Non-Linear Function Computation Broadcast

Mohammad Reza Deylam Salehi, Vijith Kumar Kizhakke Purakkal, Derya Malak

Communication Systems Department, EURECOM, Biot Sophia Antipolis, France

{deylam, kizhakke, malak}@eurecom.fr

Abstract—This work addresses the K-user computation broadcast problem consisting of a master node, which holds all datasets, and users for a general class of function demands, including linear and non-linear functions, over finite fields. The master node sends a broadcast message to enable each of K distributed users to compute its demanded function in an asymptotically lossless manner with user's side information. We derive bounds on the optimal K-user computation broadcast rate that allows the users to compute their demanded functions by capturing the structures of the computations and available side information. Our achievability scheme involves the design of a novel graph-based coding model to build a broadcast message to meet each user's demand, by leveraging the structural dependencies among the datasets, the user demands, and the side information of each user, drawing on Körner's characteristic graph framework. The converse uses the structures of the demands and the side information available at K users to yield a tight lower bound on the broadcast rate. With the help of examples, we demonstrate our scheme achieves a better communication rate than the existing state of the art.

Index Terms—Computation broadcast, distributed non-linear function computation, characteristic graphs, and union graphs.

I. Introduction

In the presence of large-scale data, communication networks are being increasingly deployed for various distributed computation scenarios, including distributed learning [1], [2], over-the-air computing [3], [4], private-key encryption [5], [6], and coded distributed computing [7], [8], to name a few. In such scenarios, meeting the growing and diverse demands of distributed users is exacerbated by the limited communication resources, privacy constraints, and bounded memory and computational capabilities of users. This motivates us to study the problem of computation broadcast, in which source data is communicated to recover functions of data at distributed users, which finds applications in radio broadcasting, digital video broadcasting, and cryptography [9]–[11].

A. Motivation and Contributions

Prior works in computation broadcast have focused on linear function demands [12]–[15], which allow for exploiting algebraic techniques to encode the source data, or complementary function demands [16]. However, these approaches often

depend on the specific demand structures or the number of users, limiting their general applicability. Furthermore, to the best of our knowledge, no general bound on the broadcast rate has been established for non-linear demands.

This paper tackles the problem of function computation broadcast, accommodating general function demands (linear, non-separable, and non-linear) over finite fields. We develop a master-users framework using *characteristic graphs* to capture demand structures. The master node, has access to all datasets and computation capabilities, transmits information to multiple users. Each user, having a fraction of the master's datasets as side information, computes a general function of the datasets.

Our main contributions are summarized as follows.

- We devise a coding scheme for the K-user computation broadcast that tackles general data distributions, and general non-linear function demands over $\mathbb{F}_q$, denoting a finite field of order $q \geq 2$. This framework, as detailed in Section II, emerges from using characteristic graphs devised by Körner [17] to identify a collection of equivalent source classes. The master builds a broadcast message to enable all users to decode their demanded functions simultaneously in an asymptotically lossless manner.

- This framework reduces communication costs by performing part of the computation at the master node and transmitting functions of datasets as the broadcast message, incorporating the structure of datasets and K user demands (Sections II and III). For general non-linear demands, our approach outperforms [18]. In a 3-user Boolean computation broadcast model, the achievable rate is 1.5 bits, improving on the 2 bits of [18] (Example 1, Section III). Similarly, for a 3-user linear computation broadcast model, we achieve 1.42 bits, outperforming [14] and [18] with rates of 1.5 and 2 bits, respectively (Example 2, Section III).

- We derive an achievability scheme and a converse on the communication rate for $K \geq 2$-user computation broadcast with linear and non-linear demands in Section IV[1].

B. State of the Art

Communication complexity: Communication complexity, which quantifies the minimum bits required for transmission among multiple parties, was introduced for distributed computation of Boolean functions [20] and later extended to approximating continuous functions [21], computing rational functions [22], and estimating correlations [23]. Inter-

[1]The supplementary proofs of the results in Section IV are given in [19].

active communication models have demonstrated significant communication savings compared to one-way models [24], [25]. While previous studies have explored communication complexity across various multi-party communication models, characterizing this complexity for computation broadcast problems with general function demands remains an open problem.

Compression for computing: Structured codes reduce rates in distributed computation for tasks like modulo-q sum [26], [27], linear functions [28], and matrix multiplication [29]. Unstructured codes address point-to-point lossy recovery [30], function computation with side information [31], and multi-terminal setups [18], [32]–[34]. Körner's characteristic graph framework [17] is used to compress sources for distributed computing [35]–[43]. Extending this to broadcast models is difficult due to the need to handle function and side information dependencies in message design.

Coded distributed computing: To meet growing computational demands and speed up computing via distributing tasks among multiple servers, techniques like distributed machine learning [44], federated learning [45], and massive parallelization algorithms, e.g., MapReduce [46] and Spark [47], have been proposed. Most strategies exploit the separable nature of functions, e.g., matrix multiplication [48] and gradient descent [49], and design codes to distribute tasks, focusing on straggler mitigation [50], communication complexity [8], [49], and communication-computation tradeoffs [48], [51]. However, they fall short for general non-linear demands.

Computation broadcast: The computation broadcast framework generalizes the index coding problem [52], where users demand source variables instead of their functions, and multiple description coding [53], a special case where all users demand the same distorted source function over $\mathbb{F}_q$ with varying distortion criteria. Linear computation broadcast has been studied for two and more users settings [14], [15]. However, for non-linear computation broadcast, only the case of $K = 2$ users with complementary side information has been explored [16]. Here, we extend [16] to $K \geq 2$ users without structural assumptions on side information.

