{|B|} \left( \frac{n - k + r - 1}{2r - 1} \right) \right\rfloor.$$

Plugging this choice of $s$ into the bandwidth bound of Claim 1 coupled with Observation 1 below finishes the proof of Theorem 2.

*Observation 1: Given $r' > r$ if $\mathcal{C}$ admits an exact centralized repair scheme for $r'$ failures with bandwidth $b$ then it also admits an exact centralized repair scheme for $r$ failures with bandwidth $b$.*

□

### B. Alternate Construction for Small Block Sizes

In this section we give another generalization of the one-failure scheme from Proposition 1. This scheme is worse asymptotically, but has better performance for small $n$, so it may be of practical value. The basic idea directly generalizes (and improves upon) that of [4]; the multiple-repair matrix $M$ is formed by concatenating $r$ separate repair matrices $M_1, \ldots, M_r$ from Proposition 1. In fact, Theorem 1 immediately implies that this is a nontrivial repair scheme, but we can do better by choosing multipliers $\delta_1, \ldots, \delta_r \in F$, and using the repair matrix formed by concatenating $\delta_1 M_1, \ldots, \delta_r M_r$. We will show how to choose the multipliers $\delta_1, \ldots, \delta_r$ so that (a) the rank of $M[I, :]$ is

not affected, but (b) the rank of the other rows $M[j, :]$ for $j \notin I$ is reduced.

We will prove the following theorem.

*Theorem 3: Let $n = |F|$, and let $B \subseteq F$ be a subfield so that $F$ has degree $t$ over $B$. Let $\mathcal{C} \subseteq F^n$ be the Reed-Solomon code of dimension $k = n - n/|B|$ with evaluation points $F = \{\alpha_1, \ldots, \alpha_n\}$. Choose $r \geq 2$ so that*

$$t > \binom{r}{2} + \log_{|B|}(r \cdot (r + \binom{r}{2}(|B| - 1)) \cdot (|B| - 1) + 1).$$

*Then for all $I \subset [n]$ of size $r$, there is a matrix $M \in F^{n \times rt}$ so that $M$ is a multiple repair matrix for $I$, with bandwidth*

$$b \leq (n - r) \cdot r - (|B| - 1)\binom{r}{2}.$$

*Remark 4 (Bandwidth Guarantee): Observe that the naive scheme (which contacts any $k$ remaining nodes) has bandwidth $kt$. Thus, the guarantee that $b \leq (n-r) \cdot r - \binom{r}{2}(|B|-1)$ improves this. Moreover, when $r = 1$, this collapses to the result of [11] that $b \leq n - 1$. For $r = 2, 3$, this improves over the result $b \leq (n - r) \cdot r$ of [5]. A comparison of Theorem 3 with Theorem 2 is shown in Figure 1.*

*Remark 5 (Large $r$): Notice that Theorem 3 allows $r$ to grow slightly with $n$. However, since we have $t = \log_{|B|}(n)$ since $n = |F|$, the requirement on $t$ implies that, for the result to hold, we need*

$$\log_{|B|}\left(\frac{n}{r \cdot (r + \binom{r}{2}(|B| - 1)) \cdot (|B| - 1) + 1}\right) > \binom{r}{2}$$

*or $r \lesssim \sqrt{\log(n)}$.*

The rest of this section is devoted to the proof of Theorem 3. We begin with a lemma which shows that, if the multipliers $\delta_1, \ldots, \delta_r$ are picked appropriately, then the matrix formed by concatenating $r$ copies of the single-repair matrices of Proposition 1 form a good multiple-repair matrix.

*Lemma 1: Let $n = |F|$, and let $B \subseteq F$ be a subfield so that $F$ has degree $t$ over $B$. Let $\mathcal{C} \subseteq F^n$ be the Reed-Solomon code of dimension $k = n - n/|B|$ with evaluation points $F = \{\alpha_1, \ldots, \alpha_n\}$. Suppose[2] that $I = [r]$ and the evaluation points corresponding to the failed nodes are $\{\alpha_1, \ldots, \alpha_r\}$. Let $\zeta_1, \ldots, \zeta_t$ be any basis for $F$ over $B$. Choose $\delta_1, \ldots, \delta_r$ so that for all $j = 1, \ldots, r$, for all $\ell > j$ and for all $s > j$, we have*

$$\mathrm{tr}_{F/B}\left(\frac{\delta_\ell}{\delta_j} \cdot \frac{\alpha_s - \alpha_j}{\alpha_j - \alpha_\ell}\right) = 0. \tag{6}$$

*Let*

$$M_i := M\left(\delta_i, \frac{\zeta_1}{\delta_i}, \frac{\zeta_2}{\delta_i}, \ldots, \frac{\zeta_t}{\delta_i}\right)$$

