

# Computation in Multicast Networks: Function Alignment and Converse Theorems

Changho Suh, *Member, IEEE*, Naveen Goela, and Michael Gastpar

**Abstract**—The classical problem in a network coding theory considers communication over multicast networks. Multiple transmitters send independent messages to multiple receivers that decode the same set of messages. In this paper, computation over multicast networks is considered: each receiver decodes an identical function of the original messages. For a countably infinite class of two-transmitter two-receiver single-hop linear deterministic networks, the computation capacity is characterized for a linear function (modulo-2 sum) of Bernoulli sources. A new upper bound is derived that is tighter than cut-set-based and genie-aided bounds. A matching inner bound is established via the development of a network decomposition theorem, which identifies elementary parallel subnetworks that can constitute an original network without loss of optimality. The decomposition theorem provides a conceptually simple proof of achievability that generalizes to  $L$ -transmitter  $L$ -receiver networks.

**Index Terms**—Computation capacity, function alignment, network decomposition theorem.

## I. INTRODUCTION

RECENTLY coding for computation in networks has received considerable attention with applications in sensor networks [1] and cloud computing scenarios [2], [3]. In a sensor network, a fusion node may be interested in computing a relevant function of the measurements from various data nodes. In a cloud computing scenario, a client may download a function or part of the original source information that is distributed (e.g. using a maximum distance separable code) across multiple data nodes.

The simplest setting for computation in networks consists of multiple sources transmitting information to a *single*

receiver which computes a function of the original sources. Appuswamy *et al.* study the fundamental limits of computation for linear and general target function classes for single-receiver networks [4]. While limited progress has been made for general target functions, the problem of linear function computation in single-receiver networks has been solved in part due to a duality theorem establishing an equivalence to the classical problem of communication over multicast networks [5].

Several results over the past decade have contributed to the understanding of classical communication in multicast networks in which the task is to transmit raw messages from transmitters to a set of receivers with identical message demands. The celebrated work of Ahlswede *et al.* [5] established that the cut-set bound is tight for multicast communication. Subsequent research developed practical linear network coding strategies ranging from random linear codes to deterministic polynomial-time code constructions [6]–[9]. The success of traditional multicast communication motivates us to explore the fundamental limits of multicasting a linear function in *multiple-receiver* networks as a natural next step. For this open problem, some facts are known based on example networks: (a) Random codes are insufficient in achieving capacity limits, and structured codes achieve higher computation rates [10]; (b) Linear codes are insufficient in general for computation over multi-receiver networks (cf. both [11] and [12]) and non-linear codes may achieve higher computation rates.

To make progress on the problem of multicasting a function, we consider a multiple-receiver network scenario in which all of the receivers compute a linear function (modulo-2 sum) of the independent Bernoulli sources generated at transmitters. Specifically, we consider the Avestimehr-Diggavi-Tse (ADT) deterministic network model [13] which well abstracts wireless Gaussian networks. In the context of classical communication, it has been well known that ADT networks can approximate wireless Gaussian networks within a constant gap to the optimality in capacity [13], [14]. Recently a similar approximation result has been established for the problem of computation in which a single receiver wishes to compute a linear function of multiple Gaussian sources [15]. Specifically [15] employs lattice codes to show that a multiple-source single-destination Gaussian network can be approximated to a class of linear deterministic networks (which includes the ADT network as a special case), within a constant factor of the optimal performance w.r.t. the distortion for computing the sum of the Gaussian sources. In this work, we intend to extend this approximation approach to more

Manuscript received October 14, 2012; revised May 11, 2015; accepted February 15, 2016. Date of publication February 23, 2016; date of current version March 16, 2016. C. Suh was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIP) (No.R0190-15-2030, Reliable crypto-system standards and core technology development for secure quantum key distribution network). This paper was presented at the 2012 IEEE Information Theory Workshop and the 2012 Allerton Conference.

C. Suh is with the School of Electrical Engineering, Korea Advanced Institute of Science and Technology, Daejeon 34141, South Korea (e-mail: chsuh@kaist.ac.kr).

N. Goela is with the Technicolor Research, Los Altos, CA 94022 USA (e-mail: naveen.goela@gmail.com).

M. Gastpar is with the School of Computer and Communication Sciences, Ecole Polytechnique Fédérale de Lausanne, Lausanne 1015, Switzerland, and also with the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, Berkeley, CA 94720 USA (e-mail: michael.gastpar@epfl.ch).

Communicated by Y. Steinberg, Associate Editor at Large.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIT.2016.2533611

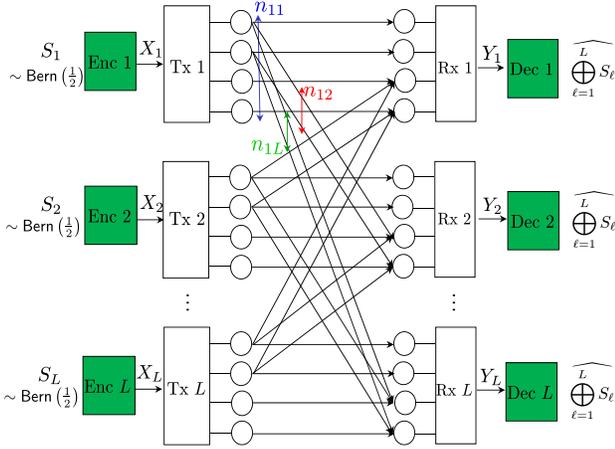


Fig. 1.  $L$ -transmitter  $L$ -receiver Avestimehr-Diggavi-Tse (ADT) deterministic network.

general computation scenarios in Gaussian networks. As an initial effort, we consider an  $L$ -transmitter  $L$ -receiver ADT network with the function multicast demand.

In this paper, we derive a new upper bound that is tighter than cut-set based bounds and genie-aided bounds. Especially in the case of  $L = 2$ , we establish a matching inner bound to characterize the computation capacity. The achievability builds upon our development of a *network decomposition theorem* which identifies elementary parallel subnetworks that can constitute an original network without loss of optimality. The network decomposition offers a conceptually simple achievability proof which we use to generalize to an arbitrary value of  $L$ . In the  $L$ -user case, we show the optimality of our achievability in the limit of  $L$ . Our achievable scheme is intimately related to the concept of *interference alignment* although the purpose of alignment is different. In our problem, the alignment idea is employed to compute a desired function with a smaller number of linear-subspace signal dimension than the number of sources involved in the function.

*Related Work:* In [11], [16], and [17], the computation capacity for multicasting a sum of sources is explored for arbitrary multiple-source multiple-destination networks. Rai and Dey [11] proved that there exists a linear solvably equivalent sum-network for any multiple-unicast network and vice-versa. Ramamoorthy and Langberg [17] characterized necessary and sufficient conditions for communicating sums of sources of two-source  $L$ -destination (or  $L$ -source two-destination) networks, when the entropy of each source is limited by 1. On the other hand, we consider sources without entropy constraints and establish the exact computation capacity of an ADT multiple-receiver network.

## II. MODEL

Consider an  $L$ -transmitter  $L$ -receiver ADT deterministic network depicted in Fig. 1. This network is described by integer parameters  $n_{ij}$ 's, each indicating the number of signal bit levels from transmitter  $i$  ( $i = 1, \dots, L$ ) to receiver  $j$  ( $j = 1, \dots, L$ ). In the example of Fig. 1,  $n_{11} = 4$ ,  $n_{12} = 2$  and  $n_{1L} = 2$ . Let  $X_\ell \in \mathbb{F}_2^q$  be transmitter  $\ell$ 's encoded signal

where  $q = \max_{ij} n_{ij}$ . The received signal at Rx  $\ell$  is then given by

$$Y_\ell = \mathbf{G}^{q-n_{1\ell}} X_1 \oplus \mathbf{G}^{q-n_{2\ell}} X_2 \oplus \dots \oplus \mathbf{G}^{q-n_{L\ell}} X_L, \quad (1)$$

where  $\mathbf{G}$  is the  $q$ -by- $q$  shift matrix, i.e.,  $[\mathbf{G}]_{ij} = \mathbf{1}\{i = j + 1\}$  ( $1 \leq i \leq q$ ;  $1 \leq j \leq q$ ), and operations are performed in  $\mathbb{F}_2$ . Here  $\oplus$  indicates the bit wise XOR.

Each receiver wishes to compute modulo-2 sums of the  $L$  Bernoulli sources ( $S_1^K, \dots, S_L^K$ ), generated at the  $L$  transmitters, with  $N$  uses of the network. Here we use shorthand notation to indicate the sequence up to  $K$ , e.g.,  $S_\ell^K := (S_{\ell 1}, \dots, S_{\ell K})$ . We assume that  $(S_1^K, \dots, S_L^K)$  are independent and identically distributed with  $\text{Bern}(\frac{1}{2})$ . Transmitter  $\ell$  uses its encoding function to map  $S_\ell^K$  to a length- $N$  codeword  $X_\ell^N$ . Receiver  $\ell$  uses a decoding function  $d_\ell$  to estimate  $\bigoplus_{\ell=1}^L S_\ell^K$  from its received signal  $Y_\ell^N$ . An error occurs whenever  $d_\ell \neq \bigoplus_{\ell=1}^L S_\ell^K$ . The average probabilities of error are given by  $\lambda_\ell = \mathbb{E}[\Pr(d_\ell \neq \bigoplus_{\ell=1}^L S_\ell^K)]$ ,  $\ell = 1, \dots, L$ .

We say that the computation rate  $R_{\text{comp}} = \frac{K}{N}$  is achievable if there exists a family of codebooks and encoder/decoder functions such that the average decoding error probabilities of  $(\lambda_1, \dots, \lambda_L)$  go to zero as code length  $N$  tends to infinity. The computation capacity  $C_{\text{comp}}$  is the supremum of the achievable computation rates.

In the case of  $L = 2$ , we classify networks into two classes depending on a channel parameter condition.

*Definition 1:* A network is said to be *degenerate* if  $n_{11} - n_{12} = n_{21} - n_{22}$ . A network is said to be *non-degenerate* if  $n_{11} - n_{12} \neq n_{21} - n_{22}$ .

*Remark 1:* Suppose that  $n_{11} - n_{12} = n_{21} - n_{22} \geq 0$ . In this case, one can see that

$$\begin{aligned} \mathbf{G}^{n_{21}-n_{22}} Y_1 &= \mathbf{G}^{q-n_{11}+n_{21}-n_{22}} X_1 \oplus \mathbf{G}^{q-n_{22}} X_2 \\ &= \mathbf{G}^{q-n_{12}} X_1 \oplus \mathbf{G}^{q-n_{22}} X_2 = Y_2. \end{aligned}$$

implying that  $Y_2$  is a *degenerated* version of  $Y_1$ . For the other case of  $n_{11} - n_{12} = n_{21} - n_{22} \leq 0$ , one can readily see that  $Y_1$  is a *degenerated* version of  $Y_2$  as  $Y_1 = \mathbf{G}^{n_{12}-n_{11}} Y_2$ .  $\square$

## III. MAIN RESULTS

### A. 2-by-2 Network

*Theorem 1:*

$$C_{\text{comp}} \leq \min\{n_{11}, n_{12}, n_{22}, n_{21}\}. \quad (2)$$

For degenerate networks:  $n_{11} - n_{12} = n_{21} - n_{22}$ , this upper bound is achievable.