II. System Model

This section outlines our K-user computation broadcast model, depicted in Figure 1. The model consists of a master node, that stores $N \in \mathbb{Z}^+$ datasets, denoted by $X_{[N]} \triangleq \{X_j\}_{j \in [N]}$, where $X_j \in \mathbb{F}_q$ for $j \in [N] \triangleq \{1, \cdots, N\}$ and $q \geq 2$, and K users. Each of $[K]$ distributed users is equipped with side information, $\mathcal{S}_i \subseteq X_{[N]}$. Users are interested in computing a set of (possibly distinct) functions $\{f_i\}_{i \in [K]}$ using the shorthand notation $f_i = f_i(X_1, \ldots, X_N)$.

The master node, using n i.i.d. realizations of $X_j^n = \{X_{j,t}\}_{t \in [n]} \in \mathbb{F}_q^n$, broadcasts a message $M = \psi(X_1^n, \ldots, X_N^n)$, designed as a function of the master node's data and the set of demanded functions. The objective of the computation broadcast problem is to minimize the communication rate from the master node, characterized by the entropy of M, denoted by $H_q(M) = \mathbb{E}[-\log_q P_M]$, that ensures each user $i \in [K]$ to
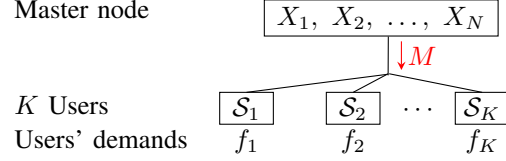

Fig. 1: Computation broadcast model for K users, each with side information $\mathcal{S}_i$ and demand f_i, for $i \in [K]$ [13].

asymptotically losslessly reconstruct the length n sequence of functions $f_i^n \triangleq \{f_i(X_{1,t}, \cdots, X_{N,t})\}_{t \in [n]}$, using the received message M and the length n realization of $\mathcal{S}_i$, i.e., $\mathcal{S}_i^n = \{\mathcal{S}_{i,t}\}_{t \in [n]}$, i.e., the following condition is met:

$$H_q(f_i^n \mid M, \mathcal{S}_i^n) = 0 , \quad \forall i \in [K] . \tag{1}$$

To capture general function demands in the proposed K-user computation broadcast model, we exploit the characteristic graph framework in [17], which has been used to characterize the communication rate for a point-to-point model, where a source holds a random variable X and a receiver has side information Y, and the receiver's goal is to reconstruct a function $f(X, Y)$ in an asymptotically lossless manner [37]. To motivate our approach, we next define characteristic graphs.

Definition 1 (**Source characteristic graphs** [37])**.** *Consider a point-to-point model with a master node having access to a source random variable $X \in \mathcal{X}$ and a user with side information $Y \in \mathcal{Y}$, aiming to compute $f(X, Y)$. The characteristic graph that the master node builds using X for $f(X, Y)$, given Y, is defined as $G_f = G(\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \mathcal{X}$ denotes the vertex set, and $\mathcal{E}$ the set of edges such that given two distinct vertices $x, x' \in \mathcal{V}$, an edge $(x, x') \in \mathcal{E}$ exists if and only if the following conditions hold[2]: i) $f(x, y) \neq f(x', y)$ and ii) $P_{X,Y}(x, y) \cdot P_{X,Y}(x', y) > 0$ for some $y \in \mathcal{Y}$.*

Characteristic graphs identify which source values can share the same codeword without introducing ambiguity to the user, enabling distributed computation in various topologies, as explored by prior literature, see e.g., [17], [36]–[42]. For the point-to-point scenario detailed above, the minimum rate required for a user to compute $f(X, Y)$ given side information Y with vanishing error probability is characterized as [37]

$$\min_{\substack{W - X - Y \\ X \in W \in \Gamma(G_f)}} I(X; W \mid Y) , \tag{2}$$

where $W - X - Y$ indicates a Markov chain, $I(X; W \mid Y) = H_q(W \mid Y) - H_q(W \mid X, Y)$, and $X \in W \in \Gamma(G_f)$ means that the minimization is over all joint PMFs $P_{W,X}(w, x) > 0$ such that W is an independent set[3] of G_f, and $\Gamma(G_f)$ is the collection of all MISs of G_f.

The paper extends the characteristic graph framework to address the computation broadcast problem by designing a

---

[2]Given random variables X and Y, the notation P_X denotes the probability mass function (PMF) of X, and $P_{X,Y}$ denotes the joint PMF of X and Y.

[3]An independent set (IS) is a subset of vertices in G_f where no two vertices are connected by an edge, ensuring that they can be encoded without ambiguity. A maximal independent set (MIS), denoted by $W \in \Gamma(G_f)$, is an independent set that cannot be extended by including additional vertices [54].
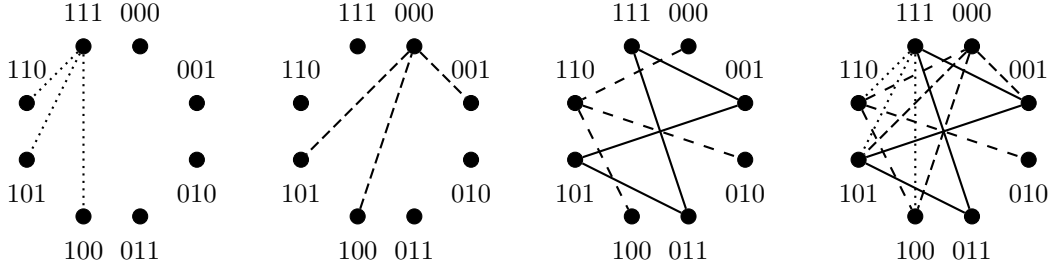
Fig. 2: Individual characteristic graphs and broadcast graph for Example 1: (Left) G_{f_1}, (Middle left) G_{f_2}, and (Middle right) G_{f_3}, based on users' demands f_1, f_2, and f_3, and the side information $\mathcal{S}_i$ for each $i \in [3]$, respectively, and (Right) broadcast graph $G_{\cup_{i=1}^3 f_i}$.

broadcast message, M, that captures dependencies among the master's datasets $X_{[N]}$, the structure of f_i, and each user's side information $\mathcal{S}_i$ for $i \in [K]$. This approach serves all users' demands simultaneously without imposing structural constraints (e.g., linearity, separability), unlike prior works [13]–[16]. To jointly capture the distinguishability requirements of K users, the framework leverages union graphs, as defined next.