*be as in Proposition 1. Then the matrix $M \in F^{n \times rt}$ given by*

$$M = [M_1 | M_2 | \cdots | M_r] \tag{7}$$

*is a multiple-repair matrix for $I$.*

Notice that Lemma 1 does not make any claims about the bandwidth of this scheme; we will show below how to choose

[2]Since we will never use anything about the ordering of the evaluation points, this assumption is without loss of generality.

the $\delta_i$ so that (6) holds, and so that the bandwidth is also small. Because the columns of $M$ are columns of the $M_i$ and we have already established that these are dual codewords, the only thing left to prove is that Property 2 of Definition 4 holds; that is, that the $r \times rt$ matrix $M[I, :]$ has a trivial right kernel over $B$. We will prove this in Section IV-B1, but first, we will show how to use Lemma 1 in order to prove Theorem 3.

Before we begin the proof of Theorem 3, we include an overview. The idea is to choose the parameters $\delta_1, \ldots, \delta_r$ successively so that Lemma 1 applies, and keep track of the bandwidth of the resulting repair matrix.

Note that this approach—even without keeping track of the bandwidth—would immediately imply that $M$ is a multiple-repair matrix for $I$ with bandwidth at most $(n - r) \cdot r$; indeed, for all $i \in [n] \setminus I$,

$$\dim_B \{\mathrm{span}_B \{\mathbf{set}\,(M[i, :])\}\} \leq r,$$

because $\mathbf{set}\,(M[i, :]) = \bigcup_{\ell=1}^r \mathbf{set}\,(M_\ell[i, :])$, and for each $\ell$ we have

$$\dim_B \{\mathrm{span}_B \{\mathbf{set}\,(M_\ell[i, :])\}\} \leq 1.$$

Then, Theorem 1 implies that this gives a repair scheme for $I$ with bandwidth $b \leq (n - r)r$.

The approach above would recover the results of [5] for $r = 2, 3$ and would generalize them to all $r$. However, in fact this calculation may be wasteful, and the idea of Theorem 3 is that by choosing the $\delta_i$ carefully we can improve the bandwidth. More precisely, we will try to choose $\delta_i$ so that the spans $\mathrm{span}_B \{\mathbf{set}\,(M[i, :])\}$ collide, meaning that the dimension of the union is less than the sum of the dimensions.

We first show, in Claim 3, that for any choices of $\delta_1, \ldots, \delta_{\ell-1}$, there are many choices of $\delta_\ell$ that will satisfy (6). Thus, at each step $\ell$ we have many choices of $\delta_\ell$ so that Lemma 1 will apply. Next, we need to argue that we can also choose the $\delta_\ell$ so that there are lots of collisions, in the sense described above. We do this by induction, choosing the $\delta_\ell$ sequentially and maintaining the inductive hypothesis that the number of positions $i \in [n] \setminus I$ that are already "collided on" is precisely $\binom{\ell}{2}(|B| - 1)$; that is, there are that many indices $i$ so that $M[i, j]$ and $M[i, m]$ are linearly dependent over $B$, for some $j, m \leq \ell$. Then at the end of the induction, we will use a counting argument to convert the number of collisions into a bound on the bandwidth.

*Proof of Theorem 3:* Suppose without loss of generality that $I = [r]$ and the evaluation points corresponding to the set of failed nodes are $\{\alpha_1, \ldots, \alpha_r\}$. Let $\zeta_1, \ldots, \zeta_t$ be any basis of $F$ over $B$. We will choose the parameters $\delta_1, \ldots, \delta_r$ successively so that Lemma 1 applies, and keep track of the bandwidth of the resulting repair matrix. We begin with a claim that will help us make sure that Lemma 1 applies.

*Claim 3: Let $\ell < r$ and suppose that $\delta_1, \ldots, \delta_{\ell-1}$ have been chosen. Then for at least $|B|^{t - (\ell-1) \cdot r + \ell(\ell-1)/2} - 1$ choices of $\delta \in F^*$, setting $\delta_\ell \leftarrow \delta$ satisfies*

$$\mathrm{tr}_{F/B}\left(\frac{\delta}{\delta_j} \cdot \frac{\alpha_s - \alpha_j}{\alpha_j - \alpha_\ell}\right) = 0$$

*for all $j < \ell$ and all $s > j$.*

*Proof:* For each $j < \ell$ and $s > j$, the above gives a linear requirement on $\delta_\ell$. There are at most

$$\sum_{j=1}^{\ell-1} (r - j) = (\ell - 1) \cdot r - \frac{\ell(\ell - 1)}{2}$$

such pairs $(j, s)$, so there are that many linear constraints. Since $F$ is a vector space over $B$ of dimension $t$, this proves the claim. $\qquad\square$

