Fiftieth Annual Allerton Conference
Allerton House, UIUC, Illinois, USA
October 1 - 5, 2012

Computation in Multicast Networks: Function Alignment and Converse Theorems

Changho Suh
KAIST, Daejeon, South Korea
chsuh@kaist.ac.kr

Naveen Goela
[♯]UC-Berkeley, Berkeley, CA, USA
ngoela@eecs.berkeley.edu

Michael Gastpar[♯]
EPFL, Lausanne, Switzerland
michael.gastpar@epfl.ch

***Abstract*—We characterize the computing capacity of a two-transmitter two-receiver linear deterministic network where both receivers wish to compute a modulo-2 sum of two Bernoulli sources generated at the two transmitters. We develop a new achievable scheme that we call *function alignment*, inspired by the concept of *interference alignment*, and derive a new upper bound to establish the computing capacity. As a consequence, we find that unlike the single-receiver function-unicasting case, the cutset-based bound is not tight in general when multicasting a linear function. Moreover we develop a *network decomposition theorem* to find elementary subnetworks that can constitute an original network without loss of optimality. This serves to provide a conceptually-simpler achievability proof as well as generalize to L-transmitter L-receiver networks.**

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 sensors. 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.* [4] studied the fundamental limits of computation for linear and general target function classes. While limited progress has been made for general classes of functions, the problem of computing linear functions has been solved 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 Ahleswede *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 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. Specifically we ask whether the cutset-based bound is tight when multicasting a linear function, as in the single-receiver case.

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 two transmitters. Specifically we consider the Avestimehr-Diggavi-Tse (ADT) deterministic single-hop network model [10] which well captures the superposition and broadcast properties of wireless Gaussian networks. In recent work [11], we developed a linear-coding capacity result for symmetric cases of this problem, which we will refer to as *function alignment*[1], inspired by the concept of *interference alignment* [13], [14]. In the present paper, we derive a new upper bound and characterize the function multicasting capacity. We also develop a *network decomposition theorem* to identify elementary subnetworks that can constitute an original network without loss of optimality, thereby offering an alternative conceptually-simpler achievability proof.

One consequence of this result is that unlike the single-receiver case, the cutset-based bound is not achieved in general when multicasting a linear function. We find that this is due to competition for shared network resources that arises in satisfying function demands at multiple receivers. Moreover our network decomposition theorem serves to extend our result to more general networks. Specifically this leads to the computing capacity characterization of L-transmitter L-receiver networks, under linear coding strategies.

Related Work: In [15], [16], [17], the computing capacity for multicasting a sum of sources is explored for arbitrary multiple-source multiple-destination networks. Rai and Dey [16] 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

[1]Niesen-Nazer-Whiting [12] introduced a similar scheme in a different context, which they named computation alignment.

978-1-4673-4539-2/12/$31.00 ©2012 IEEE

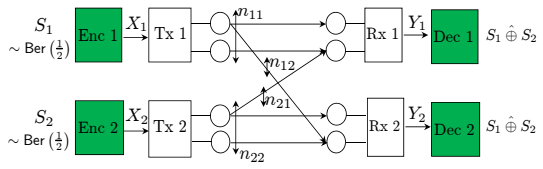

Fig. 1. Two-transmitter two-receiver Avestimehr-Diggavi-Tse (ADT) deterministic network

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 computing capacity of an ADT multi-receiver network.

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^K to estimate $S_1^K \oplus S_2^K$ from its received signal Y_ℓ^N. An error occurs whenever $d_\ell^K \neq S_1^K \oplus S_2^K$. The average probabilities of error are given by $\lambda_\ell = \mathbb{E}\left[P(d_\ell^K \neq S_1^K \oplus S_2^K)\right], \ell = 1, 2$.

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 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 proving an upper bound on the computing capacity in Theorem 1.

Definition 1 (Degenerate Networks): 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 the full version of this paper [18]. ■

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 upper bounds for the following two cases: (a) degenerate networks; (b) symmetric networks which are described as 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. ■

Remark 1: In [11], we established that Theorem 3 characterizes the linear-coding computation capacity. The present paper establishes that this is the true computation capacity.

We interpret our results with a focus on symmetric networks, which capture the essence of our function-multicasting problems. Specifically we will show that our function-multicasting scheme outperforms the separation approach where both receivers decode all of the sources and then compute modulo-2 sums of the sources. We will also show that the cutset bound is loose when multicasting linear functions, unlike a single-receiver function-unicasting case.

