On Function Computation over a Cascade Network

Milad Sefidgaran and Aslan Tchamkerten

***Abstract*—A transmitter has access to X, a relay has access to Y, and a receiver has access to Z and wants to compute a given function $f(X,Y,Z)$. How many bits must be transmitted from the transmitter to the relay and from the relay to the receiver so that the latter can reliably recover $f(X,Y,Z)$?**

The main result is an inner bound to the rate region of this problem which is tight when $X-Y-Z$ forms a Markov chain.

I. Introduction

Consider a processor with a readable/rewritable memory which observes a finite data stream sequentially. At each step, what it can write on the memory depends on its current observation and what is already stored in the memory. The problem is to find the minimum required memory such that at the end of the stream, the processor can compute a given function of the entire stream.

Interestingly, a general formulation of this problem is equivalent to the problem of function computation in a cascade setting as shown by Viswanathan [7]. For the case of three nodes (which corresponds to observing a stream of size three in the previous setup) this setting is depicted in Fig. 1. A transmitter observes X, a relay observes Y, and a receiver observes Z. Communication is performed into two steps. First the transmitter sends a message to the relay given its observation X. Based on the received message and Y, the relay sends a message to the receiver such that the latter can compute the function $f(X,Y,Z)$ reliably.

This problem was first considered by Cuff, Su, El-Gamal [1] for which they derived the rate region in the case where Z is a constant. Later, Viswanathan [7] proposed a general rate region outer bound based on the point-to-point result of Orlitsky and Roche [3] and investigated the case where $X-Y-Z$ forms a Markov chain [6]. Finally, a similar problem has been considered in [5] for the case where Z is a constant and where there is an additional direct link between the transmitter and the receiver.

In this paper, first we extend the scheme used in [7] for arbitrary X, Y, and Z. This inner bound is always tight when $X-Y-Z$ forms a Markov. This inner bound is such that R_Y can always be taken to be equal to the R_Y lower bound derived by Viswanathan. But this need not be the case for R_X. We then propose a second inner bound where the opposite holds, namely where R_X can always be taken to be equal to the R_X lower bound, but not necessarily R_Y. Finally we propose a third inner bound which always includes the convex hull of two inner bounds, and in certain cases this inclusion is strict.

This work was supported in part by a "Future et Rupture" grant from the Institut Telecom, and by an Excellence Chair Grant from the French National Research Agency (ACE project).

M. Sefidgaran and A. Tchamkerten are with the Department of Communications and Electronics, Telecom ParisTech, 46 Rue Barrault, 75634 Paris Cedex 13, France. Emails: {sefidgaran,aslan.tchamkerten}@telecom-paristech.fr.

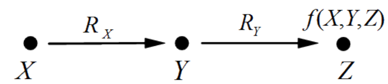

Fig. 1. Cascade Multi-terminal Network.

Section II contains the formal problem statement and some background material, Section III contains our results and Section IV deals with the proof sketchs of the main results.

II. Problem Statement and Preliminaries

The formal problem formulation given below is tailored for the cascade netword configuration of Fig. 1.

Let $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$, and $\mathcal{F}$ be finite sets, and $f:\mathcal{X}\times\mathcal{Y}\times\mathcal{Z}\to\mathcal{F}$. Let $\{(x_i,y_i,z_i)\}_{i=1}^{\infty}$ be independent instances of random variables (X,Y,Z) taking values over $\mathcal{X}\times\mathcal{Y}\times\mathcal{Z}$ and distributed according to $p(x,y,z)$.

Definition 1 (Code). An (n,R_X,R_Y) code consists of two encoding functions

$$\varphi_X:\mathcal{X}^n\to\{1,2,..,2^{nR_X}\}$$
$$\varphi_Y:\mathcal{Y}^n\times\{1,2,..,2^{nR_X}\}\to\{1,2,..,2^{nR_Y}\}$$

and a decoding function

$$\psi:\{1,2,..,2^{nR_Y}\}\times\mathcal{Z}^n\to\mathcal{F}^n\,.$$

The error probability of a code is defined as

$$P(\psi(\varphi_Y(\varphi_X(\mathbf{X}),\mathbf{Y}),\mathbf{Z})\neq f(\mathbf{X},\mathbf{Y},\mathbf{Z})),$$

where $\mathbf{X}\overset{\text{def}}{=}X_1,\ldots,X_n$ and

$$f(\mathbf{X},\mathbf{Y},\mathbf{Z})\overset{\text{def}}{=}\{f(X_1,Y_1,Z_1),...,f(X_n,Y_n,Z_n)\}\,.$$

