

2019

Distributed Matrix-Vector Multiplication: A Convolutional Coding Approach

Anindya B. Das

Iowa State University, abd149@iastate.edu

Aditya Ramamoorthy

Iowa State University, adityar@iastate.edu

Follow this and additional works at: https://lib.dr.iastate.edu/ece_pubs

 Part of the [Databases and Information Systems Commons](#), [Signal Processing Commons](#), and the [Systems and Communications Commons](#)

The complete bibliographic information for this item can be found at https://lib.dr.iastate.edu/ece_pubs/206. For information on how to cite this item, please visit <http://lib.dr.iastate.edu/howtocite.html>.

This Article is brought to you for free and open access by the Electrical and Computer Engineering at Iowa State University Digital Repository. It has been accepted for inclusion in Electrical and Computer Engineering Publications by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

Distributed Matrix-Vector Multiplication: A Convolutional Coding Approach

Abstract

Distributed computing systems are well-known to suffer from the problem of slow or failed nodes; these are referred to as stragglers. Straggler mitigation (for distributed matrix computations) has recently been investigated from the standpoint of erasure coding in several works. In this work we present a strategy for distributed matrix-vector multiplication based on convolutional coding. Our scheme can be decoded using a low-complexity peeling decoder. The recovery process enjoys excellent numerical stability as compared to Reed-Solomon coding based approaches (which exhibit significant problems owing their badly conditioned decoding matrices). Finally, our schemes are better matched to the practically important case of sparse matrix-vector multiplication as compared to many previous schemes. Extensive simulation results corroborate our findings.

Disciplines

Databases and Information Systems | Signal Processing | Systems and Communications

Comments

This is a pre-print of the article Das, Anindya Bijoy and Aditya Ramamoorthy. "Distributed Matrix-Vector Multiplication: A Convolutional Coding Approach." *arXiv preprint arXiv:1901.08716* (2019). Posted with permission.

Distributed Matrix-Vector Multiplication: A Convolutional Coding Approach

Anindya B. Das and Aditya Ramamoorthy
Department of Electrical and Computer Engineering
Iowa State University
Ames, IA 50010, U.S.A.
{abd149, adityar}@iastate.edu

Abstract—Distributed computing systems are well-known to suffer from the problem of slow or failed nodes; these are referred to as stragglers. Straggler mitigation (for distributed matrix computations) has recently been investigated from the standpoint of erasure coding in several works. In this work we present a strategy for distributed matrix-vector multiplication based on convolutional coding. Our scheme can be decoded using a low-complexity peeling decoder. The recovery process enjoys excellent numerical stability as compared to Reed-Solomon coding based approaches (which exhibit significant problems owing their badly conditioned decoding matrices). Finally, our schemes are better matched to the practically important case of sparse matrix-vector multiplication as compared to many previous schemes. Extensive simulation results corroborate our findings.

Index Terms—Distributed Computation, Stragglers, Cross Parity Check Convolutional Code, Reed-Solomon Coding

I. INTRODUCTION

Distributed computation plays a major role in several problems in optimization and machine learning. For example, large scale gradient descent often requires us to repeatedly calculate matrix-vector products. In high-dimensional problems, time and storage constraints necessitate the splitting of these computations across multiple nodes.

Distributed systems are well known to suffer from the issue of slow or faulty processors, which are referred to as stragglers. Several methods have been developed recently for straggler mitigation by using ideas from erasure coding [1]–[3]. For example, suppose that we want to compute the product of matrix $\mathbf{A} \in \mathbb{R}^{r \times t}$ and $\mathbf{x} \in \mathbb{R}^t$ in a distributed fashion. As proposed by [4], we can first split matrix \mathbf{A} into submatrices with equal number of rows as $\mathbf{A}^T = [\mathbf{A}_0^T \ \mathbf{A}_1^T]$, and assign three different worker nodes the jobs of computing $\mathbf{A}_0\mathbf{x}$, $\mathbf{A}_1\mathbf{x}$ and $(\mathbf{A}_0 + \mathbf{A}_1)\mathbf{x}$ so that the computational load on each of them is half of the original. It is evident that the master node can recover $\mathbf{A}\mathbf{x}$, as soon as it receives results from any two workers, i.e., the system is resilient to one straggler. This can be generalized by using Reed-Solomon (RS) code-like ideas. Reference [2] incorporates similar ideas also for matrix-matrix multiplication. In both these cases, the proposed solutions have the property that the master node can recover the final result as soon as it receives the results from any τ workers; τ is called the recovery threshold.

This work was supported in part by the National Science Foundation (NSF) under grant CCF-1718470.

While these solutions for distributed matrix-vector multiplication are optimal with respect to the recovery threshold, they neglect certain important issues that exist in practical scenarios. For instance, in many machine learning problems, the matrix \mathbf{A} is sparse. RS based approaches require dense linear combinations of the submatrices of \mathbf{A} . This may cause the computation time of the submatrix-vector products to go up. RS based approaches also suffer from numerical stability issues owing to the high condition number of the corresponding Vandermonde matrices. A high condition number results in the decoded value changing by a large amount even if the computed submatrix-vector products change by a small amount. This is especially relevant in the machine learning context where gradient computations can often be approximate. We note here that this issue with respect to polynomial interpolation is well recognized in the numerical analysis literature [5] and several papers discuss appropriate choices of interpolation points for numerical robustness [6]. However, in the straggler mitigation context, we need to be able to perform recovery from any large enough subset of interpolation points. This causes the worst case condition number to be quite bad.

In this work, we present a class of distributed matrix-vector multiplication schemes by leveraging (binary) cross parity check convolutional codes [7], that were originally proposed for distributed storage systems. In our context, the generator matrices specify the assignment of jobs to the worker nodes.

