Computation in Multicast Networks: Function Alignment and Converse Theorems

Changho Suh, Naveen Goela and Michael Gastpar

***Abstract*—The classical problem in network coding theory considers *communication* over multicast networks. Multiple transmitters send independent messages to multiple receivers which decode the same set of messages. In this work, *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 computing capacity is characterized for a linear function (modulo-2 sum) of Bernoulli sources. Inspired by the geometric concept of interference alignment in networks, a new achievable coding scheme called *function alignment* is introduced. A new converse theorem is established that is tighter than cut-set based and genie-aided bounds. Computation (vs. communication) over multicast networks requires additional analysis to account for multiple receivers sharing a network's computational resources. We also develop 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-simpler algebraic proof of achievability that generalizes to L-transmitter L-receiver networks.**

***Index Terms*—Computing 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]. As a consequence, it was shown that the cut-set based bound is tight in the single-receiver case.

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], [7], [8], [9]. The success of traditional multicast communication motivates us to explore the fundamental limits of multicasting a linear function in *multi-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 in multi-receiver networks, we consider the simplest two-transmitter two-receiver network in which both receivers compute a linear function (modulo-2 sum) of two independent Bernoulli sources generated at the transmitters. Specifically, we consider the Avestimehr-Diggavi-Tse (ADT) deterministic single-hop network model [13] which captures superposition and broadcast properties of wireless Gaussian networks and is a generalization of networks of orthogonal links. We develop a new achievable coding scheme termed *function alignment*[1], inspired by the concept of *interference alignment* [15], [16]. We also derive a new converse theorem that is tighter than cut-set based bounds and genie-aided bounds. As a consequence of this capacity result, we find that unlike the single-receiver case, the cut-set based bound is not achieved due to competition for shared network resources that arise in satisfying function demands at multiple receivers.

As a byproduct of our analysis, we develop a *network decomposition theorem* to identify elementary parallel subnetworks that can constitute an original network without loss of optimality for in-network computation. The network decomposition approach offers a conceptually simpler proof of achievability which we use to establish the linear computing capacity of L-transmitter L-receiver single-hop networks. In

This work was presented in part at the IEEE Information Theory Workshop 2012, and the Allerton Conference 2012.

C. Suh is with the Department of Electrical Engineering, KAIST, South Korea (Email: chsuh@kaist.ac.kr).

N. Goela and M. Gastpar are with the School of Computer and Communication Sciences, EPFL, Switzerland (Email: {naveen.goela, michael.gastpar}@epfl.ch), and with the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, Berkeley, USA (Email: {ngoela, gastpar}@eecs.berkeley.edu)

[1]Niesen-Nazer-Whiting [14] introduced a similar scheme in a different context called computation alignment.

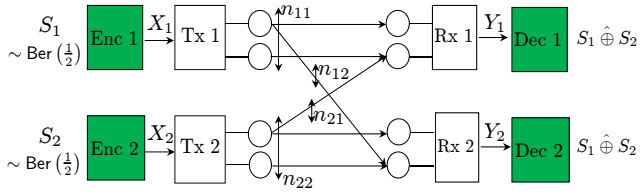

Fig. 1. Two-transmitter two-receiver Avestimehr-Diggavi-Tse (ADT) deterministic network

addition, the approach has potential for the design of structured computation codes in larger multi-hop networks.

Related Work: In [17], [11], [18], the computing 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 [18] 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, our work considers sources without entropy constraints and establishes the exact capacity of an ADT multi-receiver network which is a generalization of traditional network coding models with orthogonal links.

II. Model

We focus on a two-transmitter two-receiver ADT deterministic network. Section VI includes our results for L-transmitter L receiver networks. As shown in Fig. 1, this network is described by four integer parameters n_{ij} which indicates the number of signal bit levels from transmitter i $(i = 1, 2)$ to receiver j $(j = 1, 2)$. Let $X_\ell \in \mathbb{F}_2^q$ be transmitter ℓ's encoded signal where $q = \max_{ij} n_{ij}$. The received signals are then given by

$$\begin{aligned} Y_1 &= \mathbf{G}^{q-n_{11}} X_1 \oplus \mathbf{G}^{q-n_{21}} X_2, \\ Y_2 &= \mathbf{G}^{q-n_{12}} X_1 \oplus \mathbf{G}^{q-n_{22}} X_2, \end{aligned} \tag{1}$$

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

Each receiver wishes to compute modulo-2 sums of the two Bernoulli sources S_1^K and S_2^K, generated at the two transmitters, with N uses of the network. Here we use shorthand notation to indicate the sequence up to K, e.g., $S_1^K := (S_{11}, \cdots, S_{1K})$. We assume that S_1^K and S_2^K are independent and identically distributed with $\mathsf{Bern}(\frac{1}{2})$. Transmitter ℓ uses its encoding function to map S_ℓ^K to a length-N codeword X_ℓ^N. Receiver ℓ uses a decoding function d_ℓ to estimate $S_1^K \oplus S_2^K$ from its received signal Y_ℓ^N. An error occurs whenever $d_\ell \neq S_1^K \oplus S_2^K$. The average probabilities of error are given by $\lambda_\ell = \mathbb{E}\left[P(d_\ell \neq S_1^K \oplus S_2^K)\right], \ell = 1, 2$.

We say that the computing rate $R_{\mathsf{comp}} = \frac{K}{N}$ is achievable if there exists a family of codebooks and encoder/decoder functions such that the average decoding error probabilities of λ_1 and λ_2 go to zero as code length N tends to infinity. We will also need the notion of linear computing capacity $C_{\mathsf{comp}}^{\mathsf{lin}}$, where we restrict both the encoders and the decoders to be linear mappings. In line with the standard network coding literature, when referring to the linear computing capacity, we will assume a zero-error framework rather than the framework of negligible error we use in the context of the regular computing capacity.

We classify networks into two classes, depending on a reconstructability condition that will be specified in the sequel. The reconstructability turns out to be the key property that classifies networks. This will be clarified when we prove an upper bound on the computing capacity in Theorem 1.

Definition 1: A network is said to be *degenerate* if none of $\mathbf{G}^{q-n_{ij}} X_i$ can be reconstructed from (Y_1, Y_2) for all i, j. A network is said to be *non-degenerate* if there exists (i, j) such that $\mathbf{G}^{q-n_{ij}} X_i$ can be reconstructed from (Y_1, Y_2).

Lemma 1: A network is degenerate if and only if $n_{11} - n_{12} = n_{21} - n_{22}$. As a direct consequence, a network is non-degenerate if and only if $n_{11} - n_{12} \neq n_{21} - n_{22}$.

Proof: See Appendix A. ∎

III. Main Results

Theorem 1 (Upper Bound on Computing Capacity): The computing capacity is upper-bounded by

$$C_{\mathsf{comp}} \le \min\left\{n_{11}, n_{12}, n_{22}, n_{21}\right\}. \tag{2}$$

For non-degenerate networks where $n_{11} - n_{12} \neq n_{21} - n_{22}$,

$$C_{\mathsf{comp}} \le \frac{\max(n_{11}, n_{21}) + \max(n_{22}, n_{12})}{3}. \tag{3}$$

Proof: See Section III-A. ∎

We show the tightness of the above bounds for the following two cases: (a) degenerate networks; (b) symmetric networks characterized by two parameters of $n := n_{11} = n_{22}$ and $m := n_{12} = n_{21}$.

Theorem 2 (Degenerate Networks): For degenerate networks where $n_{11} - n_{12} = n_{21} - n_{22}$,

$$C_{\mathsf{comp}} = \min\{n_{11}, n_{12}, n_{22}, n_{21}\}. \tag{4}$$

Proof: The converse proof is immediate from Theorem 1. See Section III-B for the achievability proof. ∎

Theorem 3 (Symmetric Networks): For symmetric networks where $n := n_{11} = n_{22}$ and $m := n_{12} = n_{21}$,

$$C_{\mathsf{comp}} = \begin{cases} \min\left\{m, n, \frac{2}{3}\max(m, n)\right\}, & m \neq n; \\ n, & m = n. \end{cases} \tag{5}$$

Proof: The converse proof is immediate from Theorem 1. See Section IV for the achievability proof. ∎

Our results are interpreted with a focus on symmetric networks. Specifically, it will be shown that our scheme outperforms the separation scheme where both receivers decode all of the sources and then compute modulo-2 sums of the sources. It will also be revealed that in contrast to a single-receiver function-unicasting case, the cut-set based bound is not tight when multicasting linear functions.

For illustrative purpose, consider the normalized computing capacity as follows:

$$\frac{C_{\mathsf{comp}}}{q} = \begin{cases} \min\left\{\alpha, \frac{2}{3}\right\}, & \alpha < 1; \\ 1, & \alpha = 1, \end{cases} \tag{6}$$

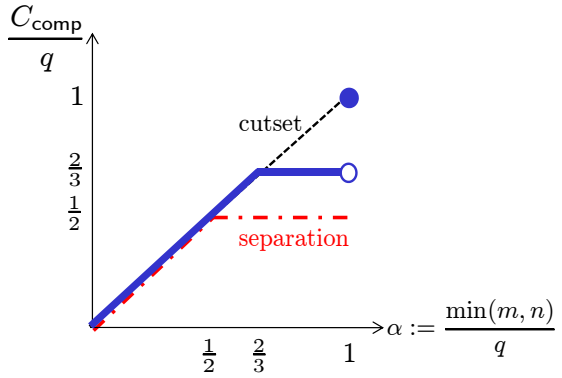

Fig. 2. Normalized computing capacity. Here $n := n_{11} = n_{22}$ and $m := n_{12} = n_{21}$. The parameter α in x-axis captures a signal-strength similarity between m and n.

where $q = \max(m,n)$ and $\alpha := \frac{\min(m,n)}{q}$.