Definition 2 (Rate Region). A rate pair (R_X,R_Y) is achievable if, for any $\epsilon>0$ and all n large enough, there exists an (n,R_X,R_Y) code whose error probability is no larger than ε. The rate region is the closure of the set of achievable rate pairs (R_X,R_Y).

The problem we consider in this paper is to characterize the rate region for given f and $p(x,y,z)$.

Notation. Given two random variables X and V, where X ranges over $\mathcal{X}$ and V over *subsets* of $\mathcal{X}$,[1] we write $X\in V$ whenever $P(X\in V)=1$.

[1] *I.e.*, a sample of V is a subset of $\mathcal{X}$.

Definition 3 (Generalized Conditional Characteristic Graph, [4])**.** Given $(V, X, Y) \sim p(v, x, y)$ such that $X \in V$ and given $f(X, Y)$, the conditional characteristic graph $G_{V|Y}$ of V given Y is the (undirected) graph whose vertex set is $\mathcal{V}$ and whose edge set $E(G_{V|Y})$ consists of the set of all (v_i, v_j) for which there exists $x_i \in v_i$, $x_j \in v_j$ and $y \in \mathcal{Y}$ such that

i. $p(v_i, x_i, y) \cdot p(v_j, x_j, y) > 0$,
ii. $f(x_i, y) \neq f(x_j, y)$.

Recall that an independent set of a graph G is a subset of vertices no two of which are connected. A maximal independent set is an independent set that is not included in any other independent set. The set of independent sets of G and the set of maximal independent sets of G are denoted by $\Gamma(G)$ and $\Gamma^*(G)$, respectively.

Definition 4 (Conditional Graph Entropy [3])**.** The conditional entropy of a graph is defined as[2]

$$\begin{aligned} H_{G_{V|Y}}(V|Y) &\stackrel{\text{def}}{=} \min_{\substack{U-V-Y \\ V\in U\in\Gamma(G_{V|Y})}} I(U;V|Y) \\ &= \min_{\substack{U-V-Y \\ V\in U\in\Gamma^*(G_{V|Y})}} I(U;V|Y). \end{aligned}$$

III. Results

The following outer bound is obtained by revealing Z (for R_X bound) and X (for R_Y bound) to the relay. In this case, the problem reduces to the two point-to-point sub-problems $X \to (Y, Z)$ and $(X, Y) \to Z$ for which [3, Theorem 1] can be applied:

Theorem 1 (Outer bound, [3],[7])**.** *If (R_X, R_Y) is achievable then*

$$\begin{aligned} R_X &\geq H_{G_{X|Y,Z}}(X|Y,Z), \\ R_Y &\geq H_{G_{X,Y|Z}}(X,Y|Z). \end{aligned}$$

The above outer bound is not tight in general. For instance, it can be checked that when $Z = X$ the conditions of Theorem 1 become

$$\begin{aligned} R_X &\geq 0, \\ R_Y &\geq H(f(X,Y)|X). \end{aligned}$$

But when $R_X = 0$, the problem for the relay-to-receiver link reduces to the standard point-to-point problem for which the minimum rate is given by

$$R_Y = H_{G_{Y|X}}(Y|X),$$

which can be strictly greater than $H(f(X,Y)|X)$.

In the particular case where Z is a constant, the conditions of Theorem 1 become

$$\begin{aligned} R_X &\geq H_{G_{X|Y}}(X|Y) \\ R_Y &\geq H(f(X,Y)), \end{aligned}$$

[2]We use the notation $U - V - W$ whenever random variables (U, V, W) form a Markov chain.

and this region is actually tight as shown by Cuff, Su, and El-Gamal [1]. Indeed, in this case if the receiver is able to compute the function then so is the relay. Therefore, the problem reduces to two point-to-point problems which can be treated independently: a computation problem between the transmitter and the relay whose solution is given by [3], and a classical single source coding problem between the relay and the receiver.