A. Main Contributions

- While the codes in [7] can result in generator matrices that are recursive, we show that our schemes always result in feed-forward encoders. This is important in our setting, because our underlying field of operation is \mathbb{R} . Furthermore, we show that in the setting when the number of stragglers is two, then our encoder only has coefficients in $\{-1, 0, 1\}$. When the number of stragglers is three, we show an upper bound on the absolute value of the coefficients.
- We demonstrate that our schemes can be decoded at the master node by a low-complexity peeling decoder. This implies that the master node can operate in iterations such that in each iteration, it solves an equation where there is only one unknown.

- Our experimental results indicate the numerical robustness of our scheme, and also shows its advantage in computation speed when \mathbf{A} is sparse.

II. CROSS PARITY CHECK CONVOLUTIONAL CODES

Consider the set of real infinite sequences $\{c_r, c_{r+1}, \dots\}$ for $r \in \mathbb{Z}$ that start at some finite integer index and continue thereafter. These sequences can be treated as elements of the formal Laurent series [7] in indeterminate D with coefficients from \mathbb{R} , i.e. $\sum_{i=r}^{\infty} c_i D^i$. Let us denote the ring of formal Laurent series over \mathbb{R} as $\mathbb{R}((x))$ under the normal addition and multiplication of formal power series. It can be shown that $\mathbb{R}((x))$ forms a field, i.e., each non-zero element of $\mathbb{R}((x))$ has a corresponding inverse.

In this work, we will consider n infinite strips that can be visualized as columns that start at index $r = 0$ and continue indefinitely downward. We denote the infinite sequence as $\{c_{0,j}, c_{1,j}, c_{2,j}, \dots\}$ for each column j , and so, we can represent the j -th strip by the formal series as

$$C_j(D) = \sum_{i=0}^{\infty} c_{i,j} D^i$$

for $0 \leq j \leq n-1$. These sequences obey the “geometric” constraint

$$\sum_{j=0}^{n-1} c_{i-m,j} = 0 \quad \text{for } i \geq 0, \quad (1)$$

which indicates the lines of slope m . For each value of $m = 0, 1, \dots, (n-k-1)$, the sequences sum to zero along the lines of the corresponding slope. The value $c_{i,j} = 0$ if $i < 0$ for any j .

Let $CP(n, k)$ denote the set of all sequences¹ that satisfy the constraints in (1). This can equivalently be expressed as

$$\sum_{j=0}^{n-1} C_j(D) D^{mj} = 0 \quad (2)$$

for $m = 0, 1, \dots, (n-k-1)$. Let

$$\underline{C}(D) = (C_0(D), C_1(D), C_2(D), \dots, C_{n-1}(D)).$$

Then, we can express the condition succinctly as

$$\underline{C}(D) \mathbf{H}_{n,k}^T(D) = 0,$$

where $\mathbf{H}_{n,k}(D)$ is the $(n-k) \times n$ matrix (analogous to a parity-check matrix), which can be obtained from (2) and written as

$$\mathbf{H}_{n,k}(D) = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & D & D^2 & \dots & D^{(n-1)} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & D^{(s-1)} & D^{2(s-1)} & \dots & D^{(n-1)(s-1)} \end{bmatrix}.$$

Since every $(n-k) \times (n-k)$ submatrix of $H_{n,k}(D)$ is a Vandermonde matrix evaluated at distinct powers of the indeterminate D , its determinant will be a non-zero polynomial

¹We will refer to this as the $CP(n, k)$ code

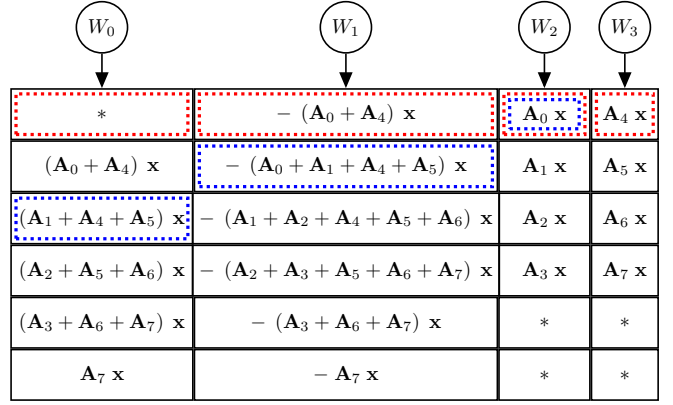


Fig. 1: Distributed Matrix-vector Multiplication embedded into a $CP(4, 2)$ code. The assigned jobs in W_0 are downshifted and its first job is denoted by the placeholder $*$. This is only to make it easy to see that the geometric constraints are satisfied. In reality, W_0 will start executing its first job, i.e., $(\mathbf{A}_0 + \mathbf{A}_4)\mathbf{x}$ right away and proceed sequentially downward. Here blue and red dotted blocks indicate examples of two constraint lines with slopes 1 and 0, respectively.

in D and hence invertible over $\mathbb{R}((x))$. Thus, $\underline{C}(D)$ can be recovered even if any $n-k$ columns are lost.

The key idea underlying our work is that distributed matrix vector multiplication can be embedded into the class of $CP(n, k)$ codes, where a given column i represents (upon appropriate interpretation) the computation assigned to worker node W_i , which sequentially processes its assigned jobs from top to bottom. The result $\mathbf{A}\mathbf{x}$ can be recovered even if any $n-k$ worker nodes fail. Crucially, $\mathbf{A}\mathbf{x}$ can be decoded using a peeling decoder. This significantly reduces the overall computation load at the master node and provides for a scheme that enjoys excellent numerical stability.

Example 1. Let \mathbf{A} be a large matrix that is split into eight block-rows, $\mathbf{A}_0, \mathbf{A}_1, \dots, \mathbf{A}_7$. Fig. 1 shows an example where the distributed computation of $\mathbf{A}\mathbf{x}$ is embedded into a $CP(4, 2)$ code. Each cell in the figure shows the responsibility assigned to the corresponding worker node (from top to bottom). It can be observed that the geometric condition in (1) is satisfied by the cell contents and that $\mathbf{A}\mathbf{x}$ can be recovered even if any two of the worker nodes fail. Furthermore, this can be achieved by a peeling decoder with only addition/subtraction operations.