Remark 1 (Comparison to Separation Scheme): The computing rate of the separation scheme can be derived from the multicast capacity. Note that the multicast capacity is the intersection of the two individual MAC capacities: the set $\mathcal{C}_{\text{mult}}$ of (R_1, R_2) such that $R_1 \le \min(m,n)$, $R_2 \le \min(m,n)$ and $R_1 + R_2 \le \max(m,n)$. Therefore, this gives

$$\frac{R_{\text{sep}}}{q} = \frac{C_{\text{sym}}}{q} = \min\left\{\alpha, \frac{1}{2}\right\}, \tag{7}$$

where $C_{\text{sym}} := \sup\{R : (R,R) \in \mathcal{C}_{\text{mult}}\}$. While this separation approach provides the optimal strategy for $0 \le \alpha \le \frac{1}{2}$, it is suboptimal for the other regime $\frac{1}{2} < \alpha \le 1$. Note that for $\frac{1}{2} < \alpha \le 1$, more-than-half of signal levels at receivers naturally form the mod-sum function of our interest. It turns out that this natural matching can provide higher computing rates. Details will be explained in Section IV. □

Remark 2 (Comparison to a Single-Receiver Case): In a single-receiver case, the computing capacity achieves the cut-set based upper bound, which will be formally proven to be $\min(m,n)$ in the next section. On the other hand, the cut-set bound is not tight when multicasting a function. Notice the non-zero gap between the function-unicasting and function-multicasting capacities when $\frac{2}{3} \le \alpha < 1$ (see Fig. 2). This comes from the tension that arises in satisfying the same demand at multiple receivers. We will clarify this while presenting our achievability in Section IV. □

A. Proof of Theorem 1

The proof of the bound (2) is based on the standard cut-set argument. The main focus is to prove the second bound (3).

Proof of (2): Starting with Fano's inequality, we get

$$\begin{aligned}
&N(R_{\text{comp}} - \epsilon_N) \le I(S_1^K \oplus S_2^K; Y_1^N) \\
&\le I(S_1^K \oplus S_2^K; Y_1^N, S_2^K) \\
&\overset{(a)}{=} I(S_1^K \oplus S_2^K; Y_1^N | S_2^K) \\
&\overset{(b)}{=} I(S_1^K \oplus S_2^K; Y_1^N | S_2^K, X_2^N) \\
&= H(Y_1^N | S_2^K, X_2^N) \overset{(c)}{\le} \sum H(Y_{1i}|X_{2i}) \le N n_{11}
\end{aligned}$$

where (a) follows from the fact that S_2^K is independent of $S_1^K \oplus S_2^K$; (b) follows from the fact that X_2^N is a function of S_2^K; (c) follows from the fact that conditioning reduces entropy. If R_{comp} is achievable, then $\epsilon_N \to 0$ as N tends to infinity. So we get $R_{\text{comp}} \le n_{11}$. Similarly we can show that $R_{\text{comp}} \le \min\{H(Y_2|X_2), H(Y_1|X_1), H(Y_2|X_1)\} \le \min\{n_{12}, n_{21}, n_{22}\}$.

Proof of (3): For non-degenerate networks, by definition, there exists (i,j) such that $\mathbf{G}^{q-n_{ij}} X_i$ can be reconstructed from (Y_1, Y_2). Without loss of generality, assume that $\mathbf{G}^{q-n_{12}} X_1$ is a function of (Y_1, Y_2).

Starting with Fano's inequality, we get

$$\begin{aligned}
&N(3R_{\text{comp}} - \epsilon_N) \\
&\le I(S_1^K \oplus S_2^K; Y_1^N) + I(S_1^K \oplus S_2^K; Y_2^N) + I(S_1^K \oplus S_2^K; Y_2^N) \\
&\overset{(a)}{\le} [H(Y_1^N) - H(Y_1^N | S_1^K \oplus S_2^K)] \\
&\quad + [H(Y_2^N) - H(Y_2^N | S_1^K \oplus S_2^K, Y_1^N)] + I(S_1^K \oplus S_2^K; Y_2^N) \\
&\le H(Y_1^N) + H(Y_2^N) \\
&\quad - H(Y_1^N, Y_2^N | S_1^K \oplus S_2^K) + I(S_1^K \oplus S_2^K; Y_2^N, S_2^K) \\
&\overset{(b)}{=} H(Y_1^N) + H(Y_2^N) \\
&\quad - H(Y_1^N, Y_2^N | S_1^K \oplus S_2^K) + I(S_1^K \oplus S_2^K; Y_2^N | S_2^K) \\
&\overset{(c)}{=} H(Y_1^N) + H(Y_2^N) \\
&\quad - H(Y_1^N, Y_2^N | S_1^K \oplus S_2^K) + H(\mathbf{T}_{12} X_1^N | S_2^K) \\
&\overset{(d)}{=} H(Y_1^N) + H(Y_2^N) \\
&\quad - H(Y_1^N, Y_2^N, \mathbf{T}_{12} X_1^N | S_1^K \oplus S_2^K) + H(\mathbf{T}_{12} X_1^N | S_2^K) \\
&\le H(Y_1^N) + H(Y_2^N) \\
&\quad - H(\mathbf{T}_{12} X_1^N | S_1^K \oplus S_2^K) + H(\mathbf{T}_{12} X_1^N | S_2^K) \\
&\overset{(e)}{=} H(Y_1^N) + H(Y_2^N) - H(\mathbf{T}_{12} X_1^N) + H(\mathbf{T}_{12} X_1^N | S_2^K) \\
&\overset{(f)}{\le} \sum [H(Y_{1i}) + H(Y_{2i})] \\
&\le 2N[\max(n_{11}, n_{21}) + \max(n_{12}, n_{22})]
\end{aligned}$$

where (a) follows from the fact that conditioning reduces entropy; (b) follows from the fact that S_1^K is independent of $S_1^K \oplus S_2^K$; (c) follows from the fact that X_2^N is a function of S_2^K and that $\mathbf{T}_{12} := \mathbf{I}_N \otimes \mathbf{G}^{q-n_{12}}$; (d) follows from our hypothesis that $\mathbf{G}^{q-n_{12}} X_1$ is a function of (Y_1, Y_2); (e) follows from the fact that X_1^N is a function of S_1^K that is independent of $S_1^K \oplus S_2^K$; (f) follows from the fact that conditioning reduces entropy. This completes the proof.

B. Proof of Theorem 2

Assume that $n_{11} - n_{12} = n_{21} - n_{22} \ge 0$. Then Y_2 is a degenerated version of Y_1:

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

This shows an equivalence to a single-receiver case which concerns receiver 2's demand only. Hence, in this case, $R_{\text{comp}} = \min\{n_{12}, n_{22}\}$. Similarly for the other case of

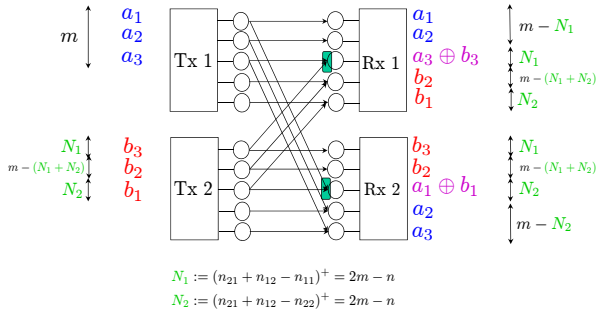

Fig. 3. [*Case I:* $\frac{1}{2} \le \alpha \le \frac{2}{3}$]: An achievable scheme for $(m, n) = (3, 5)$ and generalization to arbitrary values of (m, n).

$n_{11} - n_{12} = n_{21} - n_{22} \le 0$, one can show that Y_1 is a degenerated version of Y_2 and therefore a network becomes equivalent to a single-receiver case w.r.t receiver 1 where $R_{\mathsf{comp}} = \min\{n_{11}, n_{21}\}$.

IV. Proof of Theorem 3 via Geometric Approach

By symmetry, focus on the case of $m \le n$. The other case of $m \ge n$ is a mirror image in which transmitters 1 and 2 are swapped. As mentioned in Remark 1, the separation scheme can achieve the computing capacity for $0 \le \alpha \le \frac{1}{2}$. The case of $\alpha = 1$ is a degenerate case where the channel forms the mod-sum function by nature at both receivers. In this case, uncoded transmission can yield $R_{\mathsf{comp}} = n$. Hence, our focus is the following two non-degenerate cases.

A. Case I: $\frac{1}{2} \le \alpha \le \frac{2}{3}$

Let us explain achievability with the example $(m, n) = (3, 5)$ illustrated in Fig. 3. We will show that the cut-set bound of $\min(m, n) = 3$ can be achieved. First transmitter 1 sends the bits (a_1, a_2, a_3) on the top $3(= m)$ levels. Observe that the 3rd level at receiver 1 marked with a green square is connected with transmitter 1's upper m levels as well as transmitter 2's upper m levels. The idea is to exploit this overlapped level. Transmitter 2 sends b_3 on the top level to achieve $a_3 \oplus b_3$ on the overlapped level at receiver 1. In an arbitrary case, the number of overlapped levels is $N_1 := n_{12} + n_{21} - n_{11} = 2m - n$. On the other hand, the bit b_3 is cleanly received at receiver 2 without being interfered with by (a_1, a_2, a_3), since $N_1 + m \le n$ in the regime of $\frac{1}{2} \le \alpha \le \frac{2}{3}$. Similarly let $N_2 := n_{12} + n_{21} - n_{22} = 2m - n$ be the number of levels at receiver 2 which are connected with transmitter 1's upper m levels as well as transmitter 2's upper m levels. In this example, level 3 at receiver 2 is the overlapped level. Transmitter 2 then sends b_1 on the 3rd level so as to achieve $a_1 \oplus b_1$ on the level at receiver 2. This b_1 is cleanly received at receiver 1, since $N_2 + m \le n$ in the regime of $\frac{1}{2} \le \alpha \le \frac{2}{3}$.

Finally notice that level 2 at transmitter 2 is vacant among the top m levels. In an arbitrary case, the number of these vacant levels is $m - (N_1 + N_2)$. Transmitter 2 sends additional symbols (b_2 in this example) on the vacant $m - (N_1 + N_2)$ levels. Obviously these symbols are cleanly received at both receivers. In summary, receiver 1 can compute $a_1 \oplus b_1$, $a_2 \oplus b_2$, and $a_3 \oplus b_3$. In an arbitrary case, the total number of these computable bits is $N_1 + N_2 + \{m - (N_1 + N_2)\} = m$. Similarly receiver 2 can compute m bits. Therefore, we can achieve $R_{\mathsf{comp}} = m$.

