Secret Sharing with Optimal Decoding and Repair Bandwidth

Wentao Huang and Jehoshua Bruck

Abstract

This paper studies the communication efficiency of threshold secret sharing schemes. We construct a family of Shamir's schemes with asymptotically optimal decoding bandwidth for arbitrary parameters. We also construct a family of secret sharing schemes with both optimal decoding bandwidth and optimal repair bandwidth for arbitrary parameters. The construction also leads to a family of regenerating codes allowing centralized repair of multiple node failures with small sub-packetization.

I. Introduction

In a threshold secret sharing scheme, a secret message is encoded into n symbols so that 1) (Reliability) the message can be decoded from any $n-r$ symbols and 2) (Secrecy) any z symbols reveal no information about the message. Secret sharing schemes allow reliable and secure storage and communication of information, and are building blocks of many other important security protocols.

One of the most studied problems regarding secret sharing schemes is their storage efficiency. For threshold scheme, the optimal storage efficiency (i.e., rate) is achieved if the size of the message is $n-r-z$ symbols. The well-known Shamir's scheme [11] achieves optimal rate for the case of $n-r-z=1$, i.e., the message is a single symbol. However, Shamir's scheme has poor communication efficiency, in the sense that to decode the single message symbol (using the standard decoding algorithm), one needs to download $n-r$ encoded symbols, referred to as the *decoding bandwidth*. The decoding bandwidth of secret sharing schemes are studied in, e.g., [12], [6], [1], where lower bounds are obtained and optimal schemes achieving the bounds are proposed. However, Shamir's scheme is extensively used due to its simplicity and it remains an important problem that whether it is possible to reduce the decoding bandwidth of Shamir's scheme.

In this paper we construct a family of Shamir's schemes of arbitrary parameters that is asymptotically optimal in the decoding bandwidth. Specifically, as opposed to the original $n-r$ symbols, the decoding

bandwidth reduces to $n/(1+r)$ symbols as the field size increases. The decoding algorithm follows the framework proposed in [5] of interpolating polynomials by querying partial polynomial evaluation. Our scheme is inspired by the family of Reed-Solomon (RS) codes constructed in [14] which has asymptotically optimal repair bandwidth. Decoding Shamir's scheme and repairing RS codes are related because both are essentially the problem of determining the evaluation of a polynomial at a point, given the evaluations of the polynomial at other points. Decoding Shamir's scheme in this sense is a simpler problem because it requires finding the evaluation at a single point, while repairing RS codes requires finding the evaluation at different points, depending on which symbol is being repaired. This simplification allows us to greatly reduce the field size. Specifically, while the codes in [14] requires the extension degree of the field to be exponential in n, our scheme only requires an extension degree of $O(n(n-z)^3)$, which makes it quite practical.

In addition to the decoding bandwidth, another important aspect of communication efficiency is the repair bandwidth. Repair bandwidth of secret sharing schemes are studied in, e.g., [8], [9], where lower bounds are obtained and optimal schemes are proposed. A natural and important question is that whether it is possible to construct a scheme with both optimal decoding and repair bandwidth. Rawat et al. [10], by observing that decoding the secret sharing scheme can be viewed as repairing the message symbols in a regenerating code, propose schemes that are bandwidth efficient in both repair and decoding. However, their construction is quite restricted in parameters if rate-optimality is required.

In this paper, by formalizing the connection between regenerating codes and secret sharing schemes, and then applying the connection to the regenerating codes in [13], we obtain rate-optimal schemes with optimal decoding and repair bandwidth for arbitrary parameters. However, the schemes are not practical as they require an extremely large level of sub-packetization that is doubly exponential in n. To reduce sub-packetization, we use the fact that all message symbols are "repaired" together in a centralized manner during decoding. Therefore we essentially need a regenerating code that allows a hybrid mode of repair: centralized repair of the set of message symbols, and individual repair of the remaining symbols. We generalize the codes in [13] to this model which result in secret sharing schemes with a much smaller sub-packetization level due to the centralized repair pattern. Our generalization also leads to a family of regenerating codes that supports centralized repair of groups of nodes of flexible sizes with reduced sub-packetization level, which is a result of separate interest.

II. General Construction Framework

In an (n, k, r, z) secret sharing scheme, a message $\boldsymbol{m}$ of length k over some alphabet $\mathcal{A}$ is encoded into n shares over $\mathcal{A}$ such that: 1) $\boldsymbol{m}$ can be decoded from any subset of shares of size $\geq n - r$. 2) Any subset of shares of size $\leq z$ do not reveal information on $\boldsymbol{m}$. For the setting of distributed storage, each node will store one share. We focus on schemes with optimal rate, i.e., with $k = n - r - z$. Assume in general that an element of $\mathcal{A}$ is a vector of l symbols over some base alphabet $\mathcal{B}$. l is often referred to as the level of sub-packetization. Let d_2 be the number of nodes participating in decoding $\boldsymbol{m}$, the *decoding bandwidth* is the number of symbols over $\mathcal{B}$ to be transmitted from the d_2 nodes to the decoder. The optimal decoding bandwidth equals $\frac{kd_2l}{d_2-z}$ [6], referred to as the d_2-optimal decoding bandwidth. Similarly, in the case that h nodes are failed, let d_1 be the number of available nodes participating in the repair process, then the *repair bandwidth* is the number of symbols over $\mathcal{B}$ to be transmitted from the d_1 nodes. The (d_1, h)-optimal repair bandwidth equals $\frac{hd_1l}{h+d_1-k-z}$ [3]. When $h = 1$, we refer to it as the d_1-optimal repair bandwidth. We first formalize a connection between MDS codes and secret sharing schemes. The following theorem is a generalization of the result in [2] to the case of $k > 1$.