More generally, when $X - Y - Z$ forms a Markov chain we have:

Theorem 2 (Rate Region-Markov, [6])**.** *When $X - Y - Z$ forms a Markov chain, the rate region is given by*

$$\begin{aligned} R_X &\geq H_{G_{X|Y,Z}}(X|Y,Z) = \min_{\substack{U-X-(Y,Z) \\ X\in U\in\Gamma^*(G_{X|Y,Z})}} I(U;X|Y), \\ R_Y &\geq H_{G_{X,Y|Z}}(X,Y|Z). \end{aligned}$$

Theorem 2 may be referred to as [6, Theorem 3]. However, the claim of [6, Theorem 3] is not correct and its proof contains a few glitches. The proof of Theorem 4 to come provides an alternative derivation of this theorem.

In the achievable scheme that yields Theorem 2, the transmitter first constructs the characteristic graph $G_{X|Y,Z}$. By definition, the knowledge of an independent set that includes X together with (Y, Z) uniquely determines $f(X, Y, Z)$. Given $\mathbf{X}$, the transmitter first selects $\mathbf{U}(\mathbf{X})$, where for all $1 \leq i \leq n$, $U_i(\mathbf{X})$ is a maximal independent set in $G_{X|Y,Z}$ that includes X_i. Then it reliably sends $\mathbf{U}(\mathbf{X})$ to the relay by considering $\mathbf{Y}$ as side information. The relay, having access to $(\mathbf{U}(\mathbf{X}), \mathbf{Y})$ selects $\mathbf{W}(\mathbf{U}(\mathbf{X}), \mathbf{Y})$, where for all $1 \leq i \leq n$, $W_i(\mathbf{U}(\mathbf{X}), \mathbf{Y})$ is a maximal independent set in $G_{X,Y|Z}$ that includes (X_i, Y_i)—this step uses the property that $X - Y - Z$ forms a Markov chain. The knowledge of $\mathbf{W}(\mathbf{U}(\mathbf{X}), \mathbf{Y})$ and $\mathbf{Z}$ uniquely determines $f(\mathbf{X}, \mathbf{Y}, \mathbf{Z})$. The relay then reliably sends $\mathbf{W}(\mathbf{U}(\mathbf{X}), \mathbf{Y})$ to the receiver by considering $\mathbf{Z}$ as side information.

Theorem 2 has been generalized to the case where there are M relay nodes [6]. Here we state the correct rate region expression for this problem. Assume that the transmitter has access to X_0, relay $i, 1 \leq i \leq M$ has access to X_i and the receiver has access to X_{M+1}. The communication is performed in $M+1$ steps. At step 1, the transmitter communicates with relay 1 at rate R_0, at step $i, 2 \leq i \leq M$, relay $i-1$ communicates with relay i at rate R_{i-1}, and finally at step $M+1$, relay M communicates with the receiver at rate R_M.

Theorem 3 (Rate Region-Markov with M relays, [6])**.** *When $X_0 - X_1 - \cdots X_M - X_{M+1}$ forms a Markov chain, the rate region is*

$$\begin{aligned} R_i &\geq H_{G_{X^i|X_{i+1}^{M+1}}}(X^i|X_{i+1}^{M+1}) \\ &= H_{G_{X^i|X_{i+1}^{M+1}}}(X^i|X_{i+1}), 0 \leq i \leq M. \end{aligned}$$

The first inner bound that we propose is an immediate extension of the inner bound that yields Theorem 2 which is valid for arbitrary sources (X, Y, Z). When $X - Y - Z$

forms a Markov chain, this inner bound is tight and we recover Theorem 2.

Theorem 4 (Inner Bound I)**.** *(R_X, R_Y) is achievable whenever*

$$R_X \geq I(U;X|Y),$$
$$R_Y \geq H_{G_{U,Y|Z}}(U,Y|Z)\,,$$

where

$$X \in U \in \Gamma(G_{X|Y,Z})$$
$$U - X - (Y,Z)\,.$$

Moreover, when $X - Y - Z$ forms a Markov chain, the above inner bound is tight and becomes the same as Theorem 2.