*Proof:* The standard cut-set argument establishes the upper bound. Using Fano's inequality:  $H(S_1^K \oplus S_2^K | Y_1^N) \leq 1 + \Pr(d_1 \neq S_1^K \oplus S_2^K) K$ , and denoting  $\epsilon_N := 1/N + \Pr(d_1 \neq S_1^K \oplus S_2^K) K/N$ , we get:

$$\begin{aligned} N(R_{\text{comp}} - \epsilon_N) &\leq I(S_1^K \oplus S_2^K; Y_1^N) \\ &\leq I(S_1^K \oplus S_2^K; Y_1^N | S_2^K, X_2^N) \\ &\leq H(Y_1^N | X_2^N) \leq N n_{11} \end{aligned}$$

where the second inequality follows from the non-negativity of mutual information and the fact that  $S_2^K$  is independent

of  $S_1^K \oplus S_2^K$ . This yields  $R_{\text{comp}} \leq n_{11}$ . Similarly one can prove that  $R_{\text{comp}} \leq \min\{n_{12}, n_{21}, n_{22}\}$ .

The achievability proof is as follows. Assume that  $n_{11} - n_{12} = n_{21} - n_{22} \geq 0$ . Then  $Y_2$  is a degenerated version of  $Y_1$ :  $Y_2 = \mathbf{G}^{n_{21}-n_{22}} Y_1$ . This shows an equivalence to a single-receiver case which concerns receiver 2's demand only. Hence, the computation rate in this case is the same as that of a multiple-access channel having receiver 2 as the receiver. So  $R_{\text{comp}} \geq \min\{n_{12}, n_{22}\}$  [4], [10]. Similarly for the other case of  $n_{11} - n_{12} = n_{21} - n_{22} \leq 0$ , one can show that  $Y_1$  is a degenerated version of  $Y_2$  and therefore the network becomes equivalent to a single-receiver network w.r.t. receiver 1 where  $R_{\text{comp}} \geq \min\{n_{11}, n_{21}\}$ . In the first case,  $n_{12} \leq n_{11}$  and  $n_{22} \leq n_{21}$ ; hence  $\min\{n_{12}, n_{22}\} \leq \min\{n_{11}, n_{21}\}$ . In the second case, on the other hand,  $\min\{n_{12}, n_{22}\} \geq \min\{n_{11}, n_{21}\}$ . Therefore,  $R_{\text{comp}} \geq \min\{n_{12}, n_{22}, n_{11}, n_{21}\}$ . ■

*Theorem 2 (Upper Bound for Non-Degenerate Networks):*

$$C_{\text{comp}} \leq \frac{\max(n_{11}, n_{21}) + \max(n_{22}, n_{12})}{3}. \quad (3)$$

*Proof:* See Section IV. ■

We show the tightness of the above bound for the case of  $n := n_{11} = n_{22}$  and  $m := n_{12} = n_{21}$  that we call a symmetric case.

*Theorem 3 (Symmetric Network):* For  $n := n_{11} = n_{22}$  and  $m := n_{12} = n_{21}$ ,

$$C_{\text{comp}} = \begin{cases} \min\{m, n, \frac{2}{3} \max(m, n)\}, & m \neq n; \\ n, & m = n. \end{cases} \quad (4)$$

*Proof:* The converse proof is immediate from Theorems 1 and 2. See Section V for the achievability proof. ■

*Remark 2 (Generalization to  $p$ -Ary Models):* The results in Theorems 1, 2, and 3 hold for  $p$ -ary models in which  $S_\ell$ 's and channel input/output are in  $\mathbb{F}_p$ , the desired function is modulo- $p$  addition, and the channel operation is also modulo- $p$  addition. Here  $p$  is a prime number. Specifically, in the  $p$ -ary case, the converse proof of Theorem 1 starts with a slightly different inequality, yet yielding the same result:  $N((\log_2 p)R_{\text{comp}} - \epsilon_N) \leq H(Y_1^N | X_2^N) \leq N(\log_2 p)n_{11}$ . One can make the same argument for the other theorems. This will be clearer in the detailed proof that will be presented in Section IV.

*Remark 3 (Comparison to Separation Scheme):* A baseline strategy for the problem at hand is to let both receivers first fully recover both sources,  $S_1$  and  $S_2$ , and only then apply the desired computation. Since this strategy separates communication from computation, we refer to it as the *separation scheme*. The performance of this strategy is well known: The capacity region for data transmission is simply the intersection of the capacity regions of the two multiple-access channels, one from both transmitters to Receiver 1, the other from both transmitters to Receiver 2. For symmetric models ( $n_{11} = n_{22} = n$  and  $n_{12} = n_{21} = m$ ), this evaluates to the message rate region characterized by  $R_1 \leq \min(m, n)$ ,  $R_2 \leq \min(m, n)$  and  $R_1 + R_2 \leq \max(m, n)$ . The corresponding achievable computation rate is simply the maximum

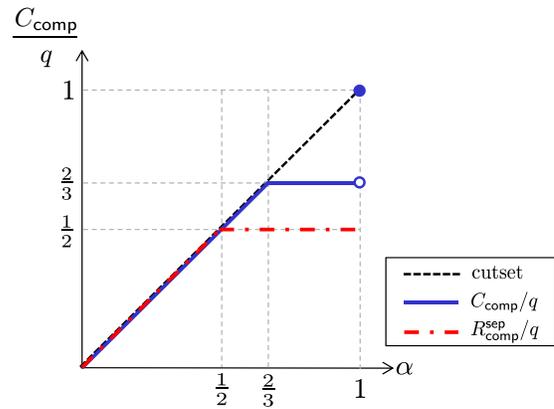


Fig. 2. Computation capacity of the symmetric network ( $n := n_{11} = n_{22}$  and  $m := n_{12} = n_{21}$ ). Here  $\alpha := \frac{\min(m, n)}{\max(m, n)}$  and  $q := \max(m, n)$ .

symmetric rate point in this region, and thus,  $\frac{R_{\text{comp}}^{\text{sep}}}{q} \geq \min\{\alpha, \frac{1}{2}\}$ , where  $\alpha := \frac{\min(m, n)}{\max(m, n)}$ . This is illustrated in Figure 2. For the regime  $0 \leq \alpha \leq \frac{1}{2}$ , the separation strategy is optimal, but for  $\frac{1}{2} < \alpha \leq 1$ , it is strictly suboptimal. □

## B. $L$ -by- $L$ Network

We consider a symmetric setting where the only two integer parameters of  $(m, n)$  describe the network. Here  $n$  indicates the number of signal bit levels from transmitter  $\ell$  to receiver  $\ell$ ; and  $m$  denotes the number of signal bit levels from transmitter  $\ell$  to receiver  $\ell' (\neq \ell)$ .

*Theorem 4 ( $L$ -by- $L$  Symmetric Network):*

$$R_{\text{comp}} \geq \begin{cases} \min\{m, n, \frac{1}{2} \max(n, m)\}, & m \neq n; \\ n, & m = n. \end{cases}$$

$$C_{\text{comp}} \leq \begin{cases} \min\{m, n, \frac{L}{2L-1} \max(n, m)\}, & m \neq n; \\ n, & m = n. \end{cases}$$

*Proof:* See Section VI. ■

*Remark 4:* Note that our information-theoretic upper bound approaches the achievable rate as  $L$  tends to infinity. As in the  $L = 2$  case, this result also holds for general  $p$ -ary models. This will be clarified in Section VI. □

## IV. PROOF OF THEOREM 2

*Lemma 1:* For non-degenerate networks ( $n_{11} - n_{12} \neq n_{21} - n_{22}$ ), there exists  $(i, j)$  such that  $\mathbf{G}^{q-n_{ij}} X_i$  can be reconstructed from the pair  $(Y_1, Y_2)$ .

*Proof:* See Appendix A. ■

Using this lemma, we can assume that without loss of generality,  $\mathbf{G}^{q-n_{12}} X_1$  is a function of  $(Y_1, Y_2)$ .

Our proof includes general  $p$ -ary models. Starting with Fano's inequality, we get:

$$\begin{aligned} & N(3(\log_2 p)R_{\text{comp}} - \epsilon_N) \\ & \leq I(S_1^K \oplus S_2^K; Y_1^N) + I(S_1^K \oplus S_2^K; Y_2^N) + I(S_1^K \oplus S_2^K; Y_2^N) \\ & \stackrel{(a)}{\leq} [H(Y_1^N) - H(Y_1^N | S_1^K \oplus S_2^K)] \\ & \quad + [H(Y_2^N) - H(Y_2^N | S_1^K \oplus S_2^K, Y_1^N)] + I(S_1^K \oplus S_2^K; Y_2^N) \end{aligned}$$

$$\begin{aligned}
 &\leq H(Y_1^N) + H(Y_2^N) \\
 &\quad - H(Y_1^N, Y_2^N | S_1^K \oplus S_2^K) + I(S_1^K \oplus S_2^K; Y_2^N, S_2^K) \\
 &\stackrel{(b)}{=} H(Y_1^N) + H(Y_2^N) \\
 &\quad - H(Y_1^N, Y_2^N | S_1^K \oplus S_2^K) + H(\{\mathbf{G}^{q-n_{12}} X_{1i}\}_{i=1}^N | S_2^K) \\
 &\stackrel{(c)}{\leq} H(Y_1^N) + H(Y_2^N) \leq \sum [H(Y_{1i}) + H(Y_{2i})] \\
 &\leq N(\log_2 p)[\max(n_{11}, n_{21}) + \max(n_{12}, n_{22})]
 \end{aligned}$$

where (a) follows from the fact that conditioning reduces entropy; (b) follows from the fact that  $S_2^K$  is independent of  $S_1^K \oplus S_2^K$ , and that  $X_2^N$  is a function of  $S_2^K$ ; and (c) follows from  $H(Y_1^N, Y_2^N | S_1^K \oplus S_2^K) \geq H(\{\mathbf{G}^{q-n_{12}} X_{1i}\}_{i=1}^N | S_2^K)$  (see Claim 1 below). Hence, we get the desired bound.

