Repair Schemes with Optimal I/O Costs for Full-Length Reed-Solomon Codes with Two Parities

Hoang Dau* and Emanuele Viterbo†
Department of Electrical and Computer Systems Engineering, Monash University
Emails: *hoang.dau@monash.edu, †emanuele.viterbo@monash.edu

***Abstract*—*Network transfer* and *disk read* constitute the two most time-consuming operations in the repair process for node failures in erasure-code-based distributed storage systems. Recent developments on Reed-Solomon codes have demonstrated repair schemes that achieve optimal *network bandwidths* in the recovery of single failures, although in certain cases at the expense of a *trivially high I/O cost*, a term referring to the number of disk reads performed in a repair scheme. We are interested in the lowest I/O cost a repair scheme can achieve for Reed-Solomon codes. We establish two repair schemes for a family of Reed-Solomon codes with two parities that achieve the optimal I/O cost.**

I. Introduction

Reed-Solomon (RS) codes [1] are arguably the most popular codes used in practical storage systems, thanks to numerous advantages, including optimal storage overhead, widest range of code parameters, and simple implementations. They form core components of major distributed storage systems (DSS) such as Google's Colossus [2], Quantcast File System [3], Facebook's f4 [4], Yahoo Object Store [5], Baidu's Atlas [6], Backblaze's Vaults [7], and Hadoop Distributed File System [8].

Recent developments on *repairing RS codes* ([9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21]) established that carefully designed repair schemes can reduce the *repair bandwidth* significantly compared to the naive scheme. Here, the repair bandwidth refers to the amount of data to be transmitted from the helper nodes to the replacement node during the *repair process* that recovers the lost content of one failed node. As pointed out in [22], however, repair schemes with optimal repair bandwidths may incur a *trivial I/O cost*. By definition, the I/O cost measures the total amount of data being read from the physical disks located at the helper nodes during the repair process of a failed node. The trivial I/O cost refers to the I/O cost of the naive repair scheme, which requires the reading of the entire file from the system.

In this work, we investigate I/O-efficient repair schemes for a family of RS codes. Fig. 1 illustrates the trade-off between the repair bandwidth and the I/O cost (in bits) during the repair process of one failed node in a DSS employing an $[8,6]_8$ Reed-Solomon code. The left-most point represents the scheme with an optimal repair bandwidth but incurring the worst I/O cost ([22]). The schemes represented by the next three points, in which the I/O cost is minimized, are our object of interest. In general, for the case of full-length RS codes with two parities over fields of characteristic two, we introduce a lower bound on the I/O cost (Section III) and demonstrate two repair schemes that achieve this bound (Section IV). The second scheme improves the repair bandwidth upon the first one. These are, as far as we know, the first repair schemes with an optimal I/O cost for Reed-Solomon codes. It is an open problem to determine the lowest possible repair bandwidth given that the I/O cost is optimal.

II. Preliminaries

Let $[n]$ denote the set $\{1,2,\ldots,n\}$. Let $F=\mathbb{F}_q$ be the finite field of q elements, for some prime power q. Let $E=\mathbb{F}_{q^\ell}$ be an extension field of F, where $\ell \geq 1$, and let $E^* = E \backslash \{0\}$. We refer to the elements of E as *symbols* and the elements of F as *sub-symbols*. The field E may also be viewed as a vector space of dimension ℓ over F, i.e. $E \cong F^\ell$, and hence each symbol in E may be represented as a vector of length ℓ over F. More specifically, suppose $\mathcal{B}=\{\boldsymbol{\beta}_i\}_{i=1}^{\ell}$ is a basis of E over F, then any element $\boldsymbol{\alpha} \in E$ can be written as $\boldsymbol{\alpha} = \sum_{i=1}^{\ell} \alpha_i \boldsymbol{\beta}_i$. The unique vector $\phi_{\mathcal{B}}(\boldsymbol{\alpha}) = (\alpha_1,\ldots,\alpha_\ell) \in F^\ell$ is called the *vector representation* of $\boldsymbol{\alpha}$ with respect to the basis $\mathcal{B}$. Throughout this work we fix $\mathcal{B}$ and use $\phi(\boldsymbol{\alpha})$, i.e. dropping the subscript. We also use $\psi(\boldsymbol{\alpha}) = \big(\phi(\boldsymbol{\alpha}\boldsymbol{\beta}_1)^{\mathrm{T}},\ldots,\phi(\boldsymbol{\alpha}\boldsymbol{\beta}_\ell)^{\mathrm{T}}\big)^{\mathrm{T}}$ to denote the $\ell \times \ell$ matrix whose ith row is $\phi(\boldsymbol{\alpha}\boldsymbol{\beta}_i)$. The following properties are well known for $\phi(\cdot)$ and $\psi(\cdot)$ (see, for instance, [14]).

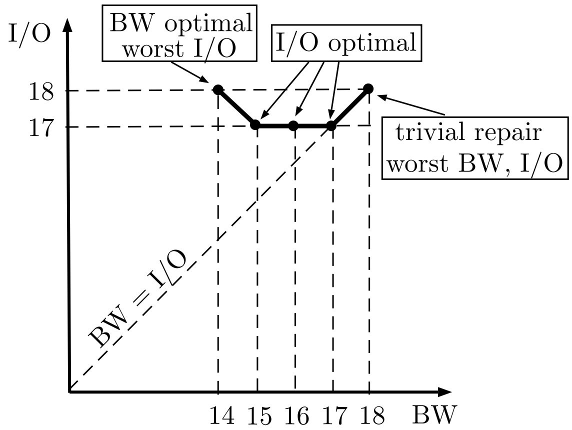

Fig. 1: Illustration of the trade-off between the repair bandwidth and the I/O cost (in bits) for an $[8,6]_8$ RS code. Given a repair bandwidth, the corresponding lowest possible I/O cost is found by an exhaustive search.

Lemma 1. *The following statements hold for $\boldsymbol{\alpha},\boldsymbol{\beta} \in E$, $b \in F$.*
(a) $\phi(\boldsymbol{\alpha}+\boldsymbol{\beta}) = \phi(\boldsymbol{\alpha}) + \phi(\boldsymbol{\beta})$, $\quad \phi(b\boldsymbol{\alpha}) = b\phi(\boldsymbol{\alpha})$,
(b) $\psi(\boldsymbol{\alpha}+\boldsymbol{\beta}) = \psi(\boldsymbol{\alpha}) + \phi(\boldsymbol{\beta})$, $\quad \psi(\boldsymbol{\alpha}\boldsymbol{\beta}) = \psi(\boldsymbol{\alpha})\psi(\boldsymbol{\beta})$,
(c) $\phi(\boldsymbol{\alpha}\boldsymbol{\beta}) = \phi(\boldsymbol{\alpha})\psi(\boldsymbol{\beta})$,
(d) $\phi(\boldsymbol{\alpha}) \neq (0,\ldots,0)$ *and* $\det_F(\psi(\boldsymbol{\alpha})) \neq 0$ *if* $\boldsymbol{\alpha} \neq 0$.

For $\boldsymbol{u},\boldsymbol{v} \in E$, if $\phi(\boldsymbol{u})\phi(\boldsymbol{v})^{\mathrm{T}} = 0$ then we say $\boldsymbol{u}$ is *orthogonal* to $\boldsymbol{v}$ and write $\boldsymbol{u} \perp \boldsymbol{v}$. For a set $U \subseteq E$, if $\boldsymbol{u} \perp \boldsymbol{v}$ for every $\boldsymbol{u} \in U$ then we also write $U \perp \boldsymbol{v}$. If U is an F-subspace of E, then $U^\perp$ denotes the *orthogonal complement* of U, which contains all elements of E orthogonal to U. We use $\mathsf{span}_F(U)$ to denote the F-subspace of E spanned by a set of elements U of E. The (field) trace of any symbol $\boldsymbol{\alpha} \in E$ over F is defined to be $\mathsf{Tr}_{E/F}(\boldsymbol{\alpha}) = \sum_{i=0}^{\ell-1} \boldsymbol{\alpha}^{q^i}$ (the subscript E/F is often omitted). The support of a vector $\boldsymbol{u} = (u_1,\ldots,u_\ell)$, denoted $\mathsf{supp}(\boldsymbol{u})$, is the set $\{j : u_j \neq 0\}$. The (Hamming) weight of $\boldsymbol{u}$, denoted $\mathsf{wt}(\boldsymbol{u})$, is $|\mathsf{supp}(\boldsymbol{u})|$. The support of a set of vectors U is $\mathsf{supp}(U) \triangleq \cup_{\boldsymbol{u}\in U}\mathsf{supp}(\boldsymbol{u})$. A *linear* $[n,k]$ *code* $\mathcal{C}$ over E is an E-subspace of E^n of dimension k. Each element of a code is referred to as a *codeword*. The *dual* of a code $\mathcal{C}$, denoted $\mathcal{C}^\perp$, is the orthogonal complement of $\mathcal{C}$ in

E^n and has dimension $r = n - k$. Elements of $\mathcal{C}^{\perp}$ are called *dual codewords*.

