Can Linear Minimum Storage Regenerating Codes be Universally Secure?

Sreechakra Goparaju
University of California, San Diego
La Jolla, CA, USA
Email: sgoparaju@ucsd.edu

Salim El Rouayheb
Illinois Institute of Technology
Chicago, IL, USA
Email: salim@iit.edu

Robert Calderbank
Duke University
Durham, NC, USA
Email: robert.calderbank@duke.edu

***Abstract*—We study the problem of making a distributed storage system information-theoretically secure against a passive eavesdropper, and aim to characterize coding schemes that are universally secure for up to a given number of eavesdropped nodes. Specifically, we consider minimum storage regenerating (MSR) codes and ask the following question: For an MSR code where a failed node is repaired using all the remaining nodes, is it possible to simultaneously be optimally secure using a single linear coding scheme? We define a pareto-optimality associated with this simultaneity and show that there exists at least one linear coding scheme that is pareto-optimal.**

I. Introduction

The adoption of cloud computing comes with a legitimate concern for cloud security. Organizations face steep penalties if data is not sufficiently protected against eavesdropping attacks and data theft. This data is typically stored in a redundant fashion on distributed storage systems (DSSs) consisting of multiple, often inexpensive, nodes. With the explosion of data (on the order of zettabytes) in recent years, triple replication, which has been the industry standard for data redundancy, no longer remains viable. New and more sophisticated erasure codes are constantly being explored [1], [2] that achieve the same reliability with a much reduced storage overhead. But these new codes present new challenges, especially when trying to secure a system against eavesdropping attacks.

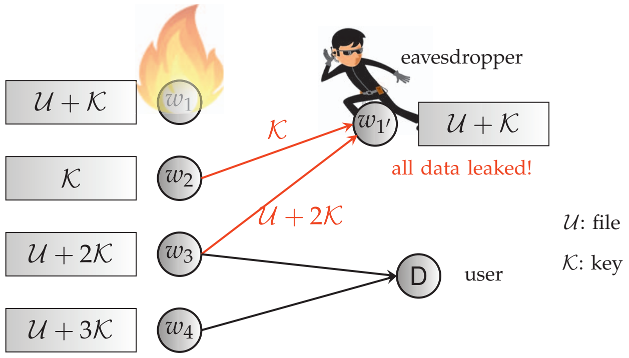

$\mathcal{U}+\mathcal{K}$
$\mathcal{K}$
$\mathcal{U}+2\mathcal{K}$
$\mathcal{U}+3\mathcal{K}$
w_1
w_2
w_3
w_4
eavesdropper
$w_{1'}$
$\mathcal{U}+\mathcal{K}$
$\mathcal{K}$
$\mathcal{U}+2\mathcal{K}$
all data leaked!
$\mathcal{U}$: file
$\mathcal{K}$: key
D
user

Fig. 1. An example of how repairing a DSS can compromise the system security. The original DSS formed of nodes $w_1, \ldots, w_4$ is secured against a single compromised node using a secret sharing scheme or a coset code. However, repairing failed nodes can break the security of the system. For instance, consider the case when node w_1 fails and is replaced by node $w_{1'}$, which is already compromised. The eavesdropper can observe all the data downloaded by node $w_{1'}$ and therefore decode the stored file $\mathcal{U}$.

Example 1. We borrow the DSS example in [3] to illustrate this; see Fig. 1. The user here intends to store a secure file $\mathcal{U}$ of one unit in a DSS of four nodes, such that no eavesdropper which can potentially observe any single node can deduce the contents of $\mathcal{U}$. In the absence of system failures and the corresponding repairs, the user can store $\mathcal{U}$ by *mixing* it with a randomly generated unit sized key $\mathcal{K}$ using the code in the figure. This code can be regarded as a secret sharing scheme [4], as a coset code for the wiretap channel II [5], or as a secure network code for the combination network [6], [7]. Notice that the user can recover $\mathcal{U}$ by downloading the data from any two nodes. This system is optimally redundant under the given circumstances. It however ceases to be so, if for instance, w_1 fails, and w_2 and w_3 aid in its repair to rebuild a replacement node 1. If the replacement node is compromised, the system ends up revealing the entire data to the eavesdropper, including the message file $\mathcal{U}$. Therefore, even if we start with a perfectly secure code, the repair process can break the system security and result in data leakage. □

In this paper, we consider information-theoretic security of a class of codes known as *minimum storage regenerating* (MSR) *codes*. Introduced in [8], these codes not only achieve the optimal redundancy for a given worst-case resilience, but also achieve the optimal repair bandwidth when repairing a failed node.