Remark 3 (Exploiting Channel Structure [10], [19]): In the regime of $\frac{1}{2} \le \alpha \le \frac{2}{3}$, more-than-half of signal levels at receivers naturally form the mod-sum function. This enables us to create a shared linear subspace. Note in the above example that at receiver 2, the symbols (a_1, b_1) share one-dimensional linear subspace spanned by $[0, 0, 1, 0, 0]^t$, where $[\cdot]^t$ indicates a transpose. This enables us to outperform the separation scheme where shared subspaces do not exist. □

Remark 4 (Connection to Interference Alignment): Note that 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* [15], [16] which has shown great potential for a variety of applications such as interference channels [16], cellular networks [20], [21], distributed storage networks [22], [23], [24] and multiple unicast networks [25], [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. To highlight this purpose, we call it *function alignment*. □

B. [Case II: $\frac{2}{3} \le \alpha < 1$]: Example

Unlike Case I, our achievability for this regime employs a vector-coding scheme. We first explain our achievability idea with the example $(m, n) = (3, 4)$ illustrated in Fig. 4. We will then invoke a geometric insight which helps generalizing to arbitrary values of (m, n). The generalization will be explained in the next section.

Our achievability idea is to *alternate* function alignment at both receivers. See Fig. 4. We first achieve function alignment $a_1 \oplus b_1$ at receiver 1. We next achieve $a_2 \oplus b_2$ at receiver 2. We repeat this until all of the resource levels are fully utilized. At the end of time 1, receiver 1 can then compute all of $a_i \oplus b_i$'s $(i = 1, 2, 3)$. However, receiver 2 can compute only $a_1 \oplus b_1$, $a_2 \oplus b_2$ and b_3. Since we start favoring receiver 1, we end up with this asymmetry.

In order to make it symmetric, we invoke the idea of vector coding. In time 2, we start by favoring receiver 2 instead and repeat the same procedure as before. We can then obtain a symmetric solution at the end of time 2. However, the solution is still inefficient. Note that b_6 is missing at receiver 1, and similarly a_3 is missing at receiver 2. To improve, we use another time slot. In time 3, we now have two purposes: (1) sending fresh source symbols; (2) delivering the b_6 and a_3 to receivers 1 and 2 respectively. We first multicast fresh symbols $a_7 \oplus b_7$ and $a_8 \oplus b_8$ with alternating function alignment. Next transmitter 1 sends a_3 (wanted by receiver 2) on the third level. But this transmission causes interference to b_8 which was already received at receiver 1. Fortunately we can resolve this conflict. Here the key observation is that $a_3 \oplus b_3$ is already obtained at receiver 1 in time 1. Hence, transmitter 2 sending b_3 on top of b_8 in time 3, we can achieve the

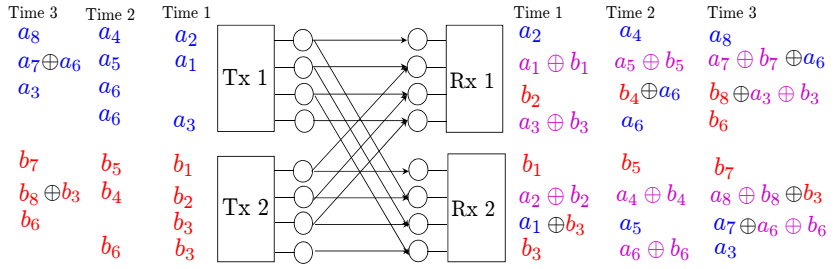

Fig. 4. [*Case II:* $\frac{2}{3} \le \alpha < 1$]: Alternating function alignment for $(m, n) = (3, 4)$.

function alignment $a_3 \oplus b_3$ at receiver 1. The $a_3 \oplus b_3$ already received in time 1 can then be exploited as *side information* to decode b_8 from $b_8 \oplus a_3 \oplus b_3$. As a result, transmitter 1 can deliver the a_3 to receiver 2 without interfering with b_8 at receiver 1. Similarly transmitter 2 can deliver the b_6 to receiver 1 without interfering with a_7 at receiver 2. Both receivers can now compute $a_i \oplus b_i$'s for $i = 1, \cdots, 8$ during 3 time slots, thus achieving $R_{\mathsf{comp}} = \frac{8}{3}$.

Geometric Interpretation: To aid generalization to arbitrary values of (m, n), we invoke geometric insights from the $(3, 4)$ example. In this example, $\mathbf{v} = [1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]^t$ can be viewed as a beamforming vector for a_1. Beamforming vector designs are closely associated with function alignment. To achieve function alignment $a_1 \oplus b_1$ at receiver 2, transmitter 2 designs its corresponding vector as $\mathbf{Tv}$, where $\mathbf{T}$ indicates the 3-time-slot equivalent channel: $\mathbf{T} := \mathbf{I}_3 \otimes \mathbf{G}^{4-3} = \mathbf{I}_3 \otimes \mathbf{G}$. With this geometric viewpoint, we can interpret the $(3, 4)$ example solution as in Fig. 5.

Let $\mathbf{a} := (a_2, a_4, a_8, a_6)^t$ and $\bar{\mathbf{a}} := (a_1, a_5, a_7, a_3)^t$; similarly $\mathbf{b} := (b_2, b_4, b_8, b_6)^t$ and $\bar{\mathbf{b}} := (b_1, b_5, b_7, b_3)^t$. Let $\mathbf{V}_1$ be a 12-by-4 beamforming matrix w.r.t. $\mathbf{a}$. Let $\mathbf{V}_2$ be a 12-by-4 beamforming matrix w.r.t. $\bar{\mathbf{b}}$. According to the code construction in Fig. 4, we have

$$\mathbf{V}_1 = \left[\begin{array}{c|c|c|c} 1&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0 \end{array}\right], \left[\begin{array}{c|c|c|c} 0&1&0&0\\0&0&0&0\\0&0&0&1\\0&0&0&1 \end{array}\right], \left[\begin{array}{c|c|c|c} 0&0&1&0\\0&0&0&1\\0&0&0&0\\0&0&0&0 \end{array}\right], \mathbf{V}_2 = \left[\begin{array}{c|c|c|c} 1&0&0&0\\0&0&0&0\\0&0&0&1\\0&0&0&1 \end{array}\right], \left[\begin{array}{c|c|c|c} 0&1&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&0 \end{array}\right], \left[\begin{array}{c|c|c|c} 0&0&1&0\\0&0&0&1\\0&0&0&0\\0&0&0&0 \end{array}\right]. \quad (8)$$

To achieve function alignment $\mathbf{a} \oplus \mathbf{b}$ at receiver 2, transmitter 2 sends $\mathbf{b}$ along with $\mathbf{TV}_1$. Similarly to achieve $\bar{\mathbf{a}} \oplus \bar{\mathbf{b}}$ at receiver 1, transmitter 1 sends $\bar{\mathbf{a}}$ along with $\mathbf{TV}_2$. Receiver 2 then gets $\mathbf{V}_2\bar{\mathbf{b}}$ and $(\mathbf{T}^2\mathbf{V}_2)\bar{\mathbf{a}}$. One can verify that $\mathsf{rank}([\mathbf{V}_1\ \mathbf{TV}_2\ \mathbf{T}^2\mathbf{V}_1]) = \mathsf{rank}([\mathbf{V}_2\ \mathbf{TV}_1\ \mathbf{T}^2\mathbf{V}_2]) = 12$. This enables both receivers to compute $\mathbf{a} \oplus \mathbf{b}$ and $\bar{\mathbf{a}} \oplus \bar{\mathbf{b}}$.

C. [Case II: $\frac{2}{3} \le \alpha < 1$]: Generalization

We now provide a code construction of $\mathbf{V}_1$ and $\mathbf{V}_2$ for arbitrary values of (m, n). Let M_1 be the column size of $\mathbf{V}_1$, i.e., the number of symbols that form function alignment at receiver 2. Similarly, let M_2 be the column size of $\mathbf{V}_2$. In the previous $(3, 4)$ example, $M_1 = M_2 = 4$. Notice that $R_{\mathsf{comp}} = \frac{M_1+M_2}{3}$ is achievable if the following matrices are full rank:

$$\mathbf{B}_1 := [\mathbf{V}_1,\ \mathbf{T}^2\mathbf{V}_1,\ \mathbf{TV}_2] \in \mathbb{F}_2^{3n\times(2M_1+M_2)}$$
$$\mathbf{B}_2 := [\mathbf{V}_2,\ \mathbf{T}^2\mathbf{V}_2,\ \mathbf{TV}_1] \in \mathbb{F}_2^{3n\times(M_1+2M_2)}.$$

We choose appropriate values of (M_1, M_2) such that $M_1 + M_2 = 2n$ and thus can yield $R_{\mathsf{comp}} = \frac{2n}{3}$. Considering the total dimension of the linear subspace at receiver 1, we get $2M_1+M_2 \le 3n$. Similarly for receiver 2, we get $M_1+2M_2 \le 3n$. This motivates us to choose $M_1 = M_2 = n$.

We construct $(\mathbf{V}_1, \mathbf{V}_2)$ such that $\mathbf{B}_1$ and $\mathbf{B}_2$ are full rank. The form of $\mathbf{V}_1$ and $\mathbf{V}_2$ in (8) inspires our construction in the general case. Note that the first three columns of $\mathbf{V}_1$ and $\mathbf{V}_2$ are the same, say $\mathbf{V}$. Inspecting more examples, we could identify the dimension of $\mathbf{V}$ as $3n$-by-$3(n-m)$:

$$\begin{aligned}\mathbf{V}_1 &= [\mathbf{V}\ \ \mathbf{P}_1] \in \mathbb{F}_2^{3n\times n}\\ \mathbf{V}_2 &= [\mathbf{V}\ \ \mathbf{P}_2] \in \mathbb{F}_2^{3n\times n}\end{aligned} \quad (9)$$

where $\mathbf{V} \in \mathbb{F}_2^{3n\times 3(n-m)}$ and $\mathbf{P}_\ell \in \mathbb{F}_2^{3n\times(3m-2n)}$, $\ell = 1, 2$. The form of (8) inspires:

$$\mathbf{V} = \mathbf{I}_3 \otimes [\mathbf{e}_1^{(n)}\ \ \cdots\ \ \mathbf{e}_{n-m}^{(n)}], \quad (10)$$

where $\mathbf{e}_i^{(n)} \in \mathbb{F}_2^n$ indicates the ith coordinate vector in an n-dimensional space. Note in (8) that $\mathbf{P}_1$ and $\mathbf{P}_2$ bear a strong similarity: the (9th-12th) rows are identical; the (1st-4th) rows of $\mathbf{P}_2$ are the same as the (5th-8th) rows of $\mathbf{P}_1$. Inspecting more examples, we could develop a construction:

$$\begin{aligned}\mathbf{P}_1 &= \mathbf{e}_3^{(3)} \otimes [\mathbf{e}_{(n-m)+1}^{(n)}\ \ \cdots\ \ \mathbf{e}_{2m-n}^{(n)}]\\ &\oplus \mathbf{e}_2^{(3)} \otimes \left\{[\mathbf{e}_{2(n-m)+1}^{(n)}\ \ \cdots\ \ \mathbf{e}_m^{(n)}] \oplus [\mathbf{e}_{3(n-m)+1}^{(n)}\ \ \cdots\ \ \mathbf{e}_n^{(n)}]\right\},\\ \mathbf{P}_2 &= \mathbf{e}_3^{(3)} \otimes [\mathbf{e}_{(n-m)+1}^{(n)}\ \ \cdots\ \ \mathbf{e}_{2m-n}^{(n)}]\\ &\oplus \mathbf{e}_1^{(3)} \otimes \left\{[\mathbf{e}_{2(n-m)+1}^{(n)}\ \ \cdots\ \ \mathbf{e}_m^{(n)}] \oplus [\mathbf{e}_{3(n-m)+1}^{(n)}\ \ \cdots\ \ \mathbf{e}_n^{(n)}]\right\}.\end{aligned} \quad (11)$$

The following lemma shows that this code ensures the full rank of $\mathbf{B}_1$ and $\mathbf{B}_2$. This completes the proof.