*Claim 1:*

$$H(Y_1^N, Y_2^N | S_1^K \oplus S_2^K) \geq H(\{\mathbf{G}^{q-n_{12}} X_{1i}\}_{i=1}^N | S_2^K). \quad (5)$$

*Proof:*

$$\begin{aligned}
 &H(Y_1^N, Y_2^N | S_1^K \oplus S_2^K) - H(\{\mathbf{G}^{q-n_{12}} X_{1i}\}_{i=1}^N | S_2^K) \\
 &\stackrel{(a)}{=} H(Y_1^N, Y_2^N, \{\mathbf{G}^{q-n_{12}} X_{1i}\}_{i=1}^N | S_1^K \oplus S_2^K) \\
 &\quad - H(\{\mathbf{G}^{q-n_{12}} X_{1i}\}_{i=1}^N | S_2^K) \\
 &\geq H(\{\mathbf{G}^{q-n_{12}} X_{1i}\}_{i=1}^N | S_1^K \oplus S_2^K) \\
 &\quad - H(\{\mathbf{G}^{q-n_{12}} X_{1i}\}_{i=1}^N | S_2^K) \\
 &\stackrel{(b)}{=} H(\{\mathbf{G}^{q-n_{12}} X_{1i}\}_{i=1}^N) - H(\{\mathbf{G}^{q-n_{12}} X_{1i}\}_{i=1}^N | S_2^K) \geq 0.
 \end{aligned}$$

where (a) follows from our hypothesis that  $\mathbf{G}^{q-n_{12}} X_1$  is a function of  $(Y_1, Y_2)$  and (b) follows from the fact that  $X_1^N$  is a function of  $S_1^K$  that is independent of  $S_1^K \oplus S_2^K$ . ■

## V. ACHIEVABILITY PROOF OF THEOREM 3

In this section, we present a network decomposition theorem that permits to decompose a network into elementary subnetworks. Our achievable strategy is then developed *separately* for each subnetwork. In general, one must expect such a decomposition to entail a loss of optimality. However, for the case  $L = 2$  (two transmitters, two receivers), our converse proof in Theorem 2 implies that our network decomposition does not lead to any loss in performance.

### A. Achievability via Network Decomposition

In the sequel, we will use the terminology  $(m, n)$  model to mean a symmetric  $L$ -user ADT network. In the case  $m < n$ , each transmitter and each receiver has  $n$  levels, and at each receiver, the last  $m$  levels are being interfered in a modulo-sum fashion by the  $m$  first levels of each interfering transmitter, as illustrated in Figure 1. In the case  $m > n$ , each transmitter and each receiver has  $m$  levels. For each transmitter, its first  $n$  levels connect to the last  $n$  levels of its corresponding receiver. Furthermore, at each receiver, the last  $n$  levels associated with its corresponding transmitter are being interfered in a modulo-sum fashion by the last  $n$  levels of each interfering transmitter; at the remaining top  $(m-n)$  levels,

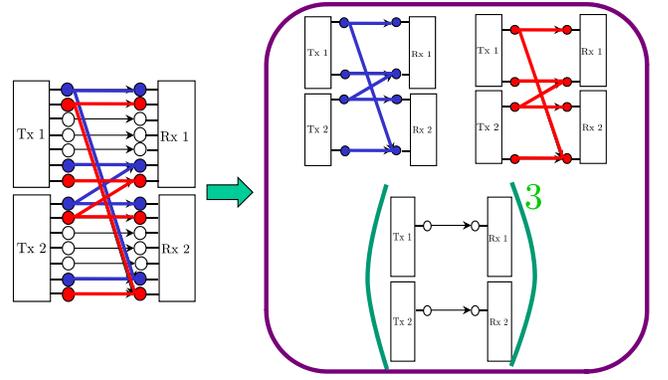


Fig. 3. A network decomposition example of an  $(m, n) = (2, 7)$  model. From (6),  $r = 0$  and  $a = 2$ ; hence, the decomposition is given by  $(2, 7) \rightarrow (0, 1)^3 \times (1, 2)^2$ .

collision occurs only across interfering transmitters. As illustrated by Figure 1, there is a natural representation of ADT models in terms of directed graphs, and we will occasionally refer to this representation in the sequel.

*Definition 2 (Network Concatenation):* The  $L$ -user concatenated model  $(m, n) \times (\tilde{m}, \tilde{n})$  is constructed from the  $L$ -user  $(m, n)$  model and the  $L$ -user  $(\tilde{m}, \tilde{n})$  by merging Transmitter  $\ell$  from the  $L$ -user  $(m, n)$  model with Transmitter  $\ell$  from the  $L$ -user  $(\tilde{m}, \tilde{n})$  model, and by also merging Receiver  $\ell$  from the  $L$ -user  $(m, n)$  model with Receiver  $\ell$  from the  $L$ -user  $(\tilde{m}, \tilde{n})$  model, for  $\ell = 1, 2, \dots, L$ . As a shorthand, we will also use the notation  $(m, n)^k = (m, n) \times (m, n) \times \dots \times (m, n)$ , where there are  $k$  terms, i.e., the concatenation of  $k$   $L$ -user  $(m, n)$  models.

Note that it is straightforward to see that network concatenation is both commutative and associative, e.g.,  $(m, n) \times (\tilde{m}, \tilde{n})$  is exactly the same model as  $(\tilde{m}, \tilde{n}) \times (m, n)$ .

*Definition 3 (Network Decomposition):* We say that the  $L$ -user  $(m, n)$  model can be decomposed into the  $L$ -user concatenated model  $(m_1, n_1) \times (m_2, n_2)$ , denoted as  $(m, n) \rightarrow (m_1, n_1) \times (m_2, n_2)$ , if the directed graph corresponding to the  $L$ -user  $(m, n)$  model is isomorphic to the directed graph corresponding to the  $L$ -user concatenated model  $(m_1, n_1) \times (m_2, n_2)$ .

An example of a network decomposition is given in Figure 3. The figure graphically proves the fact that, in our just defined notation,  $(2, 7) \rightarrow (0, 1)^3 \times (1, 2)^2$ . The following theorem establishes a number of more general statements of this type.

*Theorem 5 (Network Decomposition):* The  $L$ -transmitter  $L$ -receiver  $(m, n)$  network where  $m \neq n$ , can be decomposed into the following subnetworks:

- (1) For any  $k \in \mathbb{Z}^+$ ,

$$(km, kn) \rightarrow (m, n)^k = (m, n) \times (m, n) \times \dots \times (m, n).$$

- (2)  $(2m + 1, 2n + 1) \rightarrow (m, n) \times (m + 1, n + 1)$ .

- (3) For the arbitrary  $(m, n)$  model,

$$(m, n) \rightarrow \begin{cases} (r, r + 1)^{n-m-a} \times (r + 1, r + 2)^a, & m < n; \\ (r + 1, r)^{m-n-a} \times (r + 2, r + 1)^a, & m > n. \end{cases} \quad (6)$$

where

$$r = \left\lfloor \frac{\min\{m, n\}}{|n - m|} \right\rfloor, \quad a = \min\{m, n\} \bmod |n - m|. \quad (7)$$

The proof is given in Appendix B. Here we provide a proof idea with an  $(m, n, L) = (2, 7, 2)$  example, illustrated in Fig. 3. The idea is to use graph coloring with  $|n - m| = 5$  colors, identified by integers  $\{0, 1, 2, 3, 4\}$ . At transmitter 1, assign to level 1 and level 6 ( $= 1 + |n - m|$ ) the color 0 (blue color in this example). Use exactly the same rule to color the levels of transmitter 2 and receivers 1 and 2. The blue-colored graph represents an independent graph of model (1, 2). Next we assign the color 1 (red color in this example) to level 2 and level 7 ( $= 2 + |n - m|$ ), for all transmitters and receivers. We then obtain another independent graph of model (1, 2) and are left with model (0, 3). Obviously the model (0, 3) is decomposed into  $(0, 1)^3$ . Therefore, we get  $(2, 7) \rightarrow (1, 2)^2 \times (0, 1)^3$ .

In order to prove the achievability in Theorem 3, we will in particular leverage item (3) in Theorem 5. That is, the decomposition of any (symmetric) network into “gap-1” models, i.e., models of the form  $(r, r + 1)$  or  $(r + 1, r)$ . For those models, we can establish the following computation rates:

*Lemma 2* ( $L = 2$ ): The following computation rates are achievable:

- (1) For the model  $(0, 1)$ ,  $R_{\text{comp}} \geq 0$ .
- (2) For the model  $(1, 2)$ ,  $R_{\text{comp}} \geq 1$ .
- (3a) For the model  $(r - 1, r)$  with  $r \geq 3$ ,  $R_{\text{comp}} \geq \frac{2}{3}r$ .
- (3b) For the model  $(r, r - 1)$  with  $r \geq 3$ ,  $R_{\text{comp}} \geq \frac{2}{3}r$ .
- (4) For the model  $(r, r)$ ,  $R_{\text{comp}} \geq r$ .

*Proof:* See Section V-B. ■

By Theorem 5 and Lemma 2, we can now readily prove the achievability of Theorem 3. By symmetry, we focus on the case of  $m \leq n$ . The other case of  $m \geq n$  is a mirrored image in which the roles of transmitters 1 and 2 are swapped.

For the case of  $\alpha = 1$ ,  $R_{\text{comp}} \geq n$  by Item (4) in Lemma 2. For the case of  $0 \leq \alpha < \frac{1}{2}$ ,  $r = 0$  and  $a = m$  in (7); hence, the decomposition is given by  $(m, n) \rightarrow (0, 1)^{n-2m} \times (1, 2)^m$ . Thus, using Lemma 2, the computation rate is  $R_{\text{comp}} \geq 0 \cdot (n - 2m) + 1 \cdot m = m$ . Next, consider the case of  $\frac{1}{2} \leq \alpha < \frac{2}{3}$ . Applying the decomposition (6), we find that in this case,  $r = 1$  and  $a = 2m - n$ :  $(m, n) \rightarrow (1, 2)^{2n-3m} \times (2, 3)^{2m-n}$ . Thus, using Lemma 2, the computation rate is  $R_{\text{comp}} \geq 1 \cdot (2n - 3m) + 2 \cdot (2m - n) = m$ . Finally, consider the case of  $\alpha \geq \frac{2}{3}$ . Applying the decomposition (6), we find that in this case,  $r \geq 2$ . So we get

$$\begin{aligned} R_{\text{comp}} &\geq \frac{2}{3}(r + 1)(n - m - a) + \frac{2}{3}(r + 2)a \\ &= \frac{2}{3}\{r(n - m) + a + (n - m)\} \\ &\stackrel{(a)}{=} \frac{2}{3}\{m + (n - m)\} = \frac{2}{3}n. \end{aligned}$$

where (a) is due to (7). This completes the proof.

## B. Proof of Lemma 2

We note that Items (1), (2) and (4) are obvious, and Item (3b) follows from Item (3a), since without loss

of generality, for the multicast problem with  $L = 2$  users considered here, the case  $(m, n)$  and the case  $(n, m)$  are mirror images of each other in which the roles of transmitters 1 and 2 are swapped. We here provide an explicit proof of Item (3a), split into three cases. For notation, the symbols of Transmitter 1 will be denoted by  $a_1, a_2, a_3, \dots$  and the symbols of Transmitter 2 by  $b_1, b_2, b_3, \dots$ . The goal of both receivers is then to recover the modulo sums  $a_1 \oplus b_1, a_2 \oplus b_2, \dots$ . Moreover, we will find it convenient to collect the channel inputs used by transmitters into length- $n$  vectors denoted by  $\mathbf{x}_1$  and  $\mathbf{x}_2$ , respectively.