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

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

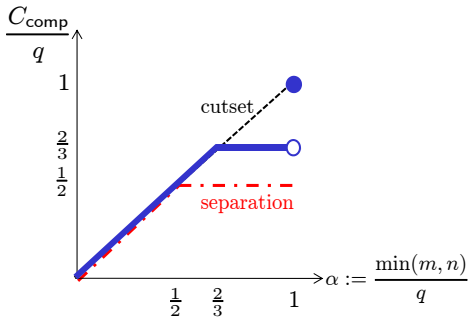

Fig. 2. Normalized computing capacity. Here $n := n_{11} = n_{22}$ and $m := n_{12} = n_{21}$. The parameter α in x-axis captures a signal-strength similarity between m and n.

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

Remark 2 (Comparison to Separation Scheme): The computing rate of the separation scheme can be easily derived from the multicast capacity. Note that the multicast capacity is the intersection of the two individual MAC capacities: the set $\mathcal{C}_{\mathsf{mult}}$ of (R_1, R_2) such that $R_1 \le \min(m,n)$, $R_2 \le \min(m,n)$ and $R_1 + R_2 \le \max(m,n)$. This gives

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

where $C_{\mathsf{sym}} := \sup\{R : (R,R) \in \mathcal{C}_{\mathsf{mult}}\}$. 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 3 (Comparison to the Single-Receiver Case): In the single-receiver case, the computing capacity achieves the cutset-based upper bound of $\min(m,n)$. A formal proof of the cutset bound will be provided shortly in the next section. On the other hand, the cutset 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 cutset argument. The main focus is to prove the second bound (3).

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

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

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 $R_{\mathsf{comp}} \le n_{11}$. Similarly we can show that $R_{\mathsf{comp}} \le \min\{H(Y_2|X_2), H(Y_1|X_1), H(Y_2|X_1)\} \le \min\{n_{12}, n_{21}, n_{22}\}$.

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

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

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:

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

So this becomes equivalent to a single-receiver case w.r.t receiver 2 where we can achieve $R_{\mathsf{comp}} = \min\{n_{12}, n_{22}\}$. Similarly for the other case of $n_{11} - n_{12} = n_{21} - n_{22} \le 0$, 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 $R_{\mathsf{comp}} = \min\{n_{11}, n_{21}\}$.

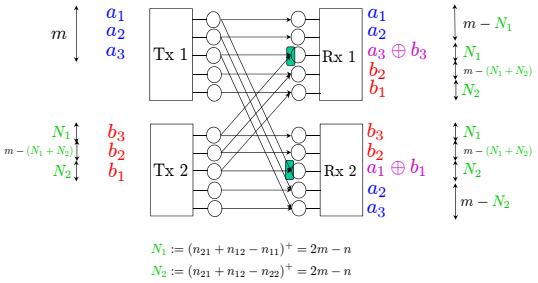

Fig. 3. [*Case I:* $\frac{1}{2} \le \alpha \le \frac{2}{3}$]: An achievable scheme for $(m, n) = (3, 5)$ and generalization to arbitrary values of (m, n).

IV. Proof of Theorem 3 via Geometric Approach

By symmetry, focus on the case of $m \le n$. Note that the case of $m \ge n$ is a mirror image of $m \le n$ in which transmitters 1 and 2 are swapped. As mentioned in Remark 2, 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 achieve $R_{\mathsf{comp}} = n$. 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 an $(m, n) = (3, 5)$ example in Fig. 3. We will show that the cutset bound of $\min(m, n) = 3$ can be achieved. First transmitter 1 sends (a_1, a_2, a_3) bits 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 connected level. Transmitter 2 sends b_3 on the top level to achieve $a_3 \oplus b_3$ on the connected level at receiver 1. In an arbitrary case, the number of these connected levels is $N_1 := n_{12} + n_{21} - n_{11} = 2m - n$. On the other hand, this 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 connected level. Transmitter 2 then sends b_1 on the 3rd level so as to achieve $a_1 \oplus b_1$ on the connected 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 4 (Exploiting Channel Structures [19], [20]): 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 5 (Connection to Interference Alignment): Note that the linear subspace with respect to a_1 is aligned with the subspace w.r.t b_1. This coincides with the popular concept of *interference alignment* [13], [14] which has shown the great potential for a variety of applications such as interference channels [14], cellular networks [21], [22], distributed storage networks [23], [24], [25] and multiple unicast networks [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.