III. EMBEDDING MATRIX-VECTOR MULTIPLICATION INTO A $CP(n, k)$ CODE

In this section, we outline the details of our proposed scheme. Towards this end, we first derive the corresponding generator matrix for the $CP(n, k)$ code and show that it can be expressed in feed-forward form. And then we discuss how the distributed computation of $\mathbf{A}\mathbf{x}$ can be mapped onto a system with n worker nodes using the $CP(n, k)$ code.

Let $s = n - k$ and $\mathbf{Y}_{a,b}$ be a $a \times b$ matrix such that

$$\mathbf{Y}_{a,b}(i, j) = D^{ij}; \quad 0 \leq i \leq a-1, \quad 0 \leq j \leq b-1;$$

and Ψ_w be a $w \times w$ diagonal matrix such that

$$\Psi_w = \text{diag} [1 \quad D \quad D^2 \quad \dots \quad D^{w-1}] .$$

Thus, we have

$$\mathbf{H}_{n,k}^T(D) = \begin{bmatrix} \mathbf{Y}_{s,s} \\ \mathbf{Y}_{k,s} \Psi_s^s \end{bmatrix} .$$

Let the systematic generator matrix be

$$\mathbf{G}_{n,k}(D) = [\mathbf{Z} \quad \mathbf{I}_k] ,$$

where \mathbf{I}_k is the $k \times k$ identity matrix. Now satisfying $\mathbf{G}_{n,k}(D) \mathbf{H}_{n,k}^T(D) = \mathbf{0}$, we obtain

$$\mathbf{Z} = -\mathbf{Y}_{k,s} \Psi_s^s \mathbf{Y}_{s,s}^{-1} . \quad (3)$$

We can express $\mathbf{W}_{k,s} = \mathbf{Y}_{k,s} \Psi_s^s$ as

$$\mathbf{W}_{k,s} = \begin{bmatrix} 1 & D^s & D^{2s} & \dots & D^{(s-1)s} \\ 1 & D^{s+1} & D^{2(s+1)} & \dots & D^{(s-1)(s+1)} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & D^{(s+k-1)} & D^{2(s+k-1)} & \dots & D^{(s-1)(s+k-1)} \end{bmatrix} .$$

Suppose that $\{y_{\ell j}\}$ are the elements of $\mathbf{Y}_{s,s}^{-1}$ and we define s different polynomials in the indeterminate σ as

$$y_j(\sigma) = \sum_{\ell=0}^{s-1} y_{\ell j} \sigma^\ell$$

for $0 \leq j \leq s$. Since $\mathbf{Y}_{s,s} \mathbf{Y}_{s,s}^{-1} = \mathbf{I}$, we can write for any j ,

$$y_j(D^\ell) = 0, \quad \text{for } \ell \neq j,$$

which indicates that any D^ℓ will be a root of the polynomial $y_j(\sigma)$ if $\ell \neq j$. On the other hand, $y_j(D^j) = 1$ so that

$$y_j(\sigma) = \sum_{\ell=0}^{s-1} y_{\ell j} \sigma^\ell = \prod_{\ell=0, \ell \neq j}^{s-1} \frac{\sigma - D^\ell}{D^j - D^\ell} .$$

Now taking the product $\mathbf{W}_{k,s} \mathbf{Y}_{s,s}^{-1}$ involves evaluating these polynomials at D^{s+i} where $i = 0, \dots, k-1$, and thus we get

$$Z_{ij} = - \prod_{\ell=0, \ell \neq j}^{s-1} \frac{D^{s+i} - D^\ell}{D^j - D^\ell} . \quad (4)$$

It is unclear whether the above expression leads to a recursive or a non-recursive $\mathbf{G}_{n,k}$. The following theorem shows that Z_{ij} can be simplified to express it as a polynomial, i.e., $\mathbf{G}_{n,k}$ can be put in feed-forward form.

Theorem 1. Any term Z_{ij} , with $0 \leq i < k$ and $0 \leq j < s$, can be written as a finite polynomial in D with integer coefficients. When $s = 2$, the coefficients of $Z_{ij} \in \{-1, 0, 1\}$ and when $s = 3$ the coefficients of Z_{ij} have absolute value at most k .

Proof. From (4), we can write

$$Z_{ij} = - \left[\frac{(D^{s+i} - D^0)(D^{s+i} - D^1) \dots (D^{s+i} - D^{j-1})}{(D^j - D^0)(D^j - D^1) \dots (D^j - D^{j-1})} \right] \\ \times \left[\frac{(D^{s+i} - D^{j+1})(D^{s+i} - D^{j+2}) \dots (D^{s+i} - D^{s-1})}{(D^j - D^{j+1})(D^j - D^{j+2}) \dots (D^j - D^{s-1})} \right]$$

which can be written as $Z_{ij} = -A_{ij}B_{ij}$. Here

$$A_{ij} = \frac{(D^{x_{ij}+j} - 1)(D^{x_{ij}+j-1} - 1) \dots (D^{x_{ij}+1} - 1)}{(D^j - 1)(D^{j-1} - 1) \dots (D^1 - 1)} ;$$

$$B_{ij} = (-1)^{e_{ij}} D^{f_{ij}} \frac{(D^{i+y_{ij}} - 1)(D^{i+y_{ij}-1} - 1) \dots (D^{i+1} - 1)}{(D^1 - 1)(D^2 - 1) \dots (D^{y_{ij}} - 1)} ;$$

where $x_{ij} = s + i - j$, $y_{ij} = s - j - 1$, $e_{ij} = s - j - 1$ and $f_{ij} \geq 0$ as $j < s$. Now consider a term W such that

$$W = \frac{(D^{x+1} - 1)(D^{x+2} - 1) \dots (D^{x+y} - 1)}{(D^1 - 1)(D^2 - 1) \dots (D^y - 1)} .$$

