

Coded Computing via Binary Linear Codes: Designs and Performance Limits

Mahdi Soleymani*, Mohammad Vahid Jamali*, and
Hessam Mahdavifar, *Member, IEEE*

Abstract—We consider the problem of coded distributed computing where a large linear computational job, such as a matrix multiplication, is divided into k smaller tasks, encoded using an (n, k) linear code, and performed over n distributed nodes. The goal is to reduce the average execution time of the computational job. We provide a connection between the problem of characterizing the average execution time of a coded distributed computing system and the problem of analyzing the error probability of codes of length n used over erasure channels. Accordingly, we present closed-form expressions for the execution time using binary random linear codes and the best execution time any linear-coded distributed computing system can achieve. It is also shown that there exist *good* binary linear codes that not only attain (asymptotically) the best performance that any linear code (not necessarily binary) can achieve but also are numerically stable against the inevitable rounding errors in practice. We then develop a low-complexity algorithm for decoding Reed-Muller (RM) codes over erasure channels. Our decoder only involves additions, subtractions, and inversion of relatively small matrices of dimensions at most $\log n + 1$, and enables coded computation over real-valued data. Extensive numerical analysis of the fundamental results as well as RM- and polar-coded computing schemes demonstrate the excellence of the RM-coded computation in achieving close-to-optimal performance while having a low-complexity decoding and explicit construction. The proposed framework in this paper enables efficient designs of distributed computing systems given the rich literature in the channel coding theory.

I. INTRODUCTION

There has been an increasing interest in recent years toward applying ideas from coding theory to improve the performance of various computation, communication, and networking applications. For example, ideas from repetition coding has been applied to several setups in computer networks, e.g., by running a request over multiple servers and waiting for the first completion of the request by discarding the rest of the request duplicates [2]–[4]. Another direction is to investigate the application of coding theory in cloud networks and distributing computing systems [5], [6]. In general, coding techniques can be applied to improve the run-time performance of distributed computing systems.

Distributed computing refers to the problem of performing a large computational job over many, say n , nodes with limited

processing capabilities. A coded computing scheme aims to divide the job to $k < n$ tasks and then to introduce $n - k$ redundant tasks using an (n, k) code, in order to alleviate the effect of slower nodes, also referred to as *stragglers*. In such a setup, it is often assumed that each node is assigned one task and hence, the total number of *encoded tasks* is n equal to the number of nodes.

Recently, there has been extensive research activities to leverage coding schemes in order to boost the performance of distributed computing systems [6]–[18]. However, most of the work in the literature focus on the application of maximum distance separable (MDS) codes. This is while encoding and decoding of MDS codes over real numbers, especially when the number of servers is large, e.g., more than 100, face several barriers, such as numerical stability issues and decoding complexity. In particular, decoding of MDS codes is not robust against unavoidable rounding errors when used over real numbers [19]. Quantizing the real-valued data and mapping them to a finite field over which the computations are carried out [20] can be an alternative approach. However, performing computations over finite fields imposes further numerical barriers due to overflow errors when used over real-valued data.

As we will show in Section III, MDS codes are *theoretically* optimal in terms of minimizing the average execution time of any linear-coded distributed computing system. However, as discussed above, their application comes with some practical impediments, either when used over real-valued inputs or large finite fields, in most of coded computing applications comprised of large number of local nodes. A sub-optimal yet practically interesting approach is to apply binary linear codes, with generator matrices consisting of -1 's and 1 's, and then perform the computation over real numbers. In this case, there is no need for the quantization as the encoded tasks sent to the worker nodes are obtained from a linear combination of the uncoded tasks merely involving additions and subtractions. Inspired by this, in this paper, we consider (n, k) binary linear codes where all computations are performed over real-valued data inputs. To this end, we first derive several fundamental limits to characterize the performance of coded computing schemes employing binary linear codes. We then investigate Reed-Muller (RM) coded computation enabled by our proposed low-complexity algorithm for decoding RM codes over erasure channels. Our decoding algorithm is specifically designed to work over real-valued data and only involves additions, subtractions, and inversion of relatively small matrices of dimensions at most $1 + \log n$.

* Authors of equal contribution.

The material in this paper was presented in part at the IEEE Information Theory Workshop (ITW), Visby, Sweden, Aug. 2019 [1].

The authors are with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109, USA (e-mail: mahdy@umich.edu, mvjamali@umich.edu, hessam@umich.edu).

This work was supported by the National Science Foundation under grants CCF-1763348, CCF-1909771, and CCF-1941633.

A. Related Work

Coded computing paradigm provides a framework to address critical issues that arise in large-scale distributed computing and learning problems such as stragglers, security, and privacy by combining coding theory and distributed computing. For instance, in [6], [8], [10]–[12], [16], [21], [22], coding theoretic techniques have been utilized to combat the deteriorating effects of stragglers in coded computing schemes. The adaptation of such protocols to the analog domain often results in numerical instability, i.e., the accuracy in the computation outcome drops significantly when the number of servers grows large. Furthermore, coded computing schemes have been proposed that enable data privacy and security [23]–[29]. However, in these prior works the data is first quantized and then mapped to a finite field where the tools from coding theory over finite fields can be applied. The performance of such schemes also drops sharply when the dataset size passes a certain threshold due to overflow errors [30], [31].

There is another line of work concerning the adoption of coded computing schemes for straggler mitigation in the analog domain [24]–[28]. Such schemes offer numerical stability but their decoding procedure often relies on inverting a certain matrix which is not scalable with the number of servers. In [32], the authors provide a framework for approximately recovering the evaluations of a function, not necessarily a polynomial, over a dataset which is numerically stable and robust against stragglers. Recently, coded computing schemes have been proposed that enable privacy in the analog domain [30], [31]. Also, codes in the analog domain have been recently studied in the context of block codes [33] as well as subspace codes [34] for analog error correction.

On the other hand, a related work to our RM-coded computing scheme is the recent work in [35] where binary polar codes are applied for distributed matrix multiplication by extending the successive cancellation (SC) decoder of polar codes for real-valued data inputs. However, to the best of our knowledge, our paper is the first to study coded computation over RM codes. As we will show in this paper (see Section VI), RM-coded computation significantly outperforms polar-coded computation in terms of the average execution time. Despite the observations of excellent performance for RM codes in various disciplines (e.g., capacity-achievability [36], [37] and scaling laws [38]), a critical aspect of RM codes is still the lack of efficient decoding algorithms that are scalable to general code parameters with low complexity. Very recently, [39] proposed a recursive projection-aggregation (RPA) algorithm for decoding RM codes over binary symmetric channels (BSCs) and general binary-input memoryless channels. However, neither [39] nor the earlier works on decoding RM codes [40]–[43] directly apply to distributed computation over real-valued data.

B. Our Contributions

In this work, we aim at making a strong connection between the problem of characterizing the average execution time of a coded distributed computing system and the fundamental problem of channel coding over erasure channels. The main

objective of this paper is twofold: 1) characterizing the fundamental performance limits of coded distributed computing systems employing binary linear codes, and 2) designing practical schemes, building upon binary linear codes, that adapt to the natural constraints imposed by the coded computation applications (e.g., operating over real-valued data) while, in the meantime, achieving very close to the fundamental performance limits with a low complexity. The main contributions of the paper are summarized as follows.

- We connect the problem of characterizing the average execution time of any coded distributed computing system to the error probability of the underlying coding scheme over n uses of erasure channels (see Lemma 1).
- Using the above connection, we characterize the performance limits of distributed computing systems such as the average execution time that any coded computation scheme can achieve (see Theorem 2), the average job completion time using binary random linear codes (see Corollary 5), and the best achievable average execution time of a coded computation scheme (see Corollary 6) that can, provably, be attained using MDS codes requiring operations over large finite fields.
- We establish the existence of binary linear codes that attain, asymptotically, the best performance of a coded computing scheme. This important result is established by studying the gap between the average execution time of binary random linear codes and the optimal performance (see Theorem 8), and then showing that the normalized gap approaches zero as $n \rightarrow \infty$ (see Corollary 9 and Corollary 10).
- By studying the numerical stability of the coded computing schemes utilizing binary linear codes, we show that there exist binary linear codes that are *numerically stable* against the inevitable rounding errors in practice while, in the meantime, having an asymptotically optimal average execution time (see Theorem 11).
- We develop an efficient low-complexity algorithm for decoding RM codes over erasure channels. Our decoding algorithm is specifically designed for distributed computation over real-valued data by avoiding any operation over finite fields. Moreover, our decoder is able to achieve very close to the performance of the optimal maximum a posteriori (MAP) decoder (see Section V-B).
- Enabled by our low-complexity decoder, we study the performance of RM-coded distributed computing systems. We also investigate polar-coded computation.
- We carry out extensive numerical analysis confirming our theoretical observations and demonstrating the excellence of RM-coded computation using our proposed decoder.

The rest of the paper is organized as follows. In Section II, we provide the system model and clarify how the system of n independent distributed servers can be viewed as n independent uses of erasure channels. In Section III, by connecting the problem of coded computation to the well-established problem of channel coding over erasure channels, we characterize fundamental limits of coded computation using binary linear codes. In Section IV, we study the numerical stability of

coded computing schemes employing binary linear codes. In Section V, we investigate RM- and polar-coded computation following the presentation of our low-complexity algorithm for decoding RM codes over erasure channels. Finally, we present comprehensive numerical results in Section VI, and conclude the paper in Section VII.

II. SYSTEM MODEL

We consider a distributed computing system consisting of n local nodes with the same computational capabilities. The run time T_i of each local node i is modeled using a shifted-exponential random variable (RV), mainly adopted in the literature [6], [9], [44]. Then, when the computational job is equally divided to k tasks, the cumulative distribution function (CDF) of T_i is given by

$$\Pr(T_i \leq t) = 1 - \exp(-\mu(kt - 1)), \quad \forall t \geq 1/k, \quad (1)$$

where μ is the exponential rate of each local node, also called the straggling parameter. Using (1) one can observe that the probability of the task assigned to the i -th server not being completed (equivalent to erasure) until time $t \geq 1/k$ is

$$\epsilon(t) \triangleq \Pr(T_i > t) = \exp(-\mu(kt - 1)), \quad (2)$$

and is one for $t < 1/k$. Therefore, given any time t , the problem of computing k parts of the computational job over n servers can be interpreted as the traditional problem of transmitting k symbols, using an (n, k) code, over n independent-and-identically-distributed (i.i.d.) erasure channels. Note that the form of the CDF in (1) suggests that $t_0 \triangleq 1/k$ is the (normalized) deterministic time required for each server to process its assigned $1/k$ portion of the total job (all tasks are erased before t_0), while any time elapsed after t_0 refers to the stochastic time as a result of servers' statistical behavior (tasks are not completed with probability $\epsilon(t)$ for $t \geq t_0$).

Given a certain code and a corresponding decoder over erasure channels, a *decodable* set of tasks refers to a pattern of unerased symbols resulting in a successful decoding with probability 1. Then, $P_e(\epsilon, n)$ is defined as the probability of decoding failure over an erasure channel with erasure probability ϵ . For instance, $P_e(\epsilon, 1) = \epsilon$ for a $(1, 1)$ code. Note that the reason to keep n in the notation is to specify that the number of servers, when the code is used in distributed computation, is also n . Finally, the total job completion time T is defined as the time at which a decodable set of tasks/outputs is obtained from the servers.

III. FUNDAMENTAL LIMITS

In this section, we first connect the problem of characterizing the average execution time of any coded distributed computing system to the error probability of the underlying coding scheme over n uses of erasure channels. We then derive several performance limits of coded computing systems such as their average execution time, their average job completion time using binary random linear codes, and their best achievable performance. Finally, we study the gap between the average execution time of binary random linear codes and the optimal performance to establish the existence of binary linear codes

that attain, asymptotically, the best performance of a coded computing scheme.