By choosing $U = X$ in Theorem 4, the constraint for R_Y becomes the same as the constraint for R_Y in the outer bound given by Theorem 1. By contrast, there need not be a specific choice of U in Theorem 4 such that the R_X constraint becomes the R_X outerbound constraint.

An alternative achievable scheme for which the opposite is true, *i.e.*, it is always possible to "minimize" R_X while not necessarily R_Y is the following. Similarly as for the achievable schemes that yield Theorems 2 and 4, the transmitter first selects an independent set $\mathbf{U}(\mathbf{X})$ given the observation of $\mathbf{X}$. However, the information that the transmitter now sends to the relay allows the relay to decode $\mathbf{U}(\mathbf{X})$ only if, in addition to $\mathbf{Y}$, it were to know $\mathbf{Z}$. This modification of the previous scheme allows to reduce R_X at the expense of R_Y and the resulting inner bound is given by the following theorem:

Theorem 5 (Inner Bound II)**.** *(R_X, R_Y) is achievable whenever*

$$R_X \geq H_{G_{X|Y,Z}}(X|Y,Z),$$
$$R_Y \geq H_{G_{X|Y,Z}}(X|Y,Z) + H(Y|Z).$$

By time sharing, the convex hull of the rate regions of Theorems 4 and 5 is achievable. The following theorem, based on a non-trivial combination of the schemes that yield Theorems 4 and 5, provides an inner bound that always includes this convex hull and, in certain cases, this inclusion is strict.

Theorem 6 (Inner bound III)**.** *(R_X, R_Y) is achievable whenever*

$$R_X \geq I(X;U|Y) + I(V;X|W,Z)$$
$$R_Y \geq I(V;X|W,Z) + I(U,Y;W|Z),$$

for some U, V, and W that satisfy

$$U - X - (Y,Z)$$
$$V - X - (U,Y,Z)$$
$$(V,Z) - (U,Y) - W,$$

and

$$X \in V \in \Gamma(G_{X|U,Y,Z})$$
$$(U,Y) \in W \in \Gamma(G_{U,Y|V,Z}).$$

When $f(X,Y,Z) = X$, by letting $V = X$ in Theorem 6 we recover [2, Theorem 1].

We now first show how Theorems 4 and 5 can be derived from Theorem 6, from which the inclusion of the convex hull of them follows.

In Theorem 6, let V be the constant set-valued variable equal to $\mathcal{X}$, let U be a random variable such that

$$U - X - (Y,Z),$$
$$X \in U \in \Gamma(G_{X|Y,Z}),$$

and let W be a random variable that minimizes $I(U,Y;W|Z)$ among all W's such that

$$(U,Y) \in W \in \Gamma^*(G_{U,Y|Z}),$$
$$Z - (U,Y) - W.$$

Theorem 6 then gives the achievable rate region of Theorem 4.

In Theorem 6, let U be a constant, $W = (U,Y)$, and V be a random variable that minimizes $I(V;X|Y,Z)$ among all V's such that

$$V - X - (Y,Z)$$
$$X \in V \in \Gamma^*(G_{X|Y,Z}).$$

Theorem 6 then gives the achievable rate region of Theorem 5.

We now show, via an example, that Theorem 6 may yield rate pairs (R_X, R_Y) strictly outside the convex hull of Theorems 4 and 5.

Example. Let

$$(X,Y,Z) = ((X_1,X_2),(Y_1,Y_2),(Z_1,Z_2))$$

with $Z_1 = X_1$. Let $X_1 = \text{Bern}(\alpha), 0 < \alpha < \frac{1}{2}$ and let Y_1 be a random variable that takes values uniformly over $\{1,2,3,4\}$. The random variables X_2, Y_2, and Z_2 can take values in arbitrary sets. The joint probability distribution of (X_1,X_2,Y_1,Y_2,Z_2) is

$$p(x_1,x_2,y_1,y_2,z_2) = p(x_1)p(y_1)p(x_2,y_2,z_2),$$

and the function to be computed is

$$f(X,Y,Z) = (f_1(X_1,Y_1), f_2(X_2,Y_2)),$$

where

$$f_1(x_1,y_1) \stackrel{\text{def}}{=} \begin{cases} 0 & \text{if} \quad (x_1,y_1) \in \{(0,0),(0,1),(1,0),(1,2)\} \\ 1 & \text{otherwise} \end{cases}$$
$$f_2(x_2,y_2) \stackrel{\text{def}}{=} (x_2,y_2)\,.$$