Definition 1. Let $E[x]$ denote the ring of polynomials over E. A Reed-Solomon code $\text{RS}(\mathcal{A}, k) \subseteq E^n$ of dimension k over a finite field E with evaluation points $\mathcal{A} = \{\boldsymbol{\alpha}_j\}_{j=1}^n \subseteq E$ is defined as

$$\text{RS}(\mathcal{A}, k) = \Big\{ \big(f(\boldsymbol{\alpha}_1), \dots, f(\boldsymbol{\alpha}_n)\big) : f \in E[x],\ \deg(f) < k \Big\}.$$

The Reed-Solomon code is *full length* if $n = |E|$. Note that the dual of a full-length Reed-Solomon code $\text{RS}(\mathcal{A}, k)$ is another Reed-Solomon code $\text{RS}(\mathcal{A}, n-k)$ (as a corollary of [23, Chp. 10, Thm. 4]).

Trace repair framework. First, note that each symbol in E can be recovered from its ℓ independent traces. More precisely, given a basis $\{\boldsymbol{\beta}_i\}_{i=1}^{\ell}$ of E over F, any $\boldsymbol{\alpha} \in E$ can be uniquely determined given the values of $\mathsf{Tr}(\boldsymbol{\beta}_i \boldsymbol{\alpha})$ for $i \in [\ell]$, i.e. $\boldsymbol{\alpha} = \sum_{i=1}^{\ell} \mathsf{Tr}(\boldsymbol{\beta}_i \boldsymbol{\alpha}) \boldsymbol{\beta}_i^*$, where $\{\boldsymbol{\beta}_i^*\}_{i=1}^{\ell}$ is the *dual (trace-orthogonal) basis* of $\{\boldsymbol{\beta}_i\}_{i=1}^{\ell}$ (see, e.g., [24, Ch. 2, Def. 2.30]).

Let $\mathcal{C}$ be an $[n, k]$ linear code over E and $\mathcal{C}^{\perp}$ be its dual. If $\boldsymbol{c} = (\boldsymbol{c}_1, \dots, \boldsymbol{c}_n) \in \mathcal{C}$ and $\boldsymbol{g} = (\boldsymbol{g}_1, \dots, \boldsymbol{g}_n) \in \mathcal{C}^{\perp}$ then $\boldsymbol{c} \cdot \boldsymbol{g} = \sum_{j=1}^{n} \boldsymbol{c}_j \boldsymbol{g}_j = 0$. Suppose $\boldsymbol{c}_{j^*}$ is erased and needs to be recovered. In the trace repair framework, choose a set of ℓ dual codewords $\boldsymbol{g}^{(1)}, \dots, \boldsymbol{g}^{(\ell)}$ such that $\dim_F \big(\{\boldsymbol{g}_{j^*}^{(i)}\}_{i=1}^{\ell}\big) = \ell$. Since trace is linear, we obtain the following ℓ equations

$$\mathsf{Tr}\big(\boldsymbol{g}_{j^*}^{(i)} \boldsymbol{c}_{j^*}\big) = -\sum_{j \neq j^*} \mathsf{Tr}\big(\boldsymbol{g}_j^{(i)} \boldsymbol{c}_j\big), \quad i \in [\ell]. \tag{1}$$

In order to recover $\boldsymbol{c}_{j^*}$, one needs to retrieve sufficient information from $\{\boldsymbol{c}_j\}_{j \neq j^*}$ to compute the right-hand sides of (1). We define, for every $j \in [n]$,

$$\mathcal{S}_{j \to j^*} \triangleq \mathsf{span}_F \left(\left\{ \boldsymbol{g}_j^{(1)}, \dots, \boldsymbol{g}_j^{(\ell)} \right\} \right) \tag{2}$$

and refer to $\mathcal{S}_{j \to j^*}$ as a *column-space* of the repair scheme when $j \neq j^*$. Then for each $j \neq j^*$, in order to determine $\mathsf{Tr}(\boldsymbol{g}_j^{(i)} \boldsymbol{c}_j)$ for all $i \in [\ell]$, it suffices to retrieve $\dim_F(\mathcal{S}_{j \to j^*})$ sub-symbols (in F) only. Indeed, suppose $\{\boldsymbol{g}_j^{(i_t)}\}_{t=1}^{s}$ is an F-basis of $\mathcal{S}_{j \to j^*}$, then by retrieving just s traces $\mathsf{Tr}(\boldsymbol{g}_j^{(i_1)} \boldsymbol{c}_j), \dots, \mathsf{Tr}(\boldsymbol{g}_j^{(i_s)} \boldsymbol{c}_j)$ of $\boldsymbol{c}_j$, all other traces $\mathsf{Tr}(\boldsymbol{g}_j^{(i)} \boldsymbol{c}_j)$ can be computed as F-linear combinations of those s traces without any knowledge of $\boldsymbol{c}$. Finally, since $\{\boldsymbol{g}_{j^*}^{(i)}\}_{i=1}^{\ell}$ is F-linearly independent, $\boldsymbol{c}_{j^*}$ can be recovered from its ℓ corresponding traces on the left-hand side of (1). We refer to such a scheme as a *repair scheme based on* $\{\boldsymbol{g}^{(i)}\}_{i=1}^{\ell}$. It was known that this type of repair schemes includes every possible linear repair scheme for RS codes [10].

Lemma 2 (Guruswami-Wootters [10])**.** *Suppose $E = \mathbb{F}_{q^\ell}$, $F = \mathbb{F}_q$, $\mathcal{C}$ is an $[n, k]$ linear code over E and $\mathcal{C}^{\perp}$ is its dual. The repair scheme for $\boldsymbol{c}_{j^*}$ based on ℓ dual codewords $\boldsymbol{g}^{(1)}, \dots, \boldsymbol{g}^{(\ell)}$, where $\dim_F \big(\{\boldsymbol{g}_{j^*}^{(i)}\}_{i=1}^{\ell}\big) = \ell$, incurs a repair bandwidth of $\sum_{j \neq j^*} \dim_F(\mathcal{S}_{j \to j^*})$ sub-symbols in F, where $\mathcal{S}_{j \to j^*}$ is defined as in* (2)*.*

I/O Cost of a Repair Scheme ([22]). Let $\mathcal{B} = \{\boldsymbol{\beta}_i\}_{i=1}^{\ell}$ be any F-basis of E. Then each element $\boldsymbol{\alpha} = \sum_{i=1}^{\ell} \alpha_i \boldsymbol{\beta}_i \in E$ can be represented by a vector $\phi(\boldsymbol{\alpha}) = (\alpha_1, \dots, \alpha_\ell) \in F^\ell$ as defined earlier. We assume throughout this work that every node uses a fixed common basis to represent and store the finite field elements. Another underlying assumption is that each sub-symbol α_i of $\boldsymbol{\alpha}$ can be read from the storage disk separately without accessing other sub-symbols. We first define the I/O cost of a function and then proceed to describe the I/O cost of a repair scheme.

Definition 2 ([22])**.** The (read) I/O cost of a function $f(\cdot)$ with respect to a basis $\mathcal{B}$ is the minimum number of sub-symbols of $\boldsymbol{\alpha} \in E$ needed for the computation of $f(\boldsymbol{\alpha})$. The I/O cost of a set of functions $\mathcal{F}$ is the minimum number of sub-symbols of $\boldsymbol{\alpha}$ needed for the computation of $\{f(\boldsymbol{\alpha}) : f \in \mathcal{F}\}$.

Lemma 3 ([22])**.** *The following statements hold.*

(a) The I/O cost of a linear function $f_{\boldsymbol{w}}(\boldsymbol{\alpha}) \triangleq \boldsymbol{w} \cdot \boldsymbol{\alpha} = \sum_i w_i \alpha_i$ with respect to a basis $\mathcal{B}$ is $\mathsf{wt}(\boldsymbol{w}) = |\mathsf{supp}(\boldsymbol{w})|$, where $\boldsymbol{w} = (w_1, \dots, w_\ell) \in E$.

(b) The I/O cost of a set of linear functions $\boldsymbol{w}_1 \cdot \boldsymbol{\alpha}, \dots, \boldsymbol{w}_s \cdot \boldsymbol{\alpha}$ with respect to $\mathcal{B}$ is $|\cup_{j=1}^{s} \mathsf{supp}(\boldsymbol{w}_j)|$.