The following Lemma connects the average execution time of any linear-coded distributed computing system to the error probability of the underlying coding scheme over n uses of an erasure channel.

Lemma 1. *The average execution time of a coded distributed computing system using a given (n, k) linear code can be characterized as*

$$T_{\text{avg}} \triangleq \mathbb{E}[T] = \int_0^\infty P_e(\epsilon(\tau), n) d\tau \quad (3)$$

$$= \frac{1}{k} + \frac{1}{\mu k} \int_0^1 \frac{P_e(\epsilon, n)}{\epsilon} d\epsilon, \quad (4)$$

where $\epsilon(\tau)$ is defined in (2).

Proof: It is well-known that the expected value of any non-negative RV T is related to its CDF $F_T(\tau)$ as $\mathbb{E}[T] = \int_0^\infty (1 - F_T(\tau)) d\tau$. Note that $1 - F_T(\tau) = \Pr(T > \tau)$ is the probability of the event that the job is not completed until some time τ . Therefore, using the system model in Section II, we can interpret $\Pr(T > \tau)$ as the probability of decoding failure $P_e(\epsilon(\tau), n)$ of the code when used over n i.i.d. erasure channels with the erasure probability $\epsilon(\tau)$. This completes the proof of (3). Now given that for the shifted-exponential distribution $d\epsilon(\tau)/d\tau = -\mu k \epsilon(\tau)$, and that $P_e(\epsilon(\tau), n) = 1$ for all $\tau \leq 1/k$, we have (4) by the change of variables. ■

Remark 1. Note that (3) holds given any model for the distribution of the run time of the servers, while (4) is obtained under shifted-exponential distribution, with servers having a same straggling parameter μ , and can be extended to other distributions in a similar approach.

Theorem 2. *The average execution time of any coded distributed computing system can be expressed as*

$$T_{\text{avg}} = \frac{1}{k} \left[1 + \sum_{i=n-k+1}^n \frac{1}{i\mu} \right] + \frac{1}{\mu k} \sum_{i=1}^{n-k} \frac{1}{i} p_{n,k}(i), \quad (5)$$

where $p_{n,k}(i)$ is the average conditional probability of decoding failure of an (n, k) linear code, for an underlying decoder, given that i encoded symbols are erased at random where the average is taken over all possible erasure patterns with i erased symbols.

Proof: Using the law of total probability and the definition of $p_{n,k}(i)$ we have

$$P_e(\epsilon, n) = \sum_{i=1}^n \binom{n}{i} \epsilon^i (1 - \epsilon)^{n-i} p_{n,k}(i). \quad (6)$$

Accordingly, characterizing T_{avg} requires computing integrals of the form $f_i \triangleq \int_0^1 \epsilon^{i-1} (1 - \epsilon)^{n-i} d\epsilon$ for $i = 1, 2, \dots, n$. Using part-by-part integration, one can find the recursive relation $f_{i+1} = \frac{i}{n-i} f_i$ which results in $1/f_i = i \binom{n}{i}$. Note that $p_{n,k}(i) = 1$ for $i > n - k$, since one cannot extract the k parts of the original job from less than k encoded symbols. Then plugging (6) into (4) leads to (5). ■

Next, we characterize the average execution time using a random ensemble of binary linear codes over $\{\pm 1\}^n$. In this

paper, when referred to binary linear codes over real numbers, we always consider codes whose generator matrices only contain ± 1 entries. The aforementioned random ensemble, denoted by $\mathcal{R}(n, k)$, is obtained by picking entries of the $k \times n$ generator matrix independently and uniformly at random followed by removing those matrices that do not have a full row rank from the ensemble.

Remark 2. Note that (6) together with the integral form in (4) suggest that a coded computing system should always encode with a full-rank generator matrix. Otherwise, the average execution time does not converge. This is the reason behind picking the particular ensemble described above. Note that this is in contrast with the conventional block coding, where we can get an arbitrarily small average probability of error over a random ensemble of all $k \times n$ binary generator matrices.

The following lemma provides an upper bound on the probability that a vector picked from $\{\pm 1\}^n$ at random lies in a given subspace of \mathbb{R}^n . We utilize this result later to characterize $p_{n,k}(i)$, defined in Theorem 2, for a random code chosen from $\mathcal{R}(n, k)$.

Lemma 3. ([45, Corollary 4]) For a subspace V of \mathbb{R}^n and r chosen uniformly at random from $\{\pm 1\}^n$ we have

$$\Pr\{\mathbf{r} \in V\} \leq 2^{-\dim(V^\perp)}, \quad (7)$$

where V^\perp is the orthogonal complement of V , and $\dim(V^\perp)$ denotes the dimension of V^\perp .

Next, this result is utilized to provide an upper bound on $p_{n,k}(i)$ for a code whose generator matrix is picked from $\mathcal{R}(n, k)$ uniformly at random.

Lemma 4. The probability that the generator matrix of a code picked from $\mathcal{R}(n, k)$ does not remain full row rank after erasing i columns uniformly at random, denoted by $p_{n,k}^{\mathcal{R}}(i)$, can be upper bounded as

$$p_{n,k}^{\mathcal{R}}(i) \leq 1 - \prod_{j=1}^k (1 - 2^{j-1-n+i}). \quad (8)$$

Proof: Define $l(m, k)$, $k \leq m$, as the probability of k signed Bernoulli uniform random vectors $\mathbf{v}_i \in \{\pm 1\}^m$ being linearly independent. Let V_j denote the subspace spanned by $\mathbf{v}_1, \dots, \mathbf{v}_j$. Then one can write

$$l(m, j+1) = \dots \quad (9)$$

$$l(m, j) \Pr[\mathbf{v}_{j+1} \notin V_j | \mathbf{v}_1, \dots, \mathbf{v}_j \text{ are linearly independent}] \quad (10)$$

$$\geq l(m, j)(1 - 2^{j-m}), \quad (11)$$

where (10) is by the law of total probability and (11) is by the result of Lemma 3. Note also that $l(m, 1) = 1 > 1 - 2^{-m}$. Combining this together with (11) results in

$$l(m, k) \geq \prod_{i=1}^k (1 - 2^{i-1-m}). \quad (12)$$

Note that $p_{n,k}^{\mathcal{R}}(i) \leq 1 - l(n-i, k)$, since rank-deficient $k \times n$ matrices are already excluded from $\mathcal{R}(n, k)$. Combining this together with (12) completes the proof. ■

Corollary 5. The average execution time using binary random linear codes from the ensemble $\mathcal{R}(n, k)$ under maximum a posteriori (MAP) decoding is upper bounded by (5) while replacing $p_{n,k}(i)$ in (5) by $p_{n,k}^{\mathcal{R}}(i)$, upper bounded in Lemma 4.

Proof: The proof is by noting that the optimal MAP decoder fails to recover the k input symbols given $n - i$ unerased encoded symbols if and only if the corresponding $k \times (n - i)$ submatrix of the generator matrix of the code is not full row rank which occurs with probability $p_{n,k}^{\mathcal{R}}(i)$. ■

Remark 3. Theorem 2 implies that the average execution time using linear codes consists of two terms. The first term is independent of the performance of the underlying coding scheme and is fixed given k , n , and μ . However, the second term is determined by the error performance of the coding scheme, i.e., $p_{n,k}(i)$ for $i = 1, 2, \dots, n - k$, and hence, can be minimized by properly designing the coding scheme.

The following corollary of Theorem 2 demonstrates that MDS codes, if they exist,¹ are optimal in the sense that they minimize the average execution time by eliminating the second term of the right hand side in (5). However, for a large number of servers n , the field size needs to be also large, e.g., $q > n$ for Reed-Solomon (RS) codes.

Corollary 6 (Optimality of MDS Codes). For given n , k , and underlying field size q , an (n, k) MDS code, if exists, achieves the minimum average execution time that can be attained by any (n, k) code.

Proof: MDS codes have the minimum distance of $d_{\min}^{\text{MDS}} = n - k + 1$ and can recover up to $d_{\min}^{\text{MDS}} - 1 = n - k$ erasures leading to $p_{n,k}(i) = 0$ for $i = 1, 2, \dots, n - k$. Therefore, the second term of (5) becomes zero for MDS codes and they achieve the following minimum average execution time that can be attained by any (n, k) code:

$$T_{\text{avg}}^{\text{MDS}} = \frac{1}{k} + \frac{1}{\mu k} \sum_{i=n-k+1}^n \frac{1}{i}. \quad (13)$$

Using Theorem 2 and Remark 3, and given that the generator matrix of any (n, k) linear code with minimum distance d_{\min} remains full rank after removing up to any $d_{\min} - 1$ columns, we have the following proposition for the *optimality criterion* in terms of minimizing the average execution time.

Proposition 7 (Optimality Criterion). An (n, k) linear code that minimizes $\sum_{i=d_{\min}}^{n-k} p_{n,k}(i)/i$ also minimizes the average execution time of a coded distributed computing system.

Although MDS codes meet the aforementioned optimality criterion over large field sizes, to the best of our knowledge, the optimal linear codes per Proposition 7, given the field size q and in particular for $q = 2$, are not known and have not been studied before, which calls for future studies.

In the following theorem we characterize the gap between the execution time of binary random linear codes and the

¹It is in general an open problem whether given n , k , and q , there exists an (n, k) MDS code over \mathbb{F}_q [46, Ch. 11.2]. A non-RS type MDS code construction has been proposed in [47]. More recently, a construction of MDS codes with complementary duals has been proposed in [48] which has received attention due to applications in cryptography [49], [50].

optimal execution time. Then Corollary 9 proves that binary random linear codes *asymptotically* achieve the normalized optimal execution time, thereby demonstrating the existence of *good* binary linear codes for distributed computation over real-valued data. The reason we compare the normalized $nT_{\text{avg}}^{\text{MDS}}$'s instead of T_{avg} 's is that, using (5), T_{avg} has a factor of $1/k$ and hence, $\lim_{n \rightarrow \infty} T_{\text{avg}} = 0$ for a fixed rate² $R \triangleq k/n > 0$.

Theorem 8 (Gap of Binary Random Linear Codes to the Optimal Performance). *Let $T_{\text{avg}}^{\text{BRG}}$ denote the average execution time of a coded distributed computing system using binary random linear codes. Then, for any given k, n , we have*

$$\begin{aligned} |nT_{\text{avg}}^{\text{MDS}} - nT_{\text{avg}}^{\text{BRG}}| &\leq (14) \\ \frac{1}{\mu R} \times \left[\frac{v(n)}{n - k - v(n) + 1} + \frac{nR(1 + \ln(n - k - v(n)))}{2^{v(n)}} \right], &\quad (15) \end{aligned}$$

where R is the rate and $v(n)$ is an arbitrary function of n with $0 \leq v(n) \leq n - k$.

