An Improved Cooperative Repair Scheme for Reed-Solomon Codes

Yaqian Zhang, Zhifang Zhang
KLMM, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China
School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China
Emails: zhangyaqian15@mails.ucas.ac.cn, zfz@amss.ac.cn

Abstract—Dau et al. recently extend Guruswami and Wootters’ scheme (STOC’2016) to cooperatively repair two or three erasures in Reed-Solomon (RS) codes. However, their schemes restrict to either the case that the characteristic of $F$ divides the extension degree $[F:B]$ or some special failure patterns, where $F$ is the base field of the RS code and $B$ is the subfield of the repair symbols. In this paper, we derive improved cooperative repair schemes for RS codes that remove all these restrictions. Specifically, we obtain a one-round cooperative repair scheme for any two erasures and a three-round scheme for any three erasures. When restricted to a wide class of failure patterns, we also achieve a one-round cooperative repair scheme for three erasures. The repair bandwidth of our schemes remains the same with that of Dau et al.’s schemes.

Index Terms—Distributed storage system, Reed-Solomon codes, cooperative repair.

I. INTRODUCTION

Reed-Solomon (RS) codes are extensively used in distributed storage systems (DSS) for providing the optimal trade-off between redundancy and reliability. For example, a $[14,10]$ RS code is deployed in the file system of Facebook [1]. In particular, an $[n,k]$ RS code is used in a DSS with $n$ nodes such that each node stores one symbol of the codeword encoded from the original data file consisting of $k$ symbols. When some node fails, i.e., the symbol stored in that node is erased, a self-sustaining system should be able to repair the failed node by downloading data from the surviving nodes (helper nodes). An important metric for the node repair efficiency is the repair bandwidth, namely, the total amount of data communicated during the repair process.

A naive repair approach for RS codes requires downloading the whole file to recover just one single node, which is wasteful in the repair bandwidth. By contrast, Guruswami and Wootters [2] used the trace function to design a linear repair scheme for RS codes that greatly reduces the repair bandwidth. Their scheme was further extended by Dau et al. [3] using linearized polynomials. However, there remains a gap between the repair bandwidth of the schemes in [2], [3] and the optimal repair bandwidth indicated by the cut-set bound [4]. A main reason for this gap is that the formers keep a low level of sub-packetization. Later, Tamo et al. [5] defined a class of RS codes over a sufficiently large field that permits a repair scheme with bandwidth achieving the cut-set bound. They also proved the sub-packetization is necessarily $O(n^2)$ for the optimal repair. Meanwhile, tradeoffs between the sub-packetization and the repair bandwidth for RS codes were also studied in [6], [7].

In some scenarios, however, multiple node failures are quite common. Two typical models for repairing multiple erasures are the centralized repair [8], [9], where a special node called data center is assumed to generate all the replacement nodes, and the cooperative repair [10], where the replacement nodes are generated in a distributed and cooperative way. On the one hand, any cooperative repair scheme can serve as a centralized repair scheme by merging the replacement nodes into one data center and treating the exchange between replacement nodes as inner operations of the data center. On the other hand, it is proved in [11] that an MDS code achieving the optimal bandwidth in the cooperative repair model naturally attains the optimal bandwidth in the centralized repair model. Therefore, the cooperative repair model can turns out to be stronger than the centralized model. Moreover, the cooperative repair model is more suitable for DSS because of the distributed pattern. We focus on the cooperative repair of RS codes in this work.

In a recent work [12], Dau et al. extend Guruswami and Wootters’ scheme [2] to cooperatively repair RS codes with two or three erasures. They achieve one-round cooperative repair of two erasures and three-round cooperative repair of special patterns of three erasures. But their scheme only works when the characteristic of $F$ divides the extension degree $t = [F:B]$. For the case of two erasures, although a repair scheme that works over any finite field is also derived in [12], it needs two-round exchange in the collaboration phase. Later in [9] the authors obtain a centralized repair scheme for general scalar MDS codes with $r$ erasures. When applied to RS codes, the scheme has less repair bandwidth than the scheme in [12] confined to the centralized model. But the centralized repair scheme in [9] can not be transformed into a cooperative one, so we do not know if the bandwidth of the cooperative scheme in [12] can be improved while keeping the extension degree level.

However, we do improve the schemes in [12] in other aspects. Namely, we remove the restriction on the characteristic of $F$ and simultaneously realize one-round cooperative repair of any two erasures. For the case of three erasures, we design three-round cooperative repair for any failure patterns and over any finite fields. Moreover, for certain patterns of three erasures, we also achieve the one-round cooperative repair. The repair bandwidth of our schemes remains the same with that of the schemes in [12].
The remaining of the paper is organized as follows. First, we recall some necessary preliminaries in Section II. Then the improved cooperative repair schemes for two erasures and for three erasures are presented in Section III and Section IV respectively. Finally, Section V concludes the paper.

II. PRELIMINARIES

A. Notations and definitions

Throughout the paper, we use $[n]$ to denote $\{1, 2, ..., n\}$. Let $F$ be a finite field and $B$ be a subfield of $F$ with $|F : B| = t$. The elements in $F$ are called symbols and elements in $B$ are called subsymbols.