Definition 2 (**Union graph** [55]). *Given two characteristic graphs $G_{f_1} = G(\mathcal{V}, \mathcal{E}_1)$ and $G_{f_2} = G(\mathcal{V}, \mathcal{E}_2)$ with the same vertex set $\mathcal{V}$, their union graph, $G_{f_1 \cup f_2}$, is defined as:*

$$G_{f_1 \cup f_2} \triangleq G(\mathcal{V}, \mathcal{E}_1 \cup \mathcal{E}_2) , \qquad (3)$$

where the set of edges is given by the union $\mathcal{E}_1 \cup \mathcal{E}_2$, which captures if a pair of source values need to be distinguished for at least one of the two functions f_1 and f_2.

Next, using the union graph approach, we will devise an achievability scheme for the computation broadcast problem, which involves the design of the broadcast message M. Generalizing Definition 2, we denote by $G_{\cup_{i=1}^K f_i}$ the union graph — referred to as broadcast graph — built by the master node to simultaneously meet the demands of all K users where $\Gamma(G_{\cup_{i=1}^K f_i})$ is the set of all MISs of this broadcast graph. From Definition 2, an MIS W of $G_{\cup_{i=1}^K f_i}$ satisfies the constraint (1).

Assigning a unique codeword to the MISs in $\Gamma(G_{\cup_{i=1}^K f_i})$ ensures that no confusion arises by the assigned codeword in any user $i \in [K]$ in reconstructing f_i. For asymptotically loss-less recovery of all demands, the master node builds the n-th OR power[4] $G_{\cup_{i=1}^K f_i}^n$ corresponding to the length n realization of $X_{[N]}$. To that end, the master node uses a binning strategy[5] and builds a mapping $en : \Gamma(G_{\cup_{i=1}^K f_i}^n) \rightarrow [0, q^{nR_{ach}} - 1]$, to construct the broadcast message $M = en(W^n)$ using the MISs $W^n \in \Gamma(G_{\cup_{i=1}^K f_i}^n)$, such that user $i \in [K]$ computes the non-componentwise function $\hat{f}_i^n \triangleq F_i(W^n, \mathcal{S}_i^n)$, and the condition $P(\hat{f}_i^n \neq f_i^n) < \epsilon$ is ensured asymptotically for some $\epsilon > 0$, rendering the condition in (1) true for M.

Hence, generalizing (2) to the broadcast setting, and using Slepian-Wolf coding [18], the broadcast rate satisfies

$$R_M \leq \min_{X_{[N]} \in W \in \Gamma(G_{\cup_{i'=1}^K f_{i'}})} \max_{i \in [K]} I(X_{[N]}; W \mid \mathcal{S}_i) , \qquad (4)$$

[4]The n-fold OR products of characteristic graphs are detailed in [36].
[5]Theorem 1 in Section IV details the binning strategy that achieves R_{ach}.

where the minimization is over all $P_{W, X_{[N]}}(w, x_{[N]}) > 0$ such that $W \in \Gamma(G_{\cup_{i=1}^K f_i})$ and (1) is ensured[6]. Since W is a function of $X_{[N]}$, the following substitution holds in (4):

$$I(X_{[N]}; W \mid \mathcal{S}_i) = H_q(W \mid \mathcal{S}_i) . \qquad (5)$$

We next demonstrate our coding scheme through examples.

III. 3-USER COMPUTATION BROADCAST: EXAMPLES

Next, we focus on a 3-user computation broadcast model, with $N = 3$ datasets $X_{[3]}$, where each X_j for $j \in [3]$ is i.i.d. and uniformly distributed over $\mathbb{F}_q$. Each user $i \in [3]$ has $\mathcal{S}_i = X_i$ and requests a distinct function $f_i = f_i(X_1, X_2, X_3)$. The former example examines Boolean (non-linear) function demands, while the latter focuses on linear demands.

Example 1 (3-**user Boolean computation broadcast**). *We denote user requests by $f_1 = X_1 \wedge X_2 \wedge X_3$, $f_2 = X_1 \vee X_2 \vee X_3$, and $f_3 = X_1 \wedge (X_2 \vee X_3)$, where $X_j \in \mathbb{F}_2$ for $j \in [3]$, $\mathcal{S}_i = X_i$ for $i \in [3]$, and $\wedge$ and $\vee$ denote binary AND and OR operations, respectively.*

The characteristic graph $G_{f_1} = (\mathcal{V}, \mathcal{E}_1)$ given $\mathcal{S}_1 = X_1$ is shown in Figure 2 (Left), where $\mathcal{V} = \mathbb{F}_2 \times \mathbb{F}_2 \times \mathbb{F}_2$. For $X_1 = 0$, the outcome of $f_1 = 0$ for any (X_2, X_3) pair. Thus, in G_{f_1}, we don't need to distinguish the vertices with $X_1 = 0$. When $X_1 = 1$, $f_1 = 1$ if and only if $X_2 = X_3 = 1$, and otherwise $f_1 = 0$. Thus, $\mathcal{E}_1 = \{(111, 100), (111, 101), (111, 110)\}$. Similarly, we construct G_{f_2} and G_{f_3}, as shown in Figure 2 (Middle left) and (Middle right), respectively. Using Definition 2, $G_{\cup_{i=1}^3 f_i}$ is shown in Figure 2 (Right).