(i) *The Case  $r = 3\ell$ , i.e., the  $(3\ell - 1, 3\ell)$  Model:* Let us start with the simplest case of  $(2, 3)$  model. At Transmitter 1, we send  $\mathbf{x}_1 = (a_1, a_2, 0)$  and at Transmitter 2,  $\mathbf{x}_2 = (b_2, b_1, 0)$ . Clearly, both receivers learn both modulo sums. This code can be extended to all models for the form  $(3\ell - 1, 3\ell)$  (for all positive integers  $\ell$ ), as follows. We set  $X_{1,3k-2} = a_{2k-1}, X_{1,3k-1} = a_{2k}, X_{1,3k} = 0$  and  $X_{2,3k-2} = b_{2k}, X_{2,3k-1} = b_{2k-1}, X_{2,3k} = 0$ , for  $k = 1, 2, \dots, \ell$ . Each receiver can reconstruct all  $2\ell$  sums  $a_k \oplus b_k$  and thus, the computation rate is  $2\ell = \frac{2}{3}(3\ell)$ , as claimed.<sup>1</sup>

(ii) *The Case  $r = 3\ell + 1$ , i.e., the  $(3\ell, 3\ell + 1)$  Model:* Consider the  $(3, 4)$  model. Here we consider a code for the  $(9, 12)$  model (which can be implemented over three channel uses of the  $(3, 4)$  model). The code is given in Fig. 4(a). This code enables both receivers to recover all 8 computations, which means a computation rate of  $8/3 = \frac{2}{3} \cdot 4$  per channel use, as claimed in the lemma.

Now, consider the general  $(3\ell, 3\ell + 1)$  model, for any positive integer  $\ell \geq 2$ . In this case, we find that the  $(6, 9)$  model acts as a building block, and hence we will first develop a code for the  $(6, 9)$  code.

*The  $(6, 9)$  Code:* This network falls into the  $(2, 3)$  model category as it can be decomposed into three orthogonal  $(2, 3)$  networks. Since it will have a prominent role in the remainder of the proof, we explicitly write out the code. Here, each transmitter has 6 symbols. At Transmitter 1, we send  $\mathbf{x}_1 = (a_1, a_3, a_5, a_2, a_4, a_6, 0, 0, 0)$  and at Transmitter 2,  $\mathbf{x}_2 = (b_2, b_4, b_6, b_1, b_3, b_5, 0, 0, 0)$ . It is easy to verify that all 6 modulo sums can be recovered. The key to observe is that the last 3 inputs at both transmitters are all zero.

Now, let us go back to the general case:  $(3\ell, 3\ell + 1)$  model,  $\ell \geq 2$ . Again, we code over 3 channel uses. By network decomposition, any code for the  $(9\ell, 9\ell + 3)$  model can be implemented over 3 channel uses of the  $(3\ell, 3\ell + 1)$  model. But for the  $(9\ell, 9\ell + 3)$  model, we *split* the network into multiple parts: On the first 9 vertices of each transmitter and receiver, we implement the  $(6, 9)$  code from above,

<sup>1</sup>In the solution for the  $(2, 3)$  model, the symbols  $(a_1, b_1)$  at receiver 2 share one-dimensional linear subspace spanned by  $(0, 1, 0)$ . In other words, the linear subspace with respect to  $a_1$  is *aligned* with the subspace w.r.t  $b_1$ . In this sense, it is an instance of the important concept of *interference alignment* [18], [19] which has shown the great potential for a variety of applications [19]–[26]. But the distinction w.r.t our problem comes from the purpose of alignment. In our problem, the aim of alignment is to form a desired function while minimizing the signal subspace occupied by the source symbols.

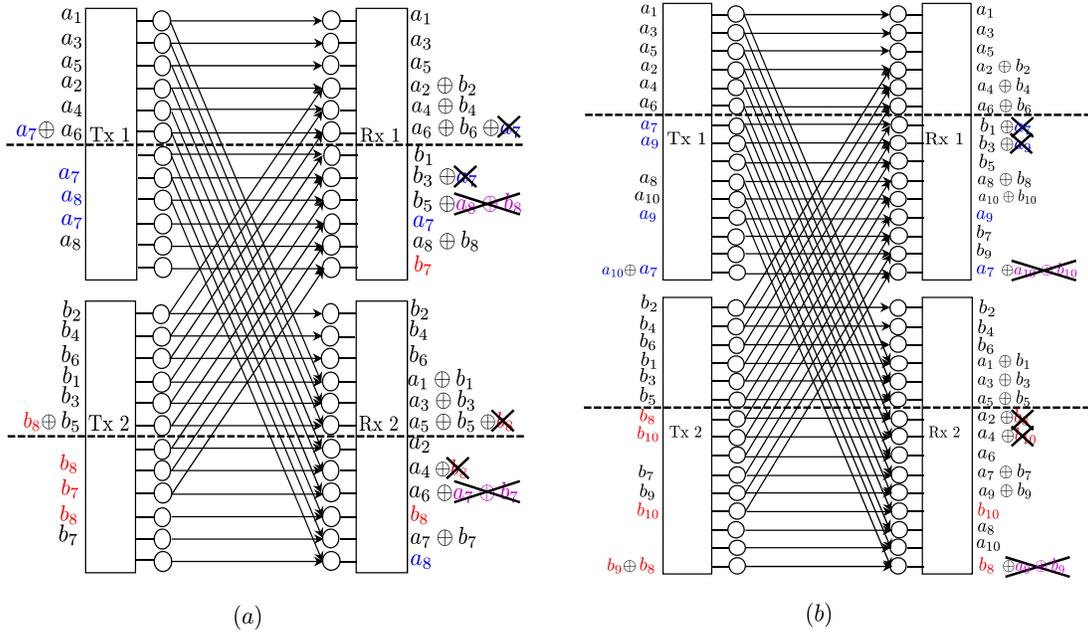


Fig. 4. Explicit code for (a) the (9, 12) model and (b) the (12, 15) model.

giving 6 computations. But since in that code, neither transmitter uses the last 3 inputs, this leaves the remaining  $9\ell + 3 - 9 = 9(\ell - 1) + 3$  vertices of the network completely unaffected at all transmitters and receivers. Hence, we can repeat this step: on the next 9 vertices, again implement the (6, 9) code from above, giving another 6 computations. We do this step exactly  $\ell - 1$  times, leading to  $6(\ell - 1)$  computations. At that point, we are left with  $9\ell + 3 - 9(\ell - 1) = 12$  vertices at each transmitter and each receiver. On these, we implement the code from Fig. 4(a), giving us another 8 computations. This gives a total of  $6(\ell - 1) + 8 = 2(3\ell + 1)$  computations. Thus, per channel use, the computation rate is  $\frac{2}{3}(3\ell + 1)$ , as claimed.

(iii) *The Case  $r = 3\ell + 2$ , i.e., the  $(3\ell + 1, 3\ell + 2)$  Model:* Consider the (4, 5) model. Here we consider a code for the (12, 15) model (which can be implemented over three channel uses of the (4, 5) model). The code is given in Fig. 4(b). This code enables both receivers to recover all 10 computations, which means a computation rate of  $10/3 = \frac{2}{3} \cdot 5$  per channel use, as claimed in the lemma.

Now, consider the general  $(3\ell + 1, 3\ell + 2)$  model, for any positive integer  $\ell \geq 2$ . Again, we code over 3 channel uses. By network decomposition, any code for the  $(9\ell + 3, 9\ell + 6)$  model can be implemented over 3 channel uses of the  $(3\ell + 1, 3\ell + 2)$  model. But for the  $(9\ell + 3, 9\ell + 6)$  model, we *split* the network into multiple parts: On the first 9 vertices of each transmitter and receiver, we implement the (6, 9) code from above, giving 6 computations. But since in that code, neither transmitter uses the last 3 inputs, this leaves the remaining  $9\ell + 6 - 9 = 9(\ell - 1) + 6$  vertices of the network completely unaffected at all transmitters and receivers. Hence, we can repeat this step: on the next 9 vertices, again implement the (6, 9) code from above, giving another 6 computations. We do this step exactly  $\ell - 1$  times, leading to  $6(\ell - 1)$  computations.

At that point, we are left with  $9\ell + 6 - 9(\ell - 1) = 15$  vertices at each transmitter and each receiver. On these, we implement the code from Fig. 4(b), giving us another 10 computations. This gives a total of  $6(\ell - 1) + 10 = 2(3\ell + 2)$  computations. Thus, per channel use, the computation rate is  $\frac{2}{3}(3\ell + 2)$ .

## VI. PROOF OF THEOREM 4

### A. Achievability Proof

The idea is to combine the network decomposition in Theorem 5 and achievability proof for elementary subnetworks.

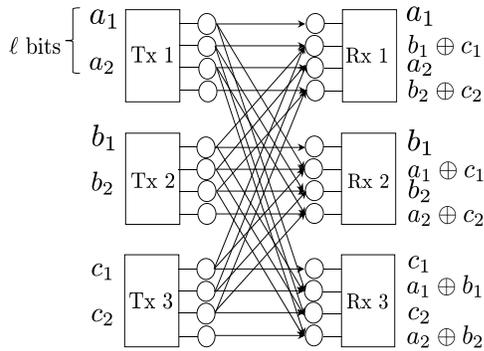
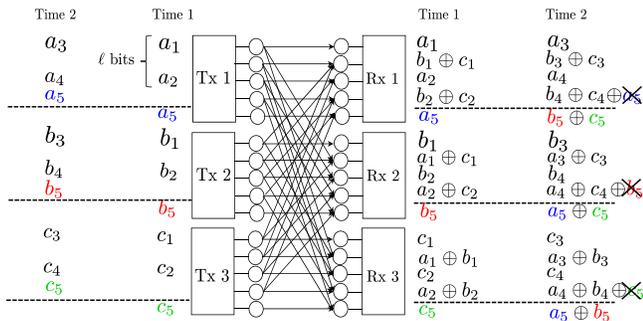
*Lemma 3 ( $L \geq 3$ ):* The following computation rates are achievable:

- (1) For the model (0, 1) or (1, 0),  $R_{\text{Comp}} \geq 0$ .
- (2a) For the model  $(r - 1, r)$  with  $r \geq 2$ ,  $R_{\text{Comp}} \geq \frac{1}{2}r$ .
- (2b) For the model  $(r, r - 1)$  with  $r \geq 2$ ,  $R_{\text{Comp}} \geq \frac{1}{2}r$ .
- (3) For the model  $(r, r)$ ,  $R_{\text{Comp}} \geq r$ .

*Proof:* The items (1) and (3) are straightforward.

(i)  *$(r - 1, r)$  Model:* We consider two cases:  $r = 2\ell$  and  $r = 2\ell + 1$ . Fig. 5 shows an achievable scheme when  $r = 2\ell = 2 \cdot 2$  and  $L = 3$ . Each transmitter uses odd-numbered vertices to send  $\ell$  symbols. The special structure of symmetric networks enables each receiver to get clean symbols on odd-numbered vertices while receiving partially-aligned functions on even-numbered ones. For example, receiver 1 gets  $(a_1, a_2)$  on the first and third vertices;  $(b_1 \oplus c_1, b_2 \oplus c_2)$  on the second and fourth vertices. Note that two resource levels are consumed to compute one desired function. Therefore, this gives a computation rate of  $\frac{1}{2}r$ . Obviously this can be applied to an arbitrary value of  $L$ .