Let $A = \{\alpha_1, ..., \alpha_n\}$ be a set of distinct elements in $F$. An $[n, k]$ RS code with the evaluator set $A$, denoted by $\text{RS}_A$, is defined as

$$\text{RS}_A = \{(f(\alpha_1), ..., f(\alpha_n)) : f \in F[x], \deg(f) < k\}.$$ 

For any $x \in F$, its trace from $F$ to $B$ is defined as

$$\text{Tr}_{F/B}(x) = x + x^{|B|} + \cdots + x^{|B|^{t-1}} \in B.$$ 

For simplicity, we use $\text{Tr}$ instead of $\text{Tr}_{F/B}$ to denote the trace function from $F$ to $B$ when the two fields are clear from the context.

B. The cooperative repair model

We simply recall the cooperative repair model introduced in [10]. Suppose $r$ replacement nodes are to be generated to replace $r$ failed nodes respectively. The process is accomplished in two phases:

1) *The download phase*: each replacement node connects to $d$ (\(\leq n - r\)) helper nodes and downloads $\beta_1$ subsymbols from each helper node.

2) *The collaboration phase*: the $r$ replacement nodes exchange $\beta_2$ subsymbols with each other.

Note that the collaboration phase may be accomplished in multiple rounds. Here we assume synchronized and simultaneous channel, namely, all $r$ nodes can send data to others simultaneously in one round. Throughout the paper, an $m$-round cooperative repair means it needs $m$-round communication in the collaboration phase. Obviously, one-round repair is preferred with respect to the round complexity.

C. Du et al.'s cooperative repair scheme for RS codes

First, we illustrate two basic facts. One is that the dual of an RS code is still a generalized RS code, i.e.,

$$\text{RS}_A^\perp = \{(\lambda_1 g(\alpha_1), ..., \lambda_n g(\alpha_n)) : g \in F[x], \deg(g) < n-k\},$$

where $\lambda_i, i \in [n]$, are nonzero elements in $F$ determined from $A$. Hereafter, a polynomial of degree less than $n-k$ is called a check polynomial of the $[n, k]$ RS code. The other fact is that every element in $F$ can be computed from its $t$ independent traces as illustrated in the following lemma.

**Lemma 1.** [2] Suppose $\{\zeta_1, ..., \zeta_t\}$ is a basis of $F$ over $B$. Then for every $\gamma \in F$, $\gamma$ can be recovered from the $t$ subsymbols $\{\text{Tr}(\zeta_1 \gamma), ..., \text{Tr}(\zeta_t \gamma)\}$, i.e., $\gamma = \sum_{i=1}^t \text{Tr}(\zeta_i \gamma) \zeta_i^\perp$, where $\{\zeta_i^\perp\}_{i \in [t]}$ is the dual basis of $\{\zeta_i\}_{i \in [t]}$.

Next, we give a brief review of the one-round cooperative repair scheme in [12] for two erasures. It requires $n - k \geq \lvert B \rvert^{t-1} - 1$ and $\text{char}(F) \mid t$.

WLOG, suppose $f(\alpha_1)$ and $f(\alpha_2)$ are erased. The two replacement nodes that recover $f(\alpha_1)$ and $f(\alpha_2)$ are called node 1 and node 2 respectively. Denote $K = \{x \in F : \text{Tr}(x) = 0\}$, and define

$$K_{1,2} = \{x \in F : \text{Tr}((\alpha_1 - \alpha_2)x) = 0\}.$$  \(1\)

Obviously, $K_{1,2} = \frac{K}{\alpha_1 - \alpha_2}$. Moreover, $K$ and $K_{1,2}$ are both $(t - 1)$-dimensional subspaces of $F$, and $K_{1,2} = K_{2,1}$. Let $\{u_i\}_{i \in [t-1]}$ be a basis of $K_{1,2}$ over $B$, and extend to $\{u_i\}_{i \in [t]}$ as a basis of $F$ over $B$.

To recover the two nodes, define $2t$ polynomials:

$$p_i(x) = \frac{\text{Tr}(u_i (x - \alpha_1))}{x - \alpha_1}, \quad q_i(x) = \frac{\text{Tr}(u_i (x - \alpha_2))}{x - \alpha_2}, \quad i \in [t].$$

Since $\deg(p_i(x)) = |B|^{t-1} - 1 < n-k$, then $p_i(x)$’s are check polynomials for $\text{RS}_A$ which define $t$ parity-check equations:

$$\lambda_1 p_1(\alpha_1) f(\alpha_1) + \lambda_2 p_2(\alpha_2) f(\alpha_2) = \sum_{j=3}^n \lambda_j p_j(\alpha_j) f(\alpha_j).$$

Apply the trace function to both sides of the above equations, one can get

$$\text{Tr}(\lambda_1 p_1(\alpha_1) f(\alpha_1)) + \text{Tr}(\lambda_2 p_2(\alpha_2) f(\alpha_2)) = \sum_{j=3}^n \text{Tr}(\lambda_j p_j(\alpha_j) f(\alpha_j)).$$ \(2\)

Note that $p_i(\alpha_2) = 0$ for $i \in [t-1]$, and $p_i(\alpha_1) = u_i$ for $i \in [t]$, and $p_i(\alpha_j) = \frac{\text{Tr}(u_i (\alpha_j - \alpha_1))}{\alpha_j - \alpha_1}$ for $j \neq 1, 2$.