Proof: Using Corollary 5 and Corollary 6, we have

$$\mathcal{S} \triangleq \mu R |nT_{\text{avg}}^{\text{MDS}} - nT_{\text{avg}}^{\text{BRG}}| = \sum_{i=1}^{n-k} \frac{1}{i} p_{n,k}^{\mathcal{R}}(i). \quad (16)$$

To prove the upper bound, the summation in (16) is split as $\mathcal{S} = \mathcal{S}_1 + \mathcal{S}_2$ where

$$\mathcal{S}_1 \triangleq \sum_{i=n-k-v(n)+1}^{n-k} \frac{1}{i} p_{n,k}^{\mathcal{R}}(i) \leq \frac{v(n)}{n - k - v(n) + 1}, \quad (17)$$

$$\mathcal{S}_2 \triangleq \sum_{i=1}^{n-k-v(n)} \frac{1}{i} p_{n,k}^{\mathcal{R}}(i). \quad (18)$$

To upper-bound \mathcal{S}_2 , we first note that the upper bound on $p_{n,k}^{\mathcal{R}}(i)$, stated in (8), is a monotonically increasing function of i . Then,

$$\mathcal{S}_2 \leq p_f(n - k - v(n), k) \sum_{i=1}^{n-k-v(n)} \frac{1}{i} \quad (19)$$

$$\leq p_f(n - k - v(n), k) (1 + \ln(n - k - v(n))), \quad (20)$$

where (20) is by the upperbound on the harmonic sum $\sum_{i=1}^n \frac{1}{i} \leq 1 + \ln(n)$. We can further upper-bound $p_f(n - k - v(n), k)$ as

$$p_f(n - k - v(n), k) \leq 1 - \prod_{j=1}^k (1 - 2^{j-1-k-v(n)}) \quad (21)$$

$$\leq 1 - \left[1 - 2^{-v(n)} \right]^k \quad (22)$$

$$\leq nR2^{-v(n)}, \quad (23)$$

where (21) is by (8), (22) follows by noting that

$$\prod_{j=1}^k (1 - 2^{j-1-k-v(n)}) = \prod_{j'=1}^k (1 - 2^{-j'-v(n)}) \geq [1 - 2^{-v(n)}]^k, \quad (24)$$

²More precisely, the coding rate over field size q is equal to $k \log_2 q/n$ but with slight abuse of terminology we have dropped the factor of $\log_2 q$ since this factor is not relevant for coded distributed computing.

and (23) follows by Bernoulli's inequality $(1 - x)^k \geq 1 - kx$ for any $0 < x < 1$ and then inserting $k = nR$. \blacksquare

Corollary 9 (Asymptotic Optimality of Binary Random Linear Codes). *The normalized average execution time $nT_{\text{avg}}^{\text{BRG}}$ approaches $nT_{\text{avg}}^{\text{MDS}}$ as n grows large. More precisely, for a given rate R , there exists a constant $c > 0$ such that for sufficiently large n , i.e., $k = nR$, we have*

$$nT_{\text{avg}}^{\text{BRG}} - nT_{\text{avg}}^{\text{MDS}} \leq c \frac{\log_2 n}{n}. \quad (25)$$

Proof: Observe that with the choice of $v(n) = 2 \log_2 n$ both terms in the right hand side of (15) become $O(\frac{\log_2 n}{n})$. Note that $n - k = n(1 - R) \geq 2 \log_2 n$, for sufficiently large n . Hence, the upper bound of (25) also holds with a proper choice of c . \blacksquare

Remark 4. For any given n , one can obtain the optimal value of k and, subsequently, the optimal value of the encoding rate R that minimizes $T_{\text{avg}}^{\text{MDS}}$ in (13). The limit of the optimal value of R when $n \rightarrow \infty$ is referred to as the asymptotically-optimal encoding rate and is denoted by R^* . Using (13) and a similar approach to [6], one can show that the asymptotically-optimal encoding rate R^* for an MDS-coded distributed computing system is the solution to

$$(1 - R^*) \ln(1 - R^*) = \mu(1 - R^*) - R^*. \quad (26)$$

Corollary 9 implies that for distributed computation using binary random linear codes, the gap between $nT_{\text{avg}}^{\text{BRG}}$ and $nT_{\text{avg}}^{\text{MDS}}$ converges to zero as n grows large. Accordingly, the optimal encoding rate also approaches R^* , described in (26).

Corollary 10. *Let T^{MDS} and T^{BRG} denote the execution time of the coded computing schemes using an MDS code and a code whose generator matrix is picked from $\mathcal{R}(n, k)$ at random, respectively. Then, there exists a constant c such that*

$$\Pr[n(T^{\text{BRG}} - T^{\text{MDS}}) \geq x] \leq \frac{c \log_2 n}{nx}. \quad (27)$$

Proof: The proof follows immediately by using the result of Corollary 9 and Markov's inequality. \blacksquare

IV. NUMERICAL STABILITY OF RANDOM BINARY LINEAR CODES

In this section, we study the numerical stability of the coded computing schemes utilizing binary linear codes. Our results indicate that there exist binary linear codes that are *numerically stable* against the inevitable rounding errors in practice which also have asymptotically optimal average execution time.

In general, in a system of linear equations $\mathbf{Ax} = \mathbf{b}$, where \mathbf{x} is a vector of unknown variables and \mathbf{A} is referred to as the *coefficient matrix*, the perturbation in the solution caused by the perturbation in \mathbf{b} is characterized as follows. Let $\hat{\mathbf{b}}$ denote a noisy version of \mathbf{b} , where the noise can be caused by round-off errors, truncation, etc. Let also $\hat{\mathbf{x}}$ denote the solution to the considered linear system when \mathbf{b} is replaced by $\hat{\mathbf{b}}$. Let $\Delta \mathbf{x} \triangleq \hat{\mathbf{x}} - \mathbf{x}$ and $\Delta \mathbf{b} \triangleq \hat{\mathbf{b}} - \mathbf{b}$ denote the perturbation in

\mathbf{x} and \mathbf{b} , respectively. Then the relative perturbations of \mathbf{x} is bounded in terms of that of \mathbf{b} as follows [51]:

$$\frac{\|\Delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq \kappa_{\mathbf{A}} \frac{\|\Delta \mathbf{b}\|}{\|\mathbf{b}\|}, \quad (28)$$

where $\kappa_{\mathbf{A}}$ is the condition number of \mathbf{A} and $\|\cdot\|$ denotes the ℓ^2 -norm.

The perturbation bound stated in (28) implies that the precision loss in the final outcome is $\log_{10} \kappa_{\mathbf{A}}$ in decimal digits, where the matrix \mathbf{A} is the submatrix of the generator whose rows correspond to non-straggling worker nodes. The precision loss in decoding procedure of the codes constructed over real and complex numbers has been studied in the literature and some codes with deterministic constructions are provided [52]–[56]. A code with random Gaussian generator matrix is a numerically stable code with high probability. This is mainly due to the fact that any submatrix of a Gaussian random matrix with i.i.d. entries is also a random Gaussian matrix and such matrices are ill-conditioned only with small probability. The random Gaussian codes are often considered as a benchmark to evaluate the numerical stability of codes over real numbers with explicit constructions [57]–[59]. The result of this section implies that the random binary linear codes also offer the same numerical stability as random Gaussian codes. The motivation for using binary linear codes instead of the random Gaussian codes or the existing codes with explicit construction over real numbers is that they can offer a better decoding complexity. This will be clarified further later in Section VI when we compare the performance of practical codes with random codes.

Let $\mathcal{G}(n, k)$ denote a random ensemble of Gaussian codes. This random ensemble is obtained by picking entries of the $k \times n$ generator matrix independently and at random from the standard normal distribution. Let \mathbf{G} denote a matrix picked randomly from $\mathcal{G}(n, k)$ and $\tilde{\mathbf{G}}$ denote a random $k \times k$ submatrix of \mathbf{G} . Note that $\tilde{\mathbf{G}}$ is also a random Gaussian matrix. The probability bounds on the condition number of a random Gaussian matrix are provided in [58]. In particular,

$$\Pr[\kappa_{\tilde{\mathbf{G}}} > nx] < \frac{1}{\sqrt{2\pi}} \frac{C}{x}, \quad (29)$$

where $5.013 < C < 6.414$. Consequently, the precision loss in recovery of the computation outcome for the coded computing system using random Gaussian codes is $O(\log_{10} k)$ with high probability.

The behavior of the largest and smallest singular values of random matrices with i.i.d. entries has been extensively studied in the literature. In particular, we use such results for random matrices with *sub-Gaussian* random variables. Recall that a random variable X is called sub-Gaussian if its tail is dominated by that of the standard normal random variable, i.e., if there exists $B > 0$ such that

$$\Pr[|X| > t] \leq 2 \exp\left(\frac{-t^2}{B^2}\right). \quad (30)$$

The minimal B is called the *Gaussian moment* of X [60]. Note that a signed Bernoulli random variable X with

$$\Pr[X = 1] = \Pr[X = -1] = \frac{1}{2}$$

is sub-Gaussian. Let λ_{\max} and λ_{\min} respectively denote the largest and the smallest singular value of a random $k \times k$ matrix whose entries are independent zero-mean sub-Gaussian random variables. Then,

$$\Pr[\lambda_{\max} > Ck^{\frac{1}{2}} + t] \leq 2 \exp(-ct^2), \quad (31)$$

where c and C are absolute constants [61]. Moreover, if the variance of the underlying sub-Gaussian random variable is at least 1, we have

$$\Pr[\lambda_{\min} \leq \epsilon k^{-\frac{1}{2}}] \leq C' \epsilon + c'^k \quad (32)$$

for all $\epsilon > 0$, where C' and c' are constants depending polynomially on the sub-Gaussian moment [60]. It is worth mentioning that the best known c' for the case of Bernoulli random variable is $\frac{1}{\sqrt{2}} + o(1)$ [62]. The bounds provided in (31) and (32) imply that the condition number of a $k \times k$ random Bernoulli matrix is also $O(n)$ with high probability. Hence, the precision loss in recovery of the outcome of a coded computing scheme utilizing random Gaussian codes and random binary linear codes are *almost* the same. This together with the result of Corollary 10 in Section III imply that there exist binary linear codes that are numerically stable with asymptotically optimal recovery time. This result is stated in the following theorem.

Theorem 11. *Let \mathbf{G} denote a $k \times n$ matrix picked from $\mathcal{R}(n, k)$ at random. Let also T^{MDS} and T^{BRC} denote the execution time of the coded computing schemes using an MDS code and a code whose generator matrix is \mathbf{G} , respectively. Then, the coded computing scheme utilizing the binary linear code generated by \mathbf{G} recovers the computation outcome with $O(\log_{10} \frac{k}{\epsilon})$ precision loss in decimal digits in $T^{\text{BRC}} \leq T^{\text{MDS}} + O(\frac{1}{n})$ time with probability $1 - O(\epsilon + \frac{\log_2 n}{n})$.*

Proof: Let $\tilde{\mathbf{G}}$ denote a random $k \times k$ submatrix of \mathbf{G} . Combining (31) with (32) together with the union bound implies

$$\Pr[\kappa_{\tilde{\mathbf{G}}} > \frac{C}{\epsilon} k + \epsilon t \sqrt{k}] \leq 2 \exp(-ct^2) + C' \epsilon + c'^k. \quad (33)$$

Combining (27) and (33) together with the union bound implies

$$\begin{aligned} \Pr[(\kappa_{\tilde{\mathbf{G}}} > \frac{C}{\epsilon} k + \epsilon t \sqrt{k}) \text{ or } (T^{\text{BRC}} - T^{\text{MDS}} > \frac{x}{n})] \\ \leq \exp(-ct^2) + C' \epsilon + c'^k + \frac{c \log_2 n}{nx} = O(\epsilon + \frac{\log_2 n}{n}) \end{aligned}$$

for all $\epsilon > 0$, which completes the proof. \blacksquare

The result of Theorem 11 implies that there exist numerically stable binary linear codes with asymptotically optimal average execution time. In the rest of the paper, we consider some coding schemes over real numbers constructed based on codes over \mathbb{F}_2 , namely, RM and polar codes, that offer lower decoding complexity than MDS codes. The numerical stability of such schemes are naturally inherited from the proposed decoding algorithms that involve additions, subtractions and, in the case of RM codes, inverting logarithmic-size matrices that are well-conditioned, as numerically verified in the next section. Moreover, their average execution times are also

compared numerically with the optimal values for a wide range of blocklengths (i.e., number of servers).