An explicit scheme for the case of  $r = 2\ell + 1 = 2 \cdot 2 + 1$  and  $L = 3$  is given in Fig. 6. If we followed the same approach

Fig. 5. Achievable scheme for the  $(2\ell - 1, 2\ell)$  model.Fig. 6. Achievable scheme for the  $(2\ell, 2\ell + 1)$  model.

as in the case of  $r = 2\ell$ , each receiver would be left with one empty vertex in the last level. In this example, receiver 1 would get  $(a_1, b_1 \oplus c_1, a_2, b_2 \oplus c_2)$  on the 1st, 2nd, 3rd, and 4th vertices, while the last bottom vertex is empty. In order to make an efficient resource utilization, we invoke an idea of vector coding. First we implement the same code in time 2. Now each transmitter will exploit the two empty vertices to provide one more computation. In time 1, on the last vertex, each transmitter sends a new fresh symbol. In this example, transmitters 1, 2, 3 send  $a_5, b_5, c_5$  respectively. However, this transmission does not give the computation of  $a_5 \oplus b_5 \oplus c_5$  yet. In time 2, on the second last vertex, each transmitter re-sends the symbol that was sent on the last vertex in time 1. Receiver 1 can then get  $b_5 \oplus c_5$ , thus achieving  $a_5 \oplus b_5 \oplus c_5$ . While the transmission of  $a_5$  causes interference to  $b_4 \oplus c_4$  on the second last vertex, we can resolve this conflict by using the  $a_5$  (that was already received in time 1) as side information. Similarly the desired function can be computed at the other receivers. This way, we can achieve  $\frac{2\ell+1}{2} = \frac{r}{2}$ . The same strategy can be applied to an arbitrary value of  $L$ .

(ii)  $(r, r - 1)$  Model: Consider two cases:  $r = 2\ell$  and  $r = 2\ell + 1$ . For both cases, the coding strategies are the same as those in the  $(r - 1, r)$  model. In the first case, each transmitter sends  $\ell$  symbols only on odd-numbered vertices. One can then readily see that all the receivers can compute  $\ell$  sums. In the second case, we use two time slots. In each time, on the upper  $2\ell$  vertices of transmitters and receivers, we implement the  $(2\ell, 2\ell - 1)$  code from above. On top of this, each transmitter sends one additional fresh symbol to provide one more computation. Specifically each transmitter sends the

$(\ell + 1)$ th symbol on the last vertex in time 1 and re-sends the symbol on the second last vertex in time 2. One can check that each receiver can compute  $2\ell + 1$  sums, thus yielding a computation rate of  $\frac{r}{2}$ . ■

Using Theorem 5 and Lemma 3, we can now prove the achievability. For the case of  $\alpha = 1$ ,  $R_{\text{comp}} \geq n$  by Item (3) in Lemma 3. For  $0 \leq \alpha < \frac{1}{2}$ , (6) gives  $r = 0$  and  $a = m$ , thus the decomposition is given by  $(m, n) \rightarrow (0, 1)^{n-2m} \times (1, 2)^m$ . Therefore, using Lemma 3,  $R_{\text{comp}} \geq 0 \cdot (n - 2m) + 1 \cdot m = m$ . Next, consider the case of  $\frac{1}{2} \leq \alpha < \frac{2}{3}$ . Using (6), we find that  $r = 1$  and  $a = 2m - n$ , hence, the decomposition is given by  $(m, n) \rightarrow (1, 2)^{2n-3m} \times (2, 3)^{2m-n}$ . Using Lemma 3,  $R_{\text{comp}} \geq 1 \cdot (2n - 3m) + \frac{3}{2} \cdot (2m - n) = \frac{1}{2}n$ . Finally, consider the case of  $\alpha \geq \frac{2}{3}$ . From (6), we know that  $r \geq 2$ . So we get  $R_{\text{comp}} \geq \frac{1}{2}n$ .

The other case of  $m > n$  similarly follows. For  $0 \leq \alpha < \frac{1}{2}$ , (6) gives  $r = 0$  and  $a = n$ , thus the decomposition is given by  $(m, n) \rightarrow (1, 0)^{m-2n} \times (2, 1)^n$ . With Lemma 3, we get  $R_{\text{comp}} \geq 0 \cdot (m - 2n) + 1 \cdot n = n$ . For  $\frac{1}{2} \leq \alpha < \frac{2}{3}$ , (6) gives  $r = 1$  and  $a = 2n - m$ . So the decomposition is given by  $(m, n) \rightarrow (2, 1)^{2m-3n} \times (3, 2)^{2n-m}$ . Using Lemma 3, we can achieve  $R_{\text{comp}} \geq 1 \cdot (2m - 3n) + \frac{3}{2} \cdot (2n - m) = \frac{1}{2}m$ . Lastly, for  $\alpha \geq \frac{2}{3}$ , (6) gives  $r \geq 2$ . So we get  $R_{\text{comp}} \geq \frac{1}{2}m$ .

### B. Proof of Upper Bound

The upper bound is a generalized version of the  $2 \times 2$  case bound (2). Hence, the bounding technique is almost the same, but it involves complicated notations. Let  $\mathcal{Y} := \{Y_\ell^N\}_{\ell=1}^L$  be a collection of all the received signals. Let  $S := \bigoplus_{\ell=1}^L S_\ell^K$  be the  $K$ -dimensional desired sum function. Let  $\bar{S}_\ell := \{S_j^K\}_{j=1}^L \setminus S_\ell^K$  be a collection of all the sources that excludes the  $\ell$ th user's source. Using these notations and starting with Fano's inequality, we get:

$$\begin{aligned}
 & N((\log_2 p)(2L - 1)R_{\text{comp}} - \epsilon_N) \\
 & \leq \sum_{\ell=1}^L I(S; Y_\ell^N) + (L - 1)I(S; Y_1^N) \\
 & \stackrel{(a)}{\leq} \sum_{\ell=1}^L \left[ H(Y_\ell^N) - H(Y_\ell^N | S, \{Y_j^N\}_{j=1}^{\ell-1}) \right] + (L - 1)I(S; Y_1^N) \\
 & \stackrel{(b)}{\leq} \sum_{\ell=1}^L H(Y_\ell^N) - H(\mathcal{Y} | S) + \sum_{\ell=2}^L I(S; Y_1^N, \bar{S}_\ell) \\
 & \stackrel{(c)}{=} \sum_{\ell=1}^L H(Y_\ell^N) - H(\mathcal{Y} | S) + \sum_{\ell=2}^L H(Y_1^N | \bar{S}_\ell) \\
 & \stackrel{(d)}{=} \sum_{\ell=1}^L H(Y_\ell^N) - H(\mathcal{Y} | S) + \sum_{\ell=2}^L H(\{\mathbf{G}^{q-n\ell} X_{\ell i}\}_{i=1}^N | \bar{S}_\ell) \\
 & \stackrel{(e)}{\leq} \sum_{\ell=1}^L H(Y_\ell^N) \leq (\log_2 p)NL \max(m, n)
 \end{aligned}$$

where (a) follows from the fact that conditioning reduces entropy; (b) follows from a chain rule and the non-negativity of mutual information; (c) follows from the fact that  $\bar{S}_\ell$  is independent of  $S$  due to the mutual independence of  $S_\ell^K$ 's;

(d) follows from the fact that  $X_{\ell i}$  is a function of  $\bar{S}_\ell \forall \ell' \neq \ell$ ; and (e) follows from  $H(\mathcal{Y}|S) \geq \sum_{\ell=2}^L H(\{\mathbf{G}^{q-n\ell} X_{\ell i}\}_{i=1}^N | \bar{S}_\ell)$  (see Claim 2 below). Hence,  $R_{\text{comp}} \leq \frac{L}{2L-1} \max(m, n)$ .

*Claim 2:*

$$H(\mathcal{Y}|S) \geq \sum_{\ell=2}^L H(\{\mathbf{G}^{q-n\ell} X_{\ell i}\}_{i=1}^N | \bar{S}_\ell). \quad (8)$$

*Proof:*

$$\begin{aligned} H(\mathcal{Y}|S) &- \sum_{\ell=2}^L H(\{\mathbf{G}^{q-n\ell} X_{\ell i}\}_{i=1}^N | \bar{S}_\ell) \\ &\stackrel{(a)}{=} H(\mathcal{Y}, \{\mathbf{G}^{q-n2\ell} X_{2i}\}_{i=1}^N, \dots, \{\mathbf{G}^{q-nL} X_{Li}\}_{i=1}^N | S) \\ &\quad - \sum_{\ell=2}^L H(\{\mathbf{G}^{q-n\ell} X_{\ell i}\}_{i=1}^N | \bar{S}_\ell) \\ &\stackrel{(b)}{\geq} H(\{\mathbf{G}^{q-n2\ell} X_{2i}\}_{i=1}^N, \dots, \{\mathbf{G}^{q-nL} X_{Li}\}_{i=1}^N | S) \\ &\quad - \sum_{\ell=2}^L H(\{\mathbf{G}^{q-n\ell} X_{\ell i}\}_{i=1}^N | \bar{S}_\ell) \\ &\stackrel{(c)}{=} \sum_{\ell=2}^L H(\{\mathbf{G}^{q-n\ell} X_{\ell i}\}_{i=1}^N) - \sum_{\ell=2}^L H(\{\mathbf{G}^{q-n\ell} X_{\ell i}\}_{i=1}^N | \bar{S}_\ell) \\ &\geq 0 \end{aligned}$$

where (a) follows from the fact that  $X_{\ell i}$  is a function of  $\{Y_j\}_{j=1}^L, \forall \ell, i$  (see Claim 3 below); (b) follows from the non-negativity of the entropy and the fact that  $(S_2^K, \dots, S_L^K)$  is independent of  $S$ ; (c) follows from a chain rule, and the fact that  $X_{\ell i}$  is independent of  $S$  and  $(X_{(\ell+1)i}, \dots, X_{Li})$ 's due to the mutual independence of  $S_\ell^K$ 's. ■

*Claim 3:* For  $m \neq n$ ,  $X_\ell$  is a function of  $\{Y_j\}_{j=1}^L, \ell = 1, \dots, L$ .