As Example 1 depicts, the optimal secure file size that can be stored in a system changes with the conditions imposed on the system. The goal here is to study the existence of MSR codes that can be used to secure a file against a given arbitrary number of eavesdropped nodes, and to investigate the optimal file size in each of those circumstances. To the best of our knowledge, this aspect of *universality* has not been studied for secure regenerating codes prior to this. In particular, we focus on the practically significant *linear* coding schemes

978-1-4673-8576-3/15/$31.00 ©2015 IEEE

(scalar and vector), and explore if a nonlinear scheme can potentially outperform them in terms of *perfect secrecy*, that is, when no information is allowed to be revealed to the eavesdropper.

Related Work: Dimakis et al. studied the information-theoretic trade-off between storage overhead and repair bandwidth in DSSs in [8], and defined the class of *regenerating codes*. One end of this trade-off corresponds to codes achieving the optimal storage overhead, and correspond to the aforementioned MSR codes. The study of information-theoretic secrecy for regenerating codes against eavesdroppers and malicious adversaries was initiated by Pawar et al. in [9] and [10]. They provided upper bounds on the system secure capacity and proved its achievability in the bandwidth-limited regime (the other end of the trade-off mentioned above) when all remaining nodes in the DSS help in the repair of a failed node ($d = n-1$). Shah et al. constructed secure codes based on a so-called product-matrix framework in [11] and [12]. These codes achieve the upper bound in [9] for the minimum-bandwidth regime. Rawat et al. gave tighter bounds on the secrecy capacity of a DSS for MSR codes [13] and proved the achievability of their bound for $d = n-1$ and for certain system parameters. Later, Goparaju et al. [3] improved these bounds and proved the optimality and universality of the code constructions in [13] and [14] for any number of eavesdropped nodes, when secrecy capacity is defined over all linear MSR codes.

Constribution: Imagine a multi-dimensional graphical representation, where the projection of a point corresponding to a code along the ℓ-th dimension represents the maximum file size that can be securely stored under the presence of ℓ eavesdropped nodes. As shown in [3], when restricted to linear MSR codes, the achievable region in this representation is cubic; in particular, it has a single optimal point, which corresponds to the existence of a universally secure linear MSR code.

In this paper, we show that in general, when all nodes help in the repair of a failed node, the optimality region may not be a point; instead, there could be a pareto-optimal plane. For instance, a code which is optimal for two eavesdropped nodes may not be optimal for three eavesdropped nodes and vice-versa. We prove however that the universally optimal linear code lies in the pareto-optimal region for all codes — linear or nonlinear. We show the result for the case of two parity nodes and the proof for the case of two eavesdropped nodes[1].

II. Problem Setting

As in [3], we consider an (n, k, d) DSS consisting of n storage nodes which collectively store a universal data file $\mathcal{F}$ of size $M = k\alpha$, such that $\mathcal{F}$ can be recovered by observing the data stored in any k of the nodes. We denote the storage capacity of each node by α. While the DSS satisfies this MDS property, it is also assumed to repair the failure of any single node by using the minimum possible repair bandwidth as described by Dimakis et al. [8]. In other words, we consider the *minimum storage regenerating* (MSR) point of the storage vs. repair bandwidth tradeoff. The repair process involves a transmission of β amount of information by each node in a *helper* set of any d still-operational nodes, where

$$\beta \quad = \quad \frac{\alpha}{d-k+1}. \tag{1}$$

Upon the completion of the repair process, a replacement node containing the *exact* data[2] stored in the previously failed node joins the remaining $n-1$ nodes in the DSS, thus reinstating the DSS to its former configuration.

A. Linear Codes

The most commonly studied scenario is that of *linear* MSR codes. Without loss of generality, we can separate the nodes in the DSS storing an MDS code into systematic and parity nodes. We designate the first k nodes as systematic, where node $i, i \in [k] := \{1, 2, \ldots, k\}$, stores the data vector w_i of column-length α. The data vector w_{k+i} stored in parity node $i, i \in [n-k]$, is given by

$$w_{k+i} \quad = \quad \sum_{j=1}^{k} A_{i,j} w_j, \tag{2}$$

where $A_{i,j} \in \mathbb{F}^{\alpha \times \alpha}$ is the *encoding matrix* corresponding to the parity node $i \in [n-k]$ and the systematic node $j \in [k]$. For an optimal bandwidth repair of a failed systematic node $i \in [k]$, all other nodes transmit β amount of information, i.e., a helper node $j \neq i$ transmits a vector of length β given by $S_j^i w_j$, where $S_j^i \in \mathbb{F}^{\beta \times \alpha}$ is the repair matrix used for the repair of node i by node j. The vector $S_j^i w_j$ can also be interpreted as a projection of w_j onto a subspace of dimension β.