Given $G_{\cup_{i=1}^3 f_i}$ that captures all user demands and side information $\{\mathcal{S}_i\}_{i \in [3]}$, the rate of the broadcast message, employing (4) and (5), is given by

$$R_M \leq \min_{X_{[3]} \in W \in \Gamma(G_{\cup_{i=1}^3 f_i})} \max_{i \in [3]} H_2(W \mid X_i) , \qquad (6)$$

where $G_{\cup_{i=1}^3 f_i}$ has the following set of MISs:

$$\Gamma(G_{\cup_{i=1}^3 f_i}) = \left\{ \begin{array}{l} \{000, 010, 011\}, \{001, 110, 011\}, \\ \{001, 010, 011, 100\}, \{101, 110\}, \\ \{010, 100, 101\}, \{000, 010, 111\} \end{array} \right\} . \qquad (7)$$

[6]The right hand side in (4) is achievable by assigning colors to $G_{\cup_{i=1}^K f_i}^n$ which achieves the entropy bound of the coloring random variable [17], [38].

The joint PMF $P_{W,X_{[3]}}$ that minimizes (6) is given in Table I, using which (6) yields $R_M \leq 1.5$ bits.

$W \setminus X_{[3]}$	000	001	010	011	100	101	110	111
$\{000, 010, 111\}$	$\frac{1}{8}$	0	0	0	0	0	0	$\frac{1}{8}$
$\{001, 010, 011, 100\}$	0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	0	0	0
$\{101, 110\}$	0	0	0	0	0	$\frac{1}{8}$	$\frac{1}{8}$	0

TABLE I: Joint PMF $P_{W,X_{[3]}}$ that solves (6).

Example 1 demonstrates that designing a broadcast message leveraging the structure of all demands and users' side information achieves a rate of $R_{ach} = 1.5$ bits. In contrast, using $\mathcal{S}_i = X_i$ at $i \in [3]$ and applying the Slepian-Wolf theorem [18], sending $(X_1 \oplus_2 X_2, X_2 \oplus_2 X_3)$ achieves the minimum rate for asymptotically lossless reconstruction at each user, as $H_2(X_1 \oplus_2 X_2, X_2 \oplus_2 X_3) = 2$ bits [18] for all $i \in [3]$, where $\oplus_q$ denotes modulo-q sum. Thus, our graph-based approach achieves a 25% rate reduction versus [18].

Next, we detail a linear computation broadcast example from [14] and show that our approach outperforms that of [14].

Example 2 (3-**user linear computation broadcast [14, Example 2]**). *We denote user requests by $f_1 = X_2 \oplus_3 X_3$, $f_2 = X_1 \oplus_3 X_3$, and $f_3 = X_1 \oplus_3 2X_2$, respectively, where $X_j \in \mathbb{F}_3$ for $j \in [3]$, and $\mathcal{S}_i = X_i$ for $i \in [3]$.*

For the proposed setup, the rate required to encode $G_{\cup_{i=1}^3 f_i}$, i.e., R_M, is upper bounded by the solution of (4). However, the broadcast graph $G_{\cup_{i=1}^K f_i}$ has $|\mathcal{V}| = q^N$ vertices, rendering the analytical solution for $\Gamma(G_{\cup_{i=1}^K f_i})$ and hence for (4) infeasible for large q^N, as compared to Example 1 with fewer vertices. However, using the sub-additivity properties of graphs, we can divide the solution space and address less complex tasks. To that end, we propose an alternative approach for designing M by considering pairs of users and defining compatible functions for each pair, which we detail below.

Compatibility function-based approach. *To build our achievability scheme, we consider any pair (i_1, i_2) of users $i_1, i_2 \in [3]$ with respective demands f_{i_1} and f_{i_2}, and with $\mathcal{S}_{i_1} = X_{i_1}$ and $\mathcal{S}_{i_2} = X_{i_2}$, respectively. The rate required to encode $G_{f_{i_1} \cup f_{i_2}}$, denoted as $R_{i_1 i_2}$, must satisfy*

$$R_{i_1 i_2} \leq \min_{X_{[3]} \in W \in \Gamma(G_{f_{i_1} \cup f_{i_2}})} \max_{i \in \{i_1, i_2\}} \{H_q(W \mid X_i)\} ,$$

where instead of explicitly determining $\Gamma(G_{f_{i_1} \cup f_{i_2}})$, we use a compatibility-based approach and demonstrate its optimality for reconstructing f_{i_1} and f_{i_2}. A given function $Z_{i_1 i_2}$ is called a compatible function for pair of users (i_1, i_2), if it yields $\phi_{i_1}(Z_{i_1 i_2}, X_{i_1}) = f_{i_1}$ and $\phi_{i_2}(Z_{i_1 i_2}, X_{i_2}) = f_{i_2}$. For users $(1, 2)$, with demands given above, a compatible function is $Z_{12} = X_1 \oplus_3 X_2 \oplus_3 X_3$. Similarly, for user pairs $(1, 3)$, and $(2, 3)$, functions $Z_{13} = X_1 \oplus_3 2X_2 \oplus_3 2X_3$ and $Z_{23} = X_1 \oplus_3 2X_2 \oplus_3 X_3$, respectively, are compatible. For any given (i_1, i_2), the rate achieved by $Z_{i_1 i_2}$ satisfies

$$R_{i_1 i_2} \leq H_3(Z_{i_1 i_2}) = H_3(f_{i_1} \mid X_{i_1}) = 1 \ \ bit, \quad i_1, i_2 \in [3] .$$