We 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. Note however that at the end of time 1, we have an asymmetric solution. Receiver 1 can compute three bits, while receiver 2 can compute only two bits.

In order to make it symmetric, we invoke the idea of vector coding. In time 2, we start by favoring receiver 2 instead. 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 function alignment $a_3 \oplus b_3$ at receiver 1. The $a_3 \oplus b_3$ already

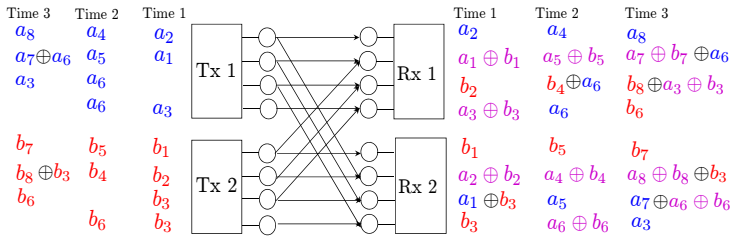

Fig. 4. [*Case II:* $\frac{2}{3} \leq \alpha < 1$]: Alternating function alignment for $(m, n) = (3, 4)$.

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 8 bits 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 $\mathbf{V}_1$ be a 12-by-4 beamforming matrix w.r.t. $\mathbf{a} := (a_2, a_4, a_8, a_6)^t$. Let $\mathbf{V}_2$ be a 12-by-4 beamforming matrix w.r.t. $\mathbf{b} := (b_1, b_5, b_7, b_3)^t$. According to the code construction in Fig. 4, we have

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

C. [*Case II:* $\frac{2}{3} \leq \alpha < 1$]: *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 receivers 2. Similar 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 offered by receiver 1, we get $2M_1 + M_2 \leq 3n$. Similarly for receiver 2, we get $M_1+2M_2 \leq 3n$. 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. 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-$3(n-m)$:

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

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.

Lemma 2:

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

Proof: See the full version [18]. ■

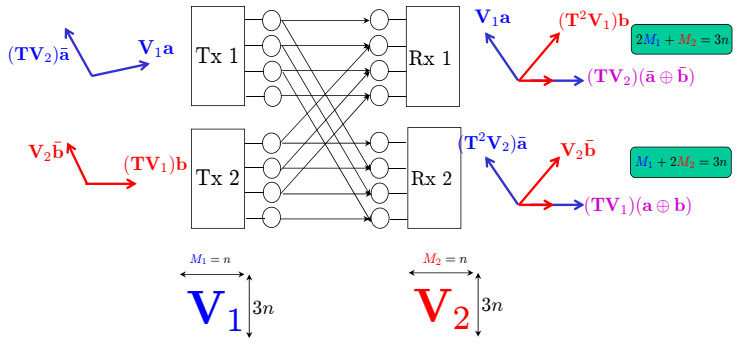

Fig. 5. Geometric interpretation of an achievable scheme. Let M_1 be be the column size of $\mathbf{V}_1$, i.e., the number of symbols that form function alignment at receivers 2. Similar let M_2 be the column size of $\mathbf{V}_2$. We choose appropriate values of (M_1, M_2) such that $M_1 + M_2 = 2n$ and thus can yield $R_{\text{comp}} = \frac{2n}{3}$. Considering the total dimension offered by receiver 1, we get $2M_1 + M_2 \leq 3n$. Similarly for receiver 2, we get $M_1 + 2M_2 \leq 3n$. This leads us to choose $M_1 = M_2 = n$.

V. Proof of Theorem 3 via Network Decomposition

In [11], we introduced a network decomposition theorem that provides elementary subnetworks such that the aggregate computation-capacity of the subnetworks is the same as the computation capacity of the original network. This theorem identifies fundamental building blocks that can constitute an original network without loss of optimality, thus establishing a *separation principle* among the building blocks as well as providing an alternative conceptually-simpler achievability proof of Theorem 3.

Here, we extend this theorem from the case of 2 users to the case of L users, which serves to establish the computing capacity of the generalized network, under linear coding strategies. This will be presented in Section VI.

Theorem 4 (Network Decomposition): Consider an L-transmitter L-receiver (m, n) network where $m \neq n$. See Section VI for details on this model. 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 arbitrary (m, n) model,

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

where

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

The proof is given in the complete version [18]. Fig. 6 illustrates the network decomposition with an $(m, n, L) = (2, 7, 2)$ example. 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&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 $(2, 7) = (1, 2)^2 \times (0, 1)^3$.