Theorem 1. *For any k, z, let $k' = k + z$ and $n' > 2k + z$, an $[n', k']$ MDS code $\mathcal{C}$ implies a $(n = n' - k, k, r = n - k', z)$ secret sharing scheme $\mathcal{S}$, obtained as follows: Encode $\mathcal{C}$ systematically so that among the k' information nodes, k of them store secret information and the remaining z nodes store uniformly distributed keys. Then discard the k nodes storing the secret information. The remaining n nodes store the n shares of $\mathcal{S}$.*

Proof. Since the minimum distance of $\mathcal{C}$ is $n' - k' + 1$, the minimum distance of the codewords of $\mathcal{S}$ is $n' - k' + 1 - k = n - k' + 1$. This proves the reliability of $\mathcal{S}$. Denote the secret information by a length-k vector $\boldsymbol{m}$ and denote the keys by a length-z vector $\boldsymbol{u}$. Let G be the encoding matrix of $\mathcal{S}$, i.e., the shares are $(\boldsymbol{m}, \boldsymbol{u})G = \boldsymbol{m}G_{\text{up}} + \boldsymbol{u}G_{\text{low}}$. Then to prove the secrecy of $\mathcal{S}$ it suffices to prove that any subset of z entries of $\boldsymbol{u}G_{\text{low}}$ is uniformly distributed. Let $G' = [I \mid P']$ be the systematic generator matrix of $\mathcal{C}$, where I is the identity matrix of order k'. Then by construction $G_{\text{low}} = [I \mid P]$ where I is the identity matrix of order z, and P is a submatrix of P'. Since $\mathcal{C}$ is MDS, by [7, Ch.11, Theorem 8], every square submatrix of P' is non-singular and therefore every square submatrix of P is also non-singular. Again by [7] this implies that G_{low} is the systematic generator matrix of a MDS code and therefore $\boldsymbol{u}$ can be decoded from any subset of z entries of $\boldsymbol{u}G_{\text{low}}$. Since $\boldsymbol{u}$ is uniformly distributed, it implies that any subset of z entries of $\boldsymbol{u}G_{\text{low}}$ is uniformly distributed, proving the theorem. □

By Theorem 1 we can construct secret sharing schemes from MDS codes. The next result formally connects the decoding and repair bandwidth of the secret sharing schemes to those of the MDS codes. A similar observation is made in [10].

Theorem 2. *If $\mathcal{C}$ allows 1) repair of individual non-discarded nodes from any $d_1 \leq n-1$ of the remaining available nodes with optimal bandwidth and 2) simultaneous repair of all k discarded nodes from any $d_2 \leq n$ of the available nodes with optimal bandwidth, then the resulting secret sharing scheme $\mathcal{S}$ achieves d_1-optimal repair bandwidth and d_2-optimal decoding bandwidth.*

Proof. For $\mathcal{S}$, 1) corresponds to the operation of repairing individual nodes and 2) corresponds to the operation of decoding. First consider 1). Note that the bandwidth of repairing a node in $\mathcal{S}$ from d_1 nodes is lower bounded by the optimal bandwidth of repairing a node from d_1 nodes in an $[n, k']$ MDS code (because $\mathcal{S}$ is a $[n, k']$ MDS code) which equals $\frac{d_1 l}{1+d_1-k'}$ symbols and is achieved by $\mathcal{C}$ by hypothesis.

Now consider 2), the optimal bandwidth to repair k nodes in $\mathcal{C}$ from d_2 nodes is $\frac{kd_2 l}{k+d_2-k'} = \frac{kd_2 l}{d_2-z}$ symbols which matches the lower bound on decoding bandwidth. This shows that $\mathcal{S}$ achieves d_2-optimal decoding bandwidth. □

By Theorem 1 and Theorem 2, we can immediately obtain secret sharing schemes with optimal repair and decoding bandwidth from the regenerating codes in [13].

Corollary 1. *For any n, r, z, there exists a rate-optimal $(n, k = n-r-z, r, z)$ secret sharing scheme with d_1-optimal repair bandwidth and d_2-optimal decoding bandwidth, universally for all $k+z \leq d_1 \leq n-1$, and $k+z \leq d_2 \leq n$.*

Proof. Apply Theorem 1 and Theorem 2 to Construction 3 or Construction 6 in [13]. □

We remark that Construction 6 in [13] is both bandwidth-optimal and access-optimal. Therefore the resulting secret sharing scheme not only achieves optimal repair and decoding bandwidth universally, but also achieves optimal *access complexity* universally during repair or decoding.

We note that, however, the above schemes obtained by directly applying Theorems 1 and 2 to Construction 3 or Construction 6 in [13] are hardly practical. This is because the degree of sub-packetization l is prohibitive. Specifically, l grows double exponentially at the speed of $O\left(2^{(k+r)^n}\right)$. In the next section we will discuss constructions of schemes with a much smaller l. We first introduce some notation. Denote $\{1, \cdots, n\}$ by $[n]$, and denote $\{i, i+1, \cdots, j\}$ by $[i:j]$.