We split a dataset X_j into two disjoint sub-variables X_j^1 and X_j^2, each with a size of 0.5 bits, similar to [14, Example 2]. Let $Z_{i_1 i_2}^l$, for $l \in [2]$, denote the compatible function sub-variables X_j^l, for $l \in [2]$, $j \in [3]$, with $H_3(Z_{i_1 i_2}^l) = 0.5$ bits. We consider three broadcast messages as follows: (i) Z_{12}^1, (ii) Z_{23}^2, and (iii) $Z_{13}^1 \oplus_3 Z_{13}^2$. From (i) and (iii), f_1 is decoded; using (i) and (ii) f_2 is decoded; and finally, f_3 is decoded given (ii) and (iii). Thus, R_M for broadcasting (i)-(iii), exploiting the independence bound on entropy, satisfies[7]

$$\begin{aligned} R_M &\leq H_3(Z_{12}^1, Z_{23}^2, Z_{13}^1 \oplus_3 Z_{13}^2) \\ &\leq H_3(Z_{12}^1) + H_3(Z_{23}^2) + H_3(Z_{13}^1 \oplus_3 Z_{13}^2) = 1.5 \ \ bits. \end{aligned}$$

Vector-based approach. *We next propose a vector-based achievability scheme, drawing on the notion of characteristic graphs, that involves designing two messages, M_1 and $\boldsymbol{M}_2$, derived from $W \in \Gamma(G_{\cup_{i=1}^3 f_i})$. When $X_2 = 0$, with a probability $1/3$, the MISs are determined by the function $X_1 \oplus_3 X_3$ and encode these MISs to the message M_1 with a rate $H_3(M_1) = H_3(X_1 \oplus_3 X_3) = 1$. Similarly, when $X_2 \in [2]$, with a probability $2/3$, the MISs are determined by $X_1 \oplus_3 X_2 \oplus_3 X_3$ and X_2. These MISs are encoded to message $\boldsymbol{M}_2$ with a rate of $H_3(\boldsymbol{M}_2) = H_3(X_1 \oplus_3 X_2 \oplus_3 X_3, X_2) = 1.63$ bits. Thus, the rate required to broadcast these messages is*

$$R_M \leq \frac{1}{3} H_3(M_1) + \frac{2}{3} H_3(\boldsymbol{M}_2) = 1.42 \ bits , \qquad (8)$$

compared to the achievable rate of 1.5 bits in [14, Example 2].

Example 2 presents two approaches for linear function demands: a compatibility function-based technique (similar to [14]) and a vector-based technique. The performance of the former approach is similar to the linear coding scheme of [14, Example 2] with a rate of 1.5 bits. The latter showed that by carefully designing the broadcast message based on the nature of the demands and the datasets, we can achieve $R_M \leq 1.42$ bits, indicating an improvement over the compatibility function-based approach that completely relies on the linear encoding of the demands. This shows that a vector-based approach, refining the characteristic graph-based broadcast message design, can surpass the state of the art. However, achieving the minimum rate heavily depends on the structure of the demanded function and lacks a generalizable form.

The next section presents our achievable and converse rates for the K-user computation broadcast model.

IV. K-USER COMPUTATION BROADCAST: ACHIEVABILITY AND CONVERSE BOUNDS

This section presents an upper bound R_{ach} and a lower bound R_{con} on the communication rate for the K-user computation broadcast model, as shown in Figure 1.

---

[7]Given a length n sequence X_j^n, we denote two disjoint splits of X_j^n by $X_j^{n/2} = \{X_{j,1}, \ldots, X_{j,n/2}\}$ and $X_j^{n/2+1:n} = \{X_{j,n/2+1}, \ldots, X_{j,n}\}$ two sub-sequences of length $\frac{n}{2}$. Combining the idea of splitting with the above-mentioned compatible function construction enables the asymptotically lossless reconstruction of $\{f_i^n : i \in [3]\}$ for sufficiently large n.

Theorem 1. *Given a K-user computation broadcast model, with N datasets $X_{[N]}$, and a broadcast graph $G_{\cup_{i=1}^K f_i}$ the communication rate for the broadcast message M satisfies:*

$$R_{ach} = \min_{X_{[N]} \in W \in \Gamma(G_{\cup_{i'=1}^K f_{i'}})} \max_{i \in [K]} I(X_{[N]}; W \mid \mathcal{S}_i) , \quad (9)$$

$$R_{con} = \min_{p(u|x_{[N]})} \max_{i \in [K]} I(X_{[N]}, U \mid \mathcal{S}_i) , \quad (10)$$

where $\Gamma(G_{\cup_{i=1}^K f_i})$ is the set of MISs of $G_{\cup_{i=1}^K f_i}$.

Proof. To derive R_{ach}, we here use the compress-bin strategy [56, Theorem 11.3]. Consider length n i.i.d realizations of all sources, i.e., $X_{[N]}^n = \{X_{1,t}, \ldots, X_{N,t}\}_{t\in[n]}$, drawn from $P_{X_{[N]}}$. Define $\hat{f}_i^n = F_i(W^n, \mathcal{S}_i^n)$, denoting an estimate of f_i^n at user $i \in [K]$, with W^n denoting an MIS of $G_{\cup_{i=1}^K f_i}^n$ built using the length n realizations $X_{[N]}^n$ and $\{\mathcal{S}_i^n\}_{i\in[K]}$.

a) Codebook generation: The master node generates $q^{nR'}$ i.i.d sequences $W^n(l)$, where $l \in [0, q^{nR'} - 1]$. For $R' > R_{ach}$, the index set $[0, q^{nR'} - 1]$ is uniformly partitioned into $q^{nR_{ach}}$ bins, $\mathcal{B}(m) = [(m-1)q^{n(R'-R_{ach})} : mq^{n(R'-R_{ach})} - 1]$, for $m \in [1, q^{nR_{ach}}]$. The codebook, denoted as $\mathcal{C} = \{m \in [1, q^{nR_{ach}}]\}$, is revealed to all users.

b) Encoding: Given $X_{[N]}^n$, the master node finds an index $l^* \in [1 : q^{nR'} - 1]$ such that $(W^n(l^*), X_{[N]}^n) \in \mathcal{T}_{\epsilon'}^{(n)}$, where $\mathcal{T}_{\epsilon'}^{(n)}$ represents ϵ'-typical sequences $(W^n, X_{[N]}^n)$ for some $\epsilon' > 0$. If there are multiple such indices, the master randomly selects one of them. If no such index exists, the master selects $l^* = 0$. The master node broadcasts m^* such that $l^* \in \mathcal{B}(m^*)$. Thus, (l^*, m^*) represents the selected indices at the master.