We briefly recall some algebra. For $\gamma \in F^*$, the *(multiplicative) coset* $\gamma \cdot B^*$ is the set

$$\gamma \cdot B^* = \left\{ \gamma \cdot b : b \in B^* \right\}.$$

We say that $\gamma \equiv_{B^*} \gamma'$ if $\gamma \in \gamma' B^*$, and it is not hard to see that $\equiv_{B^*}$ is an equivalence relation that partitions $F^*$ into $|F^*|/|B^*|$ cosets of size $|B^*|$. The group of all such cosets form the *quotient group* $F^*/B^*$. The following observations follow directly from these definitions, as well as the definition of the matrix $M_\ell$.

*Observation 2:* Suppose that $set(M[i, :]) \subseteq \gamma_1 B^* \cup \gamma_2 B^* \cup \cdots \cup \gamma_c B^*$ for $c$ different cosets $\gamma_i B^*$. Then

$$\dim_B \left\{ \mathrm{span}_B \left\{ set(M[i, :]) \right\} \right\} \leq c.$$

*Observation 3:* We have

$$\mathrm{span}_B \left\{ set(M_\ell[i, :]) \right\} \subseteq \left( \frac{\delta_\ell}{\alpha_i - \alpha_\ell} \right) B^* \cup \{0\}.$$

Say that $(j, \ell)$ **collide** at $i$ if

$$\frac{\delta_j}{\alpha_i - \alpha_j} B^* = \frac{\delta_\ell}{\alpha_i - \alpha_\ell} B^*.$$

That is, if $(j, \ell)$ collide at $i$, then $\dim_B \left\{ \mathrm{span}_B \left\{ M_\ell[i, j], M_\ell[i, \ell] \right\} \right\} = 1$ rather than 2 and this decreases the bandwidth of downloads from node $i$ by 1. Hence collisions are good for us and our aim is to choose $\delta_1, \delta_2, \ldots, \delta_r$ so as to maximize the number of useful collisions.

Notice that $(j, \ell)$ collide at $i$ if and only if

$$\delta_\ell \in \left( \frac{\alpha_i - \alpha_\ell}{\alpha_i - \alpha_j} \right) \delta_j B^*.$$

Define

$$S_\ell := \{ i \in [n] \setminus I : \exists j < m \leq \ell \text{ s.t. } (j, m) \text{ collide at } i \}.$$

Consider choosing $\delta_1, \ldots, \delta_r$ one at a time. Choose $\delta_1 = 1$, and we will proceed by induction on $\ell$, assuming that we have chosen $\delta_1, \ldots, \delta_\ell$, with the inductive hypothesis that

$$|S_\ell| = \binom{\ell}{2} \cdot (|B| - 1). \tag{8}$$

Notice that the base case when $\ell = 1$ is trivially satisfied with $\delta_1 = 1$, because $S_1 = \emptyset$. Now suppose that the inductive hypothesis (8) holds, and consider choosing $\delta_{\ell+1}$. For $j < \ell + 1$, define

$$T_j^{\ell+1} := \{ i \in [n] : (\ell + 1, j) \text{ collide at } i \}.$$

Notice that $T_j^{\ell+1}$ depends on the choice of $\delta_{\ell+1}$, as the definition of colliding depends on $\delta_{\ell+1}$.

*Claim 4:* Fix $j < \ell + 1$. For all but at most $r + \binom{\ell}{2} \cdot (|B| - 1)$ cosets $\gamma B^*$, the following holds: If $\delta_{\ell+1} \in \gamma B^*$, then

(a) $|T_j^{\ell+1}| = |B| - 1$, and

(b) $T_j^{\ell+1} \cap (I \cup S_\ell) = \emptyset$.

*Proof:* Say $\gamma$ is **good** if

$$\left( \frac{\alpha_i - \alpha_{\ell+1}}{\alpha_i - \alpha_j} \right) \delta_j \notin \gamma B^* \qquad \forall i \in I \cup S_\ell.$$

That is, $\gamma$ is good if and only if choosing $\delta_{\ell+1} \in \gamma B^*$ would mean that $(\ell + 1, j)$ do *not* collide at any $i \in I \cup S_\ell$. We first claim that there are at most $|I \cup S_\ell|$ values of $\gamma \in F$ that are not good. This follows from the fact that the map $h : z \mapsto \left( \frac{z - \alpha_{\ell+1}}{z - \alpha_j} \right) \delta_j$ is a bijection from $F$ to $F$. Given this, and the fact that $|I \cup S_\ell| = |I| + |S_\ell| = r + \binom{\ell}{2} \cdot (|B| - 1)$, we conclude that there are at most $r + \binom{\ell}{2} \cdot (|B| - 1)$ choices of $\gamma$ that are not good.