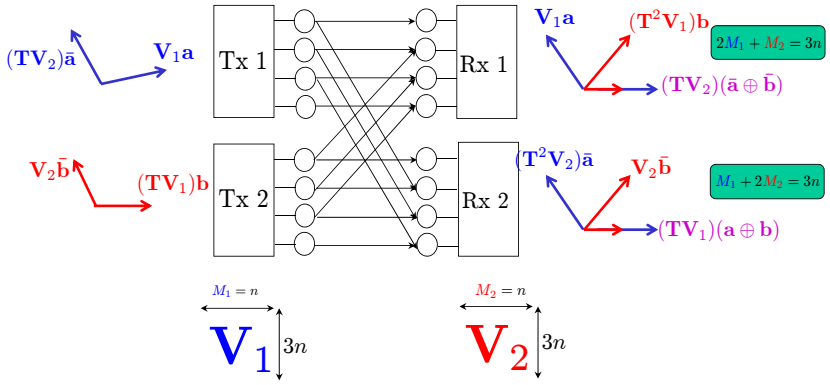

Fig. 5. Geometric interpretation of an achievable scheme.

Lemma 2:

$$\begin{aligned} \mathsf{rank}\left[\mathbf{V}_1,\ \mathbf{TV}_2,\ \mathbf{T}^2\mathbf{V}_1\right] &= 3n, \\ \mathsf{rank}\left[\mathbf{V}_2,\ \mathbf{TV}_1,\ \mathbf{T}^2\mathbf{V}_2\right] &= 3n. \end{aligned} \tag{12}$$

Proof: See Appendix B. ■

V. Proof of Theorem 3 via Network Decomposition

In this section, we present a network decomposition theorem that permits to decompose a network into elementary subnetworks. The decomposition theorem applies not only to the two-user network discussed so far, but directly extends to the L-user network, which will formally be introduced in Section VI. Using this theorem for the case of $L = 2$ users, we will provide an alternative conceptually-simpler achievability proof of Theorem 3 by coding separately over each elementary subnetwork. Interestingly, this coding strategy is sufficient to meet the converse bounds, and hence, to establish computation capacity, thus establishing a *separation principle* among the building blocks. This observation is somewhat surprising — in general interference channel problems, coding separately over parallel channels entails a significant loss in performance.

For the general case of L users, we will evaluate the performance of this coding approach in Section VI and show that it matches the upper bound for linear coding strategies.

Theorem 4 (Network Decomposition): For the L-transmitter L-receiver (m, n) network where $m \neq n$, the following network decompositions hold:[2]

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

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

(2) $(2m + 1, 2n + 1) = (m, n) \times (m + 1, n + 1)$

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

$$(m, n) = \begin{cases} (r, r+1)^{n-m-a} \times (r+1, r+2)^a, & m < n; \\ (r+1, r)^{m-n-a} \times (r+2, r+1)^a, & m > n. \end{cases} \tag{13}$$

[2]We use the symbol $\times$ for the concatenation of orthogonal models, just like in $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$.

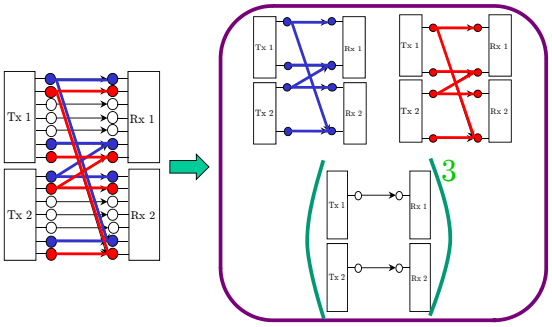

Fig. 6. A network decomposition example of an $(m, n) = (2, 7)$ model. From (13), $r = 0$ and $a = 2$; hence, the decomposition is given by $(2, 7) = (0, 1)^3 \times (1, 2)^2$.

where

$$\begin{aligned} r &= \left\lfloor \frac{\min\{m, n\}}{|n - m|} \right\rfloor, \\ a &= \min\{m, n\} \mod |n - m|. \end{aligned} \tag{14}$$

The proof is given in Appendix C. Here we provide a proof idea with an $(m, n, L) = (2, 7, 2)$ example, illustrated in Fig. 6. 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 $(2, 7) = (1, 2)^2 \times (0, 1)^3$.

Remark 5: Unlike the $L = 2$ case, for $L \geq 3$, the case $m < n$ is not symmetric with $m > n$. Nevertheless, the above symmetric decomposition holds even when $L \geq 3$. □

Remark 6: The separation principle among these decomposed subnetworks is not generally true. It is well known that for parallel interference channels, optimal performance requires joint coding across orthogonal components. □

Theorem 4 suggests that fundamental building blocks are of form $(r, r+1)$ or $(r+1, r)$, that is, "gap-1" models. Hence, we focus on the computing rates of the "gap-1" models.

Lemma 3 ($L = 2$): The following computing rates are achievable:

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

This lemma can be proved via the geometric approach in Section IV. We give a short explicit proof in Appendix D, showing that explicit codes for the $(3, 4)$ and $(4, 5)$ models (found, for example, via the method from Section IV) directly imply the general proof of the lemma.

Achievability Proof of Theorem 3: By symmetry, we focus on the case of $m < n$. For the case of $0 \leq \alpha \leq \frac{1}{2}$, $r = 0$ and $a = m$ in (14); hence, the decomposition is given by $(m, n) = (0, 1)^{n-2m} \times (1, 2)^m$. Thus, using Lemma 3, the computing rate is $R_{\mathsf{comp}} = 0 \cdot (n - 2m) + 1 \cdot m = m$. Next, consider the case of $\frac{1}{2} \leq \alpha \leq \frac{2}{3}$. Applying the decomposition (13), we find that in this case, $r = 1$ and $a = 2m - n$: $(m, n) = (1, 2)^{2n-3m} \times (2, 3)^{2m-n}$. Thus, using Lemma 3, the computing rate is $R_{\mathsf{comp}} = 1 \cdot (2n - 3m) + 2 \cdot (2m - n) = m$. Finally, consider the case of $\alpha \geq \frac{2}{3}$. Applying the decomposition (13), we find that in this case, $r \geq 2$. So we get

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

where (a) is due to (14). This completes the proof.

Remark 7: At first, it might seem that this proof is simpler than our arguments in Section IV. However, we point out that proving Lemma 3 is not straightforward, and hence, that there is no clear ordering as to which proof is simpler. Both proofs carry different intuitions and insights into the structure of the problem. □

VI. $L \times L$ Symmetric Networks

We consider an $L(\geq 3)$-transmitter L-receiver network where all of the L receivers want to compute a mod-2-sum of all of the Bernoulli sources generated at the transmitters. We consider a symmetric setting where the two integer parameters of (m, n) describe the network. Here n indicates the number of signal bit levels from transmitter ℓ to receiver ℓ; and m denotes the number of signal bit levels from transmitter ℓ to receiver $\ell' (\neq \ell)$. See Fig. 7 for an $(m, n) = (3, 4)$ example of the network. The received signal at receiver ℓ is given by

$$Y_\ell = \mathbf{G}^{q-n} X_\ell \oplus \bigoplus_{j \neq \ell} \mathbf{G}^{q-m} X_j, \tag{15}$$

for $\ell = 1, 2, \ldots, L$.

Theorem 5: The linear computing capacity is

$$C_{\mathsf{comp}}^{\mathsf{lin}} = \begin{cases} \min\left\{m, n, \frac{1}{2}\max(n, m)\right\}, & m \neq n; \\ n, & m = n. \end{cases}$$

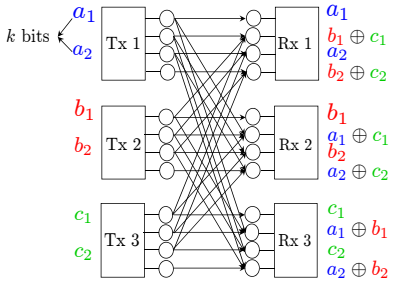

Fig. 7. Achievable scheme for the $(r-1, r)$ model where $r = 2k$.

The computing capacity is upper-bounded by

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

Proof: See Section VI-A for the achievability proof and Section VI-B for the coverse proof under linear coding strategies. See Section VI-C for the information-theoretic upper bound. ■

Remark 8: In general networks, the linear capacity is often not equal to the capacity and non-linear codes may achieve higher rates [12]. In the limit of $L \to \infty$, however, linear codes show the optimality. Note that our information-theoretic upper bound approaches the achievable rate as L tends to infinity, thus establishing the asymptotic computing capacity. □

A. Achievability Proof

The idea is to combine the network decomposition in Theorem 4 and achievability proof for elementary subnetworks.