Therefore, by downloading $\text{Tr}(\frac{\lambda_j f(\alpha_j)}{\alpha_j - \alpha_1})$ from the node storing $f(\alpha_j)$ for all $j \neq 1, 2$, node 1 can compute the left sides of (2), i.e.,

$$\begin{cases} 
\text{Tr}(u_j \lambda_1 f(\alpha_1)), & j \in [t-1] \\
\text{Tr}(u_j \lambda_1 f(\alpha_1)) + \text{Tr}(\frac{\lambda_j f(\alpha_j)}{\alpha_j - \alpha_2}), & j \notin [t-1] 
\end{cases}.$$  \(\ast\)

Similarly, from the $t$ parity-check equations defined by the $q_i(x)$’s, node 2 can obtain the following:

$$\begin{cases} 
\text{Tr}(u_j \lambda_2 f(\alpha_2)), & j \in [t-1] \\
\text{Tr}(u_j \lambda_2 f(\alpha_2)) + \text{Tr}(\frac{\lambda_j f(\alpha_j)}{\alpha_j - \alpha_2}), & j \notin [t-1] 
\end{cases}.$$  \(\ast\ast\)

Since the assumption $\text{char}(F) \mid t$ implies $\frac{1}{\alpha_1 - \alpha_2} \in K_{1,2}$, it follows that $\text{Tr}(\frac{\lambda_j f(\alpha_j)}{\alpha_j - \alpha_1})$ and $\text{Tr}(\frac{\lambda_j f(\alpha_j)}{\alpha_j - \alpha_2})$ can be respectively recovered as linear combinations of $\{\text{Tr}(u_j \lambda_2 f(\alpha_2))\}_{j \in [t-1]}$ and $\{\text{Tr}(u_j \lambda_1 f(\alpha_1))\}_{j \in [t-1]}$. Therefore, in the collaboration phase node 1 can directly sends $\text{Tr}(\frac{\lambda_j f(\alpha_j)}{\alpha_j - \alpha_1})$ to node 2, and simultaneously node 2 sends $\text{Tr}(\frac{\lambda_j f(\alpha_j)}{\alpha_j - \alpha_2})$ to node 1. After this one-round exchange, node 1 and node 2 can respectively recover $f(\alpha_1)$ and $f(\alpha_2)$ by Lemma 1.

Hereafter, we omit the $\lambda_j$’s in the parity-check equations for simplicity, because they are explicitly determined from the evaluator set $A$ and has no influence on the repair property.
III. IMPROVED SCHEME FOR TWO ERASES

We improve the one-round cooperative repair scheme for two erasures in RS\(\lambda\) by removing the assumption \(\text{char}(F) \mid t\). The improvement is achieved by the following two techniques:

1) Multiplying the check polynomials \(q_i(x)\)'s with a specially chosen parameter \(\gamma\).
2) Extending a basis \(\{u_t\}_{t \in [t-1]}\) of the subspace \(K_{1,2}\) to a basis of \(F\) over \(B\) by adding a special \(u_t\).

The details are given in Theorem 1.

Theorem 1. Let \(n-k \geq |B|^{t-1}\). WLOG, suppose \(f(\alpha_1)\) and \(f(\alpha_2)\) are erased in RS\(\lambda\). Then the two nodes admits a one-round cooperative repair. Particularly, in the downloading phase, each replacement node downloads one subsymbol from each of the remaining \(n-2\) helper nodes, and in collaboration phase, the two replacement nodes exchange one subsymbol with each other simultaneously. Thus the repair bandwidth of each failed node is \(n-1\) subsymbols in \(B\).

Proof: Using the same notations as in Section II-C. Let \(\{u_1,...,u_{t-1}\}\) be a basis of \(K_{1,2}\) over \(B\). Particularly, set \(u_t = \frac{\delta}{\alpha_1 - \alpha_2}\), where \(\delta \in F\) with \(\text{Tr}(\delta) = 1\). Then \(\{u_1,...,u_t\}\) form a basis of \(F\) over \(B\) because \(u_t \in F \setminus K_{1,2}\). Choose a nonzero element \(\gamma \in K\) and the \(2t\) check polynomials are defined as follows:

\[ p_i(x) = \frac{\text{Tr}(u_t(x - \alpha_i))}{x - \alpha_1}, \quad q_i(x) = \frac{\gamma \text{Tr}(u_t(x - \alpha_2))}{x - \alpha_2}, \quad i \in [t]. \]

It is easy to see that for \(i \in [t-1]\), \(p_i(\alpha_1) = u_t, p_i(\alpha_2) = 0\) because \(u_t \in K_{1,2}\) and \(p_1(\alpha_1) = u_t, p_1(\alpha_2) = \frac{1}{\alpha_1 - \alpha_2}\).

Similarly, \(q_i(\alpha_2) = \gamma u_t, q_i(\alpha_1) = 0\) for \(i \in [t-1]\), and \(q_0(\alpha_2) = \gamma u_t, q_0(\alpha_1) = \frac{\gamma}{\alpha_1 - \alpha_2}\).

The Download Phase. As introduced in Section II-C, the two nodes respectively use the \(t\) check polynomials \(p_i(x)\) and \(q_i(x)\), \(i \in [t]\) to get \(t\) parity-check equations, from which each of them can obtain \(t\) subsymbols by downloading one subsymbol from each helper node. The details are displayed in Table I. Note we use \(A_{-1,2}\) to denote the nodes in \(A \setminus \{\alpha_1, \alpha_2\}\). The similar notations, i.e., \(A_{-1,-2,-3}\), are also used in Table II and III.