Suppose that $\gamma$ is good, and consider any choice of $\delta_{\ell+1} \in \gamma B^*$. Since $h$ as defined above is a bijection, there are $|B^*| = |B| - 1$ elements $\alpha_i \in F$ so that

$$\left( \frac{\alpha_i - \alpha_{\ell+1}}{\alpha_i - \alpha_j} \right) \delta_j \in \gamma B^* = \delta_{\ell+1} B^*,$$

and so $(\ell + 1, j)$ collide at $i$, and by definition $T_j^{\ell+1}$ is this set. This establishes (a). Now (b) follows from the definition of good. This completes the proof of Claim 4. $\qquad\square$

Claim 4 is for a fixed $j$, and summing over all $j \leq \ell$, it implies that for all but at most $\ell \cdot (r + \binom{\ell}{2} (|B| - 1)) \cdot (|B| - 1)$ choices of $\delta_{\ell+1}$, we have that for *all* $j = 1, \ldots, \ell$, the set $T_j^{\ell+1}$ satisfies the conclusions (a) and (b) of Claim 4. We next claim that for such a choice of $\delta_{\ell+1}$, the sets $T_j^{\ell+1}$ and $T_m^{\ell+1}$ are disjoint for $j < m \leq l$.

*Claim 5:* Suppose that $\delta_{\ell+1}$ is chosen so that the conclusion (b) of Claim 4 holds for all $j = 1, \ldots, \ell$. Then for all $j < m \leq \ell$, $T_j^{\ell+1} \cap T_m^{\ell+1} = \emptyset$.

*Proof:* Suppose towards a contradiction that $i \in T_j^{\ell+1} \cap T_m^{\ell+1}$. Then by the definition of $T_j^{\ell+1}$ and $T_m^{\ell+1}$, $(\ell + 1, j)$ and $(\ell + 1, m)$ collide at $i$, which means that

$$\frac{\delta_j}{\alpha_i - \alpha_j} B^* = \frac{\delta_{\ell+1}}{\alpha_i - \alpha_{\ell+1}} B^* = \frac{\delta_m}{\alpha_i - \alpha_m} B^*.$$

This implies that $j$ and $m$ collide at $i$. However, since $j, m \leq \ell$, this implies that $i \in S_\ell$. But this contradicts the conclusion (b) of Claim 4, which states that $S_\ell \cap T_j^{\ell+1} = \emptyset$. $\qquad\square$

Now we finish the proof of Theorem 3. Suppose that $\delta_{\ell+1}$ is chosen so that the conclusions (a) and (b) of Claim 4 as well as the conditions of (6) hold for $j = 1, \ldots, \ell$. Because the conditions of (6) hold, Lemma 1 holds, and hence our construction is indeed a multiple repair matrix.

By definition we have

$$S_{\ell+1} = S_\ell \cup T_1^{\ell+1} \cup T_2^{\ell+1} \cup \cdots \cup T_\ell^{\ell+1},$$

and by conclusion (b) of Claim 4 and by Claim 5, each of these sets of disjoint. By induction, $|S_\ell| = \binom{\ell}{2}(|B| - 1)$ and

by the conclusion (a) of Claim 4 we have $|T_j^{\ell+1}| = |B| - 1$ for all $j \le \ell$. Thus,

$$
\begin{aligned}
|S_{\ell+1}| &= |S_\ell| + \sum_{j=1}^{\ell} |T_j^{\ell+1}| \\
&= \binom{\ell}{2}(|B| - 1) + \ell(|B| - 1) \\
&= \binom{\ell+1}{2}(|B| - 1).
\end{aligned}
$$

This establishes the inductive hypothesis for $\ell + 1$.

Now this means that if for all $\ell < r$, we can choose a $\delta_{\ell+1}$ so that the conclusions (a) and (b) of Claim 4 as well as the conditions of (6) hold, then by choosing $\delta_1, \ldots, \delta_r$ in this way, we can conclude by induction that

$$
|S_r| = \binom{r}{2}(|B| - 1).
$$

Since $S_r$ is the set of rows $i$ that have any collisions, each $i \in S_r$ contributes at most $r - 1$ to the bandwidth, while $i \in [n] \setminus (I \cup S_r)$ may contribute $r$. Thus, with this choice of $\delta_1, \ldots, \delta_r$, the bandwidth of the resulting scheme is at most

$$
\binom{r}{2}(|B| - 1)(r - 1) + \left( n - r - \binom{r}{2}(|B| - 1) \right) \cdot r
$$
$$
= (n - r) \cdot r - \binom{r}{2}(|B| - 1).
$$