(c) The I/O cost of the trace functional $\mathsf{Tr}_{\boldsymbol{\gamma}}(\cdot)$, defined by $\mathsf{Tr}_{\boldsymbol{\gamma}}(\boldsymbol{\alpha}) \triangleq \mathsf{Tr}(\boldsymbol{\gamma}\boldsymbol{\alpha})$, with respect to $\mathcal{B}$ is $\mathsf{wt}\big(\boldsymbol{w}^{\boldsymbol{\gamma}, \mathcal{B}}\big)$, where

$$\boldsymbol{w}^{\boldsymbol{\gamma}, \mathcal{B}} \triangleq \big(\mathsf{Tr}(\boldsymbol{\gamma}\boldsymbol{\beta}_1), \dots, \mathsf{Tr}(\boldsymbol{\gamma}\boldsymbol{\beta}_\ell)\big). \tag{3}$$

(d) The I/O cost of the set of trace functionals $\{\mathsf{Tr}_{\boldsymbol{\gamma}}(\cdot) : \boldsymbol{\gamma} \in \Gamma\}$ with respect to $\mathcal{B}$ is $|\cup_{\boldsymbol{\gamma} \in \Gamma} \mathsf{supp}(\boldsymbol{w}^{\boldsymbol{\gamma}, \mathcal{B}})|$.

The I/O cost of the repair scheme based on a set of dual codewords $\{\boldsymbol{g}^{(i)}\}_{i=1}^{\ell}$ is the minimum number of sub-symbols of $\boldsymbol{c}_j$'s, $j \neq j^*$, needed in the computation of the right-hand sides of (1). The formal definition is given below.

Definition 3 ([22])**.** The I/O cost of the repair scheme based on a set of dual codewords $\{\boldsymbol{g}^{(i)}\}_{i=1}^{\ell}$ with respect to a basis $\mathcal{B}$ is the sum of the I/O costs of the sets of trace functionals $\mathcal{F}_j = \big\{\mathsf{Tr}_{\boldsymbol{g}_j^{(i)}}(\cdot)\big\}_{i=1}^{\ell}$, $j \in [n] \setminus \{j^*\}$.

Our ultimate goal is to find the best trade-off curve between the repair bandwidth and the I/O cost. This appears to be challenging even for very particular sets of code parameters. Therefore, it is reasonable to first study the two extreme points of the curve. The first attempt along this line of research was made in [22], where the authors prove that known bandwidth-optimal repair schemes ([10], [13]) for some families of full-length RS codes actually incur a trivial I/O cost. Moreover, when the code has two parities ($r = 2$) and $F = \mathbb{F}_2$, they show that trivial I/O costs are a necessary price to pay for optimal repair bandwidths. We examine the same family of codes as above but prioritize the I/O cost. Our first repair scheme achieves the optimal I/O cost while the second one not only achieves the same I/O cost but also incurs a lower repair bandwidth, which is probably the lowest bandwidth possible.

III. A Lower Bound on the I/O Cost for Repairing Full-Length Reed-Solomon Codes with Two Parities

We establish in this section a lower bound on the I/O cost.

Theorem 1. *For a Reed-Solomon code of length $n = 2^\ell$ and $r = 2$ over $\mathbb{F}_{2^\ell}$, the I/O cost (in bits) of an arbitrary linear repair scheme always satisfies the following inequality.*

$$\mathsf{I/O} \geq (n-1)(\ell-1) + (2^{\ell-1} - 1). \tag{4}$$

Note that on the right-hand side of (4), the first term $(n-1)(\ell-1)$ corresponds to the optimal repair bandwidth of linear repair schemes, according to [13, Thm. 2], while the second term specifies the difference between the optimal repair bandwidth and the optimal I/O cost. The rest of this section is devoted to a proof of Theorem 1.

As $r = 2$, a dual codeword of the full-length RS code can be obtained by evaluating a polynomial of degree at most one at all the elements of $\mathbb{F}_{2^\ell}$. A linear repair scheme, therefore, is based on a set of ℓ polynomials $g_i(x) = \boldsymbol{a}_i x + \boldsymbol{b}_i$, $\boldsymbol{a}_i, \boldsymbol{b}_i \in \mathbb{F}_{2^\ell}$, $i \in [\ell]$. Set $A = \{\boldsymbol{a}_i\}_{i=1}^{\ell}$ and $B = \{\boldsymbol{b}_i\}_{i=1}^{\ell}$. We say the repair scheme *is defined by A and B*. Moreover, by [22, Lem. 8], it suffices to consider repairing $\boldsymbol{c}_1 = f(0)$, the first codeword component corresponding to the evaluation point $\boldsymbol{\alpha}_1 = 0$. As a repair scheme for $\boldsymbol{c}_1$, it is required that $\mathsf{rank}_{\mathbb{F}_2}(\{g_i(0)\}_{i=1}^{\ell}) = \ell$. In other words, *$B$ must be an $\mathbb{F}_2$-basis* of $\mathbb{F}_{2^\ell}$. We henceforth set $r_A \triangleq \mathsf{rank}_{\mathbb{F}_2}(A)$ and $A\boldsymbol{\gamma}+B \triangleq \{\boldsymbol{a}_i\boldsymbol{\gamma}+\boldsymbol{b}_i\}_{i=1}^{\ell}$.

It will later become clear in the proof of Lemma 8 that the set of "good" $\boldsymbol{\gamma}$ defined in Lemma 4 consists of the evaluation points where one bit of I/O can be saved in the repair scheme defined by A and B and with respect to $\boldsymbol{\beta} = \boldsymbol{\beta}_i \in \mathcal{B}$.

Lemma 4. *Suppose $A = \{\boldsymbol{a}_i\}_{i=1}^{\ell} \subset \mathbb{F}_{2^\ell}$, $B = \{\boldsymbol{b}_i\}_{i=1}^{\ell}$ is an $\mathbb{F}_2$-basis of $\mathbb{F}_{2^\ell}$, and $\boldsymbol{\beta} \in \mathbb{F}_{2^\ell}^*$. The set $G_{A,B,\boldsymbol{\beta}}$ defined as*

$$G_{A,B,\boldsymbol{\beta}} \triangleq \{\boldsymbol{\gamma} \in \mathbb{F}_{2^\ell} \colon \mathsf{Tr}((\boldsymbol{a}_i\boldsymbol{\gamma}+\boldsymbol{b}_i)\boldsymbol{\beta}) = 0,\ \forall i \in [\ell]\}\,, \quad (5)$$

which is called the set of "good" $\boldsymbol{\gamma}$, has size zero or $2^{\ell-r_A}$.

Proof. In order to show that $|G_{A,B,\boldsymbol{\beta}}|$ is zero or $2^{\ell-r_A}$, we prove that it is the solution set of a (non-homogeneous) system of linear equations where the coefficient matrix has rank r_A.

Let $K \triangleq \{\boldsymbol{\kappa} \in \mathbb{F}_{2^\ell} \colon \mathsf{Tr}(\boldsymbol{\kappa}) = 0\}$ be the kernel of the trace function. Let $\boldsymbol{u}_K$ be the normal vector of $K/\boldsymbol{\beta}$, that is, $\boldsymbol{u}_K \neq 0$ and $\boldsymbol{u}_K \perp K/\boldsymbol{\beta}$. We can rewrite $G_{A,B,\boldsymbol{\beta}}$ as follows.

$$\begin{aligned} G_{A,B,\boldsymbol{\beta}} &= \{\boldsymbol{\gamma} \in \mathbb{F}_{2^\ell} \colon \boldsymbol{a}_i\boldsymbol{\gamma}+\boldsymbol{b}_i \in K/\boldsymbol{\beta},\ \forall i \in [\ell]\} \\ &= \left\{\boldsymbol{\gamma} \in \mathbb{F}_{2^\ell} \colon \phi(\boldsymbol{a}_i\boldsymbol{\gamma}+\boldsymbol{b}_i)\phi(\boldsymbol{u}_K)^{\mathrm{T}} = 0,\ \forall i \in [\ell]\right\}. \end{aligned}$$

Therefore, $\boldsymbol{\gamma} \in G_{A,B,\boldsymbol{\beta}}$ if and only if it is a solution to the system $\phi(\boldsymbol{\gamma})\boldsymbol{C} = \boldsymbol{v}$, where

$$\boldsymbol{C} = \left(\psi(\boldsymbol{a}_1)\phi(\boldsymbol{u}_K)^{\mathrm{T}} \mid \cdots \mid \psi(\boldsymbol{a}_\ell)\phi(\boldsymbol{u}_K)^{\mathrm{T}}\right)$$

and $\boldsymbol{v} = \left(\phi(\boldsymbol{b}_1)\phi(\boldsymbol{u}_K)^{\mathrm{T}} \mid \cdots \mid \phi(\boldsymbol{b}_\ell)\phi(\boldsymbol{u}_K)^{\mathrm{T}}\right)$, due to Lemma 1. Note that since B is a basis, the vectors $\phi(\boldsymbol{b}_1), \ldots, \phi(\boldsymbol{b}_\ell)$ are linearly independent over $\mathbb{F}_2$. Due to the linearity of ψ, if $\mathsf{rank}_{\mathbb{F}_2}\{\boldsymbol{a}_{i_1}, \ldots, \boldsymbol{a}_{i_{r_A}}\} = r_A$ then the corresponding columns $\{\psi(\boldsymbol{a}_{i_j})\phi(\boldsymbol{u}_K)^{\mathrm{T}})\}_{j=1}^{r_A}$ in the coefficient matrix $\boldsymbol{C}$ spans the column space of $\boldsymbol{C}$. Therefore, $\mathsf{rank}_{\mathbb{F}_2}(\boldsymbol{C}) \leq r_A$. For every $(\eta_1, \ldots, \eta_{r_A}) \neq (0, \ldots, 0)$, $\eta_i \in \mathbb{F}_2$, since $\sum_{j=1}^{r_A} \eta_j\boldsymbol{a}_{i_j} \neq 0$, by Lemma 1, $\psi(\sum_{j=1}^{r_A} \eta_j\boldsymbol{a}_{i_j})$ is invertible. Furthermore, as $\boldsymbol{u}_K \neq 0$, it follows that