A polynomial of the form $D^q - 1$ can be written as a product of some cyclotomic polynomials [8], as

$$D^q - 1 = \prod_{d|q} \Phi_d(D)$$

where $\Phi_d(D)$ is the d^{th} cyclotomic polynomial. Thus, we can expand the denominator of W as

$$\text{den}(W) = [\Phi_1(D)]^{i_1} \times [\Phi_2(D)]^{i_2} \times \dots \times [\Phi_y(D)]^{i_y}$$

where $i_j = \lfloor y/j \rfloor$ for $j = 1, 2, \dots, y$. The numerator of W is a product of y different polynomials, and the maximum exponent of D in those y polynomials are consecutive y numbers. This indicates that we must have $\prod_{v=1}^y [\Phi_v(D)]^{i_v}$ as a factor of the numerator. But the numerator is, in fact, a product of different cyclotomic polynomials, according to its definition. Thus, after the cancellation by the denominator and expansion, all the exponents and corresponding coefficients of D in W would be integers. It indicates that the terms A_{ij} and B_{ij} are also finite polynomials of D with integer coefficients, and thus the first part of the proof is complete. The proof regarding the corresponding coefficients of the assigned submatrix-vector block products is given in Appendix A. ■

Example 2. When $n = 4, k = 2$, we can obtain $\mathbf{G}_{4,2}(D)$ as

$$\mathbf{G}_{4,2}(D) = \begin{bmatrix} D & -D-1 & 1 & 0 \\ D^2+D & -D^2-D-1 & 0 & 1 \end{bmatrix} \quad (5)$$

where we note that all coefficients of the polynomials in $\mathbf{G}_{4,2}(D)$ belong to the set $\{-1, 0, 1\}$.

Next, we discuss the usage of $\mathbf{G}_{n,k}$ in a distributed matrix-vector multiplication context. First we partition \mathbf{A} row-wise into Δ (we assume that k divides Δ) equal sized block-rows. These are denoted $\mathbf{A}_0, \mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_{\Delta-1}$. Let $u_j = \mathbf{A}_j \mathbf{x}$ where $(0 \leq j \leq \Delta - 1)$. The next step is to form k different polynomials with coefficients from the u_j 's. These are given by

$$\tilde{u}_i(D) = u_{iq} + u_{iq+1}D + \dots + u_{(i+1)q-1}D^{q-1}$$

for $0 \leq i \leq k-1$ where $q = \frac{\Delta}{k}$. Now the submatrices assigned to all the workers are determined by

$$U(D) = [\tilde{u}_0(D) \quad \tilde{u}_1(D) \quad \dots \quad \tilde{u}_{k-1}(D)] \mathbf{G}_{n,k}(D) .$$

Suppose that each worker node can store at most the equivalent of γ fraction of the rows of \mathbf{A} . Thus, if we assign ℓ_j jobs to

worker W_j , then it needs to satisfy $\frac{\ell_i}{\Delta} \leq \gamma$. The number of jobs assigned to worker node W_i , ($0 \leq i \leq n-1$), depends on the entries of the corresponding column of $\mathbf{G}_{n,k}(D)$. Let d_i denote the difference between the maximum and the minimum exponent of D in the i -th column of $\mathbf{G}_{n,k}(D)$, and let $\lambda = \max_i d_i$. Then, for satisfying the storage constraint we require

$$\frac{\Delta}{k} + \lambda \leq \gamma \Delta \quad \text{which leads to} \quad \Delta \geq \frac{\lambda}{\gamma - \frac{1}{k}}. \quad (6)$$

Example 3. We consider the same scenario as mentioned in Example 2 with $n = 4$ workers, and we need to develop a scheme that is resilient to $s = 2$ stragglers, so $k = n - s = 2$. If $\gamma = \frac{3}{4}$, then we can set $\Delta = \frac{2}{\frac{3}{4} - \frac{1}{2}} = 8$, as it is divisible k . The relevant polynomials are

$$\begin{aligned} \tilde{u}_0(D) &= u_0 + u_1 D + u_2 D^2 + u_3 D^3, \text{ and} \\ \tilde{u}_1(D) &= u_4 + u_5 D + u_6 D^2 + u_7 D^3. \end{aligned}$$

Forming $U(D) = \begin{bmatrix} \tilde{u}_0(D) & \tilde{u}_1(D) \end{bmatrix} \mathbf{G}_{4,2}(D)$, we obtain the scheme shown in Fig. 1.

Remark 1. The above procedure demonstrates the importance of a feed-forward $\mathbf{G}_{n,k}$. Indeed, if $\mathbf{G}_{n,k}$ had been recursive, the number of terms in the output would have been infinite.

IV. DECODING OF A $CP(n, k)$ -SCHEME

A major advantage of our proposed scheme is a low-complexity decoding procedure. Note that our scheme is in one-to-one correspondence with the $CP(n, k)$ code. Hence, we describe the decoding procedure for the $CP(n, k)$ code; the adaptation to recovering \mathbf{Ax} follows naturally.

Recall that the symbols are denoted by $c_{i,j}$ where $0 \leq i \leq \ell_j - 1$ and $0 \leq j \leq n - 1$. Let the indices of the straggler nodes be $0 \leq t_0 < t_1 < t_2 < \dots < t_{s-1} \leq n - 1$.