To ensure that for all $\ell < r$ we can indeed choose a $\delta_{\ell+1}$ so that the conclusions (a) and (b) of Claim 4 as well as the conditions of (6) hold, we need, also using Claim 3, that

$$
|B|^{(t - \binom{r}{2})} - 1 > r \cdot \left( r + \binom{r}{2}(|B| - 1) \right) \cdot (|B| - 1)
$$

or equivalently that $t$ is larger than

$$
\binom{r}{2} + \log_{|B|} \left( r \cdot \left( r + \binom{r}{2}(|B| - 1) \right) \cdot (|B| - 1) + 1 \right),
$$

which is exactly the condition in the statement of the Theorem 3. This completes the proof of Theorem 3. $\square$

*1) Proof of Lemma 1:* In this Section, we prove Lemma 1. Because we have made no claims about the bandwidth, we only need to show that the sub-matrix $M[I, :]$ has full rank, in the sense that for all nonzero $y \in B^{rt}$, $M[I, :]y \ne 0$. To save on notation, for the rest of this proof, let $A \in F^{r \times rt}$ denote the matrix $M[I, :]$.

As in the proof of Theorem 1, it suffices to show that the $B$-linear map $\varphi_A : F^r \to B^{rt}$ given by

$$
\varphi_A(x) = \begin{pmatrix} \mathrm{tr}_{F/B}(\langle x, A[:, 1]\rangle) \\ \mathrm{tr}_{F/B}(\langle x, A[:, 2]\rangle) \\ \vdots \\ \mathrm{tr}_{F/B}(\langle x, A[:, rt]\rangle) \end{pmatrix}
$$

is invertible, where $\mathrm{tr}_{F/B}$ is applied coordinate-wise. Our proof basically follows from analyzing the $LU$-decomposition of this matrix $A$. That is, we will give an algorithm, consisting of

row operations, which preserve the invertibility of $\varphi_M$ and after which $M$ will become a matrix of the form

$$
M' = \begin{bmatrix} \vec{\zeta}^T & * & * & \cdots & * \\ \vec{0} & \vec{\zeta}^T & * & \cdots & * \\ \vec{0} & \vec{0} & \vec{\zeta}^T & \cdots & * \\ & & \vdots & & \\ \vec{0} & \vec{0} & \vec{0} & \cdots & \vec{\zeta}^T \end{bmatrix}, \tag{9}
$$

where $\vec{0} \in F^t$ denotes the vector of $t$ zeros and $\vec{\zeta} = (\zeta_1, \zeta_2, \ldots, \zeta_t)$. Recall that $\zeta_1, \ldots, \zeta_t$ is the basis chosen in the statement of the lemma. The matrix $M'$ as in (9) clearly has the desired property ($\varphi_{M'}$ is invertible), and so this will finish the proof.

*Remark 6 (Picking a Basis): In what follows, we just go through the LU decomposition of the matrix A, and show that the assumption (6) implies that the result has the form (9). Unfortunately, this familiar argument may seem less familiar because we have not picked a basis for F over B, and are instead working with the trace functional. If the reader prefers, they may imagine picking a basis for F over B, and working with square $rt \times rt$ matrices over B. However, for this exposition we choose to leave things as they are to avoid the additional level of subscripts that picking a basis would require.*

Given $\gamma \in F$, define the map $f_\gamma : F \to F$ by

$$
f_\gamma(x) = \gamma \cdot \mathrm{tr}_{F/B}\left( \frac{x}{\gamma} \right).
$$

Then extend $f_\gamma$ to $f_\gamma : F^{rt} \to F^{rt}$ by acting coordinate-wise.

*Observation 4:* Fix any $i, j \in [r]$. Let $A \in F^{r \times rt}$, and suppose that $A'$ is obtained from $A$ by adding $f_\gamma(A[i, :])$ to $A[j, :]$, so

$$
A'[j, :] = A[j, :] + f_\gamma(A[i, :])
$$

and for $\ell \ne j$, we have $A'[\ell, :] = A[\ell, :]$. Then $\{\varphi_A(x) : x \in F^r\} = \{\varphi_{A'}(x) : x \in F^r\}$. In particular, $\varphi_A$ is invertible if and only if $\varphi_{A'}$ is invertible. (As before, above $\mathrm{tr}_{F/B}$ is applied coordinatewise).