That is, after the download phase, each node (eg. node \(1\)) obtains \(t-1\) independent traces (eg. \(\text{Tr}(u_j f(\alpha_1))\)\(\{i \in [t-1]\}\)) and one mixed term (eg. \(\text{Tr}(u_t f(\alpha_1)) + \text{Tr}(\gamma f(\alpha_2))\)). In the following, we show that after introducing the multiplier \(\gamma\) in \(q_i(x)\)'s, the two nodes can accomplish the recovery in a one-round collaboration.

The Collaboration Phase. Since \(\gamma \in K\), then \(\frac{\gamma}{\alpha_1 - \alpha_2} \in K_{1,2} = \text{Span}\{u_j : j \in [t-1]\}\). Thus node 2 can finish the recovery by downloading \(\text{Tr}(\frac{\gamma f(\alpha_2)}{\alpha_1 - \alpha_2})\) from node 1.

On the other hand, suppose

\[ \frac{1}{\alpha_1 - \alpha_2} = \sum_{i=1}^{t} a_i \gamma u_t, \quad a_i \in B. \tag{3} \]

Then using \(\{a_i\}_{i \in [t]}\) as linear combination coefficients of the \(t\) terms node 2 obtains in the download phase, it can derive \(\text{Tr}(\frac{f(\alpha_2)}{\alpha_1 - \alpha_2}) + a_t \text{Tr}(\frac{\gamma f(\alpha_2)}{\alpha_1 - \alpha_2})\) which is exactly the subsymbol that node 2 sends to node 1 in the collaboration phase. We will show this transmission also makes node 1 finish its recovery.

Because subtracting this subsymbol from the mixed term obtained in the download phase, node 1 gets \(\text{Tr}(u_t f(\alpha_1)) - a_t \text{Tr}(\frac{f(\alpha_2)}{\alpha_1 - \alpha_2} + \frac{\gamma f(\alpha_2)}{\alpha_1 - \alpha_2}) = \text{Tr}(\frac{\delta - a_t \gamma}{\alpha_1 - \alpha_2} f(\alpha_1))\). However,

\[ \dim B\{u_1,...,u_t-1,\delta - a_t \gamma, \alpha_1 - \alpha_2\} = t, \tag{4} \]

which follows from \(\delta - a_t \gamma \notin K\) and thus \(\frac{\delta - a_t \gamma}{\alpha_1 - \alpha_2} \notin K_{1,2}\). As a result, node 1 can recover \(f(\alpha_1)\) by Lemma 1.

Remark 1. Note the technique of multiplying a parameter \(\gamma\) is also used in [12] for the purpose of removing the assumption \(\text{char}(F) \mid t\). However, they did not notice the facts (3) and (4), thus they derived a two-round repair scheme even after the restriction \(\text{char}(F) \mid t\) is removed. Note the selection of the parameter \(u_t\) is not unique. Here we set \(u_t = \frac{1}{\alpha_1 - \alpha_2}\) with \(\text{Tr}(\delta) = 1\) for simplicity. Actually, one can choose any \(\delta \in F\) so long as \(\text{Tr}(\delta) \neq 0\).

IV. IMPROVED SCHEME FOR THREE ERASES

Now we discuss the cooperative repair of three erasures in RS\(\lambda\). Similar with that in Section III, we modify the parity-check equations in [12] by multiplying two elements \(\gamma_1, \gamma_2\). Later we show how to specify \(\gamma_1, \gamma_2\) in order to remove the restrictions\(^1\) of [12].

For simplicity, a failure pattern \(\{\alpha, \alpha', \alpha''\}\) means the nodes storing \(f(\alpha), f(\alpha'), f(\alpha'')\) are failed. Let \(X\) denote the set of 3-subsets of the evaluator set \(A\). That is, \(X\) contains all the failure patterns of three erasures in RS\(\lambda\). Define

\[ \Omega_1 = \{\alpha, \alpha', \alpha''\} \in X : \frac{\alpha - \alpha''}{\alpha - \alpha'} \in B^*\}, \]

\[ \Omega_2 = X \setminus \Omega_1. \]

One can easily verify that the definition of \(\Omega_1\) is independent of the order of \(\alpha, \alpha', \alpha''\).

Theorem 2. Suppose \(n-k \geq |B|^{t-1}\). For any failure pattern \(\{\alpha, \alpha', \alpha''\} \in \Omega_1\), the three failed nodes can be recovered through a one-round cooperative repair. The repair bandwidth for repairing each failed node is \(n-1\) subsymbols in \(B\).

Theorem 3. Suppose \(n-k \geq |B|^{t-1}\) and \(t > 3\). For any failure pattern \(\{\alpha, \alpha', \alpha''\} \in \Omega_2\), the three failed nodes can be recovered only if \(\text{char}(F) \mid t\) and \(\{\frac{\alpha - \alpha''}{\alpha - \alpha'}, \frac{\alpha - \alpha'}{\alpha - \alpha''}, \frac{\alpha - \alpha''}{\alpha - \alpha'}\} \cap K \neq \emptyset\).

\(^1\)In [12], the three erasures \(f(\alpha), f(\alpha'), f(\alpha'')\) can be recovered only if \(\text{char}(F) \mid t\) and \(\{\frac{\alpha - \alpha''}{\alpha - \alpha'}, \frac{\alpha - \alpha'}{\alpha - \alpha''}, \frac{\alpha - \alpha''}{\alpha - \alpha'}\} \cap K \neq \emptyset\).
recovered through a three-round cooperative repair. The repair bandwidth for repairing each failed node is \( n - 1 \) subsymbols in \( B \).