B. All Codes ($\supseteq$ Non-Linear Codes)

To consider all possible codes, we represent the transmitted and stored messages in terms of random variables. Let $\mathcal{M}$ be a random variable uniformly distributed over $\mathbb{F}^M$, representing an incompressible universal data file $\mathcal{F}$ with $H(\mathcal{M}) = M = k\alpha$. Let W_i denote the random variable corresponding to the data w_i stored in node i, $i \in [n]$. Let us assume that a set $\mathcal{D}$ of d helper nodes aid in the repair of node i. We denote the random variable corresponding to the data transmitted by a helper node $m \in \mathcal{D}$ for the repair of node i by $S_m^i(\mathcal{D})$, and the total repair data downloaded by node i by $S_{\mathcal{D}}^i$. We drop the $\mathcal{D}$

[1] We defer the full theorem and proof for the extended version of the paper.

[2] The repair process can also lead to the *functional* repair of a node, in which the new configuration of the DSS retains the MDS and the optimal repair properties, but may not necessarily contain the same data as stored in the old configuration. We only consider *exact repair*.

in the notation when the context is clear. More formally, we define an MSR code in terms of the above random variables.

Definition 1 (MSR Code). *A code stored in an (n,k,d) DSS is defined as a* minimum storage regenerating (MSR) code *(or an* optimal bandwidth MDS code*) if it satisfies the following conditions on the random variables defined above:*

$$\begin{aligned} H(\mathcal{M}) &= k\alpha, & \\ H(W_i) &= \alpha, & \forall\, i \in [n], \\ H(S_i^j) &= \beta, & \forall\, i,j \in [n], i \neq j, \\ H(\mathcal{M}|W_{i_1}, W_{i_2}, \ldots, W_{i_k}) &= 0, & i_1, \ldots, i_k \in [n], \\ H(S_i^j|W_i) &= 0, & \textit{for all } i \neq j \in [n], \\ H(W_j|S_{\mathcal{D}}^j) &= 0, & \end{aligned}$$

for all helper sets $\mathcal{D}$ of cardinality d and $j \in [n]$. Note that $i_1, \ldots, i_k$ are distinct. □

We consider the code to be symmetric[3] in its various entropies, that is, the joint entropy of any subset of storage random variables W_m's and any subset of repair random variables S_i^j's is equal to the joint entropy of the corresponding variables $W_{\sigma(m)}$'s and $S_{\sigma(i)}^{\sigma(j)}$'s, obtained by applying some permutation σ on $[n]$.

C. Eve, the Eavesdropper

We introduce now a passive eavesdropper – Eve – that has access to some ℓ nodes in the DSS, unknown to the legitimate users of the system. We assume that when Eve has access to a node, it can observe both the stored data as well as the data used to generate a replacement node to that node. Notice that it is sufficient to assume the eavesdropping of the repair data.

Let U be a random vector uniformly distributed over $\mathbb{F}^{M^{(s)}}$, representing an incompressible data file with $H(U) = M^{(s)}$. Let $\mathcal{E}$ denote the set of nodes which Eve has access to. Thus, S^i represents the total data revealed to Eve when accessing a node $i \in \mathcal{E}$. Notice that the stored data W_i is a function of the downloaded data S^i. For convenience let us denote $\{W_i : i \in \mathcal{A}\}$ by $W_{\mathcal{A}}$, $\{S_i^j : j \in \mathcal{A}\}$ by $S_i^{\mathcal{A}}$, and $\{S^i : i \in \mathcal{A}\}$ by $S^{\mathcal{A}}$.

The MDS property of the DSS can be written as

$$H\left(U\,|W_{\mathcal{A}}\right) \quad = \quad 0, \tag{3}$$

for all $\mathcal{A} \subseteq [n]$, such that $|\mathcal{A}| = k$. To store a file U on the DSS perfectly secured from the eavesdropper Eve, we have the *perfect secrecy* condition,

$$H\left(U\middle|S^{\mathcal{E}}\right) \quad = \quad H\left(U\right), \tag{4}$$

for all $\mathcal{E} \subseteq [n]$, and $|\mathcal{E}| < k$.