*Proof:* Given $x \in F^r$, consider $x'$ given by $(x')_i = x_i + \mathrm{tr}_{F/B}(\gamma x_j)/\gamma$. Then

$$
\varphi_{A'}(x) = \begin{pmatrix} \mathrm{tr}_{F/B}(\langle x, A[:, 1]\rangle) + x_j \cdot \gamma \cdot \mathrm{tr}_{F/B}(\frac{A[i,1]}{\gamma})) \\ \vdots \\ \mathrm{tr}_{F/B}(\langle x, A[:, rt]\rangle) + x_j \cdot \gamma \cdot \mathrm{tr}_{F/B}(\frac{A[i,rt]}{\gamma})) \end{pmatrix}
$$
$$
= \varphi_A(x')
$$

Since the map from $x$ to $x'$ is injective as well as surjective, we have $\{\varphi_A(x) : x \in F^r\} = \{\varphi_{A'}(x) : x \in F^r\}$. $\square$

Now consider Algorithm 1.

Because of Observation 4, we see that $\varphi_{A^{(r)}}$ is invertible if and only if $\varphi_A$ is invertible. Moreover, we claim that, if (6) is met, then $A^{(r)}$ has the form (9). To see this we proceed

---

**Algorithm 1** Algorithm **LU(A)**

---

$A^{(0)} \leftarrow A$
for $j = 1, \ldots, r - 1$:
    for $s = j + 1, \ldots, r$:
        $A^{(j)}[s, :] \leftarrow$
        $A^{(j-1)}[s, :] + f_{\delta_j/(\alpha_s - \alpha_j)}(A^{(j-1)}[j, :])$
    **end for**
**end for**
**return** $A^{(r)}$.

---

by induction. Write $A = [A_1 | A_2 | \cdots | A_r]$, where $A_j \in F^{r \times t}$, and similarly write

$$A^{(j)} = [A_1^{(j)} | A_2^{(j)} | \cdots | A_r^{(j)}].$$

Then the inductive hypothesis is that

1) For all $i \leq j$, $A_i^{(j)}$ is equal to $A_i$ on the first $i$ rows.
2) For all $i \leq j$, $A_i^{(j)}$ is equal to 0 on rows $i+1, i+2, \ldots, r$. That is, the first $j$ blocks of $A^{(j)}$ have zeros in the form of (9), and all nonzero entries are the same as in $A$. The base case is immediate for $j = 0$, with the notational assumption that any statement about $M[\ell, c]$ for $c \leq 0$ is vacuously true.

Now assuming that this holds for $j - 1$, we establish it for $j$. First notice that, because of the inductive hypothesis, the first $j - 1$ blocks do not change. For block $j$, and a row $s > j$, we update

$$A_j^{(j)}[s, :] \leftarrow A_j^{(j-1)}[s, :] + f_{\delta_j/(\alpha_s - \alpha_j)}(A_j^{(j-1)}[j, :])$$
$$= A_j[s, :] + f_{\delta_j/(\alpha_s - \alpha_j)}(A_j[j, :])$$

again using the inductive hypothesis. By construction, for all $w \in [t]$,

$$A_j[s, w] = \frac{\delta_j}{\alpha_s - \alpha_j} \cdot \mathrm{tr}_{F/B}(\zeta_w \cdot (\alpha_s - \alpha_j)/\delta_j).$$

Then the update is

$$f_{\delta_j/(\alpha_s - \alpha_j)}(A_j[j, w])$$
$$= \frac{\delta_j}{\alpha_s - \alpha_j} \cdot \mathrm{tr}_{F/B}\left(\frac{\alpha_s - \alpha_j}{\delta_j} \cdot A_j[j, w]\right)$$
$$= \frac{\delta_j}{\alpha_s - \alpha_j} \cdot \mathrm{tr}_{F/B}(\zeta_w \cdot (\alpha_s - \alpha_j)/\delta_j)$$
$$= A_j[s, w],$$

and so $A_j[s, w] - f_{\delta_j/(\alpha_s - \alpha_j)}(A_j[j, w]) = 0$ for all $s > j$. Thus, this operation indeed zeros out all but the first $j$ rows of $A_j$.

Next, consider a block $\ell > j$ and a row $s > j$. We need to show that $A_\ell^{(j-1)}[s, w]$ does not change in the $j$'th iteration, namely that the update is zero. Now we have, by induction and by definition, respectively, that

$$A_\ell^{(j-1)}[j, w] = A_\ell[j, w] = \frac{\delta_\ell}{\alpha_j - \alpha_\ell} \cdot \mathrm{tr}_{F/B}(\zeta_w \cdot (\alpha_j - \alpha_\ell)/\delta_\ell).$$

A computation similar to the one above establishes that $f_{\delta_j/(\alpha_s - \alpha_j)}(A_\ell[j, w])$ is equal to

$$\frac{\delta_j}{\alpha_s - \alpha_j} \mathrm{tr}_{F/B}\left(\frac{\zeta_w(\alpha_j - \alpha_\ell)}{\delta_\ell}\right) \mathrm{tr}_{F/B}\left(\frac{\delta_\ell}{\delta_j} \cdot \frac{\alpha_s - \alpha_j}{\alpha_j - \alpha_\ell}\right),$$

which is 0 by (6).