V. PRACTICAL CODED COMPUTING SCHEMES

In this section, we explore RM- and polar-coded distributed computation. First, we briefly review RM codes and polar codes, two closely-connected classes of codes, in Section V-A. Then, in Section V-B, we present our proposed low-complexity algorithm for decoding RM codes over erasure channels that enables RM-coded distributed computing over real-valued data. Finally, we present polar-coded computation in Section V-C.

A. Brief Review of RM and Polar Codes

Let k and n be the code dimension and blocklength, respectively, and let $m \triangleq \log_2 n$ be a design parameter. Then, the r -th order RM code of length 2^m , denoted by $\mathcal{RM}(m, r)$, is defined by the following set of vectors as the basis

$$\{\mathbf{v}_m(\mathcal{A}) : \mathcal{A} \subseteq [m], |\mathcal{A}| \leq r\}, \quad (34)$$

where $|\mathcal{A}|$ denotes the size of the set \mathcal{A} , and $[m] \triangleq \{1, 2, \dots, m\}$. Moreover, $\mathbf{v}_m(\mathcal{A})$ is a row vector of length 2^m whose components are indexed by binary vectors $\mathbf{z} = (z_1, z_2, \dots, z_m) \in \{0, 1\}^m$. Each component of $\mathbf{v}_m(\mathcal{A})$ is obtained as $\mathbf{v}_m(\mathcal{A}, \mathbf{z}) = \prod_{i \in \mathcal{A}} z_i$. In other words, considering a polynomial ring $\mathbb{F}_2[Z_1, Z_2, \dots, Z_m]$ of m variables, the components of $\mathbf{v}_m(\mathcal{A})$ are the evaluations of the monomial $\prod_{i \in \mathcal{A}} Z_i$ at points \mathbf{z} in the vector space $\mathbb{E} \triangleq \mathbb{F}_2^m$. It is easy to observe from (34) that there are $\sum_{i=0}^r \binom{m}{i}$ basis (equivalently, \mathcal{A} 's) in total, and thus an $\mathcal{RM}(m, r)$ code has a dimension of $k = \sum_{i=0}^r \binom{m}{i}$.

Finally, given the set of basis in (34), the (codebook of) $\mathcal{RM}(m, r)$ code can be defined as the following set of 2^k binary vectors

$$\mathcal{RM}(m, r) \triangleq \left\{ \sum_{\mathcal{A} \subseteq [m], |\mathcal{A}| \leq r} u(\mathcal{A}) \mathbf{v}_m(\mathcal{A}) : u(\mathcal{A}) \in \{0, 1\} \right\}. \quad (35)$$

Therefore, each codeword $\mathbf{c} = (\mathbf{c}(\mathbf{z}), \mathbf{z} \in \mathbb{E}) \in \mathcal{RM}(m, r)$, that is indexed by the binary vectors \mathbf{z} , is defined as the evaluations of an m -variate polynomial with degree at most r at points $\mathbf{z} \in \mathbb{E}$.

While RM codes have a universal construction, the construction of polar codes, on the other hand, is *channel-specific*.

Consider Arikan's $n \times n$ polarization matrix $\mathbf{G}_n = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{\otimes m}$, where $m = \log_2 n$ and $\mathbf{A}^{\otimes m}$ denotes the m -th Kronecker power of \mathbf{A} . The encoding of polar codes is obtained from the aforementioned polarization matrix \mathbf{G}_n in a channel-specific manner. Particularly, in the case of binary erasure channels (BECs), a design parameter ϵ_d is picked, as specified later in Section VI. Then the polarization transform \mathbf{G}_n is applied to a BEC with erasure probability ϵ_d , $\text{BEC}(\epsilon_d)$. The erasure probabilities of the polarized bit-channels, denoted by $\{Z_i\}_{i=1}^n$, are sorted and the k rows of \mathbf{G}_n corresponding to

the indices of the k smallest Z_i 's are picked to construct the $k \times n$ generator matrix \mathbf{G} .

One can also obtain an equivalent encoding of $\mathcal{RM}(m, r)$, similar to that of polar codes, by selecting rows of the square matrix \mathbf{G}_n that have a Hamming weight of at least 2^{m-r} . In this case, the resulting generator matrix $\mathbf{G}_{k \times n}$ will have $\binom{m}{i}$ rows of Hamming weight $n/2^i$, for $i = 0, 1, \dots, r$.

B. RM-Coded Distributed Computation

It has recently been shown that RM codes are capacity achieving over BECs [36], though under bit-MAP decoding, and numerical results suggest that they actually achieve the capacity with *almost* optimal scaling [38]. RM codes also achieve the capacity of BSCs at extreme rates, i.e., at rates converging to zero or one [37]. They are also conjectured to have characteristics similar to those of random codes in terms of both scaling laws [38] and weight enumeration [63]. Despite all these excellent properties, RM codes still lack efficient low-complexity decoders for general code dimensions and blocklengths. Very recently, Ye and Abbe [39] proposed a recursive projection-aggregation (RPA) algorithm for decoding RM codes over BSCs and general binary-input memoryless channels. The RPA algorithm is comprised of three main steps: 1) projecting the received corrupted codeword onto the cosets defined by each projection subspace, 2) recursively decoding the projected codewords, and 3) aggregating the decoded codewords at the next layer with the current observation to finally decode the original RM codeword.

In this section, we propose an efficient RPA-like algorithm for decoding RM codes over erasure channels. Our decoding algorithm has three major novelties. First, it only involves additions, subtractions, and inverting relatively small matrices of size no more than $\log n + 1$. We need to emphasize that the RPA algorithms proposed in [39] work over the binary field and do not directly apply to real-valued inputs. For example, as detailed in Section V-B1, the projection step in the original RPA algorithms requires addition of the received bits over the binary field (i.e., XOR'ing them). Therefore, our proposed decoding algorithm generalizes the RPA algorithms to the case of erasure channels while avoiding operations over finite fields. Second, our decoding algorithm has a low complexity achieved by carefully selecting a small fraction of the total number of projections, i.e., only $\binom{m}{r-1}$ projection subspaces of dimension $s = r - 1$ are selected to decode an $\mathcal{RM}(m, r)$ code. Therefore, our decoding algorithm enables decoding RM codes of higher orders and lengths with a manageable complexity. Third, our simulation results suggest that our decoding algorithm is able to achieve very close to the performance of optimal MAP decoding while maintaining a low complexity.

In the following, we explain our decoding algorithm that is comprised of three main steps separately described in Sections V-B1, V-B2, and V-B3. Assuming an $\mathcal{RM}(m, r)$ code, our decoding algorithm only applies one layer of projection using s -dimensional subspaces with $s = r - 1$. It then applies MAP decoding (block-MAP in conjunction with bit-MAP) to decode the projected codewords, and finally aggregates the results to

recover the codeword. As we show in Section V-B1, a smaller value of s increases the complexity of the decodings performed at the bottom layer but it can also result in a better decoding performance since the MAP decoding over projected vectors will be utilized to a greater extent (recall that there is only a single layer of projection). It will be clarified later that this choice of s can result in a manageable decoding complexity while not much sacrificing the decoding performance.

1) *Projection*: The decoding algorithm starts by projecting the received corrupted codeword onto the cosets defined by the projection subspaces. More specifically, let \mathbb{B}_i be a s -dimensional subspace of \mathbb{E} , with $s \leq r$ and $i \in [\mathcal{I}]$ where \mathcal{I} is the total number of projection subspaces. The quotient space \mathbb{E}/\mathbb{B}_i contains all the cosets of \mathbb{B}_i in \mathbb{E} where each coset τ has the form $\tau = \mathbf{z} + \mathbb{B}$ for some $\mathbf{z} \in \mathbb{E}$. Given a length- n codeword and a s -dimensional subspace, the objective of the projection step is to obtain a length- $n/2^s$ vector whose each component, that corresponds to one of the $n/2^s$ cosets, is an *appropriate* representative of the bits indexed by the 2^s elements of that coset. Building upon this notion and the definition of the log-likelihood ratio (LLR), the expressions for the projection of the channel LLRs over general binary-input memoryless channels are derived in [39] for subspaces of dimension $s = 1$ and 2. By applying the same principles, one can show that the projection of the length- n vector of the corrupted codeword over a BEC can be obtained by adding the bits indexed by the cosets over the binary field (i.e., XOR'ing) while assuming the addition of bits 0 or 1 with an erasure is also an erasure. In other words, considering $\mathbf{y} = (\mathbf{y}(\mathbf{z}), \mathbf{z} \in \mathbb{E})$ as the length- n vector of the corrupted codeword over a BEC, the projected vector onto a subspace \mathbb{B} can be obtained as $\mathbf{y}_{/\mathbb{B}} \triangleq (\mathbf{y}_{/\mathbb{B}}(\tau), \tau \in \mathbb{E}/\mathbb{B})$ such that $\mathbf{y}_{/\mathbb{B}}(\tau) \triangleq \bigoplus_{\mathbf{z} \in \tau} \mathbf{y}(\mathbf{z})$, where \bigoplus denotes the coordinate-wise addition in \mathbb{F}_2 while assuming that additions with an erasure is also an erasure.

While the projection rule $\mathbf{y}_{/\mathbb{B}}(\tau) = \bigoplus_{\mathbf{z} \in \tau} \mathbf{y}(\mathbf{z})$ works perfectly over the binary field, we are looking for a low-complexity decoder that works over real numbers to avoid numerical issues (caused by working over finite fields). To this end, considering the received vector as $\mathbf{y} = (\mathbf{y}(\mathbf{z}), \mathbf{z} \in \mathbb{E})$, we can obtain the projected vector by linearly combining (over real numbers) the entries indexed by the cosets to obtain $\mathbf{y}_{/\mathbb{B}} \triangleq (\mathbf{y}_{/\mathbb{B}}(\tau), \tau \in \mathbb{E}/\mathbb{B})$ such that

$$\mathbf{y}_{/\mathbb{B}}(\tau) \triangleq \sum_{\{i: \mathbf{z}_i \in \tau\}} \gamma_i \mathbf{y}(\mathbf{z}_i), \quad (36)$$

where γ_i 's are some properly-chosen real-valued combining coefficients, as clarified later. As clarified in [64, Remark 1], the result of the projection operation can be thought of as obtaining a generator matrix that is formed by merging the columns of the original code generator matrix indexed by the cosets of the projection subspace. In other words, given a $k \times n$ generator matrix $\mathbf{G}_{k \times n}$, we can define matrices of dimension $k \times n/2^s$, referred to as *projected generator matrices*, each obtained by merging the columns of the original generator matrix indexed by the cosets of each projection subspace. In this paper, assuming an $\mathcal{RM}(m, r)$ code, we work based on projection subspaces of dimension $s = r - 1$. In the binary field, this choice of s will result in order-1

RM codes, i.e., $\mathcal{RM}(m - r + 1, 1)$ codes, at the bottom layer after the projection [39]. These codes all have the same dimension of $m - r + 2$ that is also equal to the rank of the projected generator matrices. However, for our decoder that works on real-valued inputs the aforementioned projected generator matrices can have different ranks and they no longer (necessarily) correspond to the generator matrices of the lower order (and lower length) RM codes.