We consider the minimum sum rate $R_X + R_Y$. It can be shown that the minimum sum rate that can be obtained in Theorem 4 is

$$R_X + R_Y = 1 + H_b(\alpha) + H(X_2,Y_2|Z) + H(X_2|Y_2)$$

and this is achieved by letting $U = (X_1, X_2)$.

The (minimum) sum rate given in Theorem 5 is

$$R_X + R_Y = 2 + H(X_2, Y_2|Z) + H(X_2|Y_2, Z_2)\,.$$

Hence, the minimum sum rate of the convex hull of the regions given by Theorems 4 and 5 is

$$\begin{aligned} R_X + R_Y = \min(&1 + H_b(\alpha) + H(X_2, Y_2|Z) + H(X_2|Y_2), \\ &2 + H(X_2, Y_2|Z) + H(X_2|Y_2, Z_2))\,. \end{aligned} \quad (1)$$

On the other hand, in Theorem 6, by letting $U = X_1$, $V = (\mathcal{X}_1, X_2)$, $W = (W_1, W_2)$ where $W_1 \in \{A_0, A_1\}$ with

$$A_i \stackrel{\text{def}}{=} \{(x_1, y_1) \in \mathcal{X}_1 \times \mathcal{Y}_1 : f(x_1, y_1) = i\}, i \in \{0, 1\},$$

and $W_2 = (\mathcal{X}_2, Y_2)$, we achieve the sum rate

$$R_X + R_Y = 1 + H_b(\alpha) + H(X_2, Y_2|Z) + H(X_2|Y_2, Z_2)$$

which is smaller than both terms of (1) whenever the strict inequality $H(X_2|Y_2, Z_2) < H(X_2|Y_2, Z_2)$ holds. In this case Theorem 6 strictly includes the convex hull of Theorems 4 and 5.

IV. Analysis

Proof of the second part of Theorem 4: We show that when $X - Y - Z$ forms a Markov chain, Theorem 4 gives the rate region. In Theorem 4, let U^* be the random variable that achieves $H_{G_{X|Y,Z}}(X|Y, Z)$, *i.e.*, the one that minimizes $I(U; X|Y, Z) = I(U; X|Y)$ among all U's such that

$$\begin{aligned} X \in U \in \Gamma^*(G_{X|Y,Z}), \\ U - X - Y - Z. \end{aligned} \quad (2)$$

The conditional probability of u^* given x is denoted by $p_1(u^*|x)$.

Theorem 4 then gives the achievable rate region

$$\begin{aligned} R_X &\geq H_{G_{X|Y,Z}}(X|Y, Z), \\ R_Y &\geq H_{G_{U^*,Y|Z}}(U^*, Y|Z)\,. \end{aligned} \quad (3)$$

We now show that $H_{G_{X,Y|Z}}(X, Y|Z) \geq H_{G_{U^*,Y|Z}}(U^*, Y|Z)$. This, by Theorem 1, shows that (3) characterizes the rate region.

Let W^* be the random variable that achieves $H_{G_{X,Y|Z}}(X, Y|Z)$, *i.e.*, the one that minimizes $I(W; X, Y|Z)$ among all W's such that

$$\begin{aligned} (X, Y) \in W \in \Gamma^*(G_{X,Y|Z}), \\ W - (X, Y) - Z. \end{aligned} \quad (4)$$

The conditional probability of w^* given (x, y) is denoted by $p_2(w^*|x, y)$.