D. Secrecy Capacity and Universality

Defining the secrecy capacity of an (n,k,d) DSS is somewhat subjective. In this paper, we follow the often implicitly followed assumption of a *black box model*, where a certain fixed regenerating code is assumed to be used[4] across the DSS. The *secure file* represented by U is then precoded into the file $\mathcal{F}$ to be stored on the DSS. A black box model is well suited for applications which might not always be susceptible to an eavesdropper, or which use predesigned coded systems. In this paper, the fixed regenerating code guiding the definition is an MSR code.

Definition 2 (Secrecy Capacity for an MSR Code). *Given an (n,k,d) DSS with ℓ compromised nodes (as described above), its* secrecy capacity *$C_s(\alpha)$, is defined to be the maximum file size $H(U)$ that can be stored in the DSS using an MSR code (as defined in Definition 1) for exact repair, such that the reconstruction property and the perfect secrecy condition simultaneously hold, i.e.,*

$$C_s(\alpha) := \sup_{\substack{\mathcal{A},\mathcal{E}: \\ (3),(4)\text{ hold}}} H(U). \tag{5}$$

The black box model also offers a smooth transition into the concept of *universality*.

Definition 3 (Universally Secure MSR Code). *An (n,k,d) MSR code is said to be* universally secure *if it achieves the secrecy capacity, as defined in Definition 2, simultaneously, for all values of $\ell \in [k]$.*

E. The Question

We ask the following question: *For an (n,k,d) DSS with $d = n-1$ and a storage capacity of α per node, can we obtain a linear MSR code that is universally secure?*

Define the *linear coding secrecy capacity for an MSR code*, $C_{s,\text{linear}}(\alpha)$, of an (n,k,d) DSS with ℓ compromised nodes, as in Definition 2, but instead over all *linear* MSR codes. It was shown in [3] that for an $(n,k,d=n-1)$ DSS, the linear coding secrecy capacity for an MSR code, such that *any systematic node is exact repairable*, is achievable for $\alpha = (n-k)^k$ and is given by

$$C_{s,\text{linear}}(\alpha) \quad = \quad (k-\ell)\left(1 - \frac{1}{n-k}\right)^{\ell}\alpha. \tag{6}$$

Moreover, the capacity is achievable for all ℓ. In other words, there does exist a linear MSR code that is universally secure in the sense of linear coding secrecy capacity.

III. Results

The answer to the question in Section II-E is dependent on the amount (and the geometry) of information contained in the repair messages emanating from a single

[3] To read more about why such an assumption is justifiable, refer to the paper by Tian [15].

[4] Recently, Tandon et al. [16] have looked at a non black-box version of the security problem.

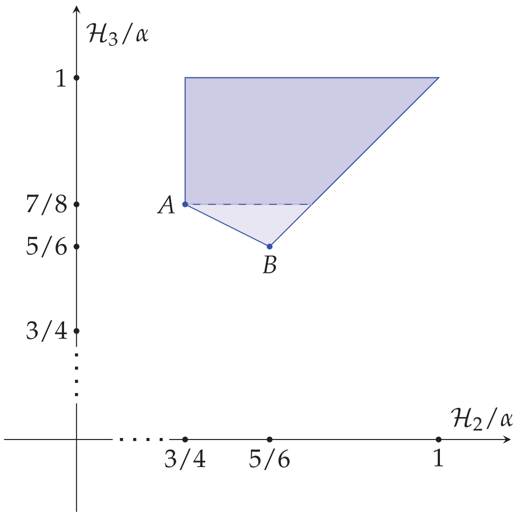


Fig. 2. The converse regions for $\mathcal{H}_2$ and $\mathcal{H}_3$ for linear (darker shade of blue) and all coding schemes (the entire blue quadrangle). All achievable codes must lie in these regions respectively. The point A is known to be achievable [3] and is the only pareto-optimal point for linear codes. It is unknown whether the pareto-optimal line A—B is achievable for some nonlinear code.