$$\begin{aligned} &\left(\psi(\boldsymbol{a}_{i_1})\phi(\boldsymbol{u}_K)^{\mathrm{T}} \mid \cdots \mid \psi(\boldsymbol{a}_{i_{r_A}})\phi(\boldsymbol{u}_K)^{\mathrm{T}}\right)(\eta_1, \ldots, \eta_{r_A})^{\mathrm{T}} \\ &\qquad = \psi\Big(\sum_{j=1}^{r_A} \eta_j\boldsymbol{a}_{i_j}\Big)\phi(\boldsymbol{u}_K)^{\mathrm{T}} \neq (0, \ldots, 0)^{\mathrm{T}}. \quad (6) \end{aligned}$$

Hence, this set of r_A columns of $\boldsymbol{C}$ is linearly independent. Therefore, $\mathsf{rank}_{\mathbb{F}_2}(\boldsymbol{C}) = r_A$, as desired. ■

Lemma 5. *Let K be an arbitrary $(\ell-1)$-dimensional $\mathbb{F}_2$-subspace of $\mathbb{F}_{2^\ell}$ and $\boldsymbol{u}_K$ its normal vector, that is $\boldsymbol{u}_K \perp K$. Then for every $\boldsymbol{\beta} \neq 0$, the normal vector of the subspace $K/\boldsymbol{\beta}$ is $\tau(\boldsymbol{\beta})\boldsymbol{u}_K$, where $\tau(\boldsymbol{\beta})$ is the unique element in $\mathbb{F}_{2^\ell}^*$ satisfying $\psi(\tau(\boldsymbol{\beta})) = \psi(\boldsymbol{\beta})^{\mathrm{T}}$.*

Proof. For every $\boldsymbol{\kappa} \in K$ we have

$$\begin{aligned} \phi(\boldsymbol{\kappa}/\boldsymbol{\beta})\phi(\tau(\boldsymbol{\beta})\boldsymbol{u}_K)^{\mathrm{T}} &= \phi(\boldsymbol{\kappa})\psi(\boldsymbol{\beta}^{-1})(\phi(\boldsymbol{u}_K)\psi(\tau(\boldsymbol{\beta})))^{\mathrm{T}} \\ = \phi(\boldsymbol{\kappa})\psi(\boldsymbol{\beta}^{-1})\psi(\tau(\boldsymbol{\beta}))^{\mathrm{T}}\phi(\boldsymbol{u}_K)^{\mathrm{T}} &= \phi(\boldsymbol{\kappa})\psi(\boldsymbol{\beta}^{-1})\psi(\boldsymbol{\beta})\phi(\boldsymbol{u}_K)^{\mathrm{T}} \\ &= \phi(\boldsymbol{\kappa})\psi(1)\phi(\boldsymbol{u}_K)^{\mathrm{T}} = \phi(\boldsymbol{\kappa})\phi(\boldsymbol{u}_K)^{\mathrm{T}} = 0. \end{aligned}$$

Thus, $\tau(\boldsymbol{\beta})\boldsymbol{u}_K$ is the normal vector of $K/\boldsymbol{\beta}$. ■

Lemma 6. *Suppose $A = \{\boldsymbol{a}_i\}_{i=1}^{\ell} \subset \mathbb{F}_{2^\ell}$ and $B = \{\boldsymbol{b}_i\}_{i=1}^{\ell}$ is an $\mathbb{F}_2$-basis of $\mathbb{F}_{2^\ell}$. Set*

$$U_{A,B} = \{\boldsymbol{u} \in \mathbb{F}_{2^\ell}^* \colon \exists \boldsymbol{\gamma} \in \mathbb{F}_{2^\ell} \text{ satisfying } (A\boldsymbol{\gamma}+B) \perp \boldsymbol{u}\}. \quad (7)$$

Then $\mathsf{rank}_{\mathbb{F}_2}(U_{A,B}) \leq r_A$.

Proof. Relabeling the elements of A and B if necessary, we may assume that $\mathsf{rank}_{\mathbb{F}_2}(\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{r_A}\}) = \mathsf{rank}_{\mathbb{F}_2}(A)$. Let $\boldsymbol{P}$ be the $(\ell - r_A) \times r_A$ binary matrix satisfying

$$\boldsymbol{P}\begin{pmatrix} \boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{a}_{r_A} \end{pmatrix} = \begin{pmatrix} \boldsymbol{a}_{r_A+1} \\ \vdots \\ \boldsymbol{a}_\ell \end{pmatrix}$$

Then, let $S_{A,B}$ be the $\mathbb{F}_2$-subspace of $\mathbb{F}_{2^\ell}$ spanned by the $\ell - r_A$ entries of the following column vector

$$\left(\boldsymbol{P} \mid \boldsymbol{I}_{\ell-r_A}\right)\begin{pmatrix} \boldsymbol{a}_1\boldsymbol{\gamma}+\boldsymbol{b}_1 \\ \vdots \\ \boldsymbol{a}_\ell\boldsymbol{\gamma}+\boldsymbol{b}_\ell \end{pmatrix} = \boldsymbol{P}\begin{pmatrix} \boldsymbol{b}_1 \\ \vdots \\ \boldsymbol{b}_{r_A} \end{pmatrix} + \begin{pmatrix} \boldsymbol{b}_{r_A+1} \\ \vdots \\ \boldsymbol{b}_\ell \end{pmatrix},$$

where $\boldsymbol{I}_{\ell-r_A}$ is the identity matrix of order $\ell - r_A$. Clearly, $S_{A,B}$ does not depend on $\boldsymbol{\gamma}$ and $S_{A,B} \subseteq S_{A,B,\boldsymbol{\gamma}} \triangleq \mathsf{span}_{\mathbb{F}_2}\{A\boldsymbol{\gamma}+B\}$, for every $\boldsymbol{\gamma} \in \mathbb{F}_{2^\ell}$. Moreover, as B is a basis, $\dim_{\mathbb{F}_2}(S_{A,B}) = \ell - r_A$. For each $u \in U_{A,B}$, there exists $\boldsymbol{\gamma} \in \mathbb{F}_{2^\ell}$ such that $\boldsymbol{u}$ is orthogonal to $A\boldsymbol{\gamma}+B$, and hence to all elements in $S_{A,B,\boldsymbol{\gamma}} \supseteq S_{A,B}$. Therefore, $\boldsymbol{u} \in S_{A,B}^{\perp}$. Hence, $U_{A,B} \subseteq S_{A,B}^{\perp}$. As a consequence, $\mathsf{rank}_{\mathbb{F}_2}(U_{A,B}) \leq \dim_{\mathbb{F}_2}(S_{A,B}^{\perp}) = r_A$. ■

As we shall see in the proof of Lemma 8, the set of "good" $\boldsymbol{\beta}$ defined as in Lemma 7 determines how many $\boldsymbol{\beta}_i \in \mathcal{B}$ can give rise to the saving of $2^{\ell-r_A}$ bits in I/O (see Lemma 4) in a repair scheme defined by A and B.