Define the joint probability distribution of (U^*, X, Y, W^*, Z) as

$$p_3(u^*, x, y, w^*, z) \stackrel{\text{def}}{=} p(x, y, z)p_1(u^*|x)p_2(w^*|x, y).$$

Note that, by definition

$$p_3(x, y, z) = p(x, y, z)$$

$$p_3(u^*|x, y) = p_3(u^*|x) = p_1(u^*|x)$$

$$p_3(w^*|x, y, z) = p_3(w^*|x, y) = p_2(w^*|x, y)\,,$$

which means that the joint probability distribution $p_3(u^*, x, y, w^*, z)$ is consistent with (2) and (4).

From $p_3(u^*, x, y, w^*, z)$ we have

$$W^* - (X, Y) - Z,$$

$$W^* - (X, Y) - (U^*, Y),$$

$$W^* - (U^*, Y) - Z,$$

from which it follows that

$$I(W^*; U^*, Y|Z) \leq I(W^*; X, Y|Z). \quad (5)$$

To conclude, we need the following result:

Proposition 1. *The following relations holds*

I. $(U^*, Y) \in W^* \in \Gamma(G_{U^*,Y|Z})$.
II. $W^* - (U^*, Y) - Z$.

By Proposition 1 and the definition of conditional characteristic graph (Definition 4) we have

$$H_{G_{U^*,Y|Z}}(U^*, Y|Z) \leq I(W^*; U^*, Y|Z)\,. \quad (6)$$

Hence, since by definition of W^*

$$I(W^*; X, Y|Z) = H_{G_{X,Y|Z}}(X, Y|Z), \quad (7)$$

from (5), (6) and (7) we conlude that

$$H_{G_{U^*,Y|Z}}(U^*, Y|Z) \leq H_{G_{X,Y|Z}}(X, Y|Z)$$

which yields the desired result. ■

Proof of Proposition 1:

I. For any $y \in Y$ and any $x_1, x_2 \in u^* \in U^*$, (x_1, y) and (x_2, y) are not connected in $G_{X,Y|Z}$ due to the fact that $u^* \in \Gamma(G_{X|Y,Z})$. Now using Lemma 1 (stated hereafter), any maximal independent set $w^* \in W^* \in \Gamma^*(G_{X,Y|Z})$ that includes (x_1, y) should include (x_2, y) too, so should also include (u^*, y).

II.

$$\begin{aligned} p_3(w^*|u^*, y, z) &= \sum_x p_3(w^*|u^*, x, y, z)p_3(x|u^*, y, z) \\ &\stackrel{(a)}{=} \sum_x p_3(w^*|u^*, x, y)p_3(x|u^*, y, z) \\ &\stackrel{(b)}{=} \sum_x p_3(w^*|u^*, x, y)p_3(x|u^*, y) \\ &= p_3(w^*|u^*, y) \end{aligned}$$

where (a) holds since $W^* - (U^*, X, Y) - (U^*, Z)$ under $p_3(u^*, x, y, w^*, z)$. Equality (b) holds since $X - (U^*, Y) - Z$ under $p_3(u^*, x, y, w^*, z)$. To see this, note that

$$p_3(x|u^*, y, z) = \frac{p_3(x, u^*|y, z)}{p_3(u^*|y, z)} = \frac{p_3(x|y)p(u^*|x)}{p_3(u^*|y, z)},$$

and note that $p_3(u^*|y,z)$ is equal to

$$\begin{aligned} p_3(u^*|y,z) &= \sum_x p_3(u^*|x,y,z)p_3(x|y,z) \\ &= \sum_x p_3(u^*|x,y)p_3(x|y) \\ &= p_3(u^*|y). \end{aligned}$$ ■

Lemma 1. *Suppose that* $(x_1,y),(x_2,y),(x',y') \in \mathcal{X}\times\mathcal{Y}$*, that* (x_1,y) *and* (x_2,y) *are not connected in* $G_{X,Y|Z}$*, and that*

$$p(x_1,y)p(x_2,y)p(x',y') > 0.$$

Then, in the graph $G_{X,Y|Z}$*,* (x',y') *and* (x_1,y) *are connected if and only if* (x',y') *and* (x_2,y) *are connected.*