Next, we discuss how to choose γ_i 's in (36). In our proposed decoder, we pick $\gamma_i \in \{-1, +1\}$ such that combining coefficients result in the same rank of $m - r + 2$ as the lower order RM code $\mathcal{RM}(m - r + 1, 1)$ after an $(r - 1)$ -dimensional projection of $\mathcal{RM}(m, r)$. To illustrate this, and for simplicity, we consider the one-dimensional projections on $\mathcal{RM}(m, r)$. Note that the argument for general s -dimensional projections, and, consequently, for $s = r - 1$, would follow naturally by cascading s one-dimensional projections. Note that there are, at least, m projections of $\mathcal{RM}(m, r)$ over real numbers that lead to the $\mathcal{RM}(m - 1, r - 1)$ code. These projections correspond to standard basis vectors \mathbf{e}_j 's, where \mathbf{e}_j , for $j \in [m]$, is a length- m vector with a 1 in the j -th position and 0 in all other $m - 1$ positions. For instance, the projection corresponding to $\mathbf{e}_m = (0, 0, \dots, 1)$ leads to pairing of consecutive column indices of the generator matrix of $\mathcal{RM}(m, r)$, i.e., $(0, 1), (2, 3), \dots, (n - 2, n - 1)$. The main observation, which enables this whole process, is that for each pair $(2i, 2i + 1)$, the support of the column $2i + 1$ is a subset of the support of the column $2i$ (when considering entries to be 0's and 1's). Therefore, if we subtract column $2i + 1$ from column $2i$, in the real domain, that would correspond to XORing them in the binary domain. Hence, we get the generator matrix of $\mathcal{RM}(m - 1, r - 1)$ with this particular choice of γ_i 's for each of these m projections. Also, a simple normalization by 2 is done in order to keep matrix entries belonging to the set $\{-1, +1\}$. For general s , one can choose $\mathcal{I} = \binom{m}{s}$ s -dimensional subspaces \mathbb{B}_i each obtained by the span of s standard basis vectors \mathbf{e}_j 's and apply the aforementioned process recursively. In particular, for $s = r - 1$, we can efficiently find the choices of γ_i 's in $\{-1, +1\}$ that lead to projection of $\mathcal{RM}(m, r)$ into $\mathcal{RM}(m - r + 1, 1)$ using $\binom{m}{r-1}$ projections.

Note that, given an $\mathcal{RM}(m, r)$ code for the coded distributed computation, the set of the projection subspaces, the cosets, the combining coefficients (that result in the rank of $m - r + 2$), and the corresponding projected generator matrices will be computed before hand to lower the complexity of the decoder and prevent any operation over finite fields during the decoding process. It is worth mentioning at the end that the specific selection of the combining coefficients as described above has several advantages. First, having the combining coefficients being either -1 or $+1$ renders a projection step that only involves additions and subtractions (and no multiplications). It also equally weights all the entries to be combined which in turn can prevent numerical stability issues that may arise from some of the weights being very large or very small. Second, the corresponding generator matrices all will have the smallest possible rank. This is a twofold gain: 1) it lowers the complexity of the MAP decoding over the projected vectors,

and 2) it can also improve the error rate performance of the decodings at the bottom layer by increasing the chances of getting the same rank for the projected generator matrices after erasing some of the columns (see Section V-B2).

2) *Decoding of the Projected Vectors:* Once the projection step is completed, we end up with \mathcal{I} length- 2^{m-r+1} projected vectors $\mathbf{y}_{/\mathbb{B}_i}$'s, $i \in [\mathcal{I}]$, each corresponding to a $k \times 2^{m-r+1}$ projected generator matrix of rank $m-r+2$. Given that we are working over erasure channels, we either know an entry perfectly (without error) or we do not know it at all. The objective of this step is to decode the projected vectors by decoding either all the erased entries (block-MAP) or a fraction of them (bit-MAP) in each of these projected vectors. For each $\mathbf{y}_{/\mathbb{B}_i}$, let \mathcal{Y}_i^e and \mathcal{Y}_i^{ne} be the sets of the indices of the erased and non-erased entries, respectively. Also let $\mathbf{G}_{p,i}^{ne}$ denote the $k \times |\mathcal{Y}_i^{ne}|$ submatrix of the projected generator matrix (over the subspace \mathbb{B}_i) comprised of the $|\mathcal{Y}_i^{ne}|$ columns corresponding to non-erased entries in $\mathbf{y}_{/\mathbb{B}_i}$. Then if the rank of $\mathbf{G}_{p,i}^{ne}$ is equal to $m-r+2$ (i.e., if erasing $|\mathcal{Y}_i^e|$ columns of the projected generator matrix does not change its rank), the block-MAP decoder can be applied to decode all the erased entries. This is equivalent to say that each of the erased entries can be obtained as a linear combination of the non-erased entries. On the other hand, if the block-MAP decoder fails, there is still a chance that the bit-MAP decoder can decode some of the entries. In particular, if rank remains the same after adding to $\mathbf{G}_{p,i}^{ne}$, a column corresponding to one of the $|\mathcal{Y}_i^e|$ indices, then that particular erased entry can be recovered as a linear combination of non-erased entries. By performing the same procedure for all $\mathbf{y}_{/\mathbb{B}_i}$'s, we obtain the decoded projected vectors $\hat{\mathbf{y}}_{/\mathbb{B}_i}$'s for all $i \in [\mathcal{I}]$.

3) *Aggregation:* The objective in this step is to combine the observation from the channel output, i.e., the corrupted vector \mathbf{y} , with that of the projected vectors, i.e., $\hat{\mathbf{y}}_{/\mathbb{B}_i}$'s, to obtain the decoded vector $\hat{\mathbf{y}}$. Let \mathcal{Y}^e be the set of the indices of the erased entries in \mathbf{y} . In the following, we describe our aggregation method assuming a given projection subspace \mathbb{B}_i , $i \in [\mathcal{I}]$, and an erased index $l \in \mathcal{Y}^e$. By applying the same procedure for all projection subspaces and all erased indices, we can obtain the decoded vector $\hat{\mathbf{y}}$.

Let \mathbf{z} be the binary vector indexing the l -th position which is an erasure. Also, let $[\mathbf{z} + \mathbb{B}_i]$ denote the cost of \mathbb{B}_i that contains \mathbf{z} . Since \mathbb{B}_i is an s -dimensional subspace (where $s = r-1$ in this paper), there are $2^s - 1$ elements, denoted by $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_{2^s-1}$, in $[\mathbf{z} + \mathbb{B}_i]$ apart from \mathbf{z} itself. Given the decoded projected vector $\hat{\mathbf{y}}_{/\mathbb{B}_i}$, we know from the projection step that $\hat{\mathbf{y}}_{/\mathbb{B}_i}([\mathbf{z} + \mathbb{B}_i])$ is an estimation of $\gamma \mathbf{y}(\mathbf{z}) + \sum_{j=1}^{2^s-1} \gamma_j \mathbf{y}(\mathbf{z}_j)$, where $\{\gamma, \gamma_1, \dots, \gamma_{2^s-1}\}$ is the set of the predefined combining coefficients for this particular projection. Now, if $\hat{\mathbf{y}}_{/\mathbb{B}_i}([\mathbf{z} + \mathbb{B}_i])$ and all $\mathbf{y}(\mathbf{z}_j)$'s, for $j = 1, \dots, 2^s - 1$, are known (i.e., none of them are erased), we can recover $\mathbf{y}(\mathbf{z})$ as

$$\hat{\mathbf{y}}(\mathbf{z}) = \frac{1}{\gamma} \left[\hat{\mathbf{y}}_{/\mathbb{B}_i}([\mathbf{z} + \mathbb{B}_i]) - \sum_{j=1}^{2^s-1} \gamma_j \mathbf{y}(\mathbf{z}_j) \right]. \quad (37)$$

Note that since we are working over erasure channels, if any projection satisfies the above condition, we do not need to

check the other projections for decoding the erased entry at the l -th index.

Example 1. Consider a corrupted codeword of $\mathcal{RM}(3, 2)$, denoted by $\mathbf{y} = (y_1, y_2, \dots, y_8)$, with an erasure at the fifth position. Also, we consider three projection subspaces as $\mathbb{B}_1 = \{(0, 0, 0), (1, 0, 0)\}$, $\mathbb{B}_2 = \{(0, 0, 0), (0, 1, 0)\}$, and $\mathbb{B}_3 = \{(0, 0, 0), (0, 0, 1)\}$. Consider projection in the direction of \mathbb{B}_1 . For this projection, the quotient space is $\mathbb{E}/\mathbb{B}_1 = \{(1, 5), (2, 6), (3, 7), (4, 8)\}$. Additionally, one can observe that the choice of combining coefficients $\gamma_1 = -1$ and $\gamma_2 = 1$ results in the minimum rank of $m-r+2 = 3$ for all three projected generator matrices. Accordingly, the projected received vector \mathbf{y}/\mathbb{B}_1 is of the form $(y_5 - y_1, y_6 - y_2, y_7 - y_3, y_8 - y_4)$ which will have an erasure only at the first position. Since the corresponding projected generator matrix remains rank-3 after removing the first column, the block-MAP decoder will be able to recover the single erasure in \mathbf{y}/\mathbb{B}_1 . Similarly, one can show that the other two projected received vectors \mathbf{y}/\mathbb{B}_2 and \mathbf{y}/\mathbb{B}_3 will both have a single erasure at the third position. Given that their corresponding projected generator matrices also remains rank-3 after removing the third column, the decoder will be able to recover the erased bits at the third indices. Finally, given the successful decoding of the projected vectors, one can successfully decode the corrupted codeword after aggregation.

Figure 1 compares the block error rate (BLER) results for our projective decoder with that of optimal MAP decoder for various RM codes of interest to distributed computing (see Table II for the rationale behind the code parameters in this figure). It is remarkable that our decoder can achieve almost the same performance as that of the MAP decoder with only $\binom{m}{r-1}$ projections for an $\mathcal{RM}(m, r)$ code³. In fact, only $\mathcal{I} = 3, 4, 10$, and 15 projection subspaces are selected for $\mathcal{RM}(3, 2)$, $\mathcal{RM}(4, 2)$, $\mathcal{RM}(5, 3)$, and $\mathcal{RM}(6, 3)$, respectively, which are significantly less than the full number of projections for RPA-like decoding of these codes. Note that the decoder may also iterate the whole process, described in this section, a few times to ensure the convergence of the algorithm. The convergence here means that there is no difference between what is known about the codeword, in terms of the corrected symbols, at the end of the current iteration with that of the previous iteration. Note that the number of outer iterations, denoted by N_{\max} , is not much to cause a serious complexity issue for our algorithm. In particular, in Figure 1, the maximum number of outer iterations is chosen to be $N_{\max} = 1, 2, 2$, and 3 for $\mathcal{RM}(3, 2)$, $\mathcal{RM}(4, 2)$, $\mathcal{RM}(5, 3)$, and $\mathcal{RM}(6, 3)$, respectively.

Next, we discuss the numerical stability of the proposed RM-coded computing system. As discussed earlier, the decoding algorithm only involves additions and subtractions, as well as inverting small matrices of size at most $\log n + 1$. To make sure that these matrices are well-conditioned, we did extensive numerical analysis by looking at the condition

³We have observed that the gap between the performance of our low-complexity decoding algorithm and that of MAP increases as we increase n (e.g. in $\mathcal{RM}(7, 4)$ and $\mathcal{RM}(8, 5)$ codes that have close code parameters to the RM subcodes considered in Table I). However, very large n 's can be less relevant in the context of distributed computing as n here corresponds to the number of servers.