*Proof:* By symmetry, it suffices to consider the case of  $\ell = 1$ . Consider the case of  $m < n$ . From (1), we get

$$\begin{aligned} \sum_{j=1}^L Y_j &= \{\mathbf{I} + (L-1)\mathbf{G}^{n-m}\} \sum_{j=1}^L X_j; \\ \sum_{j=2}^L Y_j &= \{\mathbf{I} + (L-1)\mathbf{G}^{n-m}\} \sum_{j=1}^L X_j \\ &\quad - \left( X_1 + \mathbf{G}^{n-m} \sum_{j=2}^L X_j \right) \\ &= (L-1)\mathbf{G}^{n-m} \sum_{j=1}^L X_j + \{\mathbf{I} - \mathbf{G}^{n-m}\} \sum_{j=2}^L X_j. \end{aligned}$$

The above two equations can be written as follows respectively:

$$\begin{aligned} \sum_{j=1}^L X_j &= \mathbf{A}^{-1} \sum_{j=1}^L Y_j \\ \sum_{j=2}^L X_j &= \mathbf{B}^{-1} \left[ \sum_{j=2}^L Y_j - (L-1)\mathbf{G}^{n-m} \left\{ \mathbf{A}^{-1} \sum_{j=1}^L Y_j \right\} \right] \end{aligned}$$

where  $\mathbf{A} := \mathbf{I} + (L-1)\mathbf{G}^{n-m}$  and  $\mathbf{B} := \mathbf{I} - \mathbf{G}^{n-m}$ . Note that when  $m \neq n$ , both of  $\mathbf{A}$  and  $\mathbf{B}$  are invertible. So  $X_1$  can be computed from the above two equations, and thus  $X_1$  is a function of  $\{Y_j\}_{j=1}^L$ . Similarly we can show this for the case of  $m > n$ . Notice that this proof holds for the general  $p$ -ary model. ■

## VII. DISCUSSION

### A. On the Optimality of Network Decomposition

A recent work in [27] has shown the optimality of our network-decomposition-based approach for two other problem settings under the 2-user ADT symmetric network: (1) the two-unicast problem in which each receiver wishes to decode a message from its corresponding transmitter; (2) the classical multicast problem in which each receiver wants to decode every message. It would be interesting to explore the optimality of the decomposition approach for more general problem settings.

### B. Asymmetric Networks

For the general asymmetric network, the computation capacity is not characterized even for the  $L = 2$  case. This is mainly because we found it difficult to develop a network decomposition theorem for the general asymmetric network. Moreover, we conjecture that our upper bound may be loose for some asymmetric network. This conjecture comes from our inspection on the example of  $(n_{12}, n_{11}; n_{21}, n_{22}) = (3, 3; 4, 5)$ . In this example, our upper bound gives  $C_{\text{comp}} \leq \frac{\max(n_{11}+n_{21})+\max(n_{22}, n_{12})}{3} = 3$ . On the other hand, when applying all of the achievability techniques developed in Theorem 3, we could achieve only a computation rate of  $\frac{8}{3}$ . It is expected that both of new inner and upper bounds are needed for the asymmetric network.

### C. A Class of Linear Deterministic Networks

As mentioned earlier, although the ADT network setting that we considered is somewhat specialized, the results in Theorems 1, 2, 3, and 4 apply to general  $p$ -ary models. Furthermore, we expect that the results can be useful for a more general class of linear deterministic networks in which even the channel transfer matrices are of an arbitrary form. In particular, the upper-bound technique can be readily applied to the general setting, although it may not guarantee the optimality. As for achievability, it would be an interesting future work to discover elementary subnetworks (if any) that can constitute an original network without loss of optimality.

### D. Multi-Hop Networks

In [11], [16], and [17], function multicasting has been explored in the context of multi-hop networks. In particular, Rai and Dey in [11] found some interesting equivalence relationship between sum-networks and multiple-unicast networks. Due the relationship, it has been believed that the sum-network problem is as hard as the multiple-unicast problem, and indeed the computation capacity of the sum-network has been open. For two-source  $L$ -destination or  $L$ -source two-destination networks, the computation capacity

was established only when the entropy of each source is constrained to be 1 [17].

One natural next step is exploiting the insights developed in this work, to characterize necessary and sufficient conditions of two-source two-destination multi-hop networks when the entropy of each source is limited by 2.

### E. Role of Feedback for Computation

The role of feedback for computation has initially been studied in [28] where it is shown that feedback can increase the computation rate. Interestingly the feedback gain is shown to be significant - qualitatively similar to the gain in the two-user Gaussian interference channel [29]. However, the result of [28] relies on a separation approach that naturally comes in the course of characterizing the feedback multicast capacity. Recently [30] has established the exact feedback computation capacity of the 2-by-2 ADT network considered herein to show that the feedback gain can be more significant. It would be interesting to explore this feedback gain under more realistic scenarios where feedback is offered through rate-limited bit-piped links [31] or through the corresponding backward communication network [32].

## VIII. CONCLUSION

We have established the computation capacity of a two-transmitter two-receiver ADT symmetric network where each receiver wishes to compute a modulo-2-sum function of two Bernoulli sources generated at the two transmitters. For the  $L$ -user case, we also characterized the computation capacity in the limit of  $L$ . In the process of obtaining these results, we derived new upper bounds and established a network decomposition theorem that provides a conceptually-simple achievability proof. We expect that the network-decomposition-based framework would give insights into solving more general network problems.

### APPENDIX A

#### PROOF OF LEMMA 1

Consider the following cases: (1) ( $n_{12} \leq n_{11}, n_{21} \leq n_{22}$ ); (2) ( $n_{12} \geq n_{11}, n_{21} \geq n_{22}$ ); (3) ( $n_{12} \leq n_{11}, n_{21} \geq n_{22}$ ); and (4) ( $n_{12} \geq n_{11}, n_{21} \leq n_{22}$ ). Note that Case (1) and Case (2) are symmetric, so are Case (3) and Case (4). Hence, only for Case (1) and (3), we will show that there exists  $(i, j)$  such that  $\mathbf{G}^{q-n_{ij}} X_i$  can be reconstructed from the pair  $(Y_1, Y_2)$ .

*Case (1) ( $n_{12} \leq n_{11}, n_{21} \leq n_{22}$ ):* Since  $n_{11} \geq n_{12}$ , there is no reason to send information on the last  $q - n_{11}$  levels in  $X_1$ . Note that any signal on the levels would be wiped out in  $\mathbf{G}^{q-n_{11}} X_1$ . Hence we focus on the case in which  $X_1 = [\tilde{X}_1; \mathbf{0}_{q-n_{11}}]$ . Now let us consider

$$\begin{aligned} Y_1 \oplus \mathbf{G}^{n_{22}-n_{21}} Y_2 &= (\mathbf{G}^{q-n_{11}} \oplus \mathbf{G}^{q+n_{22}-n_{21}-n_{12}}) X_1 \\ &= \left[ \begin{array}{c} \mathbf{0}_{q-n_{11}} \\ \left\{ \mathbf{I}_{n_{11}} \oplus \mathbf{G}_{n_{11}}^{n_{11}-n_{12}-(n_{21}-n_{22})} \right\} \tilde{X}_1 \end{array} \right] \end{aligned}$$

where the second equality follows from  $X_1 = [\tilde{X}_1; \mathbf{0}_{q-n_{11}}]$ . Here  $\mathbf{G}_{n_{11}}$  denotes the  $n_{11}$ -by- $n_{11}$  shift matrix. Since  $n_{11} - n_{12} \neq n_{21} - n_{22}$ ,  $\mathbf{G}_{n_{11}}^{n_{11}-n_{12}-(n_{21}-n_{22})} \neq \mathbf{I}_{n_{11}}$  and therefore

$\mathbf{I}_{n_{11}} \oplus \mathbf{G}_{n_{11}}^{n_{11}-n_{12}-(n_{21}-n_{22})}$  is invertible, implying that  $\tilde{X}_1$  can be reconstructed from  $(Y_1, Y_2)$ . Hence,  $\mathbf{G}^{q-n_{11}} X_1$  can also be reconstructed from  $(Y_1, Y_2)$ .

*Case (3) ( $n_{12} \leq n_{11}, n_{21} \geq n_{22}$ ):* First consider the case of  $n_{21} - n_{22} > n_{11} - n_{12}$ . In this case, we get:

$$\begin{aligned} \mathbf{G}^{n_{21}-n_{22}} Y_1 \oplus Y_2 &= (\mathbf{G}^{q-n_{11}+n_{21}-n_{22}} \oplus \mathbf{G}^{q-n_{12}}) X_1 \\ &= \left[ \begin{array}{c} \mathbf{0}_{q-n_{11}} \\ \left( \mathbf{G}_{n_{11}}^{n_{21}-n_{22}} \oplus \mathbf{G}_{n_{11}}^{n_{11}-n_{12}} \right) \tilde{X}_1 \end{array} \right]. \end{aligned}$$

Since  $n_{11} - n_{12} \neq n_{21} - n_{22}$ ,  $\mathbf{G}_{n_{11}}^{n_{21}-n_{22}} \oplus \mathbf{G}_{n_{11}}^{n_{11}-n_{12}} \neq \mathbf{0}$ . So  $\mathbf{G}_{n_{11}}^{n_{21}-n_{22}} \tilde{X}_1$  can be reconstructed from  $(Y_1, Y_2)$ . Therefore,  $\mathbf{G}^{q-n_{12}} X_1$  can also be reconstructed from  $(Y_1, Y_2)$ . For the other case  $n_{21} - n_{22} < n_{11} - n_{12}$ , we get:

$$\begin{aligned} \mathbf{G}^{n_{11}-n_{12}} Y_1 \oplus Y_2 &= (\mathbf{G}^{q-n_{21}+n_{11}-n_{12}} \oplus \mathbf{G}^{q-n_{22}}) X_2 \\ &= \left[ \begin{array}{c} \mathbf{0}_{q-n_{21}} \\ \left( \mathbf{G}_{n_{21}}^{n_{11}-n_{12}} \oplus \mathbf{G}_{n_{21}}^{n_{21}-n_{22}} \right) \tilde{X}_2 \end{array} \right]. \end{aligned}$$

Since  $n_{11} - n_{12} \neq n_{21} - n_{22}$ ,  $\mathbf{G}_{n_{21}}^{n_{11}-n_{12}} \oplus \mathbf{G}_{n_{21}}^{n_{21}-n_{22}} \neq \mathbf{0}$ . So  $\mathbf{G}_{n_{21}}^{n_{11}-n_{12}} \tilde{X}_2$  can be reconstructed from  $(Y_1, Y_2)$ . Therefore  $\mathbf{G}^{q-n_{22}} X_2$  can also be reconstructed from  $(Y_1, Y_2)$ .

## APPENDIX B

### PROOF OF THEOREM 5

The  $L$ -user  $(m, n)$  network studied in this paper is naturally represented by a bi-partite graph as in Figure 1. Each transmitter and each receiver is represented by  $\max(n, m)$  vertices. For the argument developed here, we consider the undirected version of this graph. First, we attach labels to each vertex. In particular, at each transmitter and each receiver separately, we label the vertices by the integers  $\{0, 1, \dots, \max(n, m) - 1\}$  from top to bottom. The following observation is key:

*Lemma 4:* In an  $L$ -user  $(m, n)$  model, for any two vertices, if their respective labels  $u$  and  $v$  satisfy  $(u - v) \bmod |n - m| \neq 0$ , then there is no path between them.