For failure patterns in \( \Omega_2 \), Theorem 3 gives a three-round cooperative repair scheme. However, one-round repair can be achieved if we restrict to the failure patterns in a subset of \( \Omega_2 \).

**Theorem 4.** Suppose \( n - k \geq |B|^{t-1} \) and \( t^2 \neq 1 \) mod \( p \), where \( p \) is the characteristic of \( F \). For any failure pattern \( \{ \alpha, \alpha', \alpha'' \} \in \Omega_2 \) with \( \{ \frac{a - b}{a - c}, \frac{a - d}{a - c}, \frac{a - e}{a - c} \} \subseteq K \), the three failed nodes can be recovered through a one-round cooperative repair. The repair bandwidth for repairing each failed node is \( n - 1 \) subsymbols in \( B \).

We give precise proofs of the three theorems in Section IV-A, IV-B and IV-C respectively. Before that, some lemmas and corollaries are needed.

WLOG, hereafter for three distinct elements \( \alpha, \alpha', \alpha'' \in A \), we can always suppose they are \( \alpha_1, \alpha_2, \alpha_3 \) respectively. Using the same notations as in (1), it has \( K_{ij} = \frac{a - b}{a - c} \) for \( i, j \in \{1, 2, 3\}, i \neq j \). Define \( K_{1,2,3} = K_{1,2} \cap K_{2,3} \cap K_{1,3} \), then \( K_{1,2,3} \) is a linear space over \( B \). Actually, \( K_{1,2,3} \) is the intersection of any two of the three spaces because \( K_{i,j} \cap K_{i,k} \subseteq K_{j,k} \) for any \( \{i, j, k\} = [3] \). Moreover, for any \( x \in K_{1,2,3} \), \( \text{Tr}(\alpha) x = \text{Tr}(\alpha_2) x = \text{Tr}(\alpha_3) x \).

**Lemma 2.** Suppose \( K = \text{Ker}(\text{Tr}) \), and \( S \) is a subspace of \( F \) with dimension \( s \), then \( s - 1 \leq \dim_B(S \cap K) \leq s \).

**Proof:** It is obvious that \( \dim_B(S \cap K) \leq s \). Let \( \text{Tr}_{|S} \) denote the trace function from \( F \) to \( B \) restricted to the subspace \( S \). Then \( \text{Tr}_{|S} \) is a linear map from \( S \) to \( B \). Since \( \dim_B(\text{Tr}_{|S}) \leq 1 \), then \( \dim_B(\text{Ker}(\text{Tr}_{|S})) = \dim_B(S \cap K) \geq s - 1 \).

**Corollary 1.** \( t - 2 \leq \dim_B K_{1,2,3} \leq t - 1 \).

The corollary follows from Lemma 2 and the facts \( K_{1,2,3} = K_{1,2} \cap K_{2,3} \) and \( K_{1,2,3} = K_{1,2} \cap K_{2,3} = \frac{a_2 - a_3}{a_1 - a_2} K \cap K \).

**Lemma 3.** For any \( \sigma \in F \), \( \sigma K = K \) iff \( \sigma \in B^* \), where \( K = \text{Ker}(\text{Tr}) \) and \( B^* \) denotes the set of nonzero elements in \( B \).

**Proof:** If \( \sigma \in B^* \), it is obvious that \( \sigma K = K \). Conversely, suppose \( \sigma K = K \). It is evident that \( \sigma \neq 0 \). Let \( \{z_1, ..., z_t\} \) be a basis of \( K \) over \( B \), then for \( j \in \{t - 1, 1\}, \sigma z_j \in K \). We extend \( \{z_j\}_{j \in \{t - 1\}} \) to a basis of \( F \) over \( B \). Let \( Z = \{z_1, ..., z_t\} \) be the dual basis of \( Z \). We claim that \( z_i = (\text{Tr}(z_i))^{-1} \) since \( \text{Tr}(z_i (\text{Tr}(z_i))^{-1}) = 0, i \in \{t - 1\} \), and \( \text{Tr}(z_i (\text{Tr}(z_i))^{-1}) = 1 \). By the uniqueness of dual basis, it follows \( z_i = (\text{Tr}(z_i))^{-1} \). Now let \( \sigma = \sum_{i=1}^{t-1} a_i z_i \), with \( a_i \in B \). From Lemma 1, we know \( a_i = \text{Tr}(\sigma z_i) \in \{t\} \). Thus \( a_i = 0 \) for \( i \in \{t - 1\} \) because \( \sigma z_j \in K \), while \( a_i = \text{Tr}(\sigma z_i) \neq 0 \). Therefore, \( \sigma = a_i z_i = \text{Tr}(\sigma z_i) (\text{Tr}(z_i))^{-1} \in B^* \).