c) Decoding: Let $\epsilon > \epsilon'$. User $i \in [K]$ receives m^*, and finds the unique index $\hat{l}_i \in \mathcal{B}(m^*)$, such that $(W^n(\hat{l}_i), \mathcal{S}_i^n) \in \mathcal{T}_{\epsilon}^{(n)}$. User $i \in [K]$ then computes $F_i(W^n(\hat{l}_i), \mathcal{S}_i^n)$. It declares an error if there is no such $\hat{l}_i$.

d) Analysis of expected error: Let $W^n(\hat{l}_i)$ be the message decoded by user $i \in [K]$. An error event at user $i \in [K]$ is denoted as $\mathcal{E}r_i = \{(W^n(\hat{l}_i), X_{[N]}^n) \notin \mathcal{T}_{\epsilon}^{(n)}\} = \{\hat{l}_i \neq l^*\}$, which can be categorized into three events:

$$\mathcal{E}r_{i1} = \{(W^n(l), X_{[N]}^n) \notin \mathcal{T}_{\epsilon'}^{(n)}, \text{ for all } l \in [0 : q^{nR'} - 1]\} ,$$

$$\mathcal{E}r_{i2} = \{(W^n(l^*), X_{[N]}^n) \notin \mathcal{T}_{\epsilon}^{(n)}\} ,$$

$$\mathcal{E}r_{i3} = \{(W^n(l), \mathcal{S}_i^n) \in \mathcal{T}_{\epsilon}^{(n)}, \text{ for some } l \in \mathcal{B}(m^*), l \neq l^*\} .$$

As $n \to \infty$, by covering lemma (cf. [56, Lemma 3.3]), $P(\mathcal{E}r_{i1}) \to 0$ if $R' > I(X_{[N]}; W) + \delta(\epsilon')$, where $\delta(\epsilon') > 0$ tends to zero as $\epsilon' \to 0$, by the conditional typically lemma [56, Section 2.5] $P(\mathcal{E}r_{i2} \cap \mathcal{E}r_{i1}^c) \to 0$, and by packing lemma [56, Lemma 3.1], $P(\mathcal{E}r_{i3}) \to 0$ if $R' - R_{ach} < I(W; \mathcal{S}_i) - \delta(\epsilon)$. Using these lemmas yields $R_{ach} > I(X_{[N]}; W \mid \mathcal{S}_i)$ for any $i \in [K]$. Hence, it must hold that

$$R_{ach} \geq \max_{i \in [K]} I(X_{[N]}; W \mid \mathcal{S}_i) , \quad (11)$$

where similar techniques exist for lossy source coding with side information, see e.g., the achievability scheme of [57].

Next, we derive the lower bound R_{con}. Given M, user $i \in [K]$ computes $\hat{f}_i^n$ upon receiving M and $\mathcal{S}_i^n$. The t-th

realization $f_{i,t}$, for $t \in [n]$, is defined as $f_{i,t} = g_{i,t}(U_t, \mathcal{S}_{i,t})$ for some function $g_{i,t}$ where U_t is an auxiliary variable defined as $U_t = (M, \mathcal{S}_i^{t-1})$. Letting $t = n$, we obtain

$$\begin{aligned}
nR_{ach} &\geq H_q(M) \overset{(a)}{\geq} H_q(M \mid \mathcal{S}_i^n) \overset{(b)}{=} I(X_{[N]}^n; M \mid \mathcal{S}_i^n) , \\
&\overset{(c)}{=} \sum_{t=1}^n I(X_{[N],t}; M \mid \mathcal{S}_i^n, X_{[N]}^{t-1}) , \\
&\overset{(d)}{=} \sum_{t=1}^n I(X_{[N],t}; M, X_{[N]}^{t-1}, \mathcal{S}_i^{t-1} \mid \mathcal{S}_{i,t}) \\
&\overset{(e)}{=} \sum_{t=1}^n I(X_{[N],t}; U_t \mid \mathcal{S}_{i,t}) ,
\end{aligned}$$

where (a) follows from conditioning, (b) from noting that M is function of $X_{[N]}^n$, i.e., $H_q(M \mid \mathcal{S}_i^n, X_{[N]}^n) = 0$, (c) from the chain rule of mutual information, (d) is due to the independence of $X_{[N],t}$ from $(X_{[N]}^{t-1}, \mathcal{S}_i^{t-1}, \mathcal{S}_{i,t+1}^n)$, and (e) from setting the auxiliary variable $U_t = (M, X_{[N]}^{t-1}, \mathcal{S}_i^{t-1})$. $\square$

We next introduce a converse bound, different from R_{con} in Theorem 1, derived using the joint PMF of each of the demanded functions and the available side information.

Proposition 1. *The lower bound on the communication rate for satisfying all users' demands, given side information $\mathcal{S}_i$ is*

$$H_q(f_1', f_2', \cdots f_K') , \quad (12)$$

where $f_i' = f_i|\mathcal{S}_i$ is a conditional random variable capturing the demand f_i of user $i \in [K]$ given $\mathcal{S}_i$, where $f_i' \in \mathbb{F}_q$.

Proof. The proof follows from expanding $I(f_1', \cdots, f_K'; M)$ in two different ways and noting that $H(f_i' \mid M) = 0$, $i \in [K]$. For details, see [19, Proof of Proposition 1]. $\square$

For comparison, we generalize the genie-aided lower bound in [58] to our model, where given the set of users in $[i-1]$ with $i \in \{2, \ldots, K\}$, a genie provides the side information $\{\mathcal{S}_{i'}\}_{i'\in[i-1]}$ and function demands $\{f_{i'}\}_{i'\in[i-1]}$ to user i, resulting in a lower bound of 1.155 bits (see [19, Proposition 2]). The application of the lower bound (12) in Proposition 1 to Example 1 yields 1.45 bits, while from (6) $R_{ach} = 1.5$ bits. This comparison shows the tightness of (12) compared to the bound in [58] for serving this type of non-linear demands.