Lemma 7. *Given $A = \{\boldsymbol{a}_i\}_{i=1}^{\ell} \subset \mathbb{F}_{2^\ell}$ and $B = \{\boldsymbol{b}_i\}_{i=1}^{\ell}$ an $\mathbb{F}_2$-basis of $\mathbb{F}_{2^\ell}$, we define the set of "good" $\boldsymbol{\beta}$ as*

$$G_{A,B} \triangleq \{\boldsymbol{\beta} \in \mathbb{F}_{2^\ell}^* \colon |G_{A,B,\boldsymbol{\beta}}| = 2^{\ell-r_A}\}, \quad (8)$$

where $G_{A,B,\boldsymbol{\beta}}$ is the set of "good" $\boldsymbol{\gamma}$ defined as in (5)*. Then $\mathsf{rank}_{\mathbb{F}_2}(G_{A,B}) \leq r_A$.*

Proof. By Lemma 4 and Lemma 5, we can rewrite the set $G_{A,B}$ as follows

$$\begin{aligned} G_{A,B} &= \{\boldsymbol{\beta} \in \mathbb{F}_{2^\ell}^* \colon \exists \boldsymbol{\gamma} \in \mathbb{F}_{2^\ell} \text{ so that } (A\boldsymbol{\gamma}+B) \subseteq K/\boldsymbol{\beta}\} \\ &= \{\boldsymbol{\beta} \in \mathbb{F}_{2^\ell}^* \colon \exists \boldsymbol{\gamma} \in \mathbb{F}_{2^\ell} \text{ so that } (A\boldsymbol{\gamma}+B) \perp \tau(\boldsymbol{\beta})\boldsymbol{u}_K\}, \end{aligned}$$

where K is the kernel of the trace function, $\boldsymbol{u}_K$ is K's normal vector, and $\tau(\boldsymbol{\beta})$ is the unique element in $\mathbb{F}_{2^\ell}^*$ satisfying $\psi(\tau(\boldsymbol{\beta})) = \psi(\boldsymbol{\beta})^{\mathrm{T}}$. We now define a map $\pi \colon \mathbb{F}_{2^\ell} \to \mathbb{F}_{2^\ell}$, $\pi(\boldsymbol{\beta}) = \tau(\boldsymbol{\beta})\boldsymbol{u}_K$. As τ is linear and one-to-one, so is π. Moreover, $\pi(G_{A,B}) = U_{A,B}$, defined in (7). Therefore,

$$\mathsf{rank}_{\mathbb{F}_2}(G_{A,B}) = \mathsf{rank}_{\mathbb{F}_2}(\pi(G_{A,B})) = \mathsf{rank}_{\mathbb{F}_2}(U_{A,B}) \leq r_A,$$

where the last inequality is by Lemma 6. ■

Lemma 8. *For the Reed-Solomon code of length $n = 2^\ell$ and $r = 2$ over $\mathbb{F}_{2^\ell}$, the I/O cost (in bits) of an arbitrary linear repair scheme always satisfies the following inequality.*

$$\text{I/O} \geq (n-1)\ell - \max_{0 \leq r_A \leq \ell}\{r_A 2^{\ell-r_A}\}. \quad (9)$$

Proof. Consider a scheme repairing the first codeword component $\boldsymbol{c}_1 = f(0)$ defined by $A = \{\boldsymbol{a}_i\}_{i=1}^{\ell}$ and $B = \{\boldsymbol{b}_i\}_{i=1}^{\ell}$, where B is an $\mathbb{F}_2$-basis of $\mathbb{F}_{2^\ell}$. The corresponding dual codewords are obtained by evaluating the polynomials $g_i(x) = \boldsymbol{a}_i x + \boldsymbol{b}_i$, $i \in [\ell]$ at all $\boldsymbol{\gamma} \in \mathbb{F}_{2^\ell}$. By Definition 3, the I/O cost of this scheme is the sum of the I/O costs of the sets of trace functionals $\mathcal{F}_{\boldsymbol{\gamma}} = \{\mathsf{Tr}_{\boldsymbol{a}_i\boldsymbol{\gamma}+\boldsymbol{b}_i}(\cdot)\}$, $\boldsymbol{\gamma} \in \mathbb{F}_{2^\ell}^*$. By Lemma 3 (c)(d), the I/O cost of $\mathcal{F}_{\boldsymbol{\gamma}}$ with respect to a fixed basis $\mathcal{B}$ is $|\cup_{i=1}^{\ell} \mathsf{supp}(\boldsymbol{w}^{\boldsymbol{a}_i\boldsymbol{\gamma}+\boldsymbol{b}_i,\mathcal{B}})|$. Recall that

$$\boldsymbol{w}^{\boldsymbol{a}_i\boldsymbol{\gamma}+\boldsymbol{b}_i,\mathcal{B}} = \left(\mathsf{Tr}((\boldsymbol{a}_i\boldsymbol{\gamma}+\boldsymbol{b}_i)\boldsymbol{\beta}_1), \ldots, \mathsf{Tr}((\boldsymbol{a}_i\boldsymbol{\gamma}+\boldsymbol{b}_i)\boldsymbol{\beta}_\ell)\right) \in \mathbb{F}_2^{\ell}.$$

Therefore, the I/O cost of $\mathcal{F}_\gamma$ with respect to $\mathcal{B}$ is precisely the number of *nonzero columns* in the $\ell \times \ell$ matrix $\boldsymbol{W}_\gamma$ whose rows are $\boldsymbol{w}^{\boldsymbol{a}_i\boldsymbol{\gamma}+\boldsymbol{b}_i,\mathcal{B}}$, $i \in [\ell]$,

$$\boldsymbol{W}_\gamma \triangleq \left(\begin{array}{c} \boldsymbol{w}^{\boldsymbol{a}_1\boldsymbol{\gamma}+\boldsymbol{b}_1,\mathcal{B}} \\ \hline \boldsymbol{w}^{\boldsymbol{a}_2\boldsymbol{\gamma}+\boldsymbol{b}_2,\mathcal{B}} \\ \hline \vdots \\ \hline \boldsymbol{w}^{\boldsymbol{a}_\ell\boldsymbol{\gamma}+\boldsymbol{b}_\ell,\mathcal{B}} \end{array}\right) =$$

$$\begin{pmatrix} \mathsf{Tr}((\boldsymbol{a}_1\boldsymbol{\gamma}+\boldsymbol{b}_1)\boldsymbol{\beta}_1)\cdots\mathsf{Tr}((\boldsymbol{a}_1\boldsymbol{\gamma}+\boldsymbol{b}_1)\boldsymbol{\beta}_i)\cdots\mathsf{Tr}((\boldsymbol{a}_1\boldsymbol{\gamma}+\boldsymbol{b}_1)\boldsymbol{\beta}_\ell) \\ \mathsf{Tr}((\boldsymbol{a}_2\boldsymbol{\gamma}+\boldsymbol{b}_2)\boldsymbol{\beta}_1)\cdots\mathsf{Tr}((\boldsymbol{a}_2\boldsymbol{\gamma}+\boldsymbol{b}_2)\boldsymbol{\beta}_i)\cdots\mathsf{Tr}((\boldsymbol{a}_2\boldsymbol{\gamma}+\boldsymbol{b}_2)\boldsymbol{\beta}_\ell) \\ \vdots \quad \ddots \quad \vdots \quad \ddots \quad \vdots \\ \mathsf{Tr}((\boldsymbol{a}_\ell\boldsymbol{\gamma}+\boldsymbol{b}_\ell)\boldsymbol{\beta}_1)\cdots\mathsf{Tr}((\boldsymbol{a}_\ell\boldsymbol{\gamma}+\boldsymbol{b}_\ell)\boldsymbol{\beta}_i)\cdots\mathsf{Tr}((\boldsymbol{a}_\ell\boldsymbol{\gamma}+\boldsymbol{b}_\ell)\boldsymbol{\beta}_\ell) \end{pmatrix}.$$

Note that if the codeword component $\boldsymbol{c}_\gamma = f(\boldsymbol{\gamma})$ is read in full, ℓ bits will need to be accessed. However, in a repair scheme, it is possible that less than ℓ bits of $\boldsymbol{c}_\gamma$ actually need to be read. Each *all-zero* column of $\boldsymbol{W}_\gamma$ indicates a bit in the vector representation of $\boldsymbol{c}_\gamma$ that does not need to be read, therefore, leads to a *saving of one bit* in I/O when $\boldsymbol{c}_\gamma$ is read in the repair scheme defined by A and B. The more all-zero columns, the larger the saving, and hence the lower the I/O cost.

Instead of counting the number of all-zero columns in each $\boldsymbol{W}_\gamma$ and sum that up for all $\boldsymbol{\gamma} \in \mathbb{F}_{q^\ell}^*$, we first fix an index i and count the number of all-zero i-th columns of $\boldsymbol{W}_\gamma$, $\boldsymbol{\gamma} \in \mathbb{F}_{2^\ell}^*$, and then sum that up over all $i \in [\ell]$. By Lemma 4, for each $\boldsymbol{\beta}_i \in \mathcal{B}$, the number of all-zero i-th columns in $\boldsymbol{W}_\gamma$, $\boldsymbol{\gamma} \in \mathbb{F}_{2^\ell}^*$, is $|G_{A,B,\boldsymbol{\beta}_i}|$, which is either zero or $2^{\ell - r_A}$. The set $G_{A,B,\boldsymbol{\beta}_i}$ consists of those "good" $\boldsymbol{\gamma}$ where the i-th column of $\boldsymbol{W}_\gamma$ is all-zero, which corresponds to a saving of one bit in I/O when $\boldsymbol{c}_\gamma$ is accessed. If $|G_{A,B,\boldsymbol{\beta}_i}| = 2^{\ell - r_A}$ then $\boldsymbol{\beta}_i$ is called "good". The set $G_{A,B}$ as defined in Lemma 7 consists of those "good" $\boldsymbol{\beta}$. Furthermore, as $\mathsf{rank}_{\mathbb{F}_2}(G_{A,B}) \leq r_A$, the basis $\mathcal{B}$ contains at most r_A "good" $\boldsymbol{\beta}_i$. In addition, if $\boldsymbol{\beta}_i$ is not "good", then there does not exists any $\boldsymbol{\gamma} \in \mathbb{F}_{2^\ell}$ such that the i-th column of $\boldsymbol{W}_\gamma$ is all-zero. Thus, the I/O cost of the repair scheme defined by A and B satisfies the following inequality

$$\text{I/O} \geq (n-1)\ell - \max_{0 \leq r_A \leq \ell} \{r_A 2^{\ell - r_A}\},$$

where the term $r_A 2^{\ell - r_A}$ corresponds to the maximum possible reduction/saving in I/O if $\mathsf{rank}_{\mathbb{F}_2}(A) = r_A$. The factor r_A accounts for the maximum number of "good" $\boldsymbol{\beta}_i \in \mathcal{B}$ while $2^{\ell - r_A}$ accounts for the saving in I/O with respect to each of such "good" $\boldsymbol{\beta}_i$. That concludes the proof. ■

Theorem 1 follows as a corollary of Lemma 8 by observing that $r_A = 1$ or $r_A = 2$ maximizes the term $r_A 2^{\ell - r_A}$.