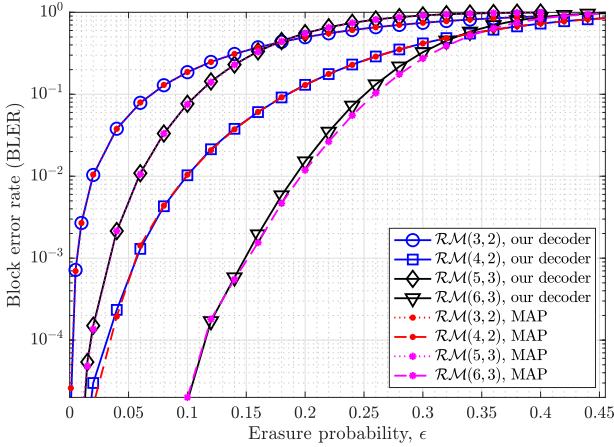


Figure 1. Block error rate (BLER) results for the MAP decoder and our projective decoder for various RM codes of interest to distributed computing. The maximum number of outer iterations is chosen to be $N_{\max} = 1, 2, 2$, and 3 for $\mathcal{RM}(3, 2)$, $\mathcal{RM}(4, 2)$, $\mathcal{RM}(5, 3)$, and $\mathcal{RM}(6, 3)$, respectively.

numbers of the matrices involved in the inversion. As explained in the paper, if a given projected generator matrix has rank $m - r + 2$ after removing the columns specified by the projected corrupted codeword, then the block-MAP decoder will be able to recover the erased indices in the projected vectors. To do so, there are several methods and, in our decoder, we compute the inverse of the matrix $\tilde{\mathbf{G}}\tilde{\mathbf{G}}^T$, where \cdot^T denotes the matrix transpose operation and $\tilde{\mathbf{G}}$ is a full-rank sub-matrix of the projected generator matrix (excluding the erased columns) obtained by selecting a subset of $m - r + 2$ linearly independent rows. Our numerical analysis indicates that the matrices $\tilde{\mathbf{G}}\tilde{\mathbf{G}}^T$ involved in the inversion have relatively small condition numbers and thus are well-conditioned. For instance, we considered decoding $\mathcal{RM}(6, 3)$ using $\binom{6}{3-1} = 15$ two-dimensional projections as explained in the paper. We also considered 60 equidistant erasure probabilities in the interval $[0.01, 0.6]$, and examined 1000 random erasure patterns for each of the erasure probabilities. We then projected each erasure pattern (i.e., corrupted codeword) onto the 15 projections, and computed the condition number of full-rank 5×5 matrices $\tilde{\mathbf{G}}\tilde{\mathbf{G}}^T$ after each projection. In our numerical results the maximum, i.e., the worst case, condition number was 428.36 among all these $1000 \times 60 \times 15 = 9 \times 10^5$, which is then equivalent to loosing no more than 3 precision digits in decimal floating-point representation. Also, the average of the condition numbers among all 1000 random trials for each given erasure probability and projection subspaces was smaller than 100.

C. Polar-Coded Distributed Computation

Binary polar codes are capacity-achieving linear codes with explicit constructions and low-complexity encoding and decoding [65]. Also, the low-complexity $O(n \log n)$ encoding and decoding of polar codes can be adapted to work over real-valued data when dealing with erasures as in coded computation systems. Given the close connection of RM

and polar codes, for the sake of the completeness of our study, we also explore polar-coded computation, which was first considered in [35]. However, our simulation results in Section VI demonstrate a significantly superior performance for the RM-coded computation enabled by our proposed low-complexity decoder in Section V-B. Next, we briefly explain the encoding and decoding procedure of real-valued data using binary polar codes.

1) *Encoding Procedure*: In Section V-A, we briefly explained the construction of the generator matrix for polar codes over BECs. The encoding procedure using the resulting $k \times n$ generator matrix \mathbf{G} , which also applies to any (n, k) binary linear code operating over real-valued data, is as follows. First, the computational job is divided into k smaller tasks. Then the j -th encoded task which will be sent to the j -th node, for $j = 1, 2, \dots, n$, is the linear combination of all tasks according to the j -th column of \mathbf{G} . Throughout the paper (including both RM- and polar-coded computation), we apply the transform $\mathbf{G} \rightarrow 2\mathbf{G} - 1$ to convert the entries of the generator matrix from $\{0, 1\}$ to $\{-1, +1\}$.

2) *Decoding Procedure*: The recursive structure of polar codes can be applied for low-complexity detection/decoding of real-valued data using parallel processing for more speedups [66], [67]. To this end, one can apply the decoding algorithms in [35] for polar-coded computing. It is well-known that in the case of SC decoding over erasure channels, the probability of decoding failure of polar codes is $P_e^{\text{SC}}(\epsilon, n) = 1 - \prod_{i \in \mathcal{B}} (1 - Z_i)$, where \mathcal{B} denotes the set of indices of the selected rows.

VI. SIMULATION RESULTS

In this section, simulation results for the execution time of various coded distributed computing schemes are presented. In particular, their gap to the optimal performance are shown and also, their performance gains are compared with the uncoded computation. In addition to the fundamental results presented in Section III, we also provide the simulation results for RM- and polar-coded distributed computation, explained in the previous section. We assume $\mu = 1$ for all numerical results in this section.

For MDS and random linear codes, T_{avg} is calculated using (13) and Corollary 5, respectively. For the polar-coded computation with SC decoding, we apply $P_e^{\text{SC}}(\epsilon, n)$, explained in Section V-C2, together with (4) to evaluate T_{avg} via the numerical integration. Similarly, for the RM-coded computation with our decoder, we first apply our decoding algorithm, presented in Section V-B to numerically obtain $P_e(\epsilon, n)$ given ϵ and n . We then apply (4) to numerically calculate T_{avg} . Note that, for each scheme, we searched over all possible values of k to obtain the optimal k^* that minimizes T_{avg} . For the RM-coded computation, we also include the results for the optimal MAP decoder. To do so, given k and n , we first construct the generator matrix of the RM (sub-) code by selecting the k rows of the polarization matrix \mathbf{G}_n that have the highest Hamming weights. We then apply Monte-Carlo numerical simulation to obtain $P_e(\epsilon, n)$ (we erase the columns of the resulting generator matrix with probability ϵ and then declare an error if the resulting matrix is not full rank) for each

Table I

AVERAGE EXECUTION TIME AND OPTIMAL k^* VALUES FOR DIFFERENT CODING SCHEMES AS WELL AS THEIR GAP g_{opt} TO THE OPTIMAL PERFORMANCE AND THEIR PERFORMANCE IMPROVEMENT GAIN G_{cod} COMPARED TO THE UNCODED COMPUTING.

n	Uncoded	MDS coding	Binary random coding	Polar coding with SC decoding	RM coding with our proposed decoder	RM coding with optimal MAP decoding
	$(T_{\text{avg}}, g_{\text{opt}})$	$(T_{\text{avg}}, k^*, G_{\text{cod}})$	$(T_{\text{avg}}, k^*, g_{\text{opt}}, G_{\text{cod}})$			
8	(0.4647, 25%)	(0.370, 6, 20%)	(0.460, 7, 25%, 1.1%)	(0.412, 7, 11%, 12%)	(0.389, 7, 5.1%, 16%)	(0.389, 7, 5.1%, 16%)
16	(0.2738, 44%)	(0.191, 11, 31%)	(0.226, 11, 18%, 18%)	(0.217, 11, 14%, 21%)	(0.198, 11, 3.6%, 28%)	(0.198, 11, 3.6%, 28%)
32	(0.1581, 63%)	(0.0968, 22, 39%)	(0.105, 21, 8.6%, 34%)	(0.114, 24, 18%, 28%)	(0.104, 26, 7.2%, 34%)	(0.104, 26, 7.2%, 34%)
64	(0.0897, 84%)	(0.0488, 44, 46%)	(0.051, 43, 3.9%, 44%)	(0.0584, 44, 20%, 35%)	(0.0506, 42, 3.7%, 44%)	(0.050, 42, 2.6%, 44%)
128	(0.0503, 105%)	(0.0245, 88, 51%)	(0.025, 87, 1.9%, 50%)	(0.0293, 88, 19%, 42%)	—	(0.0252, 97, 2.8%, 50%)
256	(0.0278, 127%)	(0.0123, 175, 56%)	(0.0124, 174, 0.9%, 56%)	(0.0146, 182, 19%, 48%)	—	(0.0123, 166, 0.6%, 56%)
512	(0.0153, 149%)	(0.0061, 350, 60%)	(0.0062, 349, 0.5%, 60%)	(0.0073, 388, 19%, 52%)	—	(0.0061, 353, 0.1%, 60%)

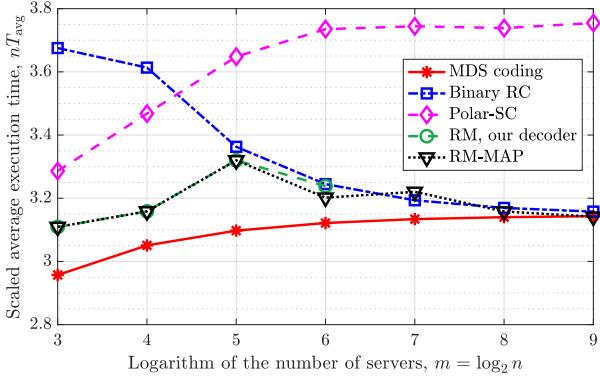


Figure 2. Scaled average execution time of a homogeneous distributed computing system with $\mu = 1$ using various coding schemes for finite number of servers $n = 8, 16, 32, 64, 128, 256$, and 512.

value of ϵ considered for the numerical evaluation of T_{avg} via (4). It is worth mentioning that we used the same k^* values for the RM-coded computation with our decoding algorithm as that of the MAP decoding. In fact, we only evaluate the performance of our decoder if, given $n = 2^m$, the value of k^* is of the form $\sum_{i=0}^r \binom{m}{i}$ for some $r \leq m$. This is because our decoding algorithm is specifically designed for RM codes and not their subcodes. Therefore, one needs to build upon the methods in this paper and the intuitions in [64] to extend our decoding algorithm for general RM subcodes (with a carefully designed generator matrix) that admit any code dimension k . Given the close performance of our decoder for RM codes to that of the optimal MAP decoder (and also similarities in decoding RM codes and their subcodes), we expect the generalized version of our decoder to the case of RM subcodes to also achieve very close to the optimal performance.

Numerical results for the performance of the coded distributed computing systems utilizing MDS codes, binary random linear codes, polar codes, and RM codes, are presented in Table II and are compared with the uncoded scenario over small-to-moderate blocklengths. We designed the polar code with $\epsilon_d = 0.1$, which is observed to be good enough for the range of blocklengths in Table II. One can also attain slightly better performance for polar codes by optimizing over ϵ_d specifically for each n . Characterizing the best ϵ_d as a function of blocklength n is left for the future work. In Table I, G_{cod} is defined as the percentage of the gain in T_{avg} compared to the uncoded scenario and g_{opt} is defined as the gap of T_{avg} for the underlying coding scheme to that of MDS codes, in

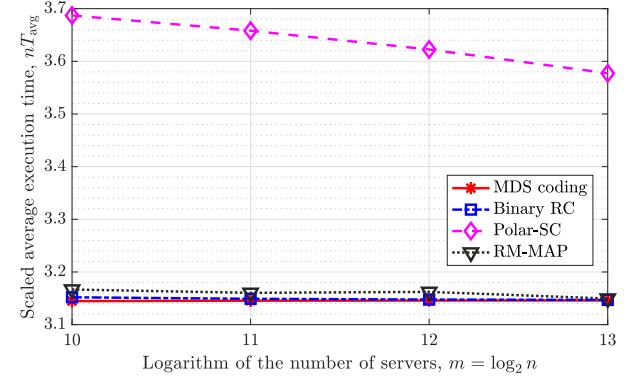


Figure 3. Scaled average execution time of a homogeneous distributed computing system with $\mu = 1$ using various coding schemes for asymptotically large number of servers $n = 1024, 2048, 4096$, and 8192.

percentage. Intuitively, G_{cod} for a coding scheme determines *how much gain* this scheme attains and g_{opt} indicates *how close* this scheme is to the optimal solution. Observe that RM-coded computation with the optimal MAP decoder is able to achieve very close to the performance of optimal MDS-coded computation, and also outperform binary RC for relatively small number of servers. Our low-complexity decoding algorithm is also able to achieve the same performance as that of the MAP decoder for the number of servers $n = 8, 16, 32$, and 64 where the optimal values of k^* correspond to RM codes. For the larger values of n , the optimal k^* 's correspond to RM sub-codes and one needs to extend our algorithm to, possibly, achieve close to the performance of the optimal MAP decoder with a low complexity. Figure 2 shows that random linear codes have weak performance in the beginning but they quickly approach the optimal T_{avg} so that they have small gaps to the optimal values, e.g., $g_{\text{opt}} = 0.5\%$ for $n = 512$. Also, observe that RM codes significantly outperform polar codes.