*Proof:* Consider any edge in the graph representation of our network (as illustrated e.g. in Figure 1). Suppose that the two vertices connected by this edge have labels  $a$  and  $b$ . Then, the structure of the graph implies that we must have  $(a - b) \bmod |n - m| = 0$ . Now, suppose that there is a path between two arbitrary vertices with labels  $u_1$  and  $u_\ell$ , passing through vertices with labels  $u_2, u_3, \dots, u_{\ell-1}$ . Furthermore, suppose that  $(u_1 - u_\ell) \bmod |n - m| \neq 0$ . This implies that there must exist  $i$  ( $1 \leq i < \ell$ ) such that  $(u_i - u_{i+1}) \bmod |n - m| \neq 0$ . But there cannot be an edge between these two vertices, and thus, the claimed path cannot exist. This proves the lemma. ■

We now proceed to the proof of Theorem 5. We introduce the following two terms to have more compact arguments. The term *direct edges* refers to the edges connecting a transmitter with its corresponding receiver. The term *cross edges* refers to the edges connecting a transmitter with all receivers other than its corresponding receiver. Owing to the symmetry of the considered models, cross edges are the same for each receiver.

For Part (1), consider the  $(km, kn)$  model. The proof uses graph coloring with  $k$  colors, identified by integers  $\{0, 1, \dots, k - 1\}$ . A vertex with label  $u$  (where  $u = 0, 1, 2, \dots, k \max(m, n) - 1$ ) receives color  $u \bmod k$ .

Hence, if two nodes (with labels  $u$  and  $v$ ) have different colors, this means that  $(u - v) \bmod k \neq 0$ . But this also implies that  $(u - v) \bmod k|m - n| \neq 0$ . Then, Lemma 4 implies that there cannot be a path between these vertices. Hence, each color represents an independent graph. Since we used  $k$  colors, we thus have exactly  $k$  independent graphs. By the symmetry of the construction, all vertices and edges must be equally distributed, and hence, each graph represents precisely an  $(m, n)$  model. Alternatively, we now offer an explicit proof of this fact for the case  $m < n$ . The case  $m > n$  follows along the same lines. Let us first observe that in the  $(km, kn)$  model (with  $m < n$ ), there are direct edges between any transmitter vertex with label  $u$  and its corresponding receiver vertex with the same label  $u$ . Moreover, there are cross edges from transmitter vertex with label  $u$  to receiver vertex with label  $v$  if and only if  $v - u = k(n - m)$ . Now, consider the subgraph whose vertices are colored with the color identified by the integer 0. This subgraph thus contains all vertices with labels  $0, k, 2k, \dots$ . That is, in this case, we can express all vertex labels as  $\ell k$ . Clearly, there are direct edges between transmitter vertex with label  $\ell k$  and receiver vertex with the same label  $\ell k$ . Moreover, there is a cross edge between transmitter vertex with label  $\ell_1 k$  and receiver vertex with label  $\ell_2 k$  if and only if  $\ell_2 k - \ell_1 k = k(n - m)$ , which can be rewritten as  $\ell_2 - \ell_1 = n - m$ . In other words, in our subgraph, if we assign new labels to each vertex by dividing the existing label by  $k$ , meaning that a vertex previously labeled as  $\ell k$  is now simply labeled as  $\ell$ , we observe that each transmitter and each receiver has exactly  $n$  vertices, that there are direct edges between vertices with the same labels, and that there is a cross edge from transmitter vertex with label  $\ell_1$  to receiver vertex with label  $\ell_2$  if and only if  $\ell_2 - \ell_1 = n - m$ . But this is precisely an  $(m, n)$  model (with  $m < n$ ). The same argument can be applied to the remaining  $k - 1$  colors.

For Part (2), we consider the model  $(2m + 1, 2n + 1)$ . Again, we proceed by graph coloring, this time with two colors: All vertices whose label  $u$  (where  $u = 0, 1, 2, \dots, 2 \max(m, n)$ ) is an even number receive one color, and all vertices whose label  $u$  is an odd number receive the other color. Hence, if two nodes (with labels  $u$  and  $v$ ) have different colors, this means that  $(u - v) \bmod 2 \neq 0$ . But this also implies that  $(u - v) \bmod 2|n - m| \neq 0$ . Then, Lemma 4 implies that there cannot be a path between these vertices. Hence, each of the two colors represents an independent graph. The remainder of the proof is to establish the structure of these two subgraphs. We will do so for the case  $m < n$ ; the case  $m > n$  follows along the same lines. First, we observe that in the  $(2m + 1, 2n + 1)$  model (with  $m < n$ ), there is a direct edge between transmitter vertex with label  $u$  and receiver vertex with the same label  $u$ . Moreover, there is a cross edges from transmitter vertex with label  $u$  to receiver vertex with label  $v$  if and only if  $v - u = 2n + 1 - (2m + 1) = 2(n - m)$ . Now, consider the subgraph corresponding to the odd-labeled vertices. That is, in this case, we can express each label in the form  $2\ell + 1$ . We start by observing that each transmitter and each receiver has exactly  $n$  odd-labeled vertices. Moreover, there are direct edges between transmitter vertex with label  $2\ell + 1$  and receiver vertex with the same label  $2\ell + 1$ .

Additionally, there is a cross edge from transmitter vertex with label  $2\ell_1 + 1$  to receiver vertex with label  $2\ell_2 + 1$  if and only if  $2\ell_2 + 1 - (2\ell_1 + 1) = 2(n - m)$ , which can be rewritten as  $\ell_2 - \ell_1 = n - m$ . Relabeling the vertices of every transmitter and every receiver by subtracting 1 from the existing label and then dividing by two (such that the vertex previously labeled as  $2\ell + 1$  is now labeled as  $\ell$ ), each transmitter and each receiver has vertices labeled  $0, 1, \dots, n - 1$ , there are direct edges, and there is a cross edge between transmitter vertex  $u$  and receiver vertex  $v$  if and only if  $v - u = n - m$ . Hence, this is an  $(m, n)$  model (with  $m < n$ ). Next, consider the subgraph corresponding to the even-labeled vertices. That is, in this case, we can express each label in the form  $2\ell$ . We start by observing that each transmitter and each receiver has exactly  $n + 1$  even-labeled vertices. Moreover, there are direct edges between transmitter vertex with label  $2\ell$  and receiver vertex with the same label  $2\ell$ . Additionally, there is a cross edge from transmitter vertex with label  $2\ell_1$  to receiver vertex with label  $2\ell_2$  if and only if  $2\ell_2 - 2\ell_1 = 2(n - m)$ , which can be rewritten as  $\ell_2 - \ell_1 = n - m$ . Relabeling the vertices of every transmitter and every receiver by dividing by two (such that the vertex previously labeled as  $2\ell$  is now labeled as  $\ell$ ), each transmitter and each receiver has vertices labeled  $0, 1, \dots, n$ , there are direct edges, and there is a cross edge between transmitter vertex  $u$  and receiver vertex  $v$  if and only if  $v - u = n - m$ . Hence, this is an  $(m + 1, n + 1)$  model (with  $m < n$ ).

For Part (3), we use graph coloring with  $|n - m|$  colors, identified by integers  $\{0, 1, \dots, |n - m| - 1\}$ . A vertex with label  $u$  (where  $u = 0, 1, 2, \dots, \max(m, n) - 1$ ) receives color  $u \bmod |n - m|$ . Hence, if two nodes (with labels  $u$  and  $v$ ) have different colors, this means that  $(u - v) \bmod |n - m| \neq 0$ . Then, Lemma 4 implies that there cannot be a path between these vertices. Hence, each color represents an independent graph. But since there are  $|n - m|$  colors, this means that there are  $|n - m|$  independent subgraphs.

The remainder of the proof is to establish the structure of these subgraphs. For convenience, we define  $r = \lfloor \frac{\min\{m, n\}}{|n - m|} \rfloor$  and  $a = \min\{m, n\} \bmod |n - m|$ .

We explicitly consider the case  $m < n$ . Note that in this case, there is a direct edge from transmitter vertex with label  $u$  to receiver vertex with the same label  $u$ . Moreover, there is a cross edge from transmitter vertex with label  $u$  to receiver vertex with label  $v$  if and only if  $v - u = n - m$ . Also note that in this case, we have  $a = m \bmod (n - m) = n \bmod (n - m)$ , since for all integers  $n$  and  $m$ , we have  $m \bmod (n - m) = n \bmod (n - m)$ .

First, let us consider the color 0. In the corresponding subgraph, each transmitter and each receiver contains all vertices with labels  $0, n - m, 2(n - m), \dots$ . If  $a \geq 1$ , that is,  $m \bmod (n - m) \geq 1$  or equivalently,  $n \bmod (n - m) \geq 1$ , then there are  $\lceil \frac{n}{n - m} \rceil = \lceil 1 + \frac{m}{n - m} \rceil = r + 2$  such vertices. Next, consider the edges in this subgraph. Clearly, there is a direct edge from transmitter vertex with label  $\ell(n - m)$  to receiver vertex with the same label  $\ell(n - m)$ . Moreover, there is a cross edge from transmitter vertex with label  $\ell_1(n - m)$  to receiver vertex with label  $\ell_2(n - m)$  if and only if  $\ell_2(n - m) - \ell_1(n - m) = n - m$ , which can be rewritten as  $\ell_2 - \ell_1 = 1$ .

Finally, assigning new vertex labels in the subgraph by dividing the existing label by  $(n - m)$  reveals that the subgraph is exactly an  $(r + 1, r + 2)$  model. Otherwise, if  $a = 0$ , that is,  $m \bmod (n - m) = 0$  or equivalently,  $n \bmod (n - m) = 0$  (implying that both  $m$  and  $n$  are divisible by  $(n - m)$ ), then there are only  $\frac{n}{n-m} = 1 + \frac{m}{n-m} = r + 1$  vertices in each transmitter and receiver. Again, clearly, there are direct edges between each transmitter-receiver pair. Moreover, there is a cross edge from transmitter vertex with label  $\ell_1(n - m)$  to receiver vertex with label  $\ell_2(n - m)$  if and only if  $\ell_2(n - m) - \ell_1(n - m) = n - m$ , which can be rewritten as  $\ell_2 - \ell_1 = 1$ . Finally, assigning new vertex labels in the subgraph by dividing the existing labels by  $(n - m)$  reveals that the subgraph is exactly an  $(r, r + 1)$  model.