IV. Repairing Full-Length Reed-Solomon Codes with Two Parities Using Minimal I/O Costs

We demonstrate two repair schemes that achieve the lower bound on the I/O cost established in Section III for full-length RS codes with two parities over $\mathbb{F}_{2^\ell}$. As suggested by the development of the lower bound, we first look for a repair scheme defined by A and B where $\mathsf{rank}_{\mathbb{F}_2}(A) = 1$ and there is one "good" $\boldsymbol{\beta}_i \in \mathcal{B}$, that is, $\boldsymbol{\beta}_i \in G_{A,B}$ for some $i \in [\ell]$. This scheme, although achieves the optimal I/O cost, incurs the worst repair bandwidth, which is the same as the I/O cost. We subsequently develop the second repair scheme correponding to the case $\mathsf{rank}_{\mathbb{F}_2}(A) = 2$ that not only attains the optimal I/O cost but also uses a lower repair bandwidth, which is probably the lowest among all repair schemes with optimal I/O costs.

Construction I. Let $A = \{\boldsymbol{a}_i\}_{i=1}^{\ell}$, where $\boldsymbol{a}_1 = \boldsymbol{a}_2 = \cdots = \boldsymbol{a}_\ell = \boldsymbol{a} \neq 0$. Choose an arbitrary $\boldsymbol{\beta} \in \mathcal{B} = \{\boldsymbol{\beta}_i\}_{i=1}^{\ell}$. Let $\{\boldsymbol{h}_i\}_{i=1}^{\ell-1}$ be an arbitrary basis of the subspace $K/\boldsymbol{\beta}$, where K is the kernel of the trace function. Choose an arbitrary $\boldsymbol{b}_1 \in \mathbb{F}_{2^\ell} \setminus K/\boldsymbol{\beta}$, and set $\boldsymbol{b}_i = \boldsymbol{b}_1 + \boldsymbol{h}_{i-1}$, for $2 \leq i \leq \ell$. Set $B = \{\boldsymbol{b}_i\}_{i=1}^{\ell}$. The output of this construction is the repair scheme of $\boldsymbol{c}_1 = f(0)$ defined by A and B. Repair schemes with the same I/O cost for other $\boldsymbol{c}_j$ can be obtained by modifying this repair scheme (see [22, Lem. 8]).

Lemma 9. *For $\boldsymbol{\beta} \in \mathcal{B}$ and A and B chosen as in Construction I, we have $\boldsymbol{\beta} \in G_{A,B}$, where $G_{A,B}$ is defined as in* (8).

Proof. Clearly, B is an $\mathbb{F}_2$-basis of $\mathbb{F}_{2^\ell}$. By definition of $G_{A,B}$, we need to show that $|G_{A,B,\boldsymbol{\beta}}| = 2^{\ell-1}$, where $G_{A,B,\boldsymbol{\beta}}$ is defined as in (5). As $\boldsymbol{h}_i = \boldsymbol{b}_1 + \boldsymbol{b}_{i+1}$, $i \in [\ell - 1]$, the subspace $K/\boldsymbol{\beta}$ consists of field elements each of which can be written as the sum of an even number of $\boldsymbol{b}_i$'s, $i \in [\ell]$. Therefore,

$$\begin{aligned} G_{A,B,\boldsymbol{\beta}} &= \{\boldsymbol{\gamma} \in \mathbb{F}_{2^\ell} \colon \boldsymbol{a\gamma} + \boldsymbol{b}_i \in K/\boldsymbol{\beta},\ \forall i \in [\ell]\} \\ &= \left\{\boldsymbol{\gamma} \in \mathbb{F}_{2^\ell} \colon \boldsymbol{a\gamma} = \sum_{i=1}^{\ell} \eta_i \boldsymbol{b}_i, \sum_{i=1}^{\ell} \eta_i = 1\right\}. \end{aligned}$$

That is, $G_{A,B,\boldsymbol{\beta}}$ consists of $\boldsymbol{\gamma} \in \mathbb{F}_{2^\ell}$ satisfying $\boldsymbol{a\gamma}$ can be written as the sum of an odd number of $\boldsymbol{b}_i$'s. Hence, $|\mathcal{G}_{A,B,\boldsymbol{\beta}}| = 2^{\ell-1}$, as desired. Thus, $\boldsymbol{\beta} \in G_{A,B}$. ■

Theorem 2. *The repair scheme in Construction I achieves the I/O cost of $(n-1)(\ell-1) + (2^{\ell-1} - 1)$ bits, which is optimal among all linear schemes repairing $\mathbf{c}_1 = f(0)$ for the Reed-Solomon code of length $n = 2^\ell$ with two parities over $\mathbb{F}_{2^\ell}$. Moreover, its repair bandwidth equals the I/O cost.*

Proof. Suppose $\boldsymbol{\beta} = \boldsymbol{\beta}_{i^*}$, for some $i^* \in [\ell]$. As $\boldsymbol{\beta} \in G_{A,B}$, it is "good", which means there are precisely $2^{\ell-1}$ elements $\boldsymbol{\gamma} \in \mathbb{F}_{2^\ell}$ satisfying $\mathsf{Tr}((\boldsymbol{a}_i\boldsymbol{\gamma}+\boldsymbol{b}_i)\boldsymbol{\beta}_{i^*}) = 0$, for all $i \in [\ell]$. Following the proof of Lemma 8, this means there are precisely $2^{\ell-1}$ elements $\boldsymbol{\gamma} \in \mathbb{F}_{2^\ell}$ where the i^*-th column in the corresponding matrix $\boldsymbol{W}_\gamma$ is all-zero. This implies a saving of $2^{\ell-1}$ bits in the I/O cost. Therefore, the I/O cost of the repair scheme in Construction I is

$$\text{I/O} = (n-1)\ell - 2^{\ell-1} = (n-1)(\ell-1) + (2^{\ell-1} - 1),$$

which meets the lower bound established in Theorem 1. The statement on the bandwidth also follows as $\mathsf{rank}_{\mathbb{F}_2}(A\boldsymbol{\gamma} + B)$ is $t-1$ when $\boldsymbol{a\gamma}$ is the sum of an odd number of $\boldsymbol{b}_i$'s and is t otherwise, which is the same as the I/O cost. ■

Remark 1. For an $[n = 2^\ell, k = 2^\ell - 2]_{\mathbb{F}_{2^\ell}}$ RS code, the proposed I/O-optimal repair scheme saves $2^{\ell-1} - \ell$ bits in I/O compared to the naive repair scheme, which incurs an I/O cost of $k\ell = (2^\ell - 2)\ell$ bits. The saving ratio, therefore, is $(2^{\ell-1} - \ell)/((2^\ell - 2)\ell) \approx 1/2\ell$, which becomes smaller when ℓ gets larger. That is understandable, given that we consider a minimal $r = 2$, which remains constant while ℓ and n grow. For larger r, the saving is expected to be more significant.

Construction II. Let $\ell \geq 3$. Suppose $\mathcal{B} = \{\boldsymbol{\beta}_i\}_{i=1}^{\ell}$ is an $\mathbb{F}_2$-basis of $\mathbb{F}_2$ with $\boldsymbol{\beta}_1 = 1$. Set $K_1 = K/\boldsymbol{\beta}_1$, $K_2 = K/\boldsymbol{\beta}_2$, and $H = K_1 \cap K_2$. Then $\dim_2(H) = \ell - 2$ (see [22, Lem. 5]). Suppose[1] that $\boldsymbol{\beta}_2 H \neq H$. Let $\{\boldsymbol{h}_i\}_{i=1}^{\ell-2}$ be an $\mathbb{F}_2$-basis of H.

Step 1 Select $\boldsymbol{a}_1 \in (H/\boldsymbol{\beta}_2) \cap (K_2 \setminus H)$ arbitrarily.