In the case of $\mu = 1$, by numerically solving (26), we have for the asymptotically-optimal encoding rate $R^* = 0.6822$. Motivated by this fact, in Figure 3, the rate of all discussed underlying coding schemes is fixed to R^* , and nT_{avg} is plotted for large blocklengths. Therefore, T_{avg} is not optimized over rates for the results demonstrated in this figure. Additionally, the polar code is designed with $\epsilon_d = 1 - R^* = 0.3178$, which makes the code to be capacity-achieving for an erasure channel with capacity equal to R^* . Furthermore, Figure 3 suggests that RM codes approach very close to the optimal performance, and also do so relatively fast.

Table II

AVERAGE EXECUTION TIME AND OPTIMAL k^* VALUES FOR DIFFERENT CODING SCHEMES UNDER SHIFTED WEIBULL DISTRIBUTION WITH PARAMETERS $\mu = 1$ AND $\alpha = 2$.

n	MDS	Polar-SC	RM-MAP	RM with our decoder
	(k^*, T_{avg})	(k^*, T_{avg})	(k^*, T_{avg})	(k^*, T_{avg})
8	(7, 0.3163)	(7, 0.3247)	(7, 0.3163)	(7, 0.3163)
16	(14, 0.1633)	(15, 0.1676)	(15, 0.1637)	(15, 0.1637)
32	(28, 0.0832)	(31, 0.0878)	(26, 0.0856)	(26, 0.0857)
64	(55, 0.0420)	(58, 0.0446)	(57, 0.0425)	(57, 0.0426)

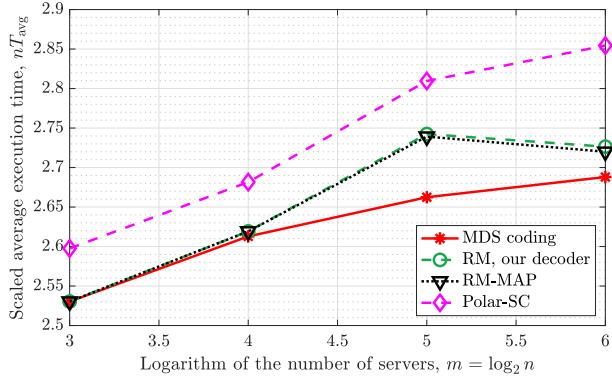


Figure 4. Scaled average execution time of various schemes under shifted Weibull distribution with parameters $\mu = 1$ and $\alpha = 2$.

In order to further demonstrate the applicability of the proposed approach, we analyzed the average execution time of coded distributed computing systems under shifted Weibull distribution which generalizes the shifted exponential distribution considered in the paper. Specifically, there is a parameter α associated with the Weibull distribution, as specified in Appendix A, and for $\alpha = 1$, the (shifted) Weibull distribution is simplified to (shifted) exponential. This in turn provides a higher flexibility in statistical modeling [68] such as the run-time of the computational servers [9]. Using Lemma 12 in Appendix A, we computed the average execution time of various coding schemes such as MDS, polar with SC, RM with MAP, and RM with our low-complexity decoder. As shown in Figure 4 and also summarized in Table II, our low-complexity decoder is able to achieve very close to the performance of MDS-coded computing even for very small number of servers.

VII. CONCLUSIONS

In this paper, we presented a coding-theoretic approach toward coded distributed computing systems by connecting the problem of characterizing their average execution time to the traditional problem of finding the error probability of a coding scheme over erasure channels. Using this connection, we provided results on the performance of coded distributed computing systems, such as their best performance bounds and asymptotic results using binary random linear codes. Accordingly, we established the existence of *good* binary linear codes that attain (asymptotically) the best performance any linear code can achieve while maintaining numerical stability against the inevitable rounding errors in practice. To enable a practical approach for coded computation, we developed

a low-complexity algorithm for decoding RM codes over erasure channels, involving only additions and subtractions (and inverting small matrices of size less than $\log n + 1$). Our RM-coded computation scheme not only has close-to-optimal performance and explicit construction but also works over real-valued data inputs with a low-complexity decoding.

APPENDIX A AVERAGE EXECUTION TIME UNDER SHIFTED WEIBULL DISTRIBUTION

Lemma 12. *The average execution time of a coded distributed computing system, using a given (n, k) linear code and assuming shifted Weibull distribution on the run-time of individual nodes, can be characterized as*

$$T_{\text{avg}} = \frac{1}{k} + \frac{1}{\mu k \alpha} \int_0^1 \frac{P_e(\epsilon, n)}{\epsilon [\ln(1/\epsilon)]^{1-1/\alpha}} d\epsilon, \quad (38)$$

where $\alpha > 0$ represents the shape parameter.

Proof: The CDF of the run time T_i of each local node i under shifted Weibull distribution and the homogeneous system model of this paper can be expressed as [68]:

$$\Pr(T_i \leq t) = 1 - \exp(-[\mu(kt - 1)]^\alpha), \quad \forall t \geq 1/k. \quad (39)$$

Equivalently, the erasure probability $\epsilon(t)$ is equal to one for $t < 1/k$ and, otherwise, is equal to $\exp(-[\mu(kt - 1)]^\alpha)$ for $t \geq 1/k$. Now using (3) and given that for the shifted Weibull distribution $d\epsilon(\tau)/d\tau = -\mu k \alpha \epsilon(\tau) [\ln(1/\epsilon(\tau))]^{1-1/\alpha}$, and that $P_e(\epsilon(\tau), n) = 1$ for all $\tau \leq 1/k$, we have (38) by the change of variables. Note that the result simplifies to the case of shifted exponential distribution for $\alpha = 1$, presented in Lemma 1. ■

REFERENCES

- [1] M. V. Jamali, M. Soleimani, and H. Mahdavifar, “Coded distributed computing: Performance limits and code designs,” in *2019 IEEE Information Theory Workshop (ITW)*, pp. 1–5.
- [2] G. Ananthanarayanan, A. Ghodsi, S. Shenker, and I. Stoica, “Effective straggler mitigation: Attack of the clones,” in *Presented as part of the 10th USENIX Symposium on Networked Systems Design and Implementation (NSDI 13)*, 2013, pp. 185–198.
- [3] A. Vulimiri, P. B. Godfrey, R. Mittal, J. Sherry, S. Ratnasamy, and S. Shenker, “Low latency via redundancy,” in *Proceedings of the ninth ACM conference on Emerging networking experiments and technologies*. ACM, 2013, pp. 283–294.
- [4] K. Gardner, S. Zbarsky, S. Doroudi, M. Harchol-Balter, and E. Hyttia, “Reducing latency via redundant requests: Exact analysis,” *ACM SIGMETRICS Performance Evaluation Review*, vol. 43, no. 1, pp. 347–360, 2015.
- [5] E. Jonas, Q. Pu, S. Venkataraman, I. Stoica, and B. Recht, “Occupy the cloud: Distributed computing for the 99%,” in *Proceedings of the 2017 Symposium on Cloud Computing*. ACM, 2017, pp. 445–451.
- [6] K. Lee, M. Lam, R. Pedarsani, D. Papailiopoulos, and K. Ramchandran, “Speeding up distributed machine learning using codes,” *IEEE Trans. Inf. Theory*, vol. 64, no. 3, pp. 1514–1529, 2018.
- [7] S. Li and S. Avestimehr, “Coded computing: Mitigating fundamental bottlenecks in large-scale distributed computing and machine learning,” *Found. Trends Commun. Inf. Theory*, vol. 17, no. 1, pp. 1–148, 2020.
- [8] S. Li, M. A. Maddah-Ali, and A. S. Avestimehr, “A unified coding framework for distributed computing with straggling servers,” in *2016 IEEE Globecom Workshops (GC Wkshps)*, pp. 1–6.
- [9] A. Reisizadeh, S. Prakash, R. Pedarsani, and A. S. Avestimehr, “Coded computation over heterogeneous clusters,” *arXiv preprint arXiv:1701.05973*, 2017.

[10] S. Li, M. A. Maddah-Ali, and A. S. Avestimehr, "Coding for distributed fog computing," *IEEE Commun. Mag.*, vol. 55, no. 4, pp. 34–40, 2017.

[11] Y. Yang, P. Grover, and S. Kar, "Computing linear transformations with unreliable components," *IEEE Trans. Inf. Theory*, vol. 63, no. 6, pp. 3729–3756, 2017.

[12] K. Lee, C. Suh, and K. Ramchandran, "High-dimensional coded matrix multiplication," in *IEEE Int. Symp. Inf. Theory (ISIT)*. IEEE, 2017, pp. 2418–2422.

[13] S. Dutta, V. Cadambe, and P. Grover, "Short-dot: Computing large linear transforms distributedly using coded short dot products," in *Adv. in Neural Info. Proc. Systems (NIPS)*, 2016, pp. 2100–2108.

[14] S. Wang, J. Liu, and N. Shroff, "Coded sparse matrix multiplication," *arXiv preprint arXiv:1802.03430*, 2018.

[15] S. Dutta, M. Fahim, F. Haddadpour, H. Jeong, V. Cadambe, and P. Grover, "On the optimal recovery threshold of coded matrix multiplication," *IEEE Trans. Inf. Theory*, vol. 66, no. 1, pp. 278–301, 2019.

[16] Q. Yu, M. A. Maddah-Ali, and A. S. Avestimehr, "Straggler mitigation in distributed matrix multiplication: Fundamental limits and optimal coding," *IEEE Trans. Inf. Theory*, vol. 66, no. 3, pp. 1920–1933, 2020.

[17] S. Prakash, S. Dhakal, M. R. Akdeniz, Y. Yona, S. Talwar, S. Avestimehr, and N. Himayat, "Coded computing for low-latency federated learning over wireless edge networks," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 1, pp. 233–250, 2020.

[18] T. Jahani-Nezhad and M. A. Maddah-Ali, "Codedsketch: A coding scheme for distributed computation of approximated matrix multiplication," *arXiv preprint arXiv:1812.10460*, 2018.

[19] N. J. Higham, *Accuracy and stability of numerical algorithms*. Siam, 2002, vol. 80.

[20] Q. Yu, S. Li, N. Raviv, S. M. M. Kalan, M. Soltanolkotabi, and S. A. Avestimehr, "Lagrange coded computing: Optimal design for resiliency, security, and privacy," in *The 22nd International Conference on Artificial Intelligence and Statistics*. PMLR, 2019, pp. 1215–1225.

[21] A. Reisizadeh, S. Prakash, R. Pedarsani, and A. S. Avestimehr, "Coded computation over heterogeneous clusters," *IEEE Trans. Inf. Theory*, vol. 65, no. 7, pp. 4227–4242, 2019.

[22] S. Li, M. A. Maddah-Ali, and A. S. Avestimehr, "Coded distributed computing: Straggler servers and multistage dataflows," in *54th Annual Allerton Conference*. IEEE, 2016, pp. 164–171.