At each step, our decoding process in Algorithm 1 exploits the geometric constraints in (1) to identify an equation where there is one unknown; it continues in a systematic fashion until all the unknowns are decoded. In the sequel, we refer to this as decoding in a peeling decoder fashion. For instance, to start decoding symbols from straggler t_0 , we can use (1) and obtain the constraint for the line with slope $s - 1$ as

$$\sum_{j=0}^{n-1} c_{i-(s-1)j,j} = 0$$

which can pass through a symbol c_{α,t_0} in straggler t_0 . So we have the constraint as

$$\sum_{j=0}^{n-1} c_{(s-1)(t_0-j)+\alpha,j} = 0 \quad (7)$$

In (7), if $j < t_0$, we assumed that these symbols are known. Now if $j > t_0$, the elements $c_{\alpha,j}$'s are also known until a constraint line passes through c_{0,t_1} with the increase of j . Thus, in the extreme case, the line can pass through c_{-1,t_1} , so we can set $j = t_1$ and

$$\alpha + (s-1)(t_0 - j) = -1.$$

Thus if $0 \leq \alpha \leq (s-1)(t_1 - t_0) - 1$, we can say that the only unknown in (7) is c_{α,t_0} . In this manner, we can obtain the first $(s-1)(t_1 - t_0)$ symbols of straggler t_0 in a peeling decoding fashion.

To better understand the behavior of the algorithm, we divide it into phases. We say that the algorithm has finished phase p , where $0 \leq p \leq s - 2$, if it has only recovered as many symbols as possible in a peeling decoding fashion from the stragglers t_i , $i \leq p$, without recovering any symbol from stragglers t_j , $j > p$. Let us denote $\eta_{p,y}$ as the number of recovered symbols in a peeling decoding fashion from straggler t_y after finishing phase p .

Lemma 1.

$$\eta_{p,y} = \begin{cases} \sum_{i=y}^p (s-1-i)(t_{i+1} - t_i), & \text{if } y \leq p; \\ 0, & \text{otherwise;} \end{cases} \quad (8)$$

for $p < s - 1$.

Proof. We are going to prove this by induction.

Base case: After finishing the phase $p = 0$, we recover as many symbols as possible from straggler t_0 . Thus we have proved earlier in this Section that

$$\begin{aligned} \eta_{0,0} &= (s-1)(t_1 - t_0), \text{ and} \\ \eta_{0,y} &= 0, \quad \text{for } y > 0. \end{aligned}$$

Inductive step: Now we assume that after finishing phase p , the number of recovered symbols by the stragglers can be represented by

$$\eta_{p,y} = \begin{cases} \sum_{i=y}^p (s-1-i)(t_{i+1} - t_i), & \text{if } y \leq p; \\ 0, & \text{otherwise;} \end{cases}$$

where $0 \leq p < s - 2$. Then after finishing phase $p + 1$, to recover symbols from the straggler t_{p+1} , we obtain the constraint for the line with slope $(s - p - 2)$ as

$$\sum_{j=0}^{n-1} c_{i-(s-p-2)j,j} = 0, \quad (9)$$

which can pass through $c_{0,t_{p+1}}$. So we obtain

$$\sum_{j=0}^{n-1} c_{(s-p-2)(t_{p+1}-j),j} = 0.$$

For $j < t_{p+1}$, if the line passes through the straggler t_q where $0 \leq q \leq p$, it will pass through the points $c_{(s-p-2)(t_{p+1}-t_q),t_q}$, but from the hypothesis we assumed $\eta_{p,q}$ symbols are known from worker t_q after phase p , and

$$\begin{aligned} \eta_{p,q} &= \sum_{i=q}^p (s-1-i)(t_{i+1} - t_i) \\ &> \sum_{i=q}^p (s-p-2)(t_{i+1} - t_i) = (s-p-2)(t_{p+1} - t_q), \end{aligned}$$

which indicates that all necessary symbols of the stragglers are already known. And, if $j > t_{p+1}$, then the row index is negative which means that those corresponding elements are zero. So, we can recover $c_{0,t_{p+1}}$ in a peeling decoding fashion.

Now, we move to straggler t_p and want to recover one more symbol, where already we recovered $\eta_{p,p}$ symbols. So, we can find the constraint line with slope $(s-p-1)$,

$$\sum_{j=0}^{n-1} c_{i-(s-p-1)j,j} = 0$$

and to recover one additional symbol from straggler t_p , this line needs to pass through $c_{(s-p-1)(t_{p+1}-t_p),t_p}$, so we obtain

$$\sum_{j=0}^{n-1} c_{(s-p-1)(t_{p+1}-j),j} = 0.$$

In this equation, the terms with $j > t_{p+1}$ are actually zero, and the term with $j = t_{p+1}$ is $c_{0,t_{p+1}}$ which is decoded before. Now if $j = t_q < t_p$, the values are already known after finishing phase p because

$$\begin{aligned} \eta_{p,q} &= \sum_{i=q}^p (s-1-i)(t_{i+1}-t_i) \\ &> \sum_{i=q}^p (s-p-1)(t_{i+1}-t_i) = (s-p-1)(t_{p+1}-t_q). \end{aligned}$$

So we can obtain $c_{\eta_{p,p},t_p}$ in a peeling decoding fashion.

Thus we need previously decoded $c_{0,t_{p+1}}$ to decode $c_{\eta_{p,p},t_p}$. Now in a similar way, we can decode all $c_{\eta_{p,q},t_q}$, where $q < p$, each of which would need the previously decoded $c_{\eta_{p,q+1},t_{q+1}}$. It should be noted that we do not need any symbols from straggler t_i to decode $c_{0,t_{p+1}}$, where $i > p+1$. Thus we do not need them to decode $c_{\eta_{p,q},t_q}$ too, ($q \leq p+1$), because of the constraint lines with increasing slopes. This argument is sketched in Appendix B.