Lemma 4 ($L \geq 3$): The following computing rates are achievable:

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

Proof: The items (1) and (3) are straightforward. For the (2a) model, we consider two cases: $r = 2k$ and $r = 2k+1$. Fig. 7 shows an achievable scheme when $r = 2k = 2 \cdot 2$ and $L = 3$. Each transmitter uses odd-numbered levels to send k symbols. The special structure of symmetric networks allows each receiver to get clean symbols on odd-numbered levels while receiving partially-satisfied functions on even-numbered levels. For example, receiver 1 gets (a_1, a_2) on the first and third levels; $(b_1 \oplus c_1, b_2 \oplus c_2)$ on the second and fourth levels. Note that two resource levels are consumed to compute one desired function. Therefore, this gives $R_{\mathsf{comp}} = \frac{1}{2}r$. Obviously this can be applied to an arbitrary value of L as well as the (2b) model.

Fig. 8 shows an achievable scheme for the case of $r = 2k+1 = 2 \cdot 2 + 1$ and $L = 3$. If we followed the same approach as in the case of $r = 2k$, each receiver would end up with having a resource hole in the last bottom 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, and 3rd levels, while the last bottom level is empty. In order to make an efficient resource utilization, we

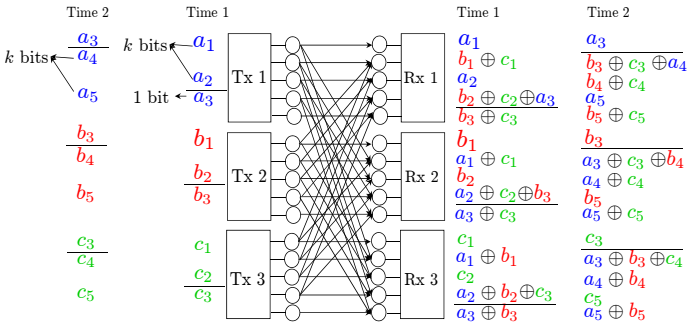

Fig. 8. Achievable scheme for the $(r-1, r)$ model where $r = 2k+1$.

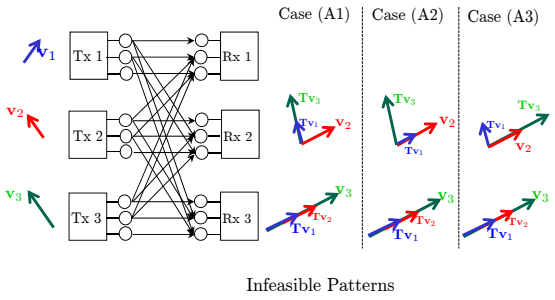

Fig. 9. Infeasible patterns of received signals for $L = 3$. Once perfect function alignment is achieved at receiver 3, any vectors cannot be aligned at the other receivers.

again invoke the vector coding idea. At the end of time 1, each transmitter sends an additional symbol on the last *even*-numbered level. This transmission causes a conflict at each receiver. For example, a_3 has a conflict with $b_2 \oplus c_2$. However, this can be resolved by using another time slot. In time 2, using the first level, each transmitter re-sends the symbol that was sent on the last even-numbered level in time 1. From the second to last levels, we repeat the same procedure as in time 1 to send fresh k symbols. Note that a_3 is cleanly received at receiver 1 in time 2. This a_3 can now be used to decode $b_2 \oplus c_2$, which was interfered with by a_3 in time 1. Also the $b_3 \oplus c_3$ that was received at receiver 1 in time 1 can be used to decode a_4 which is interfered with by $b_3 \oplus c_3$ on the second level in time 2. Therefore, we can achieve $R_{\mathsf{comp}} = \frac{k+1+k}{2} = \frac{r}{2}$. The same strategy can be applied to arbitrary values of $(L, r = 2k+1)$ as well as the $(2b)$ model. ∎

Using Theorem 4 and Lemma 4, we can now prove the achievability. We focus on the case of $m < n$. The other case of $m > n$ similarly follows. For $0 \le \alpha \le \frac{1}{2}$, (13) gives $r = 0$ and $a = m$, thus the decomposition is given by $(m, n) = (0,1)^{n-2m} \times (1,2)^m$. Therefore, using Lemma 4, we can achieve $R_{\mathsf{comp}} = 0 \cdot (n-2m) + 1 \cdot m = m$. Next, consider the case of $\frac{1}{2} \le \alpha \le \frac{2}{3}$. Using (13), we find that $r = 1$ and $a = 2m - n$, hence, the decomposition is given by $(m, n) = (1,2)^{2n-3m} \times (2,3)^{2m-n}$. Using Lemma 4, we can achieve $R_{\mathsf{comp}} = 1 \cdot (2n-3m) + \frac{3}{2} \cdot (2m-n) = \frac{1}{2}n$. Finally, consider the case of $\alpha \ge \frac{2}{3}$. From (13), we know that $r \ge 2$. So we get $R_{\mathsf{comp}} = \frac{1}{2}n$.

B. Converse Proof under Linear Coding Strategies

Straightforward cut-set arguments give $R_{\mathsf{comp}} \le \min\{m, n\}$; hence, it suffices to prove that $R_{\mathsf{comp}}^{\mathsf{lin}} \le \frac{\max(m,n)}{2}$. Consider any vector linear code over N uses of the network. Denoting the vector of K successive bits of user ℓ by S_ℓ^K, this means that the transmitted signals can be written as

$$\sum_{i=1}^{K} \mathbf{v}_{\ell,i} S_{\ell,i} = \mathbf{V}_\ell S_\ell^K,$$

where $\mathbf{v}_{\ell,j}$ are the "beamforming" vectors of length $N\max(m,n)$ to be chosen optimally, and $\mathbf{V}_\ell \in \mathbb{F}_2^{N\max(m,n) \times K}$ is the matrix composed of the K beamforming vectors of transmitter ℓ. Assuming that K computation bits are successfully decoded by linear decoding at all receivers, our aim is to prove that $R_{\mathsf{comp}}^{\mathsf{lin}} := \frac{K}{N} \le \frac{\max(m,n)}{2}$.

Our proof relies on a dimensionality argument, evaluated from the receivers' perspective. To formulate our argument, we define the following space:

$$W_{i,\ell} = \mathsf{span}\{\mathbf{T}\mathbf{v}_{1,i}, \ldots, \mathbf{T}\mathbf{v}_{\ell-1,i}, \mathbf{v}_{\ell,i}, \mathbf{T}\mathbf{v}_{\ell+1,i}, \ldots, \mathbf{T}\mathbf{v}_{L,i}\}.$$

Intuitively, this is the space taken up by the ith computed bit at receiver ℓ. First, we observe the following fact:

Lemma 5: For every receiver ℓ ($\ell = 1, 2, \ldots, L$), the subspaces $W_{i,\ell}$, for $i = 1, 2, \ldots, K$, must be linearly independent, i.e., for every i, j and ℓ, we have $W_{i,\ell} \cap W_{j,\ell} = \mathbf{0}$, where $\mathbf{0}$ is the all-zero vector.

Proof: This lemma can be proved by contradiction: suppose there exists $j \ne i$ such that $W_{j,\ell}$ and $W_{i,\ell}$ are not linearly independent subspaces. Then, it is not possible to guarantee that the computed bits i and j can be decoded without error. ∎

The following lemma relates the dimensionality of the various $W_{i,\ell}$:

Lemma 6: For any i, if there exists ℓ such that $\dim(W_{i,\ell}) = 1$, then for all $m \ne \ell$, we must have $\dim(W_{i,m}) > 2$.

Proof: See Appendix E. ∎

This lemma says that for any bit i for which function alignment is perfectly achieved at some receiver ℓ (i.e., $\mathsf{dim}(W_{i,\ell}) = 1$), then for all other receivers, this same bit i must take up at least 3 dimensions. For illustration, Fig. 9 shows some examples of these infeasible patterns when $L = 3$.

We can restate this lemma in the following way. For any bit i, one of these two alternatives must apply:

(B) : There exists ℓ such that $\mathsf{dim}(W_{i,\ell}) = 1$. Then, $\mathsf{dim}(W_{i,m}) \ge 3$ for all $m \ne \ell$.

(C) : There does not exist ℓ such that $\mathsf{dim}(W_{i,\ell}) = 1$. Then, $\mathsf{dim}(W_{i,m}) \ge 2$ for all m.

For illustration, Fig. 10 shows examples. Cases (B1)-(B3) are the ones where the dimension for any other receiver (except the perfect-alignment receiver) is exactly 3. Cases (C1)-(C3) are the ones where the dimension is exactly 2 for all of the receivers.

Let us introduce the set $\mathcal{J}$ of those indices i for which there exists ℓ such that $\mathsf{dim}(W_{i,\ell}) = 1$. For every $i \in \mathcal{J}$, we have

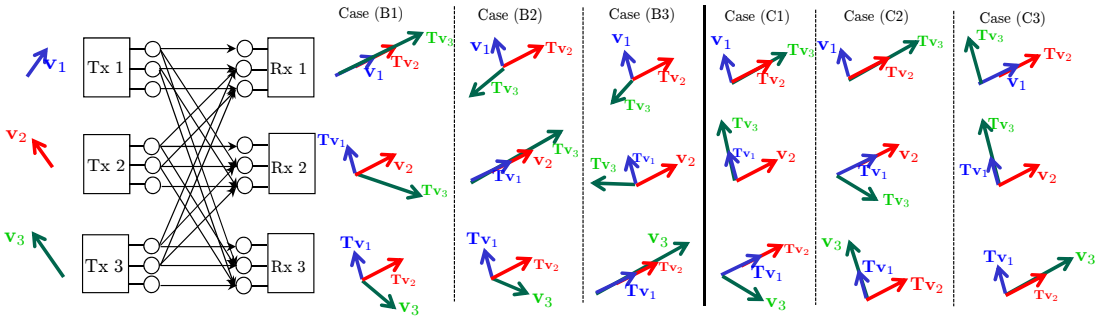

Fig. 10. Feasible patterns of received signals for $L = 3$. Cases (B1)-(B3) are the ones where the dimension of the linear subspace for any other receiver (except for the perfect-alignment receiver) is exactly 3. Cases (C1)-(C3) are the ones where the dimension is exactly 2 for all of the receivers.