[23] Q. Yu, S. Li, N. Raviv, S. M. M. Kalan, M. Soltanolkotabi, and S. A. Avestimehr, "Lagrange coded computing: Optimal design for resiliency, security, and privacy," in *The 22nd International Conference on Artificial Intelligence and Statistics*, 2019, pp. 1215–1225.

[24] Q. Yu and A. S. Avestimehr, "Entangled polynomial codes for secure, private, and batch distributed matrix multiplication: Breaking the "cubic" barrier," in *2020 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2020, pp. 245–250.

[25] M. Aliasgari, O. Simeone, and J. Kliewer, "Private and secure distributed matrix multiplication with flexible communication load," *IEEE Trans. on Information Forensics and Security*, vol. 15, pp. 2722–2734, 2020.

[26] R. G. D'Oliveira, S. El Rouayheb, and D. Karpuk, "GASP codes for secure distributed matrix multiplication," *IEEE Trans. Inf. Theory*, vol. 66, pp. 4038–4050, 2020.

[27] R. Bitar, Y. Xing, Y. Keshtkarjahromi, V. Dasari, S. El Rouayheb, and H. Seferoglu, "Private and rateless adaptive coded matrix-vector multiplication," *EURASIP Journal on Wireless Communications and Networking*, vol. 2021, no. 1, pp. 1–25, 2021.

[28] H. A. Nodehi and M. A. Maddah-Ali, "Secure coded multi-party computation for massive matrix operations," *IEEE Transactions on Information Theory*, 2021.

[29] M. Soleymani, R. E. Ali, H. Mahdavifar, and A. S. Avestimehr, "List-decodable coded computing: Breaking the adversarial toleration barrier," *arXiv preprint arXiv:2101.11653*, 2021.

[30] M. Soleymani, H. Mahdavifar, and A. S. Avestimehr, "Privacy-preserving distributed learning in the analog domain," *arXiv preprint:2007.08803*, 2020.

[31] —, "Analog Lagrange coded computing," *IEEE Journal on Selected Areas in Information Theory (JSAIT): Special issue on Privacy and Security of Information Systems*, 2021.

[32] T. Jahani-Nezhad and M. A. Maddah-Ali, "Berrut approximated coded computing: Straggler resistance beyond polynomial computing," *arXiv preprint arXiv:2009.08327*, 2020.

[33] R. M. Roth, "Analog error-correcting codes," *IEEE Transactions on Information Theory*, vol. 66, no. 7, pp. 4075–4088, 2020.

[34] M. Soleymani and H. Mahdavifar, "Analog subspace coding: A new approach to coding for non-coherent wireless networks," in *2020 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2020, pp. 31–36.

[35] B. Bartan and M. Pilanci, "Polar coded distributed matrix multiplication," *arXiv preprint arXiv:1901.06811*, 2019.

[36] S. Kudekar, S. Kumar, M. Mondelli, H. D. Pfister, E. Şaşoğlu, and R. L. Urbanke, "Reed-Muller codes achieve capacity on erasure channels," *IEEE Trans. Inf. Theory*, vol. 63, no. 7, pp. 4298–4316, 2017.

[37] E. Abbe, A. Shpilka, and A. Wigderson, "Reed-Muller codes for random erasures and errors," *IEEE Trans. Inf. Theory*, vol. 61, no. 10, pp. 5229–5252, 2015.

[38] H. Hassani, S. Kudekar, O. Ordentlich, Y. Polyanskiy, and R. Urbanke, "Almost optimal scaling of Reed-Muller codes on BEC and BSC channels," in *IEEE Int. Symp. Inf. Theory*. IEEE, 2018, pp. 311–315.

[39] M. Ye and E. Abbe, "Recursive projection-aggregation decoding of Reed-Muller codes," *IEEE Trans. Inf. Theory*, vol. 66, no. 8, pp. 4948–4965, 2020.

[40] I. Reed, "A class of multiple-error-correcting codes and the decoding scheme," *Transactions of the IRE Professional Group on Information Theory*, vol. 4, no. 4, pp. 38–49, 1954.

[41] I. Dumer and K. Shabunov, "Soft-decision decoding of Reed-Muller codes: recursive lists," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 1260–1266, 2006.

[42] R. Saptharishi, A. Shpilka, and B. L. Volk, "Efficiently decoding Reed-Muller codes from random errors," *IEEE Trans. Inf. Theory*, vol. 63, no. 4, pp. 1954–1960, 2017.

[43] E. Santi, C. Hager, and H. D. Pfister, "Decoding Reed-Muller codes using minimum-weight parity checks," in *2018 IEEE Int. Symp. Inf. Theory (ISIT)*. IEEE, 2018, pp. 1296–1300.

[44] G. Liang and U. C. Kozat, "TOFEC: achieving optimal throughput-delay trade-off of cloud storage using erasure codes," in *IEEE Conf. Computer Commun. (INFOCOM)*. IEEE, 2014, pp. 826–834.

[45] J. Kahn, J. Komlós, and E. Szemerédi, "On the probability that a random ± 1 -matrix is singular," *Journal of the American Mathematical Society*, vol. 8, no. 1, pp. 223–240, 1995.

[46] F. J. MacWilliams and N. J. A. Sloane, *The theory of error-correcting codes*. New York: North-Holland, 1977.

[47] R. M. Roth and A. Lempel, "A construction of non-Reed-Solomon type MDS codes," *IEEE transactions on information theory*, vol. 35, no. 3, pp. 655–657, 1989.

[48] P. Beelen and L. Jin, "Explicit MDS codes with complementary duals," *IEEE Transactions on Information Theory*, vol. 64, no. 11, pp. 7188–7193, 2018.

[49] B. Chen and H. Liu, "New constructions of MDS codes with complementary duals," *IEEE Transactions on Information Theory*, vol. 64, no. 8, pp. 5776–5782, 2017.

[50] C. Carlet, S. Mesnager, C. Tang, and Y. Qi, "Euclidean and Hermitian LCD MDS codes," *Designs, Codes and Cryptography*, vol. 86, no. 11, pp. 2605–2618, 2018.

[51] J. W. Demmel, *Applied numerical linear algebra*. Siam, 1997, vol. 56.

[52] D. L. Boley, R. P. Brent, G. H. Golub, and F. T. Luk, "Algorithmic fault tolerance using the Lanczos method," *SIAM Journal on Matrix Analysis and Applications*, vol. 13, no. 1, pp. 312–332, 1992.

[53] P. J. Ferreira, "Stability issues in error control coding in the complex field, interpolation, and frame bounds," *IEEE Signal Processing Letters*, vol. 7, no. 3, pp. 57–59, 2000.

[54] P. J. Ferreira and J. M. Vieira, "Stable DFT codes and frames," *IEEE Signal Processing Letters*, vol. 10, no. 2, pp. 50–53, 2003.

[55] W. Henkel, "Multiple error correction with analog codes," in *International Conference on Applied Algebra, Algebraic Algorithms, and Error-Correcting Codes*. Springer, 1988, pp. 239–249.

[56] F. Marvasti, M. Hasan, M. Echhert, and S. Talebi, "Efficient algorithms for burst error recovery using FFT and other transform kernels," *IEEE Transactions on Signal Processing*, vol. 47, no. 4, pp. 1065–1075, 1999.

[57] Z. Chen and J. Dongarra, "Numerically stable real number codes based on random matrices," in *International Conference on Computational Science*. Springer, 2005, pp. 115–122.

[58] Z. Chen and J. J. Dongarra, "Condition numbers of gaussian random matrices," *SIAM Journal on Matrix Analysis and Applications*, vol. 27, no. 3, pp. 603–620, 2005.

[59] Z. Chen, "Optimal real number codes for fault tolerant matrix operations," in *Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis*, 2009, pp. 1–10.

[60] M. Rudelson and R. Vershynin, "The littlewood-offord problem and invertibility of random matrices," *Advances in Mathematics*, vol. 218, no. 2, pp. 600–633, 2008.

[61] —, "Non-asymptotic theory of random matrices: extreme singular values," in *Proceedings of the International Congress of Mathematicians 2010 (ICM 2010) (In 4 Volumes) Vol. I: Plenary Lectures and*

Ceremonies Vols. II–IV: Invited Lectures. World Scientific, 2010, pp. 1576–1602.

[62] J. Bourgain, V. H. Vu, and P. M. Wood, “On the singularity probability of discrete random matrices,” *Journal of Functional Analysis*, vol. 258, no. 2, pp. 559–603, 2010.

[63] T. Kaufman, S. Lovett, and E. Porat, “Weight distribution and list-decoding size of Reed-Muller codes,” *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 2689–2696, 2012.

[64] M. V. Jamali, X. Liu, A. V. Makkuvu, H. Mahdavifar, S. Oh, and P. Viswanath, “Reed-Muller subcodes: Machine learning-aided design of efficient soft recursive decoding,” *arXiv preprint arXiv:2102.01671*, 2021.

[65] E. Arikan, “Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels,” *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3051–3073, 2009.

[66] M. V. Jamali and H. Mahdavifar, “A low-complexity recursive approach toward code-domain NOMA for massive communications,” in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2018, pp. 1–6.

[67] ——, “Massive coded-NOMA for low-capacity channels: A low-complexity recursive approach,” *arXiv preprint arXiv:2006.06917*, 2020.

[68] S. Coles, J. Bawa, L. Trenner, and P. Dorazio, *An introduction to statistical modeling of extreme values*. Springer, 2001, vol. 208.



Hessam Mahdavifar (S’10, M’12) is an Assistant Professor in the Department of Electrical Engineering and Computer Science at the University of Michigan Ann Arbor. He received the B.Sc. degree from the Sharif University of Technology, Tehran, Iran, in 2007, and the M.Sc. and the Ph.D. degrees from the University of California San Diego (UCSD), La Jolla, in 2009, and 2012, respectively, all in electrical engineering. He was with the Samsung US R&D between 2012 and 2016, in San Diego, US, as a staff research engineer.

He received the NSF career award in 2020. He also received Best Paper Award in 2015 IEEE International Conference on RFID, and the 2013 Samsung Best Paper Award. He also received two Silver Medals at International Mathematical Olympiad in 2002 and 2003, and two Gold Medals at Iran National Mathematical Olympiad in 2001 and 2002. His main area of research is coding and information theory with applications to wireless communications, storage systems, security, and privacy.



Mahdi Soleymani (Student Member, IEEE) received his B.S. and M.S. degrees in Electrical Engineering at Sharif University of Technology, Tehran, Iran, in 2014 and 2016, respectively. He is currently pursuing his Ph.D. degree in Electrical Engineering and Computer Science at University of Michigan, Ann Arbor. His research interests lie in the area of algebraic coding theory with applications to distributed storage systems, wireless networks, distributed computing and machine learning.



Mohammad Vahid Jamali is a Ph.D. candidate in the Department of Electrical Engineering and Computer Science at the University of Michigan, Ann Arbor, MI, USA. He received the B.Sc. degree from K.N. Toosi University of Technology, Tehran, Iran, in 2013, and the M.Sc. degree from Sharif University of Technology (SUT), Tehran, Iran, in 2015, both in electrical engineering. From 2013 to 2017, he was a member of the technical staff of the Optical Networks Research Laboratory (ONRL) at SUT. He received several awards and honors including 2021 Qualcomm Innovation Fellowship, The 2017 Best M.Sc. Thesis Award by the IEEE Iran Section, Best Paper Award of IEEE GLOBECOM Workshops on UHS5G, Abu Dhabi, UAE, 2018, and Exemplary Reviewer of the IEEE Transactions on Communications in 2019. His general research interests are in coding, information theory, wireless communications, and optics.