Proof: Assume that (x',y) and (x_1,y) are connected in $G_{X,Y|Z}$. This means that for some $z \in \mathcal{Z}$

$$p(x_1,y,z)p(x',y,z) = p(x_1,y)p(z|y)p(x',y,z) > 0, \quad (8)$$

$$f(x_1,y,z) \neq f(x',y,z). \quad (9)$$

Now, since $p(x_2,y,z) = p(x_2,y)p(z|y)$ and $p(z|y)$ is positive according to (8), we have $p(x_2,y,z) > 0$. Now, by hypothesis, (x_1,y) and (x_2,y) are not connected in $G_{X,Y|Z}$, hence $f(x_1,y,z) = f(x_2,y,z)$. Hence, from (9) we have $f(x_2,y,z) \neq f(x',y,z)$, which means that (x',y) and (x_2,y) are connected in $G_{X,Y|Z}$. ■

Proof Sketch of Theorem 6 : Assume U, V and W satisfy the hypothesis of the theorem. These random variables together with X,Y,Z are distributed according to some joint distribution $p(v,x,u,y,w,z)$.

For $v \in \Gamma(G_{X|U,Y,Z})$ and $w \in \Gamma(G_{U,Y|V,Z})$, define $\tilde{f}(v,w,z)$ to be equal to $f(x,y,z)$ for $x \in v$ and $(u,y) \in w$ whenever $p(x,u,y,z) > 0$ (Notice that all such (x,y) gives the same $f(x,y,z)$.). Further, for $\boldsymbol{v} = (v_1,\ldots,v_n)$ and $\boldsymbol{w} = (w_1,\ldots,w_n)$ define

$$\tilde{f}(\boldsymbol{v},\boldsymbol{w},\boldsymbol{z}) \stackrel{\text{def}}{=} \{\tilde{f}(v_1,w_1,z_1),\ldots,\tilde{f}(v_n,w_n,z_n)\}.$$

Randomly generate $2^{nI(X;U)}$, $2^{nI(V;X)}$, and $2^{nI(U,Y;W)}$ independent sequences

$$\boldsymbol{u}^{(i)} = (u_1^{(i)},u_2^{(i)},\ldots,u_n^{(i)}), i \in \{1,2,\ldots,2^{nI(X;U)}\},$$

$$\boldsymbol{v}^{(j)} = (v_1^{(j)},v_2^{(j)},\ldots,v_n^{(j)}), j \in \{1,2,\ldots,2^{nI(V;X)}\},$$

$$\boldsymbol{w}^{(k)} = (w_1^{(k)},w_2^{(k)},\ldots,w_n^{(k)}), k \in \{1,2,\ldots,2^{nI(U,Y;W)}\},$$

in an i.i.d. manner according to the marginal distributions $p(u)$, $p(v)$, and $p(w)$, respectively, and randomly and uniformly bin these sequences into $2^{nR_{X,1}}$, $2^{nR_{X,2}}$, and $2^{nR_{Y,1}}$ bins, respectively. Reveal the bin assignments $\phi_{X,1}$, $\phi_{X,2}$, and $\phi_{Y,1}$ to the transmitter, relay, and receiver.

Encoding: The transmitter finds the sequences $\boldsymbol{u}$ and $\boldsymbol{v}$ such that $(\boldsymbol{x},\boldsymbol{u})$ and $(\boldsymbol{v},\boldsymbol{x})$ are jointly robust typical and sends the indices of the bins that contain $\boldsymbol{u}$ and $\boldsymbol{v}$, *i.e.*, $(\phi_{X,1}(\boldsymbol{u}),\phi_{X,2}(\boldsymbol{v}))$.

The relay, upon receiving $(\phi_{X,1}(\boldsymbol{u}),\phi_{X,2}(\boldsymbol{v}))$, finds a unique $\check{\boldsymbol{u}}$ in the bin $\phi_{X,1}(\boldsymbol{u})$ that is jointly robust typical with $\boldsymbol{y}$. Then it finds a sequence $\boldsymbol{w}$ that is jointly robust typical with $(\check{\boldsymbol{u}},\boldsymbol{y})$ and finds the index of the bin that contains $\boldsymbol{w}$ *i.e.*, $\phi_{Y,1}(\boldsymbol{w})$, and sends $(\phi_{Y,1}(\boldsymbol{w}),\phi_{X,2}(\boldsymbol{v}))$.