Now we repeat the whole process until we require any symbol from straggler t_{p+2} to decode a symbol from the straggler t_{p+1} . We assume that we can recover maximum ν symbols from t_{p+1} without decoding any symbol from the straggler t_q , $q > p+1$. So, the constraint line in (9) passes through $c_{\nu,t_{p+1}}$ and $c_{-1,t_{p+2}}$, and then, we can write

$$\nu + (s-p-2)(t_{p+1}-t_{p+2}) = -1$$

which leads to $\nu = (s-p-2)(t_{p+2}-t_{p+1}) - 1$. So we can recover $(s-p-2)(t_{p+2}-t_{p+1})$ symbols from each of the stragglers, t_0, t_1, \dots, t_{p+1} . So, we can say

$$\eta_{p+1,y} = \begin{cases} \sum_{i=y}^{p+1} (s-1-i)(t_{i+1}-t_i), & \text{if } y \leq p+1; \\ 0, & \text{otherwise;} \end{cases}$$

which proves the inductive step. \blacksquare

Theorem 2. The decoding procedure in Algorithm 1 allows the recovery of all $c_{i,j}$ where $0 \leq i \leq \ell_j-1$ and $0 \leq j \leq n-1$ in a peeling decoding fashion, as long as there are at most s node failures.

Algorithm 1: Decoding of $CP(n, k)$ Scheme

Input : A $CP(n, k)$ scheme to obtain \mathbf{Ax} and the index of the stragglers

- 1 Sort the stragglers according to their index as $0 \leq t_0 < t_1 < t_2 < \dots < t_{s-1} \leq n-1$
- 2 **for** $j \leftarrow 0$ **to** $s-2$ **do**
- 3 **for** $i \leftarrow 0$ **to** $(s-1-j)(t_{j+1}-t_j)-1$ **do**
- 4 Set $q = j$;
- 5 **while** $q \geq 0$ **do**
- 6 Obtain an additional submatrix-vector block product (until ℓ_{t_q}) in straggler t_q using constraint line with slope $s-1-q$;
- 7 $q \leftarrow q-1$;
- 8 **end**
- 9 **end**
- 10 **end**
- 11 **for** $j \leftarrow 0$ **to** $\ell_{s-1}-1$ **do**
- 12 Set $q = s-1$;
- 13 **while** $q \geq 0$ **do**
- 14 Obtain an additional submatrix-vector block product (until ℓ_{t_q}) in straggler t_q using constraint line with slope $s-1-q$;
- 15 $q \leftarrow q-1$;
- 16 **end**
- 17 **end**

Output: All block products to obtain \mathbf{Ax}

Proof. From Lemma 1, we have proved that we can decode $\eta_{s-2,q}$ elements from a straggler t_q after finishing phase $s-2$, for $q < s-1$. Now if we apply the constraint line with slope zero, then we can recover $\eta_{s-2,s-2} = (t_{s-1}-t_{s-2})$ elements from straggler t_{s-1} , and similar to the proof of inductive step, applying the other slopes we can recover additional $(t_{s-1}-t_{s-2})$ elements from other stragglers, until all the symbols from a straggler are recovered. If we just continue this process, we can recover all symbols from all the workers in a peeling decoding fashion. \blacksquare

Example 4. In this example, we demonstrate how we can recover \mathbf{Ax} in a peeling decoding fashion for the same scenario in Fig. 1. In this example, we have four workers, W_0, W_1, W_2 and W_3 , among which we assume that W_2 and W_3 are stragglers. According to Algorithm 1, we utilize the slope 1 constraint lines to recover the block products of W_2 and slope 0 constraint lines to recover the block products of W_3 . For example, if we utilize the slope 1 constraint line through the blue dotted blocks in Fig. 1, then we can recover the first block of W_2 , which is $\mathbf{A}_0\mathbf{x}$. Similarly, if we utilize the slope 0 constraint line through the red dotted blocks, then we can also recover the first block of W_3 , which is $\mathbf{A}_4\mathbf{x}$. Using this same fashion, we can decode all the submatrix-vector block products from workers W_2 and W_3 , which completes the job.

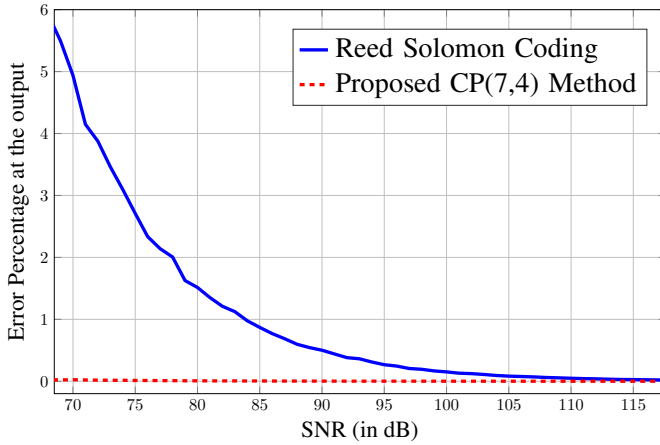


Fig. 2: Comparison between our proposed method and RS coding based method at different noise levels

V. NUMERICAL RESULTS

We consider a scenario where \mathbf{A} has dimension $8,000 \times 10,000$ and \mathbf{x} is of length 10,000. Suppose that we have $n = 7$ workers, $\gamma = \frac{3}{10}$ (storage fraction). We require the system to be resilient to $s = 3$ stragglers.

For the RS based approach we can partition \mathbf{A} in $\Delta = 10$ parts, and assign three submatrix-vector products to each worker. The evaluation points for the RS code are chosen as real numbers equally spaced in $[-1, 1]$; this is a better choice than integers [9]. On the other hand, we can embed this problem into a $CP(7, 4)$ using our proposed approach. The value of λ (cf. Section III) in this case is 8 and $\Delta = \frac{\lambda}{\gamma - \frac{1}{k}} = \frac{8}{\frac{3}{10} - \frac{1}{4}} = 160$. The assignment of jobs to each worker node can be determined by the procedure in Section III

In order to test the numerical stability of recovery at the master node, we add white Gaussian noise to each submatrix-vector product computed by the workers. Fig. 2 presents a comparison of the RS based approach and our proposed approach. The error percentage values at the output are shown for different noise levels. If the correct output is y and the recovered output by the master is \hat{y} , then the error percentage is measured as $\frac{\|y - \hat{y}\|}{\|y\|} \times 100\%$. We can see that the error percentage for our proposed method is nearly zero, whereas the RS based method has around 5% error even at an SNR = 70 dB. This is due to the high condition number of the associated real Vandermonde matrix of the RS based method.