Discussion. In this paper, we examined the K-user computation broadcast model with general data distributions and general function demands over finite fields. Using the characteristic graph framework, we proposed a novel broadcast coding scheme that reduces the communication rate while accommodating both linear and non-linear demands compared to the state of the art. Moreover, we presented achievable and converse rates for the general K-user computation broadcast problem, while more research is needed to quantify the gap between the bounds. Future directions include exploring more intelligent storing strategies to reduce communication complexity, where available side information may consist of functions of master's datasets (instead of pure datasets), and where users might have heterogeneous storage capacities.

REFERENCES

[1] J. B. Predd *et al.*, "Distributed learning in wireless sensor networks," *IEEE Signal Process. Mag.*, vol. 23, no. 4, pp. 56–69, Jul. 2006.

[2] H. Yang, T. Ding, and X. Yuan, "Federated learning with lossy distributed source coding: Analysis and optimization," *IEEE Trans. Commun.*, vol. 71, no. 8, pp. 4561–4576, May 2023.

[3] A. Şahin and R. Yang, "A survey on over-the-air computation," *IEEE Comm. Surveys & Tutorials*, vol. 25, no. 3, pp. 1877–1908, Apr. 2023.

[4] J. Montalban *et al.*, "Broadcast core-network: Converging broadcasting with the connected world," *IEEE Trans. on Broadcast.*, vol. 67, no. 3, pp. 558–569, Aug. 2021.

[5] A. Kate and I. Goldberg, "Distributed private-key generators for identity-based cryptography," in *Proc. Springer Int. Conf. Security and Crypto. for Networks*, Sep. 2010, pp. 436–453.

[6] X. Liu *et al.*, "Privacy-preserving multi-keyword searchable encryption for distributed systems," *IEEE Trans. Paral. and Dis. Systems*, vol. 32, no. 3, pp. 561–574, Sep. 2020.

[7] S. Li *et al.*, "Coded distributed computing: Fundamental limits and practical challenges," in *Proc. Asilomar Conf. Sig. Sys. Comp. (ACSSC)*, Nov. 2016, pp. 509–513.

[8] S. Dutta *et al.*, "On the optimal recovery threshold of coded matrix multiplication," *IEEE Trans. Inf. Theory*, vol. 66, no. 1, pp. 278–301, Jul. 2019.

[9] B. S. Chlebus *et al.*, "Deterministic radio broadcasting," in *Proc. Springer Auto. Lang. Prog.*, Geneva, Switzerland, Jul. 2000, pp. 717–729.

[10] U. Reimers, "Digital video broadcasting," *IEEE Communic. Mag.*, vol. 36, no. 6, pp. 104–110, Jun. 1998.

[11] A. Dupin and S. Abelard, "Broadcast Encryption using Sum-Product decomposition of Boolean functions," Cryptology ePrint Archive, Paper 2024/154, 2024.

[12] K. Wan *et al.*, "On the optimal load-memory tradeoff of cache-aided scalar linear function retrieval," *IEEE Trans. Inf. Theory*, vol. 67, no. 6, pp. 4001–4018, 2021.

[13] H. Sun and S. A. Jafar, "On the capacity of computation broadcast," *IEEE Trans. Inf. Theory*, vol. 66, no. 6, pp. 3417–3434, Dec. 2019.

[14] Y. Yao and S. A. Jafar, "The capacity of 3 user linear computation broadcast," *IEEE Trans. Inf. Theory*, Apr. 2024.

[15] ——, "On the generic capacity of K-user symmetric linear computation broadcast," *IEEE Trans. Inf. Theory*, Mar. 2024.

[16] J. Ravi and B. K. Dey, "Broadcast function computation with complementary side information," in *Proc., IEEE Global Commun. Conf. (Globecom) Wkshps*, Dec. 2016, pp. 1–6.

[17] J. Körner, "Coding of an information source having ambiguous alphabet and the entropy of graphs," in *Proc., Prague Conf. Inf. Theory*, Prague, Hungary, Sep. 1973, pp. 411–425.

[18] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. 19, no. 4, pp. 471–480, Jul. 1973.

[19] M. R. Deylam Salehi, V. K. K. Purakkal, and D. Malak, "Non-linear function computation broadcast," *preprint arXiv:2502.13688*, 2025.

[20] A. C.-C. Yao, "Some complexity questions related to distributive computing (preliminary report)," in *Proc., ACM Symp. Theory Comput.*, Atlanta, GA, Apr. 1979, pp. 209–213.

[21] J. N. Tsitsiklis and Z.-Q. Luo, "Communication complexity of convex optimization," *J. Complex.*, vol. 3, no. 3, pp. 231–243, Sep. 1987.

[22] Z.-Q. Luo, "Communication complexity of some problems in distributed computation," Ph.D. dissertation, MIT, Aug. 1989.

[23] U. Hadar, J. Liu, Y. Polyanskiy, and O. Shayevitz, "Communication complexity of estimating correlations," in *Proc., ACM Symp. Theory of Comput.*, Phoenix, AZ, USA, Jun. 2019, pp. 792–803.

[24] N. Ma, P. Ishwar, and P. Gupta, "Interactive source coding for function computation in collocated networks," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4289–4305, Apr. 2012.

[25] N. Ma and P. Ishwar, "Some results on distributed source coding for interactive function computation," *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 6180–6195, Aug. 2011.