Denote the entries of the codewords of $\mathcal{C}$ by C_i, $i = 1, \cdots, n'$. Each C_i is a column vector of length l over a finite field F. Adopting the framework in [13], we define $\mathcal{C}$ by its parity check equations:

$$\mathcal{C} = \{(C_1, \cdots, C_{n'}) : \sum_{i=1}^{n'} A_{t,i} C_i = 0,\ t = 1, \cdots, r'\} \tag{1}$$

where $r' = n' - k'$, and $A_{t,i}$, $t \in [r']$, $i \in [n']$ are $l \times l$ matrices over F. In this paper we consider array codes that resemble the structure of Vandermonde matrix, i.e., we let

$$A_{t,i} = A_i^{t-1}, \quad t \in [r'],\ i \in [n'] \tag{2}$$

where A_i, $i \in [n']$ are $l \times l$ matrices (with the convection that $A^0 = I$). In the next section we will construct $\mathcal{C}$ by designing specific A_i's.

III. Secret Sharing Schemes with Optimal Decoding and Repair Bandwidth

The d_2-optimal decoding bandwidth of a $(n, k = n - r - z, r, z)$ scheme is $\frac{k d_2 l}{d_2 - z}$ symbols, implying that each of the d_2 nodes participating in decoding will transmit a fraction of $\frac{k}{d_2 - z}$ of the symbols that they store. We start with the case that $\frac{k}{d_2 - z} = \frac{1}{\rho}$, where $\rho \geq 1$ is an integer, which allows a simplified scheme. The following construction is a generalization of Construction 2 in [13].

Construction 1. *Consider any $n, r, z, k = n - r - z$ and $k + z \leq d_1 \leq n - 1$, $k + z \leq d_2 \leq n$ such that $\frac{k}{d_2 - z} = \frac{1}{\rho}$. Let $n' = n + k$, $k' = k + z$, $r' = n' - k'$ and F be a finite field of size $|F| \geq k\rho + ns$, where $s = d_1 + 1 - k'$. Let $\{\lambda_{i,j}\}_{i \in [n], j = 0, \cdots, s-1} \bigcup \{\lambda_{n+i,j}\}_{i \in [k], j = 0, \cdots, \rho - 1}$ be distinct elements in F. Consider the code family given by (1) and (2), where $l = \rho s^n$ and*

$$A_i = \sum_{a=0}^{l-1} \lambda_{i,a_i} e_a e_a^T, \quad i = 1, \cdots, n$$

$$A_{n+i} = \sum_{a=0}^{l-1} \lambda_{n+i,a_{n+1}} e_a e_a^T \quad i = 1, \cdots, k.$$

Here $\{e_a : a = 0, \cdots, l-1\}$ is the standard basis of F^l and we represent a using the $(n+1)$-digit notation $a = (a_{n+1}, a_n, \cdots, a_1)$, where $a_{n+1} \in \{0, \cdots, \rho - 1\}$ and $a_i \in \{0, \cdots, s-1\}$, for all $i \in [n]$.

Note that A_i, $i \in [n']$ are diagonal matrices and we can expand the parity-check equations (1) coordinatewise. Let $c_{i,a}$ denote the a-th entry of C_i, we have,

$$\sum_{i=1}^{n} \lambda_{i,a_i}^t c_{i,a} + \sum_{i=1}^{k} \lambda_{n+i,a_{n+1}}^t c_{n+i,a} = 0, \tag{3}$$

for $a \in \{0, \cdots, l-1\},\ t \in \{0, \cdots, r'-1\}$.

Lemma 1. *The array code $\mathcal{C}$ given by Construction 1 is MDS.*

Proof. Writing (3) in matrix form, for all $a = 0, \cdots, l-1$, we have

$$\begin{bmatrix} 1 & \cdots & 1 & 1 & \cdots & 1 \\ \lambda_{1,a_1} & \cdots & \lambda_{n,a_n} & \lambda_{n+1,a_{n+1}} & \cdots & \lambda_{n',a_{n+1}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \lambda_{1,a_1}^{r'-1} & \cdots & \lambda_{n,a_n}^{r'-1} & \lambda_{n+1,a_{n+1}}^{r'-1} & \cdots & \lambda_{n',a_{n+1}}^{r'-1} \end{bmatrix} \begin{bmatrix} c_{1,a} \\ c_{2,a} \\ \vdots \\ c_{n',a} \end{bmatrix} = 0. \tag{4}$$

Therefore any r' columns of the parity check matrix in (4) are linearly independent, implying that from any $n' - r' = k'$ elements of $\{c_{1,a}, \cdots, c_{n',a}\}$ we can recover the whole set. This shows that we can recover the set $\{C_1, \cdots, C_{n'}\}$ from any of its k' elements. Hence $\mathcal{C}$ is an MDS array code. □

Lemma 2. *The array code $\mathcal{C}$ given by Construction 1 attains optimal bandwidth when 1) repairing a single node i, $i \in [n]$, from any d_1 nodes, and 2) repairing the set of nodes $\{n+1, \cdots, n'\}$ from any d_2 nodes.*

Proof. Let $a(i,u) = (a_{n+1}, \cdots, a_{i+1}, u, a_{i-1}, \cdots, a_1)$. To prove the first statement of the theorem we claim that for $i \in [n]$ and $a = 0, \cdots, l-1$, the set of entries $\{c_{i,a(i,0)}, \cdots, c_{i,a(i,s-1)}\}$ of C_i can be recovered from any subset of d_1 symbols of the following set of $n'-1$ symbols over F:

$$\mu_{j,i}^{(a)} \triangleq \sum_{u=0}^{s-1} c_{j,a(i,u)}, \quad j \in [n'] \backslash i$$

The claim implies that each of the d_1 nodes only needs to send one symbol in order to recover s symbols in C_i, achieving the optimal bandwidth. To prove the claim, by (3), for any $i \in [n]$, $a = 0, \cdots, l-1$, $t = 0, \cdots, r'-1$ and $u = 0, \cdots, s-1$, we have:

$$\lambda_{i,u}^t c_{i,a(i,u)} + \sum_{j \in [n] \backslash \{i\}} \lambda_{j,a_j}^t c_{j,a(i,u)} + \sum_{j=1}^{k} \lambda_{n+j,a_{n+1}}^t c_{n+j,a(i,u)} = 0. \tag{5}$$