This establishes the other part of the inductive hypothesis (that all other entries remain the same). The only thing remaining is to show that $A^{(r)}$ has diagonal entries given by $\zeta$ as in 9. To see this, observe that by our inductive hypothesis (which we have proved), the diagonal of $A^{(r)}$ is identical to the diagonal of $A^{(0)}$, which has the desired property by construction. This completes the proof.

## V. CONCLUSION

In this work, we have extended the framework of [11] to handle multiple failures, and instantiated this framework in two ways to give improved results for Reed-Solomon codes with multiple failures. However, several open problems remain. We highlight a few promising directions below.

1) We have no reason to believe that the bound in Theorem 2 is asymptotically tight. We leave it as an open question to either obtain an improved construction or a matching lower bound. As a concrete question, we may ask if it is possible to repair full-length Reed-Solomon codes from $r = \omega(n^R)$ failures with nontrivial bandwidth.
2) Our work is restricted to the centralized model for repair of multiple nodes. On the other hand, the work of [5] obtains results for Reed-Solomon codes for $r = 2, 3$ in other models where the communication between the nodes is taken into account when measuring the bandwidth; our framework does not apply there. Could our techniques be adapted to apply to this model as well?
3) Finally, we have instantiated our framework for full-length RS codes, but it may be interesting in other parameter regimes as well. Recently, Ye and Barg have given a construction of a Reed-Solomon code which achieves the cut-set bound for multiple failures [26]; however, the subpacketization is (necessarily) extremely large. Can we instantiate our framework to obtain RS codes (or other scalar MDS codes) with repair bandwidth that improves over our scheme, but which still have small (sublinear in $n$) subpacketization?

## REFERENCES

[1] V. R. Cadambe, S. A. Jafar, H. Maleki, K. Ramchandran, and C. Suh, "Asymptotic interference alignment for optimal repair of MDS codes in distributed storage," *IEEE Trans. Inf. Theory*, vol. 59, no. 5, pp. 2974–2987, May 2013.
[2] A. Chowdhury and A. Vardy, "Improved schemes for asymptotically optimal repair of MDS codes," in *Proc. 55th Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Oct. 2017, pp. 950–957.
[3] H. Dau, I. Duursma, and H. Chu. (2018). "On the I/O costs of some repair schemes for full-length Reed-Solomon codes." [Online]. Available: https://arxiv.org/abs/1801.05101
[4] H. Dau, I. Duursma, H. M. Kiah, and O. Milenkovic. (2017). "Repairing Reed-Solomon codes with two erasures." [Online]. Available: https://arxiv.org/abs/1701.07118
[5] H. Dau, I. Duursma, H. M. Kiah, and O. Milenkovic, "Repairing Reed-Solomon codes with multiple erasures," *IEEE Trans. Inf. Theory*, vol. 64, no. 10, pp. 6567–6582, Oct. 2018.
[6] H. Dau and O. Milenkovic. (2017). "Optimal repair schemes for some families of full-length Reed-Solomon codes." [Online]. Available: https://arxiv.org/abs/1701.04120

[7] A. G. Dimakis, P. B. Godfrey, Y. Wu, M. J. Wainwright, and K. Ramchandran, "Network coding for distributed storage systems," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4539–4551, Sep. 2010.

[8] A. G. Dimakis, K. Ramchandran, Y. Wu, and C. Suh, "A survey on network codes for distributed storage," *Proc. IEEE*, vol. 99, no. 3, pp. 476–489, Mar. 2011.

[9] I. Duursma and H. Dau, "Low bandwidth repair of the RS(10,4) Reed-Solomon code," in *Proc. Inf. Theory Appl. Workshop (ITA)*, Feb. 2017, pp. 1–10.

[10] J. B. Fraleigh, *A First Course in Abstract Algebra*. Bengaluru, India: Pearson, 2003.

[11] V. Guruswami and M. Wootters, "Repairing Reed-Solomon codes," *IEEE Trans. Inf. Theory*, vol. 63, no. 9, pp. 5684–5698, Sep. 2017.

[12] A.-M. Kermarrec, N. L. Scouarnec, and G. Straub, "Repairing multiple failures with coordinated and adaptive regenerating codes," in *Proc. Int. Symp. Netw. Coding (NetCod)*, Jul. 2011, pp. 1–6.