[26] J. Körner and K. Marton, "How to encode the modulo-two sum of binary sources (corresp.)," *IEEE Trans. Inf. Theory*, vol. 25, no. 2, pp. 219–221, Mar. 1979.

[27] V. Lalitha *et al.*, "Linear coding schemes for the distributed computation of subspaces," *IEEE J. Sel. Areas Com.*, vol. 31, no. 4, pp. 678–690, Mar. 2013.

[28] B. Nazer and M. Gastpar, "Computation over multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3498–3516, Sep. 2007.

[29] D. Malak, "Structured codes for distributed matrix multiplication," *arXiv preprint arXiv:2501.00371*, Jan. 2025.

[30] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inf. Theoy*, vol. 22, no. 1, pp. 1–10, Jan. 1976.

[31] H. Yamamoto, "Wyner-Ziv theory for a general function of the correlated sources," *IEEE Trans. Inf. Theory*, vol. 28, pp. 803–807, Sep. 1982.

[32] E. Tuncel, "Slepian-wolf coding over broadcast channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1469–1482, Apr. 2006.

[33] R. Gray and A. Wyner, "Source coding for a simple network," *J. Bell Sys. Tech.*, vol. 53, no. 9, pp. 1681–1721, Nov. 1974.

[34] C. T. Li and A. El Gamal, "Extended gray–wyner system with complementary causal side information," *IEEE Trans. Inf. Theory*, vol. 64, no. 8, pp. 5862–5878, 2017.

[35] T. Han and K. Kobayashi, "A dichotomy of functions F(X, Y) of correlated sources (X, Y)," *IEEE Trans. Inf. Theory*, vol. 33, no. 1, pp. 69–76, Jan. 1987.

[36] N. Alon and A. Orlitsky, "Source coding and graph entropies," *IEEE Trans. Inf. Theory*, vol. 42, no. 5, pp. 1329–39, Sep. 1996.

[37] A. Orlitsky and J. R. Roche, "Coding for computing," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 903–17, Mar. 2001.

[38] S. Feizi and M. Médard, "On network functional compression," *IEEE Trans. Inf. Theory*, vol. 60, no. 9, pp. 5387–5401, Jun. 2014.

[39] D. Malak *et al.*, "Multi-server multi-function distributed computation," *Entropy*, vol. 26, no. 6, p. 448, Jun. 2024.

[40] M. R. D. Salehi and D. Malak, "An achievable low complexity encoding scheme for coloring cyclic graphs," in *Proc., Allerton Conf. Commun., Control, and Comput.*, Sep. 2023, pp. 1–8.

[41] D. Malak, "Distributed computing of functions of structured sources with helper side information," in *Proc., IEEE Int. Wksh. Signal Proces. Advances in Wireless Commun.*, Sep. 2023, pp. 276–280.

[42] D. Malak *et al.*, "Multi-functional distributed computing," in *Proc., Allerton Conf. Commun., Control, and Comput.*, Sep. 2024, pp. 1–8.

[43] E. Özyılkan and E. Erkip, "Distributed compression in the era of machine learning: A review of recent advances," in *Proc. Conf. Info. Sci. and Sys. (CISS)*, Apr. 2024, pp. 1–6.

[44] K. Lee *et al.*, "Speeding up distributed machine learning using codes," *IEEE Trans. Inf. Theory*, vol. 64, no. 3, pp. 1514–1529, Aug. 2017.

[45] L. Cui *et al.*, "Toward optimized federated learning with compressed communications by rate adaption," *IEEE Trans. Net.*, Jan. 2025.

[46] J. Dean and S. Ghemawat, "MapReduce: Simplified data processing on large clusters," *J. Commun. ACM*, vol. 51, no. 1, pp. 107–113, Jan. 2008.

[47] M. Zaharia *et al.*, "Spark: Cluster computing with working sets," in *Proc., USENIX HotTop Cloud Comp. Works.*, Boston, MA, USA, Jun. 2010.

[48] Z. Jia and S. A. Jafar, "On the capacity of secure distributed batch matrix multiplication," *IEEE Trans. Inf. Theory*, vol. 67, pp. 20–37, Sep. 2021.

[49] R. Tandon *et al.*, "Gradient coding: Avoiding stragglers in distributed learning," in *Proc., Int. Conf. Machine Learn.* Sydney, Australia: PMLR, Jul. 2017.

[50] Q. Yu *et al.*, "Straggler mitigation in distributed matrix multiplication: Fundamental limits and optimal coding," *IEEE Trans. Inf. Theory*, vol. 66, no. 3, pp. 1920–1933, Jan. 2020.

[51] A. Khalesi and P. Elia, "Multi-user linearly-separable distributed computing," *IEEE. Trans. Inf. Theory*, vol. 69, pp. 6314–39, Jun. 2023.

[52] F. Arbabjolfaei *et al.*, "Fundamentals of index coding," *Found. and Trends. Com. and Info. Theo.*, vol. 14, no. 3-4, pp. 163–346, Oct. 2018.

[53] A. Gamal and T. Cover, "Achievable rates for multiple descriptions," *IEEE Trans. Inf. Theory*, vol. 28, no. 6, pp. 851–857, Nov. 1982.

[54] R. Beigel, "Finding maximum independent sets in sparse and general graphs," in *Proc., SODA*, vol. 99, Baltimore, MD, USA, Jan. 1999, pp. 856–857.

[55] J. Körner, "Fredman–Komlós bounds and information theory," *J. SIAM on Alg. Disc. Methods*, vol. 7, no. 4, pp. 560–570, Oct. 1986.

[56] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge university press, Dec. 2011.

[57] C. Heegard and T. Berger, "Rate distortion when side information may be absent," *IEEE Trans. Inf. Theory*, vol. 31, no. 6, pp. 727–734, Nov 1985.

[58] S. Jafar, "Capacity with causal and noncausal side information: A unified view," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5468–5474, Nov. 2006.