Assume without loss of generality that $i = 1$. By summing (5) over $u = 0, \cdots, s-1$ we obtain

$$\begin{bmatrix} 1 & \cdots & 1 \\ \lambda_{1,0} & \cdots & \lambda_{1,s-1} \\ \vdots & \vdots & \vdots \\ \lambda_{1,0}^{r'-1} & \cdots & \lambda_{1,s-1}^{r'-1} \end{bmatrix} \begin{bmatrix} c_{1,a(1,0)} \\ \vdots \\ c_{1,a(1,s-1)} \end{bmatrix} = - \begin{bmatrix} 1 & \cdots & 1 & 1 & \cdots & 1 \\ \lambda_{2,a_2} & \cdots & \lambda_{n,a_n} & \lambda_{n+1,a_{n+1}} & \cdots & \lambda_{n',a_{n+1}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \lambda_{2,a_2}^{r'-1} & \cdots & \lambda_{n,a_n}^{r'-1} & \lambda_{n+1,a_{n+1}}^{r'-1} & \cdots & \lambda_{n',a_{n+1}}^{r'-1} \end{bmatrix} \begin{bmatrix} \mu_{2,1}^{(a)} \\ \vdots \\ \mu_{n',1}^{(a)} \end{bmatrix} \tag{6}$$

Note that in (6), there are r' equations and $s+(n'-1-d_1)$ unknown variables (s from the L.H.S. and $n'-1-d_1$ from the R.H.S.). Moreover, $s+(n'-1-d_1)=(d_1+1-k')+(n'-1-d_1)=n'-k'=r'$. Therefore the number of equations equal the number of variables. Below we show that it is indeed possible to solve all variables from (6). We follow an approach similar to that in [13]. Define polynomials $p_i(x)=x^i\prod_{u=0}^{s-1}(x-\lambda_{1,u})$, $i=0,\cdots,r'-s-1$. Let $p_i(x)=\sum_{j=0}^{r'-1}p_{i,j}x^j$, and define matrix $P\triangleq(p_{i,j})_{i\in[0:r'-s-1],j\in[0:r'-1]}$. Then by construction, multiply P on the left to (6) we have

$$P\begin{bmatrix} 1 & \cdots & 1 & 1 & \cdots & 1 \\ \lambda_{2,a_2} & \cdots & \lambda_{n,a_n} & \lambda_{n+1,a_{n+1}} & \cdots & \lambda_{n',a_{n+1}} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \lambda_{2,a_2}^{r'-1} & \cdots & \lambda_{n,a_n}^{r'-1} & \lambda_{n+1,a_{n+1}}^{r'-1} & \cdots & \lambda_{n',a_{n+1}}^{r'-1} \end{bmatrix}\begin{bmatrix} \mu_{2,1}^{(a)} \\ \vdots \\ \mu_{n',1}^{(a)} \end{bmatrix}=0. \tag{7}$$

The product of the first two terms in (7) equals $Q=(q_{i,j})_{i\in[r'-s],j\in[n'-1]}$, where

$$q_{i,j}=p_0(\lambda_{j+1,a_{j+1}})\lambda_{j+1,a_{j+1}}^{i-1},\quad i\in[r'-s],\ j\in[n-1] \tag{8}$$

$$q_{i,n-1+j}=p_0(\lambda_{n+j,a_{n+1}})\lambda_{j+1,a_{n+1}}^{i-1},\quad i\in[r'-s],\ j\in[k] \tag{9}$$

Therefore Q is a Vandermonde matrix in which each column is scaled by a non-zero constant. Therefore any $r'-s$ columns of Q are linearly independent, implying that from any $n'-1-(r'-s)=d_1$ elements of $\{\mu_{2,1}^{(a)},\cdots,\mu_{n',1}^{(a)}\}$ we can recover the whole set. And then we can recover $\{c_{1,a(1,0)},\cdots,c_{1,a(1,s-1)}\}$ from (6). This proves the claim and the first statement of the theorem.

We now prove the second statement and claim that for any $a=0,\cdots,l-1$, the set of entries $\{c_{n+i,a(n+1,j)}:i\in[k],j\in[0:\rho-1]\}$ can be recovered from any subset of d_2 elements of the set $\{\mu_{j,n+1}^{(a)}:j\in[n]\}$. In other words, each of the d_2 nodes only needs to send one symbol in order to decode ρ symbols, achieving the optimal decoding bandwidth. To prove the claim, from (3) we have:

$$\sum_{j=1}^{n}\lambda_{j,a_j}^{t}c_{j,a(j,u)}+\sum_{j=1}^{k}\lambda_{n+j,u}^{t}c_{n+j,a(n+1,u)}=0, \tag{10}$$

Summing over u we have

$$\begin{bmatrix} 1 & \cdots & 1 & \cdots & 1 & \cdots & 1 \\ \lambda_{n+1,0} & \cdots & \lambda_{n+1,\rho-1} & \cdots & \lambda_{n+k,0} & \cdots & \lambda_{n+k,\rho-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \lambda_{n+1,0}^{r'-1} & \cdots & \lambda_{n+1,\rho-1}^{r'-1} & \cdots & \lambda_{n+k,0}^{r'-1} & \cdots & \lambda_{n+k,\rho-1}^{r'-1} \end{bmatrix} \begin{bmatrix} c_{n+1,a(n+1,0)} \\ \vdots \\ c_{n+1,a(n+1,\rho-1)} \\ \vdots \\ c_{n+k,a(n+1,0)} \\ \vdots \\ c_{n+k,a(n+1,\rho-1)} \end{bmatrix} = \begin{bmatrix} \mu_{1,n+1}^{(a)} \\ \vdots \\ \mu_{n,n+1}^{(a)} \end{bmatrix} \tag{11}$$