(by case (B) above)

$$\sum_{m=1}^{L} \mathsf{dim}\left(W_{i,m}\right) \geq 1 + 3(L-1) = 3L - 2. \tag{16}$$

Let us also denote the complement of the set $\mathcal{J}$ (in the set of integers between 1 and K) by $\mathcal{J}^c$. For every $i \in \mathcal{J}^c$, we have (by case (C) above)

$$\sum_{m=1}^{L} \mathsf{dim}\left(W_{i,m}\right) \geq 2L. \tag{17}$$

Now, since $K = |\mathcal{J}| + |\mathcal{J}^c|$, and by definition, $R_{\mathsf{comp}}^{\mathsf{lin}} = K/N$, we can write

$$\begin{aligned}
&2LNR_{\mathsf{comp}}^{\mathsf{lin}} = 2L\left(|\mathcal{J}| + |\mathcal{J}^c|\right) \\
&\leq |\mathcal{J}|(3L-2) + |\mathcal{J}^c|2L \\
&\overset{(a)}{\leq} \sum_{i\in\mathcal{J}}\sum_{m=1}^{L} \mathsf{dim}\left(W_{i,m}\right) + \sum_{i\in\mathcal{J}^c}\sum_{m=1}^{L} \mathsf{dim}\left(W_{i,m}\right) \\
&= \sum_{m=1}^{L}\sum_{i=1}^{K} \mathsf{dim}\left(W_{i,m}\right) \\
&\overset{(b)}{\leq} \sum_{m=1}^{L} N\max(m,n) = LN\max(m,n),
\end{aligned}$$

where (a) follows from Lemma 6 (rewritten as in Equations (16) and (17)), and (b) follows because for each receiver ℓ, the subspaces $W_{i,\ell}$ must be linearly independent (Lemma 5) and their total dimensionality cannot exceed the total number of dimensions available at receiver ℓ over N channel uses, which is $N\max(m,n)$. Therefore, we get $R_{\mathsf{comp}}^{\mathsf{lin}} \leq \frac{\max(m,n)}{2}$. This concludes the proof.

C. Proof of Upper Bound

The following upper bound is a generalized version of the 2×2 case bound (3). Starting with Fano's inequality, we get

$$\begin{aligned}
&N((2L-1)R_{\mathsf{comp}} - \epsilon_N) \\
&\leq \sum_{\ell=1}^{L} I(\bigoplus S_i^K; Y_\ell^N) + (L-1)I(\bigoplus S_i^K; Y_1^N) \\
&\overset{(a)}{\leq} \sum_{\ell=1}^{L}\left[H(Y_\ell^N) - H(Y_\ell^N|\bigoplus S_i^K, [Y_j^N]_{j=1}^{\ell-1})\right] \\
&\quad + (L-1)I(\bigoplus S_i^K; Y_1^N) \\
&\overset{(b)}{\leq} \sum_{\ell=1}^{L} H(Y_\ell^N) - \left\{H(\mathcal{Y}|\bigoplus S_i^K) + \sum_{\ell=2}^{L} I(\bigoplus S_i^K; Y_1^N, \bar{\mathcal{S}}_\ell)\right\} \\
&\overset{(c)}{=} \sum_{\ell=1}^{L} H(Y_\ell^N) - \left\{H(\mathcal{Y}|\bigoplus S_i^K) + \sum_{\ell=2}^{L} H(\mathbf{T}_{\ell 1}X_\ell^N|\bar{\mathcal{S}}_\ell)\right\} \\
&\overset{(d)}{=} \sum_{\ell=1}^{L} H(Y_\ell^N) \\
&\quad - \left\{H(\mathcal{Y}, [\mathbf{T}_{\ell 1}X_\ell^N]_{\ell=2}^{L}|\bigoplus S_i^K) + \sum_{\ell=2}^{L} H(\mathbf{T}_{\ell 1}X_\ell^N|\bar{\mathcal{S}}_\ell)\right\} \\
&\overset{(e)}{\leq} \sum_{\ell=1}^{L} H(Y_\ell^N) - \left\{H([\mathbf{T}_{\ell 1}X_\ell^N]_{\ell=2}^{L}) + \sum_{\ell=2}^{L} H(\mathbf{T}_{\ell 1}X_\ell^N|\bar{\mathcal{S}}_\ell)\right\} \\
&\overset{(f)}{=} \sum_{\ell=1}^{L} H(Y_\ell^N) - \left\{\sum_{\ell=2}^{L} H(\mathbf{T}_{\ell 1}X_\ell^N) + \sum_{\ell=2}^{L} H(\mathbf{T}_{\ell 1}X_\ell^N|\bar{\mathcal{S}}_\ell)\right\} \\
&\leq \sum_{\ell=1}^{L} H(Y_\ell^N) \leq NL\max(m,n)
\end{aligned}$$

where (a) follows from the fact that conditioning reduces entropy; (b) follows from non-negativity of mutual information, $\mathcal{Y} := [Y_j^N]_{j=1}^L$, and $\bar{\mathcal{S}}_\ell := [S_i^K]_{i=1}^L \setminus S_\ell^K$; (c) follows from the fact that S_ℓ^K's are mutually independent; (d) follows from the fact that X_ℓ is a function of $\mathcal{Y}$ (see Claim 1 below); (e) follows from the fact that $(S_2^K, \cdots, S_L^K)$ is independent of $\bigoplus_i S_i^K$; (f) follows from the fact that S_ℓ^K's are mutually independent;

Claim 1: For $m \neq n$, X_ℓ is a function of $[Y_i]_{i=1}^L$, $\ell = 1, \cdots, L$.

Proof: Consider the case of $m < n$. From (15), we get

$$\bigoplus_{i=1}^{L} Y_i = \left\{\mathbf{I} \oplus (1 \oplus L)\mathbf{G}^{n-m}\right\} \bigoplus_{i=1}^{L} X_i$$
$$\bigoplus_{i=2}^{L} Y_i = \left(\mathbf{I} \oplus \mathbf{G}^{n-m}\right) \bigoplus_{i=2}^{L} X_i \oplus \left\{(L-1)\mathbf{G}^{n-m}\right\} \bigoplus_{i=1}^{L} X_i.$$

Straightforward computation gives

$$X_1 = \mathbf{A}^{-1} \bigoplus_{i=1}^{L} Y_i \oplus \mathbf{B}^{-1} \left[\bigoplus_{i=2}^{L} Y_i \oplus \left\{(L-1)\mathbf{G}^{n-m}\right\} \left\{ \mathbf{A}^{-1} \bigoplus_{i=1}^{L} Y_i \right\} \right],$$

where $\mathbf{A} := \{\mathbf{I} \oplus (1 \oplus L)\mathbf{G}^{n-m}\}$ and $\mathbf{B} := (\mathbf{I} \oplus \mathbf{G}^{n-m})$, both of which are invertible since $m \neq n$. Hence, X_1 is a function of $[Y_i]_{i=1}^{L}$. By symmetry, X_ℓ is a function of $[Y_i]_{i=1}^{L}$, $\ell = 2, \cdots, L$. Similarly we can show this for the case of $m > n$. ∎

VII. Discussion

A. Multi-hop Networks

In [17], [11], [18], function multicasting has been explored in the context of multi-hop networks. While some interesting relationship between sum-network and multiple-unicast networks was found in [11], determining the computing capacity in general has been open. For two-source L-destination or L-source two-destination networks, the computing capacity was established only when the entropy of each source is constrained to be 1 [18].

While in this work we remove the entropy constraint of sources, the network model we consider here is somewhat specialized and also restricted to a single-hop network. But we expect that our results will shed some lights on arbitrary multi-hop networks. 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.

B. Role of Feedback for Computation

The role of feedback for function computation has initially been studied in [27] where it is shown that feedback can increase the computing rate. Interestingly the feedback gain is shown to be significant - qualitatively similar to the gain in the two-user Gaussian interference channel [28], which revealed an unbounded feedback gain: the gap between nonfeedback and feedback capacities can be arbitrarily large as the signal-to-noise ratio of each link increases. However, the result of [27] relies on a separation approach that naturally comes in the course of characterizing feedback multicast capacity.

Our future work is characterizing the feedback computing capacity of the networks considered herein to explore whether we can do better than the separation approach. This will provide a deeper understanding of the feedback gain. Moreover it would be more interesting to explore this feedback gain under more realistic scenarios where feedback is offered through rate-limited bit-piped links [29] or through the corresponding backward communication network [30]. Furthermore, we are interested in extending to more general multi-hop networks [31].

VIII. Conclusion

We have established the computing 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. We also characterized the linear computing capacity of an L-transmitter L-receiver symmetric network. We developed a new achievable scheme and derive new upper bounds. Furthermore we established a network decomposition theorem that provides an alternative but conceptually-simpler achievability proof. We expect that the network-decomposition-based framework would play a role in extending to arbitrary multi-hop networks.

Appendix A
Proof of Lemma 1

A. Direct Part →

Without loss of generality, assume that $n_{11} - n_{12} = n_{21} - n_{22} \geq 0$. We can then get:

$$\mathbf{G}^{n_{21}-n_{22}} Y_1 \oplus Y_2 = (\mathbf{G}^{q-n_{11}+n_{21}-n_{22}} \oplus \mathbf{G}^{q-n_{12}}) X_1.$$

For the obvious reason, $X_1 \in \mathbb{F}_2^q$ contains nontrivial values only on the top n_{11} levels:

$$X_1 = \begin{bmatrix} \tilde{X}_1 \\ \mathbf{0}_{q-n_{11}} \end{bmatrix} \in \mathbb{F}_2^q,$$

where $\tilde{X}_1 \in \mathbb{F}_2^{n_{11}}$. Using this expression, we can rewrite the above as:

$$\mathbf{G}^{n_{21}-n_{22}} Y_1 \oplus Y_2 = \begin{bmatrix} \mathbf{0}_{q-n_{11}} \\ (\mathbf{G}_{n_{11}}^{n_{21}-n_{22}} \oplus \mathbf{G}_{n_{11}}^{n_{11}-n_{12}}) \tilde{X}_1 \end{bmatrix}.$$

where $G_{n_{11}}$ indicates an n_{11}-by-n_{11} shift matrix. Since $n_{11} - n_{12} = n_{21} - n_{22}$, we have $\mathbf{G}_{n_{11}}^{n_{21}-n_{22}} \oplus \mathbf{G}_{n_{11}}^{n_{11}-n_{12}} = \mathbf{0}$, and therefore any $G^{q-n_{1j}} X_1$ cannot be reconstructed. Similarly one can show that any $G^{q-n_{2j}} X_2$ cannot be reconstructed from (Y_1, Y_2) by considering $\mathbf{G}^{n_{11}-n_{12}} Y_1 \oplus Y_2$. Hence, a network is degenerate.