Now, let us proceed to color  $j$ , where  $j \in \{1, 2, 3, n - m - 1\}$ . In the corresponding subgraph, each transmitter and each receiver contains all vertices with labels  $j, n - m + j, 2(n - m) + j, \dots$ . If  $a \geq j + 1$ , that is,  $m \bmod (n - m) \geq j + 1$  or equivalently,  $n \bmod (n - m) \geq j + 1$  then there are  $\lceil \frac{n}{n-m} \rceil = \lceil 1 + \frac{m}{n-m} \rceil = r + 2$  such vertices. Next, consider the edges in this subgraph. Clearly, there is a direct edge from transmitter vertex with label  $\ell(n - m) + j$  to receiver vertex with the same label  $\ell(n - m) + j$ . Moreover, there is a cross edge from transmitter vertex with label  $\ell_1(n - m) + j$  to receiver vertex with label  $\ell_2(n - m) + j$  if and only if  $\ell_2(n - m) + j - (\ell_1(n - m) + j) = n - m$ , which can be rewritten as  $\ell_2 - \ell_1 = 1$ . Finally, assigning new vertex labels in the subgraph by subtracting  $j$  from the existing label and dividing by  $(n - m)$  reveals that the subgraph is exactly an  $(r + 1, r + 2)$  model. Otherwise, if  $a \leq j$ , that is,  $m \bmod (n - m) \leq j$  or equivalently,  $n \bmod (n - m) \leq j$ , then there are only  $\lfloor \frac{n}{n-m} \rfloor = \lfloor 1 + \frac{m}{n-m} \rfloor = r + 1$  vertices in each transmitter and receiver. Again, clearly, there are direct edges between each transmitter-receiver pair. Moreover, there is a cross edge from transmitter vertex with label  $\ell_1(n - m) + j$  to receiver vertex with label  $\ell_2(n - m) + j$  if and only if  $\ell_2(n - m) + j - (\ell_1(n - m) + j) = n - m$ , or  $\ell_2 - \ell_1 = 1$ . Finally, assigning new vertex labels in the subgraph by subtracting  $j$  from the existing label and then dividing by  $(n - m)$  reveals that the subgraph is exactly an  $(r, r + 1)$  model.

In summary, we observe that color  $j$  (with  $j \in \{0, 1, 2, 3, n - m - 1\}$ ) represents an  $(r + 1, r + 2)$  model if  $a \geq j + 1$  and an  $(r, r + 1)$  model if  $a < j + 1$ . In other words, the colors with indices  $j = 0, 1, \dots, a - 1$  represent  $(r + 1, r + 2)$  models, and colors with indices  $j = a, a + 1, \dots, (n - m)$  represent  $(r, r + 1)$  models. This means that in the decomposition, we will have exactly  $a$  models of type  $(r + 1, r + 2)$  and the rest, i.e.,  $n - m - a$  models of type  $(r, r + 1)$ .

The case  $m > n$  follows along the same lines and is omitted for brevity. It is found in this case that we obtain  $a$  models of the type  $(r + 2, r + 1)$  and  $|n - m| - a$  models of the type  $(r + 1, r)$ .

## REFERENCES

- [1] A. Giridhar and P. R. Kumar, "Computing and communicating functions over sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 23, pp. 755–764, Apr. 2005.
- [2] A. G. Dimakis, P. B. Godfrey, Y. Wu, M. J. Wainwright, and K. Ramchandran, "Network coding for distributed storage systems," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4539–4551, Sep. 2010.
- [3] A. G. Dimakis, K. Ramchandran, Y. Wu, and C. Suh, "A survey on network codes for distributed storage," *Proc. IEEE*, vol. 99, no. 3, pp. 476–489, Mar. 2011.
- [4] R. Appuswamy, M. Franceschetti, N. Karamchandani, and K. Zeger, "Network coding for computing: Cut-set bounds," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 1015–1030, Feb. 2011.
- [5] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, Jul. 2000.
- [6] S.-Y. R. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Trans. Inf. Theory*, vol. 49, no. 2, pp. 371–381, Feb. 2003.
- [7] R. Koetter and M. Médard, "An algebraic approach to network coding," *IEEE/ACM Trans. Netw.*, vol. 11, no. 5, pp. 782–795, Oct. 2003.
- [8] T. Ho *et al.*, "A random linear network coding approach to multicast," *IEEE Trans. Inf. Theory*, vol. 52, no. 10, pp. 4413–4430, Oct. 2006.
- [9] S. Jaggi *et al.*, "Polynomial time algorithms for multicast network code construction," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 1973–1982, Jun. 2005.
- [10] B. Nazer and M. Gastpar, "Computation over multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3498–3516, Oct. 2007.
- [11] B. K. Rai and B. K. Dey, "On network coding for sum-networks," *IEEE Trans. Inf. Theory*, vol. 58, no. 1, pp. 50–63, Jan. 2012.
- [12] R. Dougherty, C. Freiling, and K. Zeger, "Insufficiency of linear coding in network information flow," *IEEE Trans. Inf. Theory*, vol. 51, no. 8, pp. 2745–2759, Aug. 2005.
- [13] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, "Wireless network information flow: A deterministic approach," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 1872–1905, Apr. 2011.
- [14] G. Bresler and D. N. C. Tse, "The two-user Gaussian interference channel: A deterministic view," *Eur. Trans. Telecommun.*, vol. 19, no. 4, pp. 333–354, Jun. 2008.
- [15] J. Zhan, S. Y. Park, M. Gastpar, and A. Sahai, "Linear function computation in networks: Duality and constant gap results," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 4, pp. 620–638, Apr. 2013.
- [16] A. Ramamoorthy, "Communicating the sum of sources over a network," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2008, pp. 1646–1650.
- [17] A. Ramamoorthy and M. Langberg, "Communicating the sum of sources over a network," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 4, pp. 655–665, Apr. 2013.
- [18] M. A. Maddah-Ali, A. S. Motahari, and A. K. Khandani, "Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3457–3470, Aug. 2008.
- [19] V. R. Cadambe and S. A. Jafar, "Interference alignment and the degree of freedom for the  $K$ -user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.
- [20] C. Suh and D. N. C. Tse, "Interference alignment for cellular networks," in *Proc. 46th Annu. Allerton Conf. Commun., Control, Comput.*, Sep. 2008, pp. 1037–1044.
- [21] C. Suh, M. Ho, and D. N. C. Tse, "Downlink interference alignment," *IEEE Trans. Commun.*, vol. 59, no. 9, pp. 2616–2626, Sep. 2011.
- [22] Y. Wu and A. G. Dimakis, "Reducing repair traffic for erasure coding-based storage via interference alignment," in *Proc. IEEE Int. Symp. Inf. Theory*, Seoul, South Korea, Jun./Jul. 2009, pp. 2276–2280.
- [23] N. B. Shah, K. V. Rashmi, P. V. Kumar, and K. Ramchandran, "Interference alignment in regenerating codes for distributed storage: Necessity and code constructions," *IEEE Trans. Inf. Theory*, vol. 58, no. 4, pp. 2134–2158, Apr. 2012.
- [24] C. Suh and K. Ramchandran, "Exact-repair MDS code construction using interference alignment," *IEEE Trans. Inf. Theory*, vol. 57, no. 3, pp. 1425–1442, Mar. 2011.
- [25] A. Das, S. Vishwanath, S. Jafar, and A. Markopoulou, "Network coding for multiple unicasts: An interference alignment approach," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2010, pp. 1878–1882.
- [26] A. Ramakrishnan, A. Das, H. Maleki, A. Markopoulou, S. Jafar, and S. Vishwanath, "Network coding for three unicast sessions: Interference alignment approaches," in *Proc. 48th Annu. Allerton Conf. Commun., Control, Comput.*, Sep./Oct. 2010, pp. 1054–1061.
- [27] C. Suh and M. Gastpar, "Network decomposition for function computation," in *Proc. IEEE 14th Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2013, pp. 340–344.

- [28] C. Suh, N. Goela, and M. Gastpar, "Approximate feedback capacity of the Gaussian multicast channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2012, pp. 2909–2913.
- [29] C. Suh and D. N. C. Tse, "Feedback capacity of the Gaussian interference channel to within 2 bits," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2667–2685, May 2011.
- [30] C. Suh and M. Gastpar, "Interactive function computation," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2013, pp. 2329–2333.
- [31] A. Vahid, C. Suh, and A. S. Avestimehr, "Interference channels with rate-limited feedback," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 2788–2812, May 2012.
- [32] C. Suh, I.-H. Wang, and D. N. C. Tse, "Two-way interference channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2012, pp. 2801–2805.

**Changho Suh** (S'10–M'12) is an Ewon Assistant Professor in the School of Electrical Engineering at Korea Advanced Institute of Science and Technology (KAIST) since 2012. He received the B.S. and M.S. degrees in Electrical Engineering from KAIST in 2000 and 2002 respectively, and the Ph.D. degree in Electrical Engineering and Computer Sciences from UC-Berkeley in 2011. From 2011 to 2012, he was a postdoctoral associate at the Research Laboratory of Electronics in MIT. From 2002 to 2006, he had been with the Telecommunication R&D Center, Samsung Electronics.

Dr. Suh received the 2015 Hadong Young Engineer Award from the Institute of Electronics and Information Engineers, the 2013 Stephen O. Rice Prize from the IEEE Communications Society, the David J. Sakrison Memorial Prize from the UC-Berkeley EECS Department in 2011, and the Best Student Paper Award of the IEEE International Symposium on Information Theory in 2009.

**Naveen Goela** received undergraduate degrees in computer science and mathematics, and a M. Eng. degree in computer science in 2004, all from the Massachusetts Institute of Technology (MIT) in Cambridge, MA, USA. His research at MIT focused on multi-camera visual hull reconstruction, and motion graphs. In 2013, he received the Ph.D. in electrical engineering and computer science with designated emphasis in statistics from the University of California, Berkeley. His past experience includes a research internship at Mitsubishi Electric Research Labs (MERL) in 2006, a visiting scholarship at Ecole Polytechnique Fédérale de Lausanne (EPFL) in Switzerland during 2011–2012, and a post-doctoral research position at Qualcomm Research in Berkeley, CA, USA during 2013–2014. Currently, he is a member of Technicolor Research in Los Altos, CA, USA. From 2007–2010, he was awarded the U.S. National Defense Science and Engineering Graduate (NDSEG) Fellowship. He received Qualcomm's company-wide ImpaqT award in 2014 for patents related to 5G wireless communications, and Technicolor's company-wide innovation awards in 2015 for research focusing on the signal processing of light signals. His research interests include networks, statistics, machine learning, and information theory.

**Michael Gastpar** received the Dipl. El.-Ing. degree from the Eidgenössische Technische Hochschule (ETH), Zürich, Switzerland, in 1997, the M.S. degree in electrical engineering from the University of Illinois at Urbana-Champaign, Urbana, IL, USA, in 1999, and the Doctoratès Science degree from the Ecole Polytechnique Fédérale (EPFL), Lausanne, Switzerland, in 2002. He was also a student in engineering and philosophy at the Universities of Edinburgh and Lausanne.

During the years 2003–2011, he was an Assistant and tenured Associate Professor with the Department of Electrical Engineering and Computer Sciences at the University of California, Berkeley. Since 2011, he has been a Professor in the School of Computer and Communication Sciences, Ecole Polytechnique Fédérale (EPFL), Lausanne, Switzerland. He is also a professor at Delft University of Technology, The Netherlands. He was a Researcher with the Mathematics of Communications Department, Bell Labs, Lucent Technologies, Murray Hill, NJ. His research interests are in network information theory and related coding and signal processing techniques, with applications to sensor networks and neuroscience.

Dr. Gastpar received the IEEE Communications Society and Information Theory Society Joint Paper Award in 2013 and the EPFL Best Thesis Award in 2002. He was an Information Theory Society Distinguished Lecturer (2009–2011), an Associate Editor for Shannon Theory for the IEEE TRANSACTIONS ON INFORMATION THEORY (2008–2011), and he has served as Technical Program Committee Co-Chair for the 2010 International Symposium on Information Theory, Austin, TX.