Note that in (11) we have r' equations and $k\rho+n-d_2$ unknown variables ($k\rho$ from the L.H.S. and $n-d_2$ from the R.H.S.). But $k\rho + n - d_2 = r'$ and the number of equations equals the number of variables. Similar to the way that we treat (6), we can recover $\{c_{n+i,a(n+1,j)} : i \in [k], j \in [0:\rho-1]\}$ by solving (11). This proves the claim and the second statement of the theorem. □

By Lemmas 1, 2 and Theorem 2, we have:

Theorem 3. *The secret sharing scheme obtained by applying Theorem 1 to the code given in Construction 1, where the last k nodes are discarded, is a $(n, k = n-r-z, r, z)$ scheme with d_1-optimal repair bandwidth and d_2-optimal decoding bandwidth.*

We now generalize Construction 1 to the case of arbitrary $\frac{k}{d_2-z}$.

Construction 2. *For any $n, r, z, k = n-r-z$, $k+z \le d_1 \le n-1$ and $k+z \le d_2 \le n$, let $n' = n+k$, $k' = k+z$, $r' = n'-k'$ and $s = d_1+1-k'$. Let $\theta = \gcd(k, d_2-z)$, $\tau = \frac{k}{\theta}$, $\rho = \frac{d_2-z}{\theta}$ and $\delta = \rho-\tau$. Let F be a finite field of size $|F| \ge sn + \sum_{i=1}^{\tau}(i+\delta)$, and let $\{\lambda_{i,j} : i \in [n], j = 0, \cdots, s-1\} \bigcup \{\lambda_{i+k,j} : i \in [k], j = 0, \cdots, \lfloor \frac{i-1}{\theta} \rfloor + \delta\}$ be distinct elements in F. Let $l = s^n \prod_{i=1}^{\tau}(i+\delta)$. Consider the code family given by (1) and (2), where*

$$A_i = \sum_{a=0}^{l-1} \lambda_{i,a_i} e_a e_a^T, \quad i = 1, \cdots, n$$

$$A_{n+i} = \sum_{a=0}^{l-1} \lambda_{n+i,a_{n+\lceil \frac{i}{\theta} \rceil}} e_a e_a^T \quad i = 1, \cdots, k.$$

Here $\{e_a : a = 0, \cdots, l-1\}$ is the standard basis of F^l and we represent a using the $(n+\tau)$-digit

notation $a = (a_{n+\tau}, \cdots, a_1)$*, where* $a_i \in \{0, \cdots, s-1\}$*, for* $i \in [n]$ *and* $a_{n+i} \in \{0, \cdots, i+\delta-1\}$*, for* $i \in [\tau]$.

Following an argument similar to the proof of Lemma 1, it is easy to show that the code given by Construction 2 is an MDS array code.

Lemma 3. *The array code* $\mathcal{C}$ *given by Construction 2 attains optimal bandwidth when 1) repairing a single node* i*,* $i \in [n]$*, from any* d_1 *nodes, and 2) repairing the set of nodes* $\{n+1, \cdots, n'\}$ *from any* d_2 *node.*

Proof. We omit the proof of the first statement because it is similar to the proof of Lemma 2. To prove the second statement, the key idea is to divide the $(n+1)$-th to the n'-th nodes into τ groups and repair the groups one by one iteratively. Formally, let $\mathcal{C}_R \subset \{C_1, \cdots, C_n\}$ be the set of nodes accessed, with $|\mathcal{C}_R| = d_2$. For $i = 1, \cdots, \tau$, let $\mathcal{C}_i = \{C_{n+(i-1)\theta+1}, \cdots, C_{n+(i-1)\theta+\theta}\}$, then we first use $\mathcal{C}_R$ to repair $\mathcal{C}_1$, then use $\mathcal{C}_R \cup \mathcal{C}_1$ to repair $\mathcal{C}_2$, $\cdots$, and finally use $\mathcal{C}_R \cup \mathcal{C}_1 \cup \cdots \cup \mathcal{C}_{\tau-1}$ to repair $\mathcal{C}_\tau$. By the proof of Lemma 2, we can recover $\mathcal{C}_i$ from $\{\sum_{u=0}^{\delta+i-1} c_{v,a(n+i,u)} : a_{n+i} = 0, C_v \in \mathcal{C}_R \cup \mathcal{C}_{[i-1]}\}$. Since the nodes in $\mathcal{C}_{[i-1]}$ are already recovered, it suffices to know the set of values

$$\Omega_i = \left\{ \sum_{u=0}^{\delta+i-1} c_{v,a(n+i,u)} : a_{n+i} = 0, C_v \in \mathcal{C}_R \right\}. \tag{12}$$

Therefore to recover all τ groups of nodes, it suffices to know the values in the set $\Lambda = \bigcup_{i=1}^{\tau} \Omega_i$. We remark that a value in Λ can be derived from other values in the set and so it suffices to download a spanning set of Λ. Let

$$\Lambda^* = \bigcup_{i=1}^{\tau} \left\{ \sum_{u=0}^{\delta+i-1} c_{v,a(n+i,u)} : a_{n+i} = 0, a_{n+j} < \delta + j - 1, j \in [i-1], C_v \in \mathcal{C}_R \right\} \tag{13}$$