node. This is based on the simple argument that $H(U)$ is upper bounded by the total information in the DSS, $H(\mathcal{M}) = k\alpha$, minus the total information available to the eavesdropper if the user knew the locations of the nodes $\mathcal{E}$ that are eavesdropped, that is,

$$\begin{aligned} H(U) &\le k\alpha - H(S^{\mathcal{E}}), \\ &\le k\alpha - H(W_{\mathcal{E}}) - (k-\ell)H(S_i^{\mathcal{E}}), \\ &= (k-\ell)(\alpha - H(S_i^{\mathcal{E}})), \end{aligned} \tag{7}$$

where the last two equalities are obtained by splitting the eavesdropped information into two parts: the data vectors in the eavesdropped nodes ($W^{\mathcal{E}}$), and the data transmitted from the non-eavesdropped systematic nodes to the eavesdropped nodes ($(k-\ell)(H(S_i^{\mathcal{E}})$, where the second term is due to entropic symmetry.

Note that for linear codes, the upper bound, as given in [3], is due to the geometry of the corresponding *subspaces* $S_i^{\mathcal{E}}$,

$$H(U) \quad \le \quad (k-\ell)(\alpha - \mathsf{dim}(S_i^{\mathcal{E}})).$$

Example 2. Consider an $(n,k,d) = (5,3,4)$ MSR code. For linear codes, we have from [3] and [17], the following two inequalities,

$$H\left(S_i^{\mathcal{A}}\right) \ \ge \ \frac{3}{4}\alpha, \qquad H\left(S_i^{\mathcal{B}}\right) \ \ge \ \frac{7}{8}\alpha,$$

where $\mathcal{A}, \mathcal{B} \subseteq [k]\backslash\{i\}$ and $|\mathcal{A}| = 2$ and $|\mathcal{B}| = 3$.

On the other hand, we show below that we have the following converse region for all possible codes (assuming entropic symmetry with respect of the permutations of $[n]$),

$$H\left(S_i^{\mathcal{A}}\right) \ \ge \ \frac{3}{4}\alpha, \qquad 2H\left(S_i^{\mathcal{B}}\right) + H\left(S_i^{\mathcal{A}}\right) \ \ge \ \frac{5}{2}\alpha,$$

where $\mathcal{A}, \mathcal{B} \subseteq [k]\backslash\{i\}$ and $|\mathcal{A}| = 2$ and $|\mathcal{B}| = 3$. Fig. 2 represents these regions pictorially. We represent by $\mathcal{H}_t$ the entropy corresponding to the repair messages for t nodes, that is, $\mathcal{H}_t = H\left(S_i^{\mathcal{A}}\right)$, where $i \notin \mathcal{A}$. The regions take into account the additional constraints: $\alpha \ge \mathcal{H}_3 \ge \mathcal{H}_2$. From (7), it can be seen that to achieve a high secrecy capacity, we need the values of $H(S_i^{\mathcal{A}})$ to be as low as possible. We see that there is a pareto-optimal boundary to the converse region for the case of nonlinear codes. For the special case of linear codes, this boundary constitutes just a single point which is known to be achievable by the results in [13] and the relevant citations therein. □

We arrive at the main theorem of the paper.

Theorem 1. *Consider an MSR code stored in an $(n,k,d = n-1)$ DSS having two parity nodes, $n = k+2$, with nodes having a storage capacity α. Under the assumption of entropic symmetry, for $\ell \le k$, we have,*

$$2\mathcal{H}_\ell + \sum_{i=1}^{\ell-1} \mathcal{H}_i \quad \ge \quad \ell\alpha. \tag{8}$$

Before proving this, notice that a linear MSR code exists, for example, a zigzag code [14], which satisfies the set of equations in (8) with equality, and thus lies in the pareto-optimal region defined by Theorem 1 for the vector $(\mathcal{H}_2, \ldots, \mathcal{H}_k)$, thus semi-answering the question in Section II-E.

Proof Sketch of Theorem 1. For lack of space, the proof sketch is given for the case of $\ell = 2$. We start with proving the result for the simplest case of $(n,k,d) = (4,2,3)$. Consider the random variables in bold font in Figure 3(a). The information in these transmitted messages is sufficient to reconstruct the universal data file $\mathcal{M}$ of size 2α. To show this, we have

$$\begin{aligned} H\left(S^1_{\{2,3,4\}}, S^2_{\{3,4\}}\right) &= H\left(S^1_{\{2,3,4\}}, W_1, S^2_{\{3,4\}}\right), \\ &= H\left(S^1_{\{2,3,4\}}, W_1, S^2_{\{1,3,4\}}\right), \\ &= H\left(S^1_{\{2,3,4\}}, W_1, S^2_{\{1,3,4\}}, W_2\right), \\ &= H\left(W_1, W_2\right) \ = \ 2\alpha, \end{aligned}$$