**Corollary 2.** A failure pattern \( \{\alpha_1, \alpha_2, \alpha_3\} \in \Omega_1 \) if\( \dim_B K_{1,2,3} = t - 1 \).

For simplicity, denote \( l = \dim_B K_{1,2,3} \), then \( l = t - 1 \) or \( t - 2 \). Let \( U = \{u_i\}_{i \in \{t\}}, V = \{v_i\}_{i \in \{t\}}, W = \{w_i\}_{i \in \{t\}} \) be three bases of \( K_{1,2,3} \) over \( B \) which can be the same basis. We extend \( U \) to a basis of \( K_{1,2} \), denoted by \( \{u_i\}_{i \in \{t - 1\}} \). Similarly, extend \( V \) to a basis of \( K_{2,3} \), denoted by \( \{v_i\}_{i \in \{t - 1\}} \) and extend \( W \) to a basis of \( K_{1,3} \), denoted by \( \{w_i\}_{i \in \{t - 1\}} \). Then we further extend them to three bases of \( F \) over \( B \), i.e., \( U' = \{u_i\}_{i \in \{t\}}, V' = \{v_i\}_{i \in \{t\}}, W' = \{w_i\}_{i \in \{t\}} \). Next define \( 3t \) check polynomials, i.e., for \( 1 \leq i \leq t \),

\[
p_i(x) = \frac{\text{Tr}(u_i(x - \alpha_1))}{x - \alpha_1}, \quad q_i(x) = \gamma_1 \frac{\text{Tr}(v_i(x - \alpha_2))}{x - \alpha_2}, \quad r_i(x) = \gamma_2 \frac{\text{Tr}(w_i(x - \alpha_3))}{x - \alpha_3},
\]

where \( \gamma_1, \gamma_2 \) are two nonzero elements in \( F \). In the following, we illustrate the repair schemes in Section IV-A, IV-B and IV-C w.r.t. Theorem 2, Theorem 3 and Theorem 4 respectively.

**A. One-round repair of failure patterns in \( \Omega_1 \) (Theorem 2)**

Suppose \( \{\alpha_1, \alpha_2, \alpha_3\} \in \Omega_1 \), by Corollary 2, it has \( K_{1,2,3} = K_{1,2} = K_{2,3} = K_{1,3} \). Set \( u_t = v_t = w_t = \frac{1}{\alpha_1 - \alpha_2} \), where \( \delta \) is chosen from \( F \) with \( \text{Tr}(\delta) = 1 \). By the definition of \( p_i(x), q_i(x) \) and \( r_i(x) \), it has

\[
\begin{align*}
& \{p_j(x) = u_j, \quad p_j(x) = p_j(x) = 0, \quad j \in \{t - 1\}, \quad p_0(x) = u_0, \quad p_0(x) = p_0(x) = 1, \\
& q_i(x) = \gamma_1 v_i, \quad q_j(x) = q_j(x) = 0, \quad j \in \{t - 1\}, \quad q_0(x) = 0, \quad q_0(x) = q_0(x) = 0, \\
& r_j(x) = \gamma_2 w_j, \quad r_j(x) = r_j(x) = 0, \quad j \in \{t - 1\}, \quad r_0(x) = 0, \quad r_0(x) = r_0(x) = 0.
\end{align*}
\]

Note here \( p_i(x) = 1, \quad \frac{\text{Tr}(\alpha - \alpha_1 \delta)}{\alpha_1 - \alpha_2} \), where the second equality holds because \( \frac{\alpha_1 - \alpha_2}{\alpha_1 - \alpha_2} \in B^* \). Other details of the computations are evident.

**The Download Phase.** The three nodes obtain some independent traces and mixed terms according to the parity-check equations defined by \( p_i(x), q_i(x), r_i(x) \) respectively, \( i \in \{t\} \). As in Section III we illustrate this process in Table II.

**The Collaboration Phase.** By properly choosing the parameters \( \gamma_1 \) and \( \gamma_2 \), the recovery can be accomplished through a one-round collaboration, which is illustrated in Lemma 4.

**Lemma 4.** Given \( \gamma_1, \gamma_2 \in F^* \), if the following equations on \( x_i, y_i, i \in \{3\} \), have solutions in \( B \):

\[
\begin{align*}
& x_1 + \text{Tr}(\gamma_1 \gamma_2 y_1) = \text{Tr}(\gamma_1 \gamma_2), \\
& \text{Tr}(\gamma_1 \gamma_2^{-1}) x_1 + y_1 = \text{Tr}(\gamma_1 \gamma_2^{-1}), \\
& \text{Tr}(\gamma_1) x_1 + \text{Tr}(\gamma_2) y_1 \neq 1.
\end{align*}
\]
\[ x_2 + \text{Tr}(\gamma_2)y_2 = \text{Tr}(\gamma_1), \]
\[ \text{Tr}(\gamma_2 \alpha - \gamma_2)y_2 = \text{Tr}(\gamma_1 \gamma_2), \]
\[ \text{Tr}(\gamma_1 \alpha - \gamma_2)y_2 \neq 1. \]

Then the recovery can be accomplished by each replacement node exchanging one subymbol with the other two nodes in one round.

**Proof:** The proof can be found in [13]. Due to space limitation, we omit it here. \[ \square \]

Now we are left to specify \( \gamma_1 \) and \( \gamma_2 \) such that (5-7) have solutions in \( B \). There are three cases:

1. \( t \geq 3 \): Choose \( \gamma_1 \in K^* \) and \( \gamma_2 \in K^* \cap \gamma_1 K \). We can do this because \( \dim_B(K) = t-1 \) and \( \dim_B(K \cap \gamma_1 K) \geq t-2 \geq 1 \) from Lemma 2. Then we have \( \text{Tr}(\gamma_1) = \text{Tr}(\gamma_2) = \text{Tr}(\gamma_1 \gamma_2) = 0 \). It is easy to verify that (5-7) are solvable.

2. \( t = 2 \) and \( \text{char}(F) \neq 3 \): Choose \( \gamma_1 = \gamma_2 \in K^* \). Then \( \text{Tr}(\gamma_1) = \text{Tr}(\gamma_2) = 0, \text{Tr}(\gamma_1 \gamma_2) = \text{Tr}(\gamma_2)^2 = 1 \). It is easy to verify that (5-7) are solvable in \( B \).