Step 2 Set $\boldsymbol{a}_2 = \boldsymbol{\beta}_2\boldsymbol{a}_1$ and $A = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_2\}$.

Step 3 Select $\boldsymbol{b}_1 \in \boldsymbol{a}_1 + H$ arbitrarily.

Step 4 Select $\boldsymbol{b}_2 \in \mathbb{F}_{2^\ell} \setminus (K_1 \cup K_2)$ arbitrarily, and set $B = \{\boldsymbol{b}_i\}_{i=1}^{\ell}$, where $\boldsymbol{b}_i = \boldsymbol{b}_2 + \boldsymbol{h}_{i-2}$, for $3 \leq i \leq \ell$.

The output of this construction is the repair scheme of $\boldsymbol{c}_1 = f(0)$ defined by A and B.

[1] This condition holds, e.g., when $\boldsymbol{\beta}_2$ is a primitive element with a zero trace. Indeed, as $\boldsymbol{\beta}_2, \boldsymbol{\beta}_2^2 \in K$, we have $\boldsymbol{\beta}_2 \in H$. Note that $K \not\supset \{\boldsymbol{\beta}_2^i\}_{i=1}^{\ell}$. Let $s \in \{2, \ldots, \ell\}$ be the smallest integer so that $\boldsymbol{\beta}_2^s \notin H$. Then $\boldsymbol{\beta}_2^s \in \boldsymbol{\beta}_2 H \setminus H$.

Lemma 10. *The sets* $(H/\boldsymbol{\beta}_2)\cap(K_2\setminus H)$ *and* $\mathbb{F}_{2^\ell}\setminus(K_1\cup K_2)$ *are nonempty* $(\ell\geq 3)$*. Hence, Step 1 and Step 4 in Construction II are valid. Moreover,* $\mathsf{rank}_{\mathbb{F}_2}(A)=2$ *and* $\mathsf{rank}_{\mathbb{F}_2}(B)=\ell$*.*

Proof. As $H/\boldsymbol{\beta}_2\subset K_1/\boldsymbol{\beta}_2=K/\boldsymbol{\beta}_2=K_2$ and $H/\boldsymbol{\beta}_2\neq H$ by our assumption, $(H/\boldsymbol{\beta}_2)\cap(K_2\setminus H)\neq\varnothing$. The same conclusion holds for $\mathbb{F}_{2^\ell}\setminus(K_1\cup K_2)$ as this set has size $2^\ell-(2^\ell-2^{\ell-2})=2^{\ell-2}\geq 2$. Moreover, $\mathsf{rank}_{\mathbb{F}_2}(A)=2$ since $\boldsymbol{a}_1\notin H$ while $\boldsymbol{a}_2=\boldsymbol{\beta}_2\boldsymbol{a}_1\in H$. And finally, $\mathsf{rank}_{\mathbb{F}_2}(B)=\ell$ since $\{\boldsymbol{b}_2+\boldsymbol{b}_i\}_{i=3}^{\ell}=\{\boldsymbol{h}_i\}_{i=1}^{\ell-2}$ spans H while $\boldsymbol{b}_1\notin H$ as $\boldsymbol{b}_1\in\boldsymbol{a}_1+H=K_2\setminus H$, and $\boldsymbol{b}_2\notin(K_1\cup K_2)\supset K_2=H\cup(\boldsymbol{a}_1+H)$. ■

Lemma 11. *A repair scheme of* $\boldsymbol{c}_1=f(0)$ *for an* $[2^\ell,2^\ell-2]_{2^\ell}$ *RS code defined by* $A=\{\boldsymbol{a}_1,\boldsymbol{a}_2,\ldots,\boldsymbol{a}_2\}$ *of rank two and* B *that is a basis of* $\mathbb{F}_2$ *has an I/O cost of* $(n-1)(\ell-1)+(2^{\ell-1}-1)$ *bits and a repair bandwidth of at most* $(n-1)(\ell-1)+(2^{\ell-2}-1)$ *bits if the following conditions are satisfied.*

(C1) $\mathsf{span}_{\mathbb{F}_2}(\{\boldsymbol{b}_2+\boldsymbol{b}_i\}_{i=3}^{\ell})=K_1\cap K_2$.

(C2) *The sets* $\frac{K_j+\boldsymbol{b}_1}{\boldsymbol{a}_1}\cap\frac{K_j+\boldsymbol{b}_2}{\boldsymbol{a}_2}$*,* $j=1,2$*, are nonempty.*

(C3) $\mathsf{rank}_{\mathbb{F}_2}(A\boldsymbol{\gamma}+B)=\ell$ *only for* $2^{\ell-2}$ *elements* $\boldsymbol{\gamma}\in\mathbb{F}_{2^\ell}$*.*

Proof. First, by the proof of Lemma 8, the scheme defined by A and B achieves an optimal I/O cost if $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are "good". To this end, we need to show that $|G_{A,B,\boldsymbol{\beta}_j}|=2^{\ell-2}$, $j=1,2$, which means that $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ each contributes a reduction of $2^{\ell-2}$ bits in the I/O cost. By Lemma 4, it suffices to prove that $G_{A,B,\boldsymbol{\beta}_j}\neq\varnothing$, $j=1,2$. By its definition, $G_{A,B,\boldsymbol{\beta}_j}$ consists of $\boldsymbol{\gamma}\in\mathbb{F}_{2^\ell}$ satisfying $A\boldsymbol{\gamma}+B\subset K_j$. By (C2), for each $j=1,2$, we can select $\boldsymbol{\gamma}_j\in\frac{K_j+\boldsymbol{b}_1}{\boldsymbol{a}_1}\cap\frac{K_j+\boldsymbol{b}_2}{\boldsymbol{a}_2}$. Then $\boldsymbol{a}_1\boldsymbol{\gamma}_j+\boldsymbol{b}_1\in K_j$ and $\boldsymbol{a}_2\boldsymbol{\gamma}_j+\boldsymbol{b}_2\in K_j$. Due to (C1), we have $(\boldsymbol{a}_2\boldsymbol{\gamma}_j+\boldsymbol{b}_2)+(\boldsymbol{a}_2\boldsymbol{\gamma}_j+\boldsymbol{b}_i)=\boldsymbol{b}_2+\boldsymbol{b}_i\in K_1\cap K_2$, for all $3\leq i\leq\ell$. Thus, $A\boldsymbol{\gamma}_j+B\subset K_j$, for $j=1,2$, as desired.

If (C3) holds then all but $2^{\ell-2}-1$ column-spaces $\mathcal{S}_{j\to j^*}$ of the repair scheme have dimension at most $\ell-1$ (excluding the column of $\boldsymbol{\gamma}=0$). Therefore, the scheme uses a repair bandwidth of at most $(n-1)(\ell-1)+(2^{\ell-2}-1)$ bits. ■

Theorem 3. *The repair scheme in Construction II achieves the same I/O cost as that in Construction I while incurs a lower repair bandwidth of* $(n-1)(\ell-1)+(2^{\ell-2}-1)$ *bits* $(\ell\geq 3)$*.*

Proof. In light of Lemma 11, we aim to show that A and B produced by Construction II satisfy (C1), (C2), and (C3).

First, (C1) follows immediately from the way $\boldsymbol{b}_i$, $3\leq i\leq\ell$, are defined in Step 4 of Construction II.

Second, for (C2) to hold, the idea is to show that there exists an element in $\mathbb{F}_{2^\ell}$ that belongs to neither $\frac{K_j+\boldsymbol{b}_1}{\boldsymbol{a}_1}$ nor $\frac{K_j+\boldsymbol{b}_2}{\boldsymbol{a}_2}$. As the sizes of these two sets sum up to 2^ℓ, they must intersect. Since $K_2=H\cup(\boldsymbol{a}_1+H)$, we have $(\boldsymbol{a}_1+H)\cap K_1=\varnothing$. As $\boldsymbol{b}_1\in\boldsymbol{a}_1+H$, we deduce that $\boldsymbol{b}_1\notin K_1$. Moreover, $\boldsymbol{b}_2\notin K_1$ by its definition. Therefore, $0\notin\frac{K_1+\boldsymbol{b}_1}{\boldsymbol{a}_1}\cup\frac{K_1+\boldsymbol{b}_2}{\boldsymbol{a}_2}$, as desired. Next, we show that $\frac{K_2+\boldsymbol{b}_1}{\boldsymbol{a}_1}\cap\frac{K_2+\boldsymbol{b}_2}{\boldsymbol{a}_2}\neq\varnothing$. Since $\boldsymbol{b}_1\in\boldsymbol{a}_1+H\subset K_2$, we have $K_2+\boldsymbol{b}_1=K_2$. Using the fact that $\boldsymbol{a}_2=\boldsymbol{\beta}_2\boldsymbol{a}_1$ and $K_2=K/\boldsymbol{\beta}_2$, it suffices to prove that $\boldsymbol{a}_1$ belongs to neither K nor $K_2+\boldsymbol{b}_2$. Since $\boldsymbol{a}_1\in K_2\setminus H$ and $(K_2\setminus H)\cap K_1=\varnothing$, we have $\boldsymbol{a}_1\notin K_1=K$ (here we use the assumption that $\boldsymbol{\beta}_1=1$, which implies $K_1=K$). Furthermore, since $\boldsymbol{a}_1\in K_2$ by its definition, we have $\boldsymbol{a}_1+K_2=K_2\not\ni\boldsymbol{b}_2$. Hence, $\boldsymbol{a}_1\notin K_2+\boldsymbol{b}_2$, as desired. Thus, (C2) holds.