If the transmitter or the relay doesn't find such indices, they declare an error, and if there is more than one possible index, one of them is randomly and uniformly selected.

Decoding: Given $\mathbf{z}$ and the index pair $(\phi_{Y,1}(\boldsymbol{w}),\phi_{X,2}(\boldsymbol{v}))$, declare $\tilde{f}(\hat{\boldsymbol{v}},\hat{\boldsymbol{w}},\boldsymbol{z})$ if there exists a unique jointly robust typical $(\hat{\boldsymbol{v}},\hat{\boldsymbol{w}},\boldsymbol{z})$ such that $\phi_{X,2}(\hat{\boldsymbol{v}}) = \phi_{X,2}(\boldsymbol{v})$ and $\phi_{Y,1}(\hat{\boldsymbol{w}}) = \phi_{Y,1}(\boldsymbol{w})$, and such that $\tilde{f}(\hat{\boldsymbol{v}},\hat{\boldsymbol{w}},\boldsymbol{z})$ is defined. Otherwise declare an error.

Probability of Error: There are two types of error. The first type of error occurs when no $\boldsymbol{u}$'s or $\boldsymbol{v}$'s, respectively $\boldsymbol{w}$'s, is jointly robust typical with $\boldsymbol{x}$, respectively with $(\check{\boldsymbol{u}},\boldsymbol{y})$. The probability of each of these two errors is shown to be negligible in [3] for n large enough. Hence, the probability of the first type of error can be made arbitrary small by taking n large enough.

The second type of error refers to the case when the relay cannot recover $\boldsymbol{u}$ or when the receiver cannot recover $(\boldsymbol{v},\boldsymbol{w})$. It can shown that the probability of these events vanishes for n large enough whenever

$$R_{X,1} > I(X;U|Y) \quad (10)$$

$$R_{Y,1} > I(U,Y;W|Z) \quad (11)$$

$$R_{X,2} > I(V;X|W,Z). \quad (12)$$

Note that whenever $(\check{\boldsymbol{u}},\hat{\boldsymbol{v}},\hat{\boldsymbol{w}}) = (\boldsymbol{u},\boldsymbol{v},\boldsymbol{w})$, where $(\boldsymbol{u},\boldsymbol{v})$ and $\boldsymbol{w}$ are the chosen sequences at the transmitter and at the relay, respectively, there is no error, *i.e.*, $f(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) = \tilde{f}(\boldsymbol{v},\boldsymbol{w},\boldsymbol{z})$ by definition of robust typicality and by the definitions of U, V, and W. From (10), (11), and (12) the error probability goes to zero as $n \to \infty$ whenever

$$\begin{aligned} R_X &> I(X;U|Y) + I(V;X|W,Z), \\ R_Y &> I(V;X|W,Z) + I(U,Y;W|Z), \end{aligned}$$

which concludes the proof. ■

REFERENCES

[1] P. Cuff, Han-I Su, and A. El Gamal. Cascade multiterminal source coding. In *Information Theory Proceedings (ISIT), 2009 IEEE International Symposium on*, pages 1199 –1203, June 2009.

[2] P. Ishwar and S.S. Pradhan. A relay-assisted distributed source coding problem. In *Information Theory and Applications Workshop, 2008*, pages 136 –141, February 2008.

[3] A. Orlitsky and J.R. Roche. Coding for computing. *Information Theory, IEEE Transactions on*, 47(3):903 –917, March 2001.

[4] M. Sefidgaran and A. Tchamkerten. Computing a function of correlated sources. http://arxiv.org/abs/1107.5806, 2011.

[5] M. Sefidgaran and A. Tchamkerten. On cooperation in multi-terminal computation and rate distortion. In *Information Theory Proceedings (ISIT), 2012 IEEE International Symposium on*, July 2012.

[6] K. Viswanathan. Information-theoretic analysis of function computation on streams. In *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*, pages 1147 –1152, 29 2010-oct. 1 2010.

[7] K. Viswanathan. On the memory required to compute functions of streaming data. In *Information Theory Proceedings (ISIT), 2010 IEEE International Symposium on*, pages 196 –200, June 2010.