3. \( t = 2 \) and \( \text{char}(F) = 3 \): Choose \( \gamma_1 = \gamma_2 = 1 \). Since \( \text{Tr}(2) = 1 \), we can set \( \alpha = 1 \) in particular, then \( u_1 = v_1 = w_1 = 2/\alpha - 2 \). In this case, we give a straightforward way to complete the collaboration phase without concerning the equations (5-7). Specifically, the three nodes directly exchange the mixed terms obtained in the download phase with each other. Then all of the three nodes can get:

\[
\begin{pmatrix}
\text{Tr}(u_1)f(\alpha_1) + \text{Tr}(v_1)f(\alpha_2) + \text{Tr}(w_1)f(\alpha_3) \\
\text{Tr}(u_1)v_1f(\alpha_2) + \text{Tr}(v_1)w_1f(\alpha_3) + \text{Tr}(w_1)\alpha_2 \\
\text{Tr}(w_1)f(\alpha_3)
\end{pmatrix}
\]

\[
\begin{pmatrix}
2 & 1 & 1 \\
1 & 2 & 1 \\
1 & 1 & 2
\end{pmatrix}
\begin{pmatrix}
\text{Tr}(f(\alpha_1)) \\
\text{Tr}(f(\alpha_2)) \\
\text{Tr}(f(\alpha_3))
\end{pmatrix}
\]

Since the coefficient matrix is invertible, the three nodes can directly compute the \( t \)-th independent trace for recovery.

**B. Three-round repair of failure patterns in \( \Omega_2 \) (Theorem 3)**

Suppose \( \{\alpha_1, \alpha_2, \alpha_3\} \in \Omega_2 \), by Corollary 1 and Corollary 2, it has \( \dim_B K_1,2,3 = t-2 \) and \( K_2,3, K_2,3 \), \( K_1,3 \) are distinct. We can choose \( u_{-1} \in K_1,3 \cap K_2,3 \) with \( \text{Tr}(u_{-1} (\alpha_1 - \alpha_3)) = 1 \). Similarly, choose \( v_{-1} \in K_2,3 \cap K_1,3 \) with \( \text{Tr}(v_{-1} (\alpha_2 - \alpha_3)) = 1 \) and \( w_{-1} \in K_1,3 \cap K_2,3 \) with \( \text{Tr}(w_{-1}(\alpha_1 - \alpha_2)) = 1 \). Then we have

\[
\begin{align*}
\text{Tr}(u_{-1} \alpha_1) &= \text{Tr}(u_{-1} \alpha_2) = 1 + \text{Tr}(u_{-1} \alpha_3) \\
1 + \text{Tr}(u_{-1} \alpha_1) &= \text{Tr}(u_{-1} \alpha_2) = \text{Tr}(u_{-1} \alpha_3) \\
\text{Tr}(u_{-1} \alpha_1) &= 1 + \text{Tr}(w_{-1} \alpha_2) = \text{Tr}(w_{-1} \alpha_3) \\
\end{align*}
\]

Moreover, we can set \( u_t = u_{t-1} \) since \( u_{-1} \notin K_1,2,3 \). Similarly, set \( v_t = v_{t-1} \) and \( w_t = w_{t-1} \).

The download phase is the same with that in Section IV-A, except that the data obtained here is a little different due to different selections of the bases \( U', V', W' \). The details are illustrated in Table III.

Next we specify the choice of \( \gamma_1 \) and \( \gamma_2 \) to make sure the recovery can be realized after the collaboration phase. First choose \( \gamma_2 \in F^* \) such that \( \frac{\gamma_2}{\alpha - \alpha_1} \in K_{1,2,3} \). Then choose \( \gamma_1 \in F^* \) such that \( \gamma_1 \gamma_2 \in \gamma_2 (\alpha_1 - \alpha_2) K_{1,2,3} \cap K \). We can do this because \( \dim_B (\gamma_2 (\alpha_1 - \alpha_2) K_{1,2,3} \cap K) \geq t-3 \geq 1 \) by Lemma 2 and the assumption that \( t > 3 \).

Then the collaboration phase proceeds in three rounds as displayed in Fig. 1. Specifically, round 1 can be achieved because \( \alpha_1 - \alpha_2 \in \gamma_1 K_{1,2,3} \) and \( \alpha_2 - \alpha_3 \in \gamma_2 K_{1,2,3} \) by the choice of \( \gamma_1, \gamma_2 \). Then \( f(\alpha_1) \) can be recovered after round 1. As a result, round 2 can proceed. After round 2, node 2 recovers \( \text{Tr}(\gamma_2 v_{-1} f(\alpha_2)) \) and node 3 recovers \( \text{Tr}(\gamma_2 v_{-1} f(\alpha_3)) \). Then the two nodes exchange one subymbol with each other in round 3. Actually, the round 3 is reduced to the collaboration phase when repairing two erasures. Since \( \gamma_1 \gamma_2 \in K \), a similar computation as that presented in Section III is involved here for the final recovery.