Clearly $\Lambda^* \subset \Lambda$. We claim that every value in Λ can be determined by the values in Λ^*. To prove the claim it suffices to show that for $i \in [\tau]$, $\Omega_i \subset \Lambda^*$. We prove by induction on i. Clearly $\Omega_1 \subset \Lambda^*$. Now suppose that $\Omega_1, \cdots, \Omega_{i-1} \in \Lambda^*$, and consider the set

$$\Omega_i \backslash \Lambda^* = \left\{ \sum_{u=0}^{\delta+i-1} c_{v,a(n+i,u)} : a_{n+i} = 0, a_{n+j} = \delta + j - 1, j \in [i-1], C_v \in \mathcal{C}_R \right\}. \tag{14}$$

Consider an arbitrary element $\sum_{u=0}^{\delta+i-1} c_{v,a(n+i,u)}$ of $\Omega_i \backslash \Lambda^*$, so that a satisfies $a_{n+i} = 0$ and $a_{n+j} = \delta + j - 1$ for some $j \leq i-1$. Denote by $a(j, i, x, y) = (\cdots, a_{j-1}, x, a_{j+1}, \cdots, a_{i-1}, y, a_{i+1}, \cdots)$. Since

$\Omega_j \subset \Lambda^*$, it follows that $\sum_{u_1=0}^{\delta+j-1} c_{v,a(n+j,n+i,u_1,u_2)} \in \Lambda^*$, for all $u_2 = 0, \cdots, \delta+i-1$. Therefore

$$\sum_{u_2=0}^{\delta+i-1} \sum_{u_1=0}^{\delta+j-1} c_{v,a(n+j,n+i,u_1,u_2)} \in \text{span}(\Lambda^*). \tag{15}$$

Moreover, by construction $\sum_{u_2=0}^{\delta+i-1} c_{v,a(n+j,n+i,u_1,u_2)} \in \Lambda$ for all $u_1 = 0, \cdots, \delta+j-2$. Therefore

$$\sum_{u_1=0}^{\delta+j-2} \sum_{u_2=0}^{\delta+i-1} c_{v,a(n+j,n+i,u_1,u_2)} \in \text{span}(\Lambda^*). \tag{16}$$

Subtract (16) from (15) we have

$$\sum_{u_2=0}^{\delta+i-1} c_{v,a(n+j,n+i,\delta+j-1,u_2)} = \sum_{u=0}^{\delta+i-1} c_{v,a(n+i,u)} \in \text{span}(\Lambda^*). \tag{17}$$

This proves that $\Omega_i \in \text{span}(\Lambda^*)$ and the claim. We now analyze the size of Λ^*. Let

$$\Omega_i^* = \left\{ \sum_{u=0}^{\delta+i-1} c_{v,a(n+i,u)} : a_{n+i} = 0, a_{n+j} < \delta+j-1, j \in [i-1], C_v \in \mathcal{C}_R \right\} \tag{18}$$

so that $\Lambda^* = \bigcup_{i=1}^{\tau} \Omega_i^*$. By counting the elements of the set we have

$$|\Omega_i^*| = s^n \cdot \prod_{j=1}^{i-1} (\delta+j-1) \cdot \prod_{j=i+1}^{\tau} (\delta+j). \tag{19}$$

We claim that

$$\sum_{j=1}^{i} |\Omega_j^*| = s^n \frac{i}{\delta+i} \prod_{j=1}^{\tau} (\delta+j) \tag{20}$$

We prove (20) by induction on i. Clearly $|\Omega_1^*| = s^n \prod_{j=2}^{\tau} (\delta+j) = s^n \frac{1}{\delta+1} \prod_{j=1}^{\tau} (\delta+j)$. Now suppose

that (20) is true up to $i-1$, then it follows

$$\sum_{j=1}^{i}|\Omega_j^*| = \sum_{j=1}^{i-1}|\Omega_j^*| + |\Omega_i^*| \tag{21}$$

$$= s^n \frac{i-1}{\delta+i-1}\prod_{j=1}^{\tau}(\delta+j) + s^n \prod_{j=1}^{i-1}(\delta+j-1)\prod_{j=i+1}^{\tau}(\delta+j) \tag{22}$$

$$= s^n \left(\frac{i-1}{\delta+i-1}\prod_{j=1}^{i}(\delta+j) + \prod_{j=1}^{i-1}(\delta+j-1)\right)\prod_{j=i+1}^{\tau}(\delta+j) \tag{23}$$

$$= s^n \left(\frac{i-1}{\delta+i-1}\prod_{j=1}^{i}(\delta+j) + \frac{\delta}{(\delta+i)(\delta+i-1)}\prod_{j=1}^{i}(\delta+j)\right)\prod_{j=i+1}^{\tau}(\delta+j) \tag{24}$$

$$= s^n \left(\frac{i-1}{\delta+i-1} + \frac{\delta}{(\delta+i)(\delta+i-1)}\right)\prod_{j=1}^{\tau}(\delta+j) \tag{25}$$

$$= s^n \frac{i}{\delta+i}\prod_{j=1}^{\tau}(\delta+j), \tag{26}$$

proving (20). Therefore $|\Lambda^*| = \sum_{j=1}^{\tau}|\Omega_j^*| = \frac{\tau}{\rho}s^n\prod_{j=1}^{\tau}(\delta+j)$. Note that $|\Lambda^*|$ is the number of symbols over F that need to be downloaded from $\mathcal{C}_R$ and the total number of symbols stored by $\mathcal{C}_R$ is $s^n\prod_{j=1}^{\tau}(\delta+j)$. Therefore a fraction of $\frac{\tau}{\rho}$ of the symbols are downloaded which attains the lower bound. The proof is complete. □