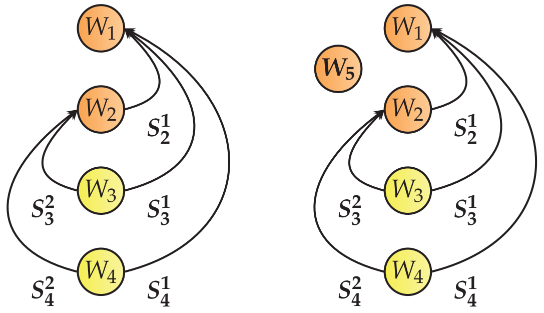
where each equality is either adding a function of the variables in the previous equality or removing the variables which are functions of the remaining variables in the current equality, thus keeping constant the entropy.

On the other hand, we can expand the entropy of the bold messages in the following manner.

$$\begin{aligned} H\left(S^1_{\{2,3,4\}}, S^2_{\{3,4\}}\right) &= H\left(S_3^{\{1,2\}}, S_4^{\{1,2\}}, S_2^1\right), \\ &\le H\left(S_3^{\{1,2\}}\right) + H\left(S_4^{\{1,2\}}\right) + H\left(S_2^1\right). \end{aligned}$$

Thus, from the above sets of equations, and the fact that the entropy of a repair message is β, we have

$$H\left(S_3^{\{1,2\}}\right) + H\left(S_4^{\{1,2\}}\right) \quad \ge \quad 2\alpha - \frac{1}{2}\alpha.$$



(a) $(n, k, d) = (4, 2, 3)$ (b) $(n, k, d) = (5, 3, 4)$

Fig. 3. Information flows relevant for the proof of Theorem 1. Notice that the representation of parity nodes (in yellow) is no longer relevant and is only shown here as a memory aid for recalling n and k.

For the particular case when the entropies are symmetric with respect to permutations of $\{1, 2, 3, 4\}$, we have $H\left(S_3^{\{1,2\}}\right) \geq 3\alpha/4$.

In general, for an $(n, k, d) = (k + 2, k, n - 1)$ DSS, the proof roughly remains the same. Figure 3(b) demonstrates the case for $n = 5$. When $n > 2$, the difference is that the remaining stored message variables, $\{W_5, \ldots, W_n\}$, are in bold too. Without loss of generality,

$$\begin{aligned} & H\left(S^1_{\{2,3,4\}}, S^2_{\{3,4\}}, W_{\{5,6,\ldots,n\}}\right) \\ = \; & H\left(S^1_{\{2,3,4\}}, W_1, S^2_{\{3,4\}}, W_{\{5,6,\ldots,n\}}\right), \\ = \; & H\left(S^1_{\{2,3,4\}}, W_1, S^2_1, S^2_{\{3,4\}}, W_{\{5,6,\ldots,n\}}\right), \\ = \; & H\left(W_1, W_2, W_{\{5,6,\ldots,n\}}\right), \\ = \; & H(\mathcal{M}) \; = \; k\alpha, \text{ and,} \end{aligned}$$

$$\begin{aligned} & H\left(S^1_{\{2,3,4\}}, S^2_{\{3,4\}}, W_{\{5,6,\ldots,n\}}\right) \\ = \; & H\left(S_3^{\{1,2\}}, S_4^{\{1,2\}}, S_2^1, W_{\{5,6,\ldots,n\}}\right), \\ \leq \; & H\left(S_3^{\{1,2\}}\right) + H\left(S_4^{\{1,2\}}\right) + H\left(S_2^1\right) + H\left(W_{\{5,6,\ldots,n\}}\right). \end{aligned}$$

Again, from the above sets of equations, and the fact that the entropy of a repair message is β, we have

$$H\left(S_3^{\{1,2\}}\right) + H\left(S_4^{\{1,2\}}\right) \quad \geq \quad k\alpha - \frac{1}{2}\alpha - (k-2)\alpha.$$

For the particular case when the entropies are symmetric with respect to permutations of $[n]$, we have $H\left(S_3^{\{1,2\}}\right) \geq 3\alpha/4$. □

IV. Concluding Remarks

We initiate the search for obtaining an MSR code that achieves secrecy capacity for any arbitrary number of eavesdropped nodes and show that a linear MSR code lies on the appropriately defined pareto-optimal region. Though the question remains as to whether there exist other pareto-optimal codes or whether the region can be tightened to obtain a single optimal point, the analysis gives an interesting insight into the similarity between the intersection results of linear subspaces and nonlinear repair variables. A further exploration of the latter connection might prove useful to solve other open problems in non-linear regenerating codes.