Next, we compare the computation times of the two schemes when the matrix \mathbf{A} is sparse. Consider a system with $n = 5$ workers with $\gamma = \frac{3}{5}$. We choose \mathbf{A} (of dimension $12,000 \times 12,000$) to be of limited bandwidth [10], i.e., it has non-zero values only in the β - diagonals (diagonals from top-left to bottom-right) and $\beta = -b, -b + 1, \dots, -1, 0, 1, \dots, b - 1, b$, where $b < 12,000$. Thus the sparsity of the matrix decreases if b increases.

In the RS based approach, we choose $\Delta = 5$ and assign 3 jobs to each of the workers. For our scheme we embed the computation into a $CP(5, 2)$ code. The parameter $\lambda = 4$, and

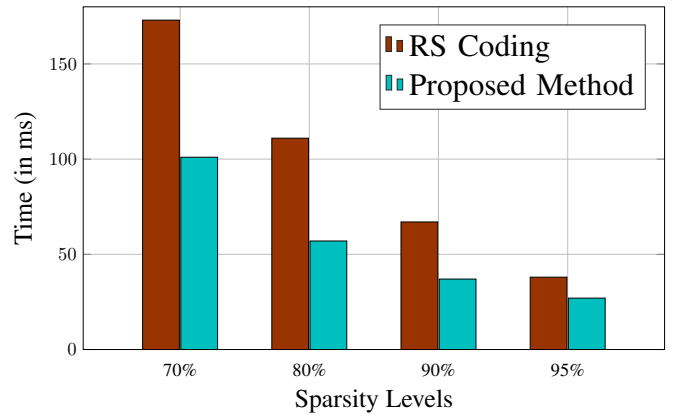


Fig. 3: Comparison between our proposed method and RS coding based method in terms of computation time needed by a worker

$\Delta = \frac{\lambda}{\gamma - \frac{1}{k}} = 40$. Note that both of the schemes are resilient to three stragglers.

Fig. 3 shows the maximum time needed by any worker for different approaches at different sparsity levels. It is evident that the workers take significantly less amount of time in the proposed method in comparison to the RS based approach. This is attributed to the fact that the sparsity level in the coded jobs can be preserved more in our proposed scheme than RS approach. For example, in our experiment, when matrix \mathbf{A} has 90% sparsity, the submatrices assigned to any worker in RS coding approach has around 51% sparsity, whereas in our proposed scheme, even in the worst case, a parity worker can enjoy, on average 70% sparsity. It should be noted that the message workers take further less time since they preserve the same level of sparsity as matrix \mathbf{A} .

VI. CONCLUSION

In this paper, we present an approach for embedding distributed matrix-vector multiplication into the class of cross-parity check convolutional codes towards the goal of straggler mitigation. Our proposed scheme has significant advantages over RS based approaches. The recovery of the intended product is performed using a low-complexity peeling decoder in our scheme as compared to polynomial interpolation in RS based approaches. Unlike RS based approaches which suffer from ill-conditioned recovery matrices, our recovery is numerically quite stable. Finally, for the case of sparse \mathbf{A} matrices, our scheme requires much sparser combinations of the block-rows of \mathbf{A} , leading to faster computation at the worker nodes. Numerical examples have confirmed these results.

REFERENCES

- [1] S. Dutta, V. Cadambe, and P. Grover, "Short-dot: Computing large linear transforms distributedly using coded short dot products," in *Proc. of Advances in Neural Information Processing Systems (NIPS)*, 2016, pp. 2100–2108.
- [2] Q. Yu, M. Maddah-Ali, and S. Avestimehr, "Polynomial codes: an optimal design for high-dimensional coded matrix multiplication," in *Proc. of Advances in Neural Information Processing Systems (NIPS)*, 2017, pp. 4403–4413.

- [3] R. Tandon, Q. Lei, A. G. Dimakis, and N. Karampatziakis, "Gradient coding: Avoiding stragglers in distributed learning," in *Proc. of Intl. Conf. on Machine Learning (ICML)*, 2017.
- [4] K. Lee, M. Lam, R. Pedarsani, D. Papailiopoulos, and K. Ramchandran, "Speeding up distributed machine learning using codes," *IEEE Trans. on Info. Th.*, vol. 64, no. 3, pp. 1514–1529, 2018.
- [5] F. S. Acton, *Numerical methods that (Usually) work*. The Mathematical Association of America, 1990.
- [6] J.-P. Berrut and L. N. Trefethen, "Barycentric lagrange interpolation," *SIAM review*, vol. 46, no. 3, pp. 501–517, 2004.
- [7] T. Fuja, C. Heegard, and M. Blaum, "Cross parity check convolutional codes," *IEEE Transactions on Information Theory*, vol. 35, no. 6, pp. 1264–1276, 1989.
- [8] T. Nagell, *Introduction to number theory*. Chelsea Pub. Co., 1964.
- [9] L. Tang, K. Konstantinidis, and A. Ramamoorthy, "Erasure coding for distributed matrix multiplication for matrices with bounded entries," *IEEE Communications Letters*, vol. 23, no. 1, pp. 8–11, Jan 2019.
- [10] G. H. Golub and C. F. Van Loan, *Matrix computations*. JHU Press, 2012, vol. 3.

APPENDIX A

Now we are going to discuss about the coefficients in the $CP(n, k)$ schemes for $s = 2$ or 3.