By Lemmas 3 and Theorem 2, we have:

Theorem 4. *The secret sharing scheme obtained by applying Theorem 1 to the code given in Construction 2, where the last k nodes are discarded, is a $(n, k = n-r-z, r, z)$ scheme with d_1-optimal repair bandwidth and d_2-optimal decoding bandwidth.*

Finally, we remark that Construction 1 and Construction 2 both has a sub-packetization level of $l = O(s^n)$, which is comparable to existing regenerating codes, e.g., [13], of the same parameters (note that our secret sharing schemes can be viewed as regenerating codes). In fact, an exponential l is shown to be necessary in order to achieve the optimal repair bandwidth [4]. This suggests that the additional optimal decoding requirement has a small impact on l. We also remark that Construction 2 naturally generalizes to a family of regenerating codes that supports centralized repair of groups of nodes of flexible sizes with reduced sub-packetization, which is a result of separate interest. A similar centralized repair problem was studied in [10] whereas the code construction therein is restricted in parameters.

IV. Shamir's Scheme with Asymptotically Optimal Decoding Bandwidth

In this section we look at a different perspective of the decoding bandwidth problem. Instead of constructing new secret sharing schemes with optimal decoding bandwidth, we study the decoding bandwidth of the classical Shamir's scheme[11]. Though schemes with optimal decoding bandwidth exist, improving the decoding bandwidth of Shamir's scheme remains an important problem as it is extensively used due to its simplicity. Below we describe a new family of Shamir's scheme with asymptotically optimal decoding bandwidth by extending the ideas recently developed in [5], [14] on repairing Reed-Solomon codes.

Consider Shamir's original scheme: let F be a finite field of size $|F| > n$, and let $c_0, \cdots, c_{t-1}$ be t elements in F, where c_0 is the secret message and $c_1, \cdots, c_{t-1}$ are randomly selected. Let $\lambda_1, \cdots, \lambda_n$ be any n distinct non-zero elements in F and let $f(x) = \sum_{i=0}^{t-1} c_i x^i$. Then the n shares are $f(\lambda_1), \cdots, f(\lambda_n)$. Shamir's scheme is an (n, t) threshold scheme, i.e., from any t shares c_0 can be decoded by polynomial interpolation, and any $t-1$ shares reveal no information about c_0. Clearly, decoding c_0 by polynomial interpolation requires communication t symbols over F. In what follows we show that by choosing F and the set of evaluation points $\lambda_1, \cdots, \lambda_n$ carefully, it is possible to reduce the decoding bandwidth (when all n shares are available) to a fraction of approximately $\frac{n}{t(n-t+1)}$ of the original bandwidth. We need to slightly generalize the way that the secret message m is encoded: rather than setting $c_0 = m$, we let $c_0 = m - \sum_{i=1}^{t-1} \lambda_0^i c_i$, for some $\lambda_0 \in F$. In other words $m = \sum_{i=0}^{t-1} \lambda_0^i c_i = f(\lambda_0)$. The corresponding scheme is an (n, t) threshold scheme as long as $\lambda_i \neq \lambda_0$, $i \in [n]$. Note that Shamir's original scheme corresponds to the case that $\lambda_0 = 0$.

To reduce the decoding bandwidth, we follow the framework proposed in [5] of interpolating polynomials by querying partial polynomial evaluation. Specifically, let F be the extension field of degree l of a subfield K. During decoding, each of the n nodes applies K-linear transforms to the share over F that it holds to obtain a set of symbols over K. The decoder collects these sets of symbols and performs K-linear transforms to them to assemble the secret message over F. Formally, viewing F as a vector space of dimension l over K, it is shown in [5] that

Lemma 4. *For a finite field K and its degree-l extension field F, let f be a polynomial over F of degree $< t$, and $f(\lambda_1), \cdots, f(\lambda_n)$ be n evaluations. Let λ_0 be an element in F, and let $g_1(x), \cdots, g_l(x)$ be l polynomials over F of degree $< n - t + 1$, such that $\{g_i(\lambda_0) : i \in [l]\}$ is a basis of F over K. Then to determine $f(\lambda_0)$, it suffices to know the set of values $\bigcup_{i\in[n]} \{\mathrm{tr}\left(g_j(\lambda_i) f(\lambda_i)\right) : j \in [l]\}$, where $\mathrm{tr} : F \to K$ is the trace function.*

The task of decoding the scheme is equivalent to determining $f(\lambda_0)$ and therefore it suffices to download the set of values $\{\text{tr}\,(g_j(\lambda_i)f(\lambda_i)) : j \in [l]\}$ from node i, for $i \in [n]$. However, due to the linearity of the trace function, we can reconstruct this set from values in $\{\text{tr}(\beta f(\lambda_i)) : \beta \in B_i\}$, where B_i is a basis (over K) of span$[\{g_j(\lambda_i) : j \in [l]\}]$. Therefore, the number of symbols over K that we need to download from node i equals $|B_i| = \dim(\text{span}[\{g_j(\lambda_i)\}_{j\in[l]}])$. We now discuss a way to select λ_i, $i = 0, \cdots, n$ as well as $g_i(x)$, $i \in [l]$, so that the condition in Lemma 4 is satisfied and that $|B_i|$, $i \in [n]$ is minimized. We remark that our construction idea is inspired by the codes in [14].