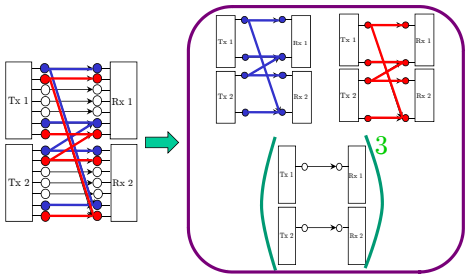

Fig. 6. A network decomposition example of an $(m, n) = (2, 7)$ model. From (13), $r = 0$ and $a = 2$; hence, the decomposition is given by $(2, 7) = (0, 1)^3 \times (1, 2)^2$.

Remark 6: 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 7: The separation principle among these decomposed subnetworks is not generally true. It is well known that for parallel interference channels, the optimal performance can be attained by coding over orthogonal components. □

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

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

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

(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 the complete version [18], 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

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

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

Remark 8: 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: 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

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

Theorem 5: The linear computing capacity is

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

The computing capacity is upper-bounded by

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

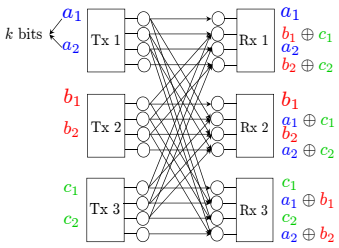

Fig. 7. Achievable scheme for the $(r-1, r)$ model where $r = 2k$.

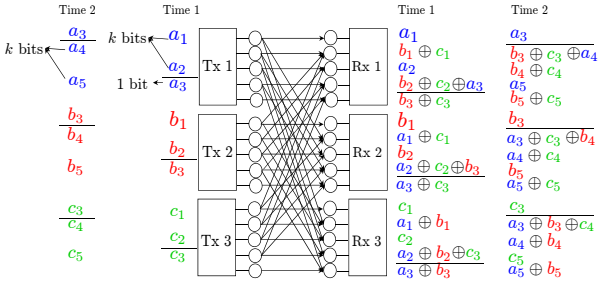

Fig. 8. Achievable scheme for the $(r-1, r)$ model where $r = 2k+1$.

Remark 9: In general networks, the linear capacity is often not equal to the capacity and non-linear codes may achieve higher rates [27]. 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. □

Proof: See the full version [18] for the coverse proof under linear coding strategies, as well as the information-theoretic upper bound.

The idea of achievability is to combine the network decomposition in Theorem 4 and the 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. Each receiver then gets 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. Since two resource levels are consumed to compute one desired function, we achieve $R_{\mathsf{comp}} = \frac{1}{2}r$. Obviously this applies 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 order to make an efficient resource utilization, we again invoke the vector coding idea. At the end of time 1, each transmitter sends an additional symbol on the last *even*-numbered level. Note that this transmission causes a conflict at each receiver. 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. This way we can achieve $R_{\mathsf{comp}} = \frac{k+1+k}{2} = \frac{r}{2}$. The same strategy applies 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 $0 \leq \alpha \leq \frac{1}{2}$, (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 $\frac{1}{2} \leq \alpha \leq \frac{2}{3}$. 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 $\alpha \geq \frac{2}{3}$. From (13), we know that $r \geq 2$. So we get $R_{\mathsf{comp}} = \frac{1}{2}n$.

■

VII. 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 significant role in extending to arbitrary multi-hop networks.

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] 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.

[11] N. Goela, C. Suh, and M. Gastpar, "Network coding with computation alignment," *Proceedings of the IEEE Information Theory Workshop, Lausanne, Switzerland*, Sept. 2012.

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

[13] 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.

[14] 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.

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

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

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

[18] C. Suh, , N. Goela, and M. Gastpar, "Computation in mutlicast networks: Function alignment and converse theorems," *arXiv:1209.3358*, Sept. 2012.

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

[20] 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.

[21] C. Suh and D. Tse, "Interference alignment for cellular networks," *Allerton Conference on Control, Computing and Communication*, Sept. 2008.

[22] C. Suh, M. Ho, and D. Tse, "Downlink interference alignment," *IEEE Transactions on Communications*, vol. 59, pp. 2616–2626, Sept. 2011.

[23] 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.

[24] 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.

[25] 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.

[26] 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.

[27] 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.