[13] R. B. Kiran, K. Tati, Y.-C. Cheng, S. Savage, and G. M. Voelker, "Total recall: System support for automated availability management," in *Proc. NSDI*, 2004, p. 25.

[14] J. Li and B. Li, "Cooperative repair with minimum-storage regenerating codes for distributed storage," in *Proc. IEEE INFOCOM*, Apr. 2014, pp. 316–324.

[15] R. Li, J. Lin, and P. P. C. Lee, "Enabling concurrent failure recovery for regenerating-coding-based storage systems: From theory to practice," *IEEE Trans. Comput.*, vol. 64, no. 7, pp. 1898–1911, Jul. 2015.

[16] W. Li, Z. Wang, and H. Jafarkhani, "A tradeoff between the sub-packetization size and the repair bandwidth for Reed-Solomon code," in *Proc. 55th Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Oct. 2017, pp. 942–949.

[17] W. Li, Z. Wang, and H. Jafarkhani. (2018). "A tradeoff between the sub-packetization size and the repair bandwidth for Reed-Solomon codes." [Online]. Available: https://arxiv.org/abs/1806.00496

[18] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error Correcting Codes*. Amsterdam, The Netherlands: Elsevier, 1977.

[19] K. Rashmi, N. B. Shah, D. Gu, H. Kuang, D. Borthakur, and K. Ramchandran, "A solution to the network challenges of data recovery in erasure-coded distributed storage systems: A study on the Facebook warehouse cluster," in *Proc. HotStorage*, 2013, pp. 1–67.

[20] A. S. Rawat, O. O. Koyluoglu, and S. Vishwanath. (2016). "Centralized repair of multiple node failures with applications to communication efficient secret sharing." [Online]. Available: https://arxiv.org/abs/1603.04822

[21] K. Shanmugam, D. S. Papailiopoulos, A. G. Dimakis, and G. Caire, "A repair framework for scalar MDS codes," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 5, pp. 998–1007, May 2014.

[22] K. W. Shum and Y. Hu, "Cooperative regenerating codes," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7229–7258, Nov. 2013.

[23] I. Tamo, M. Ye, and A. Barg. (2017). "Optimal repair of Reed-Solomon codes: Achieving the cut-set bound." [Online]. Available: https://arxiv.org/abs/1706.00112

[24] Z. Wang, I. Tamo, and J. Bruck, "Optimal rebuilding of multiple erasures in MDS codes," *IEEE Trans. Inf. Theory*, vol. 63, no. 2, pp. 1084–1101, Feb. 2017.

[25] M. Ye and A. Barg, "Explicit constructions of MDS array codes and RS codes with optimal repair bandwidth," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2016, pp. 1202–1206.

[26] M. Ye and A. Barg. (2017). "Repairing Reed-Solomon codes: Universally achieving the cut-set bound for any number of erasures." [Online]. Available: https://arxiv.org/abs/1710.07216

[27] M. Ye and A. Barg. (2018). "Optimal MDS codes for cooperative repair." [Online]. Available: https://arxiv.org/abs/1801.09665

[28] M. Zorgui and Z. Wang, "Centralized multi-node repair in distributed storage," in *Proc. 54th Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Sep. 2016, pp. 617–624.

[29] M. Zorgui and Z. Wang, "Centralized multi-node repair for minimum storage regenerating codes," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2017, pp. 2213–2217.

[30] M. Zorgui and Z. Wang. (2017). "Centralized multi-node repair regenerating codes." [Online]. Available: https://arxiv.org/abs/1706.05431

**Jay Mardia** is a PhD student in Electrical Engineering at Stanford University, where he is a Stanford Graduate Fellow for the years 2017-2020. He received a B. Tech in Electrical Engineering from the Indian Institute of Technology–Bombay in 2017, where he also minored in math.

**Burak Bartan** is a PhD student at Electrical Engineering Department at Stanford University. He received an MS degree in electrical engineering from Stanford University in 2018 and a BS degree in electrical and electronics engineering from Bilkent University in 2016. During his undergraduate studies at Bilkent University, he worked on applying fractional Fourier, and linear canonical transforms to image compression and discrete linear canonical transforms. His current academic interests include machine learning, and coding theory.

**Mary Wootters** is an assistant professor of Computer Science and Electrical Engineering at Stanford University. She received a PhD in mathematics from the University of Michigan in 2014, and a BA in math and computer science from Swarthmore College in 2008; she was an NSF postdoctoral fellow at Carnegie Mellon University from 2014 to 2016. Her research interests include randomized algorithms, coding theory, dimension reduction, matrix completion, and sparse signal processing.