B. Converse Part ←

We will show that if $n_{11} - n_{12} \neq n_{21} - n_{22}$, then a network is non-degenerate. Consider the following four cases:

$$\text{Case I}: n_{12} \leq n_{11},\ n_{21} \leq n_{22}$$
$$\text{Case II}: n_{12} \geq n_{11},\ n_{21} \geq n_{22}$$
$$\text{Case III}: n_{12} \leq n_{11},\ n_{21} \geq n_{22}$$
$$\text{Case IV}: n_{12} \geq n_{11},\ n_{21} \leq n_{22}.$$

Note that Case I and Case II are symmetric, so are Case III and Case IV. Hence, we focus on Case I and III.

Case I $(n_{12} \leq n_{11}, n_{21} \leq n_{22})$: Consider

$$Y_1 \oplus \mathbf{G}^{n_{22}-n_{21}} Y_2 = (\mathbf{G}^{q-n_{11}} \oplus \mathbf{G}^{q+n_{22}-n_{21}-n_{12}}) X_1$$
$$= \begin{bmatrix} \mathbf{0}_{q-n_{11}} \\ (\mathbf{I}_{n_{11}} \oplus \mathbf{G}_{n_{11}}^{n_{11}+n_{22}-n_{21}-n_{12}}) \tilde{X}_1 \end{bmatrix}.$$

Since $n_{11}-n_{12} \neq n_{21}-n_{22}$, $\mathbf{G}_{n_{11}}^{n_{11}+n_{22}-n_{21}-n_{12}} \neq \mathbf{I}_{n_{11}}$ and therefore $\mathbf{I}_{n_{11}} \oplus \mathbf{G}_{n_{11}}^{n_{11}+n_{22}-n_{21}-n_{12}}$ is invertible. This implies that X_1 is a function of (Y_1, Y_2). Hence, a network is non-degenerate.

Case III $(n_{12} \leq n_{11}, n_{21} \geq n_{22})$: First consider the case of $n_{21}-n_{22} > n_{11}-n_{12}$. We then get

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

Since $n_{11}-n_{12} \neq n_{21}-n_{22}$, $\mathbf{G}_{n_{11}}^{n_{21}-n_{22}} \neq \mathbf{G}_{n_{11}}^{n_{11}-n_{12}}$. This implies that $\mathbf{G}_{n_{11}}^{n_{11}-n_{12}}\tilde{X}_1$ is decodable and therefore $\mathbf{G}^{q-n_{12}}X_1$ is decodable. Hence, the network is non-degenerate. We now consider the other case of $n_{21}-n_{22} < n_{11}-n_{12}$. We then get

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

Since $n_{11}-n_{12} \neq n_{21}-n_{22}$, $\mathbf{G}_{n_{21}}^{n_{11}-n_{12}} \neq \mathbf{G}_{n_{21}}^{n_{21}-n_{22}}$. This implies that $\mathbf{G}_{n_{21}}^{n_{21}-n_{22}}\tilde{X}_2$ is decodable and therefore $\mathbf{G}^{q-n_{22}}X_2$ is decodable. Hence, the network is non-degenerate.

Appendix B
Proof of Lemma 2

Using (9), (10) and (11), we compute:

$$\begin{aligned}
\mathbf{TV} &= \mathbf{I}_3 \otimes [\mathbf{e}_{(n-m)+1}^{(n)} \ \cdots \ \mathbf{e}_{2(n-m)}^{(n)}] \\
\mathbf{T}^2\mathbf{V} &= \mathbf{I}_3 \otimes [\mathbf{e}_{2(n-m)+1}^{(n)} \ \cdots \ \mathbf{e}_{3(n-m)}^{(n)}] \\
\mathbf{TP}_1 &= \mathbf{e}_3^{(3)} \otimes [\mathbf{e}_{2(n-m)+1}^{(n)} \ \cdots \ \mathbf{e}_m^{(n)}] \\
&\quad \oplus \mathbf{e}_2^{(3)} \otimes \left\{ [\mathbf{e}_{3(n-m)+1}^{(n)} \ \cdots \ \mathbf{e}_n^{(n)}] \oplus [\mathbf{e}_{4(n-m)+1}^{(n)} \ \cdots \ \mathbf{0}] \right\} \\
\mathbf{T}^2\mathbf{P}_1 &= \mathbf{e}_3^{(3)} \otimes [\mathbf{e}_{3(n-m)+1}^{(n)} \ \cdots \ \mathbf{e}_n^{(n)}] \\
&\quad \oplus \mathbf{e}_2^{(3)} \otimes [\mathbf{e}_{4(n-m)+1}^{(n)} \ \cdots \ \mathbf{0}] \\
\mathbf{TP}_2 &= \mathbf{e}_3^{(3)} \otimes [\mathbf{e}_{2(n-m)+1}^{(n)} \ \cdots \ \mathbf{e}_m^{(n)}] \\
&\quad \oplus \mathbf{e}_1^{(3)} \otimes \left\{ [\mathbf{e}_{3(n-m)+1}^{(n)} \ \cdots \ \mathbf{e}_n^{(n)}] \oplus [\mathbf{e}_{4(n-m)+1}^{(n)} \ \cdots \ \mathbf{0}] \right\} \\
\mathbf{T}^2\mathbf{P}_2 &= \mathbf{e}_3^{(3)} \otimes [\mathbf{e}_{3(n-m)+1}^{(n)} \ \cdots \ \mathbf{e}_n^{(n)}] \\
&\quad \oplus \mathbf{e}_1^{(3)} \otimes [\mathbf{e}_{4(n-m)+1}^{(n)} \ \cdots \ \mathbf{0}].
\end{aligned}$$

With the Gaussian elimination method, we can show that

$$\begin{aligned}&\mathsf{span}\left[\mathbf{V}_1,\ \mathbf{TV}_2,\ \mathbf{T}^2\mathbf{V}_1\right] \\ &= \mathsf{span}\left[\mathbf{V}, \mathbf{TV}, \mathbf{T}^2\mathbf{V}, \mathbf{P}_1, \mathbf{T}^2\mathbf{P}_1, \mathbf{TP}_2\right] \\ &= \mathsf{span}\left[\mathbf{I}_3 \otimes \mathbf{I}_n\right].\end{aligned}$$

Hence, $\mathsf{rank}\left[\mathbf{V}_1,\ \mathbf{TV}_2,\ \mathbf{T}^2\mathbf{V}_1\right] = 3n$. Similarly we can show that $\mathsf{rank}\left[\mathbf{V}_2,\ \mathbf{TV}_1,\ \mathbf{T}^2\mathbf{V}_2\right] = 3n$.

Appendix C
Proof of Theorem 4

For Part *(1)*, consider the (km, kn) model. The proof uses graph coloring with k colors, identified by integers $\{0, 1, \ldots, k-1\}$. At transmitter 1, assign to level p (for $p = 1, 2, \ldots, k\max(m,n)$) the color $(p-1) \mod k$. Use exactly the same rule to color the vertices of receiver 1 as well as the transmitters and receivers of the remaining $(L-1)$ users. It is seen by inspection that each color represents an independent graph. Moreover, each color represents precisely an (m, n) model.

For Part *(2)*, we use graph coloring with 2 colors. At all transmitters and receivers, assign one color to the even-numbered levels and the other color to the odd-numbered levels. By inspection, it can be verified that each color represents an independent graph. Moreover, one color represents an (m, n) model and the other represents an $(m+1, n+1)$ model.

For Part *(3)*, we use graph coloring with $|n-m|$ colors, identified by integers $\{0, 1, \ldots, |n-m|-1\}$. At transmitter 1, assign to level p (for $p = 1, 2, \ldots, \max(m,n)$) the color $(p-1) \mod |n-m|$. Use exactly the same rule to color the levels of receiver 1 as well as the transmitters and receivers of the remaining $(L-1)$ users. It is seen by inspection that each color represents an independent graph. A tedious but straightforward calculation shows that of the resulting $|n-m|$ independent graphs, there are a number of models $(r+1, r+2)$ and $n-m-a$ number of models $(r, r+1)$, with the claimed values for r and a.

Appendix D
Proof of Lemma 3

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, we will find it convenient to denote the vector of (binary) channel inputs used by transmitter 1 as $(X_{1,1}, X_{1,2}, \ldots)^t$ and the one used by transmitter 2 as $(X_{2,1}, X_{2,2}, \ldots)^t$.

(i) The case $r = 3\ell - 1$, for any integer $\ell \geq 1$. For $\ell = 1$ (hence, $r = 2$), an explicit code is as follows: $X_{1,1} = a_1, X_{1,2} = a_2, X_{1,3} = 0$ and $X_{2,1} = b_2, X_{2,2} = b_1, X_{2,3} = 0$. It is straightforward to verify that both receivers can reconstruct $a_1 \oplus b_1$ and $a_2 \oplus b_2$, hence, a computation rate of 2 is attained. For the general case, 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 $k = 1, 2, \ldots, \ell$. Each receiver can reconstruct all 2ℓ sums $a_k \oplus b_k$ and thus, the computation rate is $2\ell = \frac{2}{3}(r+1)$.

(ii) The case $r = 3\ell$, for any integer $\ell \geq 1$, derives easily once we have an explicit code for $\ell = 1$, i.e., for the $(3, 4)$ model. For the $(3, 4)$ model, consider coding over 3 channel uses, which corresponds (by network decomposition) to the $(9, 12)$ model. An explicit code can be found for example via the construction (9), (10), (11) in Section IV, leading to 8 computations. For $\ell \geq 2$, we consider a vector linear code over 3 channel uses, and thus, the $(9\ell, 9\ell+3)$ model. For this model, we set $X_{1,9k-8} = a_{6k-5}, X_{1,9k-7} = a_{6k-3}, X_{1,9k-6} = a_{6k-1}, X_{1,9k-5} = a_{6k-4}, X_{1,9k-4} = a_{6k-2}, X_{1,9k-6} =$

$a_{6k}, X_{1,9k-2} = X_{1,9k-1} = X_{1,9k} = 0$ and $X_{2,9k-8} = a_{6k-4}, X_{2,9k-7} = a_{6k-2}, X_{2,9k-6} = a_{6k}, X_{2,9k-5} = a_{6k-5}, X_{2,9k-3} = a_{6k-3}, X_{2,9k-6} = a_{6k-1}, X_{2,9k-2} = X_{2,9k-1} = X_{2,9k} = 0$ for $k = 1, 2, \ldots, \ell - 1$. It is easily verified by inspection that each receiver can recover all $6(\ell-1)$ modulo sums. Additionally, we observe that this code does not involve or affect any of the last 12 positions at transmitters or receivers. These last 12 positions constitute exactly a $(9, 12)$ model, for which we already know that an additional 8 computations are achievable. This gives a total of $6(\ell-1)+8 = 2(3\ell+1)$ computations. Thus, per channel use, the computation rate is $\frac{2}{3}(3\ell+1) = \frac{2}{3}(r+1)$.