![Fig. 1. Collaboration Phase (\( \alpha_{-1} \) comes from \( \frac{\gamma_1}{\alpha - \alpha_2} = \sum_{i=1}^{k} a_i \)).](image)

**C. One-round repair of special patterns in \( \Omega_2 \) (Theorem 4)**

Let \( \{\alpha_1, \alpha_2, \alpha_3\} \) be a failure pattern as described in Theorem 4, we continue the settings in Section IV-B by letting \( \gamma_1 = \frac{\alpha_2 - \alpha_3}{\alpha - \alpha_2} \) and \( \gamma_2 = \frac{\alpha_3 - \alpha_1}{\alpha - \alpha_3} \). Then the download phase is the same as in Table III. In the following, we illustrate that the collaboration phase can be accomplished in one round under the additional conditions in Theorem 4, i.e.,

(i) \( t^2 \neq 1 \) mod \( p \).

\[
\begin{align*}
\text{download from } f(\alpha), \\
\text{obtain}
\end{align*}
\]

| node 1 | Tr(\(\frac{f(\alpha)}{\alpha - \alpha_2}\)) | Tr(\(\frac{f(\alpha)}{\alpha - \alpha_2}\)) | Tr(\(\frac{f(\alpha)}{\alpha - \alpha_2}\)) | Tr(\(\frac{f(\alpha)}{\alpha - \alpha_2}\)) |
| node 2 | Tr(\(\frac{f(\alpha)}{\alpha - \alpha_2}\)) | Tr(\(\frac{f(\alpha)}{\alpha - \alpha_2}\)) | Tr(\(\frac{f(\alpha)}{\alpha - \alpha_2}\)) | Tr(\(\frac{f(\alpha)}{\alpha - \alpha_2}\)) |
| node 3 | Tr(\(\frac{f(\alpha)}{\alpha - \alpha_2}\)) | Tr(\(\frac{f(\alpha)}{\alpha - \alpha_2}\)) | Tr(\(\frac{f(\alpha)}{\alpha - \alpha_2}\)) | Tr(\(\frac{f(\alpha)}{\alpha - \alpha_2}\)) |

\[
\begin{align*}
\text{Tr}(\gamma_1 v_{-1} f(\alpha_2)) &= \text{Tr}(\gamma_2 v_{-1} f(\alpha_3)) \\
\text{Tr}(\gamma_2 v_{-1} f(\alpha_3)) &= \text{Tr}(\gamma_2 v_{-1} f(\alpha_3))
\end{align*}
\]

\[
\text{Tr}(\gamma_2 v_{-1} f(\alpha_3)) = \text{Tr}(\gamma_2 v_{-1} f(\alpha_3))
\]
By condition (i) and (ii). Therefore, \( f(\alpha_1) \) can be recovered. Similarly, node 2 and node 3 can also be recovered. Due to limit of space, we omit the details.

V. CONCLUSIONS

We give an improved cooperative repair scheme for Reed-Solomon codes with two or three erasures, removing all the restrictions required in Dau et al.'s work [12]. An interesting problem in the future is to develop a cooperative repair scheme for any number of erasures in Reed-Solomon codes.

ACKNOWLEDGMENTS

The authors appreciate the anonymous reviewers for their valuable suggestions. This work was supported in part by the National Natural Science Foundation of China under Grant 61872353.

REFERENCES

[1] M. Sathiamoorthy, M. Asteris et al., “Xoring elephants: novel erasure codes for big data”, in Proceedings of the VLDB Endowment, vol. 6, pp. 325-336, VLDB Endowment, 2013.
[2] V. Guruswami and M. Wootters, “Repairing Reed-Solomon codes”, in 2016 IEEE 12th Int. Symp. Inform. Theory Proc. (ISIT), 2016, pp. 346-350.
[3] H. Dau, O. Milenkovic, “Optimal repair schemes for some families of full-length Reed-Solomon codes”, in Proc. 2017 IEEE Int. Conf. Comput. Commun. (INFOCOM), 2007, pp. 2000-2008.
[4] I. Tamo, M. Ye, A. Barg, “Optimal repair of Reed-Solomon codes: achieving the cut-set bound,” in Foundations of Computer Science (FOCS), 2017, pp. 216-227.
[5] W. Li, Z. Wang and H. Jafarkhani., “A tradeoff between the sub-packetization size and the repair bandwidth for Reed-Solomon codes”, in Proc. 55th Annual Allerton Conf. Comm Control Comput. (Allerton), 2017, pp. 942-949.
[6] 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, pp. 950-957.
[7] I. Tamo, M. Ye, A. Barg, “The repair problem for Reed-Solomon codes: optimal repair of single and multiple erasures, asymptotically optimal node size”, in press.
[8] J. Mardia, B. Bartan and M. Wootters, “Repairing multiple failures for scalar MDS codes”, IEEE Transactions on Information Theory, vol. 65, no. 5, pp. 2661-2672, 2018.
[9] Y. Hu, Y. Xu, X. Wang, C. Zhan, and P. Li, “Cooperative recovery of distributed storage systems from multiple losses with network coding”, in IEEE J. On Selected Areas in Commun, vol. 28, no. 2, pp. 268-275, Feb. 2010.
[10] M. Ye, A. Barg, “Cooperative repair: constructions of optimal MDS codes for all admissible parameters”, in IEEE Transactions on Information Theory, 2018, 1-1.
[11] H. Dau, I. Duursma, H. M. Kiah, O. Milenkovic, “Repairing Reed-Solomon codes with multiple erasures”, IEEE Transactions on Information Theory, vol. 64, no. 10, pp. 6567-6582, 2018.
[12] Y. Zhang, Z. Zhang, “An improved cooperative repair scheme for Reed-Solomon codes”, arXiv:1901.03142.