Finally, we demonstrate that (C3) also holds. According to Step 4 in Construction II, the set $\{(\boldsymbol{a}_2\boldsymbol{\gamma}+\boldsymbol{b}_2)+(\boldsymbol{a}_2\boldsymbol{\gamma}+\boldsymbol{b}_i)\}_{i=3}^{\ell}$ spans H. Therefore, $\mathsf{rank}_{\mathbb{F}_2}(A\boldsymbol{\gamma}+B)=\ell$ if and only if $\boldsymbol{a}_1\boldsymbol{\gamma}+\boldsymbol{b}_1\notin H$, $\boldsymbol{a}_2\boldsymbol{\gamma}+\boldsymbol{b}_2\notin H$, and $(\boldsymbol{a}_1\boldsymbol{\gamma}+\boldsymbol{b}_1)+(\boldsymbol{a}_2\boldsymbol{\gamma}+\boldsymbol{b}_2)\notin H$. Equivalently, $\mathsf{rank}_{\mathbb{F}_2}(A\boldsymbol{\gamma}+B)=\ell$ if and only if $\boldsymbol{\gamma}\in\mathbb{F}_{2^\ell}\setminus\left(\frac{H+\boldsymbol{b}_1}{\boldsymbol{a}_1}\cup\frac{H+\boldsymbol{b}_2}{\boldsymbol{a}_2}\cup\frac{H+\boldsymbol{b}_1+\boldsymbol{b}_2}{\boldsymbol{a}_1+\boldsymbol{a}_2}\right)$. Note that each of the three sets in the union has size $2^{\ell-2}$. As long as they are pairwise disjoint, our conclusion on the number of such $\boldsymbol{\gamma}$ is justified. First, we show that $\frac{H+\boldsymbol{b}_1}{\boldsymbol{a}_1}\cap\frac{H+\boldsymbol{b}_2}{\boldsymbol{a}_2}=\varnothing$. Suppose by contradiction that there exist $\boldsymbol{h}',\boldsymbol{h}''\in H$ such that $\frac{\boldsymbol{h}'+\boldsymbol{b}_1}{\boldsymbol{a}_1}=\frac{\boldsymbol{h}''+\boldsymbol{b}_2}{\boldsymbol{a}_2}$. Since $\boldsymbol{b}_1\in\boldsymbol{a}_1+H$, we can write $\boldsymbol{b}_1=\boldsymbol{a}_1+\boldsymbol{h}^*$, for some $\boldsymbol{h}^*\in H$. Set $\boldsymbol{h}=\boldsymbol{h}'+\boldsymbol{h}^*\in H$, we have $(\boldsymbol{h}/\boldsymbol{a}_1+1)\boldsymbol{a}_2=\boldsymbol{h}''+\boldsymbol{b}_2$. Since $\boldsymbol{a}_2/\boldsymbol{a}_1=\boldsymbol{\beta}_2$ and $\boldsymbol{a}_2\in H$, we deduce that $\boldsymbol{b}_2=(\boldsymbol{h}''+\boldsymbol{a}_2)+\boldsymbol{\beta}_2\boldsymbol{h}\in H+\boldsymbol{\beta}_2H\subseteq K_1$, which is impossible since $\boldsymbol{b}_2\notin K_1$ according to Step 4 in Construction II (the last inclusion follows from the fact that $\boldsymbol{\beta}_2H=\boldsymbol{\beta}_2K_1\cap\boldsymbol{\beta}_2K_2\subset K=K_1$). We can show that $\frac{H+\boldsymbol{b}_1}{\boldsymbol{a}_1}\cap\frac{H+\boldsymbol{b}_1+\boldsymbol{b}_2}{\boldsymbol{a}_1+\boldsymbol{a}_2}=\frac{H+\boldsymbol{b}_2}{\boldsymbol{a}_2}\cap\frac{H+\boldsymbol{b}_1+\boldsymbol{b}_2}{\boldsymbol{a}_1+\boldsymbol{a}_2}=\varnothing$ using similar arguments. This concludes the proof. ■

Acknowledgment

This work is supported by the 210124 ARC DECRA grant DE180100768 and the Monash 250003 Faculty Initiative Fund.

References

[1] I. S. Reed and G. Solomon, "Polynomial codes over certain finite fields," *J. Soc. Ind. Appl. Math.*, vol. 8, no. 2, pp. 300–304, 1960.

[2] A. Fikes, "Colossus, Successor to Google File System," http://static.googleusercontent.com/media/research.google.com/en/us/university/relations/facultysummit2010/storage_architecture_and_challenges.pdf.

[3] M. Ovsiannikov, S. Rus, D. Reeves, P. Sutter, S. Rao, and J. Kelly, "The Quantcast File System," in *Proc. VLDB Endow.*, vol. 6, no. 11, 2013, pp. 1092–1101.

[4] S. Muralidhar *et al.*, "f4: Facebook's warm BLOB storage system," in *Proc. 11th ACM/USENIX Symp. Oper. Syst. Des. Implementation (OSDI)*, 2014, pp. 383–398.

[5] P. Narayanan, S. Samal, and S. Nanniyur, "Yahoo cloud object store - object storage at exabyte scale," https://yahooeng.tumblr.com/post/116391291701/yahoo-cloud-object-store-object-storage-at.

[6] C. Lai, S. Jiang, L. Yang, S. Lin, G. Sun, Z. Hou, C. Cui, and J. Cong, "Atlas: Baidu's key-value storage system for cloud data," in *Proc. 31st Symp. Mass Stor. Syst. Tech.(MSST)*, 2015, pp. 1–14.

[7] B. Beach, "Backblaze Vaults: Zettabyte-scale cloud storage architecture," https://www.backblaze.com/blog/vault-cloud-storage-architecture/.

[8] R. Li, Z. Zhang, K. Zheng, and A. Wang, "Progress report: Bringing erasure coding to Apache Hadoop," http://blog.cloudera.com/blog/2016/02/progress-report-bringing-erasure-coding-to-apache-hadoop/.

[9] K. Shanmugam, D. S. Papailiopoulos, A. G. Dimakis, and G. Caire, "A repair framework for scalar MDS codes," *IEEE J. Selected Areas Comm. (JSAC)*, vol. 32, no. 5, pp. 998–1007, 2014.

[10] V. Guruswami and M. Wootters, "Repairing Reed-Solomon codes," in *Proc. Annu. Symp. Theory Comput. (STOC)*, 2016.

[11] ——, "Repairing Reed-Solomon codes," *IEEE Trans. Inform. Theory*, vol. 63, no. 9, pp. 5684–5698, 2017.

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

[13] H. Dau and O. Milenkovic, "Optimal repair schemes for some families of Reed-Solomon codes," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, 2017, pp. 346–350.

[14] I. Duursma and H. Dau, "Low bandwidth repair of the RS(10,4) Reed-Solomon code," in *Proc. Inform. Theory Applicat. Workshop (ITA)*, 2017.

[15] A. Chowdhury and A. Vardy, "Improved schemes for asymptotically optimal repair of MDS codes," in *Proc. 55th Annual Allerton Conf. Comm Control Comput. (Allerton)*, 2017.

[16] I. Tamo, M. Ye, and A. Barg, "Optimal repair of Reed-Solomon codes: Achieving the cut-set bound," in *Proc. 58th Annual IEEE Symp. Foundations Computer Sci. (FOCS)*, 2017.

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

[18] H. Dau, I. Duursma, H. M. Kiah, and O. Milenkovic, "Repairing Reed-Solomon codes with multiple erasures," *IEEE Trans. Inform. Theory*, 2018, accepted.

[19] ——, "Repairing Reed-Solomon codes with two erasures," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, 2017, pp. 351–355.

[20] B. Bartan and M. Wootters, "Repairing multiple failures for scalar MDS codes," in *Proc. 55th Annual Allerton Conf. Comm Control Comput. (Allerton)*, 2017.

[21] M. Ye and A. Barg, "Repairing Reed-Solomon codes: Universally achieving the cut-set bound for any number of erasures," available at https://arxiv.org/abs/1710.07216.

[22] H. Dau, I. Duursma, and H. Chu, "On the I/O costs of some repair schemes for full-length Reed-Solomon codes," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, 2018, pp. 1700–1704.

[23] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. Amsterdam: North-Holland, 1977.

[24] R. Lidl and H. Niederreiter, *Introduction to Finite Fields and Their Applications*. Cambridge University Press, 1986.