Case $s = 2$: In this case, from (4) we can write

$$Z_{ij} = - \prod_{\ell=0, \ell \neq j}^1 \frac{D^{2+i} - D^\ell}{D^j - D^\ell}.$$

If $j = 0$, then we get

$$Z_{i0} = - \frac{D^{2+i} - D}{1 - D} = D^{i+1} + D^i + \dots + D^2 + D.$$

And if $j = 1$, then we get

$$Z_{i1} = - \frac{D^{2+i} - 1}{D - 1} = -(D^{i+1} + D^i + \dots + D + 1).$$

So, we can say the coefficients are 0 or ± 1 .

Case $s = 3$: In this case, from (4) we can write

$$Z_{ij} = - \prod_{\ell=0, \ell \neq j}^2 \frac{D^{3+i} - D^\ell}{D^j - D^\ell}.$$

If $j = 0$, then we can write

$$Z_{i0} = - D^3 \frac{(D^{2+i} - 1)(D^{1+i} - 1)}{(1 - D)(1 - D^2)}.$$

Here, we can have two different cases. If i is even, then

$$Z_{i0} = - D^3 \left(\sum_{r=0}^{\frac{i}{2}} D^{2r} \right) \left(\sum_{r=0}^i D^r \right).$$

On the other hand, if i is odd, then

$$Z_{i0} = - D^3 \left(\sum_{r=0}^{i+1} D^r \right) \left(\sum_{r=0}^{\frac{i-1}{2}} D^{2r} \right).$$

Thus after multiplication, in terms of absolute value, the highest coefficient would be $\lfloor \frac{i}{2} \rfloor + 1$ in both cases. It will also be maximum when i is maximum, $k - 1$, and that coefficient would $\lfloor \frac{k-1}{2} \rfloor + 1$. Next, if $j = 1$, then we can write

$$Z_{i1} = - \frac{(D^{3+i} - 1)(D^{3+i} - D^2)}{(D - 1)(D - D^2)}$$

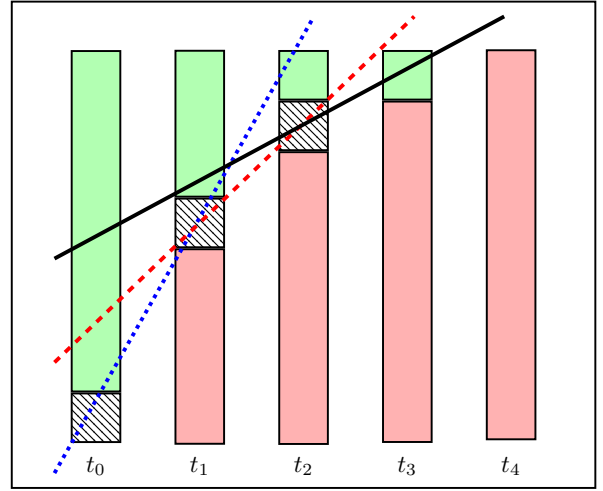


Fig. 4: Recovering blocks from the stragglers, where the green and red parts indicate the decoded and undecoded blocks, respectively

which leads to

$$Z_{i1} = (D^{i+2} + D^{i+1} + \dots + D + 1)(D^{i+1} + D^i + \dots + D).$$

Thus after multiplication, the maximum coefficient would be $i + 1$, and so, the maximum coefficient is k . Finally, if $j = 2$, then we can write

$$Z_{i2} = - \frac{(D^{3+i} - 1)(D^{2+i} - 1)}{(D^2 - 1)(D - 1)}.$$

Here, we can have two different cases again. If i is even, then

$$Z_{i2} = - (D^{i+2} + D^{i+1} + \dots + 1)(D^i + D^{i-2} + \dots + 1).$$

On the other hand, if i is odd, then

$$Z_{i2} = - (D^{i+1} + D^{i-1} + \dots + 1)(D^{i+1} + D^i + \dots + 1).$$

Thus after multiplication, in terms of absolute value, the highest coefficient would be $\lfloor \frac{i+1}{2} \rfloor + 1$ in both cases. It will also be maximum when i is maximum, $k - 1$, and that coefficient would $\lfloor \frac{k}{2} \rfloor + 1$. Thus the result follows.

APPENDIX B

Lemma 2. We assume two stragglers having index t_u and t_v , where $u < v$. We assume that two constraint lines with slopes $(s - 1 - u)$ and $(s - 1 - w)$ pass through the same point c_{α, t_u} , where $w < u$. But, these two lines will pass through two different points, c_{β, t_v} and c_{γ, t_v} in worker t_v , since the slopes are not equal. Then it can be shown that

$$\gamma = \beta - (t_v - t_u)(u - w).$$

Proof. We can find the equation for the constraint line with slope $(s - 1 - u)$ from (1) as

$$\sum_{j=0}^{n-1} c_{i-(s-1-u)j, j} = 0,$$

which passes through c_{α,t_u} and c_{β,t_v} , so we get

$$\alpha = \beta + (t_v - t_u)(s - 1 - u).$$

And the constraint line with slope $(s - 1 - w)$ is given by

$$\sum_{j=0}^{n-1} c_{i-(s-1-w)j,j} = 0,$$

which passes through c_{α,t_u} and c_{γ,t_v} , so we get

$$\alpha = \gamma + (t_v - t_u)(s - 1 - w),$$

which leads to

$$\gamma = \beta - (t_v - t_u)(u - w).$$

■

From this lemma, we can say, if we have different constraint lines with different slopes passing through a particular point α , in a straggler t_u , then the constraint line with a larger slope will pass through a prior point of a straggler t_v , if $u < v$, and vice versa.

We can see an example with five stragglers in Fig. 4, where the green and red blocks indicate that they are decoded and undecoded, respectively. The slopes of the constraint lines are higher for decoding the straggler t_0 than for t_1 . We assume that two different constraint lines from these two workers (blue for t_0 and red for t_1) intersect at a point in t_1 . Then from the lemma we can say, the blue line will always intersect with the next stragglers prior to the red line does. And in a similar way, we can say that because of the larger slope, the red line will intersect with t_3 or t_4 before the black line; since these two lines have already intersected at a point in straggler t_2 .