(iii) The case $r = 3\ell + 1$, for any integer $\ell \geq 1$, derives easily once we have an explicit code for $\ell = 1$, i.e., for the $(4, 5)$ model. For the $(4, 5)$ model, consider coding over 3 channel uses, which corresponds (by network decomposition) to the $(12, 15)$ model. An explicit code can be found for example via the construction (9), (10), (11) in Section IV, leading to 10 computations. For $\ell \geq 2$,, we consider a vector linear code over 3 channel uses, and thus, the $(9\ell+3, 9\ell+6)$ model. For this model, we set $X_{1,9k-8} = a_{6k-5}, X_{1,9k-7} = a_{6k-3}, X_{1,9k-6} = a_{6k-1}, X_{1,9k-5} = a_{6k-4}, X_{1,9k-4} = a_{6k-2}, X_{1,9k-6} = a_{6k}, X_{1,9k-2} = X_{1,9k-1} = X_{1,9k} = 0$ and $X_{2,9k-8} = a_{6k-4}, X_{2,9k-7} = a_{6k-2}, X_{2,9k-6} = a_{6k}, X_{2,9k-5} = a_{6k-5}, X_{2,9k-3} = a_{6k-3}, X_{2,9k-6} = a_{6k-1}, X_{2,9k-2} = X_{2,9k-1} = X_{2,9k} = 0$ for $k = 1, 2, \ldots, \ell - 1$. It is easily verified by inspection that each receiver can recover all $6(\ell-1)$ modulo sums. Additionally, we observe that this code does not involve or affect any of the last 15 positions at transmitters or receivers. These last 15 positions constitute exactly a $(12, 15)$ model, for which we already know that an additional 10 computations are achievable. This gives a total of $6(\ell-1)+10 = 2(3\ell+2)$ computations. Thus, per channel use, the computation rate is $\frac{2}{3}(3\ell+2) = \frac{2}{3}(r+1)$.

APPENDIX E
PROOF OF LEMMA 6

We focus on the case of $m < n$. The other case similarly follows. By assumption, for the considered i, we have that there exists ℓ such that

$$\mathsf{dim}\left(W_{i,\ell}\right) = 1.$$

To simplify notation, and without loss of generality, let us suppose that this holds for the last receiver, that is, for $\ell = L$. But this also trivially implies that

$$\mathsf{span}(\mathbf{v}_{L,i}) = \mathsf{span}(\mathbf{T}\mathbf{v}_{1,i}) = \cdots = \mathsf{span}(\mathbf{T}\mathbf{v}_{L-1,i}), \quad (18)$$

where $\mathbf{T} := \mathbf{I}_N \otimes \mathbf{G}^{n-m}$. Now, consider the subspace $W_{i,1}$ at receiver 1. We have that

$$\begin{aligned} W_{i,1} &= \mathsf{span}[\mathbf{v}_{1,i},\ \mathbf{T}\mathbf{v}_{2,i},\ \cdots, \mathbf{T}\mathbf{v}_{L-1,i},\ \mathbf{T}\mathbf{v}_{L,i}] \\ &= \mathsf{span}[\mathbf{v}_{1,i},\ \mathbf{T}\mathbf{v}_{1,i},\ \mathbf{T}^2\mathbf{v}_{1,i}] \end{aligned}$$

where the last equality is due to (18). Since we assume that $\bigoplus_\ell S_{\ell i}$ is decodable at all receivers, any individual symbol must appear at all receivers. This implies that

$$\begin{aligned} &\mathsf{dim}(\mathsf{span}(\mathbf{v}_{1,i})) \\ &= \mathsf{dim}(\mathsf{span}(\mathbf{T}\mathbf{v}_{1,i})) \\ &= \mathsf{dim}(\mathsf{span}(\mathbf{T}^2\mathbf{v}_{1,i})) \\ &= 1. \end{aligned}$$

The key observation here is that for $m \neq n$, these subspaces are linearly independent:

$$\mathsf{dim}\left(\mathsf{span}[\mathbf{v}_{1,i},\ \mathbf{T}\mathbf{v}_{1,i},\ \mathbf{T}^2\mathbf{v}_{1,i}]\right) = 3.$$

We can apply the same argument for the other receivers to complete the proof.

REFERENCES

[1] A. Giridhar and P. R. Kumar, "Computing and communicating functions over sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 23, pp. 755–764, Apr. 2005.

[2] A. G. Dimakis, P. B. Godfrey, Y. Wu, M. Wainwright, and K. Ramchandran, "Network coding for distributed storage systems," *IEEE Transactions on Information Theory*, vol. 56, pp. 4539–4551, Sept. 2010.

[3] A. G. Dimakis, K. Ramchandran, Y. Wu, and C. Suh, "A survey on network codes for distributed storage," *Proceedings of the IEEE*, vol. 99, pp. 476–489, Mar. 2011.

[4] R. Appuswamy, M. Franceschetti, N. Karamchandani, and K. Zeger, "Network coding for computing: Cut-set bounds," *IEEE Transactions on Information Theory*, vol. 57, pp. 1015–1030, Feb. 2011.

[5] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, pp. 1204–1216, July 2000.

[6] S.-Y. R. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Transactions on Information Theory*, vol. 49, pp. 371–381, Feb. 2003.

[7] R. Koetter and M. Médard, "An algebraic approach to network coding," *IEEE/ACM Transactions on Networking*, vol. 11, pp. 782–795, Oct. 2003.

[8] T. Ho, M. Médard, R. Koetter, D. R. Karger, M. Effros, J. Shi, and B. Leong, "A random linear network coding approach to multicast," *IEEE Transactions on Information Theory*, vol. 52, pp. 4413–4430, Oct. 2006.

[9] S. Jaggi, P. Sanders, P. Chou, M. Effros, S. Egner, K. Jain, and L. Tolhuizen, "Polynomial time algorithms for multicast network code construction," *IEEE Transactions on Information Theory*, vol. 51, pp. 1973–1982, June 2005.

[10] B. Nazer and M. Gastpar, "Computation over multiple-access channels," *IEEE Transactions on Information Theory*, vol. 53, pp. 3498–3516, Oct. 2007.

[11] B. Rai and B. Dey, "On network coding for sum-networks," *IEEE Transactions on Information Theory*, vol. 58, pp. 50–63, Jan. 2012.

[12] R. Dougherty, C. Freiling, and K. Zeger, "Insufficiency of linear coding in network information flow," *IEEE Transactions on Information Theory*, vol. 51, pp. 2745–2759, Aug. 2005.

[13] S. Avestimehr, S. Diggavi, and D. Tse, "Wireless network information flow: A deterministic approach," *IEEE Transactions on Information Theory*, vol. 57, pp. 1872–1905, Apr. 2011.

[14] U. Niesen, B. Nazer, and P. Whiting, "Computation alignment: Capacity approximation without noise accumulation," *arXiv:1108.6312*, Aug. 2011.

[15] M. A. Maddah-Ali, S. A. Motahari, and A. K. Khandani, "Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis," *IEEE Transactions on Information Theory*, vol. 54, pp. 3457–3470, Aug. 2008.

[16] V. R. Cadambe and S. A. Jafar, "Interference alignment and the degree of freedom for the K user interference channel," *IEEE Transactions on Information Theory*, vol. 54, pp. 3425–3441, Aug. 2008.

[17] A. Ramamoorthy, "Communicating the sum of sources over a network," *Proceedings of the IEEE International Symposium on Information Theory*, pp. 1646–1650, July 2008.

[18] A. Ramamoorthy and M. Langberg, "Communicating the sum of sources over a network," *arXiv:1001.5319*, Jan. 2010.

[19] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference through structured codes," *IEEE Transactions on Information Theory*, vol. 57, pp. 6463–6486, Oct. 2011.

[20] C. Suh and D. Tse, "Interference alignment for cellular networks," *Allerton Conference on Control, Computing and Communication*, Sept. 2008.
[21] C. Suh, M. Ho, and D. Tse, "Downlink interference alignment," *IEEE Transactions on Communications*, vol. 59, pp. 2616–2626, Sept. 2011.
[22] Y. Wu and A. G. Dimakis, "Reducing repair traffic for erasure coding-based storage via interference alignment," *Proceedings of the IEEE International Symposium on Information Theory, Seoul, Korea*, July 2009.
[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 Transactions on Information Theory*, vol. 58, pp. 2134–2158, Apr. 2012.
[24] C. Suh and K. Ramchandran, "Exact-repair MDS code construction using interference alignment," *IEEE Transactions on Information Theory*, vol. 57, pp. 1425–1442, Mar. 2011.
[25] A. Das, S. Vishwanath, S. Jafar, and A. Markopoulou, "Network coding for multiple unicasts: An interference alignment approach," *Proceedings of the IEEE International Symposium on Information Theory*, June 2010.
[26] A. Ramakrishnan, A. Das, H. Maleki, A. Markopoulou, S. A. Jafar, and S. Vishwanath, "Network coding for three unicast sessions: Interference alignment approaches," *Allerton Conference on Control, Computing and Communication*, Sept. 2010.
[27] C. Suh, N. Goela, and M. Gastpar, "Approximate feedback capacity of the Gaussian multicast channel," *Proceedings of the IEEE International Symposium on Information Theory*, July 2012.
[28] C. Suh and D. Tse, "Feedback capacity of the Gaussian interference channel to within 2 bits," *IEEE Transactions on Information Theory*, vol. 57, pp. 2667–2685, May 2011.
[29] A. Vahid, C. Suh, and A. S. Avestimehr, "Interference channels with rate-limited feedback," *IEEE Transactions on Information Theory*, vol. 58, pp. 2788–2812, May 2012.
[30] C. Suh, I.-H. Wang, and D. Tse, "Two-way interference channels," *Proceedings of the IEEE International Symposium on Information Theory*, July 2012.
[31] I.-H. Wang, "On two unicast wireless networks with destination-to-source feedback," *Proceedings of the IEEE International Symposium on Information Theory*, July 2012.