Construction 3. *For any n, t, let $s = n - t + 1$ and $l = \tau s$ for some $\tau \geq n + 1$. Let K be a finite field and $h(x) \in K[x]$ be a degree l irreducible polynomial. Let β be a root of $h(x)$, and F be the field generated by β over K. Let $\lambda_0 = \beta$, $\lambda_i = \beta^{is}$, $i \in [n]$, and let $\{g_i(x) : i \in [l]\} = \{\beta^a x^b : a = 0, s, \cdots, (\tau - 1)s,\ b = 0, \cdots, s - 1\}$.*

Theorem 5. *The (n, t) Shamir's scheme obtained by choosing F and $\{\lambda_i : i = 0, \cdots, n\}$ according to Construction 3 attains a decoding bandwidth of less than $\frac{nl}{s} + \frac{n^2 s^2}{4}$ symbols over K.*

Proof. Define $\{g_i(x) : i \in [l]\}$ according to Construction 3. Then it follows that $\{g_i(\lambda_0),\ i \in [l]\} = \{\beta^0, \beta^1, \cdots, \beta^{l-1}\}$, which is a basis of F over K. Therefore the condition of Lemma 4 is satisfied and we invoke the lemma to decode $f(\lambda_0)$. We now analyze the size of $\{g_j(\lambda_i) : j \in [l]\}$, for $i \in [n]$:

$$\{g_j(\lambda_i) : j \in [l]\} = \{\beta^a \beta^{isb} : a = us, u \in [0 : \tau - 1], b \in [0 : s - 1]\} \tag{27}$$

$$\subset \{\beta^{us} : u \in [0 : \tau - 1]\} \cup \{\beta^{a+isb} : a + isb \geq l, a = us, u \in [0 : \tau - 1], b \in [0 : s - 1]\}. \tag{28}$$

Denote the two sets in the R.H.S. of (28) by I_1 and I_2, respectively. Then $|I_1| = \tau = l/s$ and the size of I_2 is bounded by

$$|I_2| \leq \sum_{j=1}^{s-1} |\{\beta^{a+isb} : a = (\tau - (j-1)i - k)s,\ k \in [i],\ b \in [j : s - 1]\}| \tag{29}$$

$$= \sum_{j=1}^{s-1} i(s - j) = \frac{is(s-1)}{2} \tag{30}$$

Therefore,

$$\sum_{i=1}^{n} \dim\left(\text{span}[\{g_j(\lambda_i) : j \in [l]\}]\right) \leq \sum_{i=1}^{n} |\{g_j(\lambda_i) : j \in [l]\}| \tag{31}$$

$$\leq n\tau + \sum_{i=1}^{n} \frac{is(s-1)}{2} \tag{32}$$

$$= \frac{nl}{s} + \frac{ns(n+1)(s-1)}{4} \tag{33}$$

$$< \frac{nl}{s} + \frac{n^2 s^2}{4} \tag{34}$$

By the remarks after Lemma 4, no more than $\frac{nl}{s} + \frac{n^2 s^2}{4}$ symbols over K need to be downloaded, proving the theorem. □

Note that the lower bound on the decoding bandwidth is $\frac{nl}{s}$ symbols and by Theorem 5, the decoding bandwidth of proposed scheme is less than $(1 + \frac{ns^3}{4l})\frac{nl}{s}$ symbols. Therefore as $l \to \infty$, the decoding bandwidth is asymptotically optimal in the sense that the ratio between the actual decoding bandwidth and the optimal bandwidth approaches 1.

REFERENCES

[1] R. Bitar and S. El Rouayheb, "Staircase codes for secret sharing with optimal communication and read overheads," in *IEEE International Symposium on Information Theory*, 2016.

[2] G. Blakley and G. Kabatianski, "Ideal perfect threshold schemes and MDS codes," in *IEEE International Symposium on Information Theory*, vol. 4, 1995, p. 77843.

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

[4] S. Goparaju, I. Tamo, and R. Calderbank, "An improved sub-packetization bound for minimum storage regenerating codes," *IEEE Transactions on Information Theory*, vol. 60, no. 5, pp. 2770–2779, 2014.

[5] V. Guruswami and M. Wootters, "Repairing Reed-Solomon codes," in *ACM symposium on Theory of Computing*, 2016.

[6] W. Huang, M. Langberg, J. Kliewer, and J. Bruck, "Communication efficient secret sharing," *IEEE Transactions on Information Theory*, vol. 62, no. 12, pp. 7195–7206, 2016.

[7] F. MacWilliams and N. Sloane, *The theory of error-correcting codes*. North Holland Publishing, 1977.

[8] S. Pawar, S. El Rouayheb, and K. Ramchandran, "Securing dynamic distributed storage systems against eavesdropping and adversarial attacks," *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 6734–6753, 2011.

[9] K. V. Rashmi, N. B. Shah, and P. V. Kumar, "Optimal exact-regenerating codes for distributed storage at the MSR and MBR points via a product-matrix construction," *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 5227–5239, 2011.

[10] A. S. Rawat, O. O. Koyluoglu, and S. Vishwanath, "Centralized repair of multiple node failures with applications to communication efficient secret sharing," *arXiv:1603.04822*, 2016.

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

[12] H. Wang and D. Wong, "On secret reconstruction in secret sharing schemes," *IEEE Transactions on Information Theory*, vol. 54, no. 1, pp. 473–480, 2008.

[13] M. Ye and A. Barg, "Explicit constructions of high-rate mds array codes with optimal repair bandwidth," *arXiv:1604.00454*, 2016.

[14] ——, "Explicit constructions of MDS array codes and RS codes with optimal repair bandwidth," in *IEEE International Symposium on Information Theory*, 2016.