References

[1] C. Huang, H. Simitci, Y. Xu, A. Ogus, B. Calder, P. Gopalan, J. Li, and S. Yekhanin, "Erasure Coding in Windows Azure Storage," in *Proceedings of the 2012 USENIX Annual Technical Conference (ATC)*, Boston, MA, 2012.

[2] M. Sathiamoorthy, M. Asteris, D. Papailiopoulos, A. G. Dimakis, R. Vadali, S. Chen, and D. Borthakur, "XORing Elephants: Novel Erasure Codes for Big Data," in *Proceedings of the VLDB Endowment*, vol. 6, March 2013, pp. 325–336.

[3] S. Goparaju, S. E. Rouayheb, R. Calderbank, and H. V. Poor, "Data Secrecy in Distributed Storage Systems under Exact Repair," in *Proceedings of IEEE International Symposium on Network Coding (NetCod)*, June 2013.

[4] A. Shamir, "How to Share a Secret," in *Communications of the ACM*, vol. 22, no. 11, 1979, pp. 612–613.

[5] L. H. Ozarow and A. D. Wyner, "Wire-Tap Channel-II," in *AT&T Bell Laboratories Technical Journal*, vol. 63, no. 10, December 1984, pp. 2135–2157.

[6] N. Cai and R. W. Yeung, "Secure Network Coding on a Wiretap Network," in *IEEE Transactions on Information Theory*, vol. 57, no. 1, January 2011, pp. 424–435.

[7] S. El Rouayheb, E. Soljanin, and A. Sprintson, "Secure Network Coding for Wiretap Networks of Type II," in *IEEE Transactions on Information Theory*, vol. 58, no. 3, March 2012, pp. 1361–1371.

[8] A. G. Dimakis, P. B. Godfrey, Y. Wu, M. Wainwright, and K. Ramchandran, "Network Coding for Distributed Storage Systems," in *IEEE Transactions on Information Theory*, vol. 56, no. 9, September 2010, pp. 4539–4551.

[9] S. Pawar, S. El Rouayheb, and K. Ramchandran, "On Secure Distributed Data Storage under Repair Dynamics," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, June 2010, pp. 2543–2547.

[10] ——, "Securing Dynamic Distributed Storage Systems against Eavesdropping and Adversarial Attacks," in *IEEE Transactions on Information Theory*, vol. 57, October 2011, pp. 6734–6753.

[11] K. V. Rashmi, N. B. Shah, and P. V. Kumar, "Regenerating Codes for Errors and Erasures in Distributed Storage," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, July 2012, pp. 1202–1206.

[12] N. B. Shah, K. V. Rashmi, and P. V. Kumar, "Information-Theoretically Secure Regenerating Codes for Distributed Storage," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM)*, December 2011, pp. 1–5.

[13] A. S. Rawat, O. O. Koyluoglu, N. Silberstein, and S. Vishwanath, "Optimal Locally Repairable and Secure Codes for Distributed Storage Systems," in *IEEE Transactions on Information Theory*, January 2014, pp. 212–236.

[14] I. Tamo, Z. Wang, and J. Bruck, "Zigzag Codes: MDS Array Codes With Optimal Rebuilding," in *IEEE Transactions on Information Theory*, vol. 59, March 2013, pp. 1597–1616.

[15] C. Tian, "Rate Region of the $(4, 3, 3)$ Exact-Repair Regenerating Codes," in *Proceedings of IEEE International Symposium on Information Theory (ISIT)*, Jul. 2013, pp. 1426–1430.

[16] R. Tandon, S. Amuru, T. C. Clancy, and R. M. Buehrer, "Towards Optimal Secure Distributed Storage Systems with Exact Repair," in *arxiv.org*, September 2013.

[17] S. Goparaju, I. Tamo, and R. Calderbank, "An Improved Sub-Packetization Bound for Minimum Storage Regenerating Codes," in *IEEE Transactions on Information Theory*, vol. 60, no. 5, 2014, pp. 2770–2779.