Source Coding for Dependent Sources

M. Langberg
The Open University of Israel
mikel@openu.ac.il

M. Effros
California Institute of Technology
effros@caltech.edu

***Abstract*—In this work, we address the capacity region of multi-source multi-terminal network communication problems, and study the change in capacity when one moves form *independent* to *dependent* source information. Specifically, we ask whether the trade off between capacity and source independence is of continuous nature. We tie the question at hand to that of *edge removal* which has seen recent interest.**

I. Introduction

The network coding paradigm has seen significant interest in the last decade (see e.g., [1], [2], [3], [4], [5] and references therein). In the multiple source network coding problem, it is common to assume that the information available at different source nodes is *independent*. Under this assumption, several aspects of network coding have been explored, including the study of the capacity region of the multi-source multi-terminal network coding problem, e.g. [6], [7], [8], [9], [10], [11].

In this work we focus our study on the "independence" property of source information, and ask whether it is of significant importance in the study of network coding capacities. Loosely speaking, we consider the following question. All notions that appear below will be defined rigorously in the upcoming Section II.

Question 1: Given an instance to the multi-source multi-terminal network coding problem, does the capacity region differ significantly when one removes the requirement that the source information is independent?

Clearly, if the source information is highly dependent, it is not hard to devise instances to the network coding problem in which the corresponding capacity region differs significantly from that in which source information is independent. However, what is the trade off between independence and capacity? Can it be that the trade off is not *continuous*?

The main contribution of this work is in a connection we make between the question at hand and the recently studied "edge removal" problem [12], [13], [14] in which one asks to quantify the loss in capacity when removing a single edge from a given network. For all but a few special cases of networks, the effect on capacity of single edge removal is not fully understood. We show that Question 1 is closely related to the edge removal problem. In particular, we show that quantifying the rate loss in the former will imply a quantification for the latter and vice-versa.

This work was supported in part by NSF grant CCF-1018741, ISF grant 480/08 and the Open University of Israel's research fund (grant no. 46114). Work done at part while the first author was visiting Caltech.

As the edge removal problem is open, our connection does not progress in answering Question 1, but rather puts it in a broader perspective. For example, as a corollary of our equivalence, we show that removing an edge of vanishing capacity (in the block length) from a network will have a vanishing effect on the capacity of the network if and only if the trade off in Question 1 is continuous (as before, rigorous definitions will be given in Section II). Using recent results of ours from [14], this implies a similar equivalence between Question 1 and the advantage in network coding capacity when one allows an $\varepsilon > 0$ error in communication as opposed to zero error communication.

II. Model

A k-source coding problem $(\mathcal{I}, n, X)$ is defined by (a) a network instance $\mathcal{I}$, (b) a block length n, and (c) a vector of random variables X.

1) Instance $\mathcal{I} = (G, S, T, C, D)$ describes a directed, acyclic graph $G(V, E)$, a multiset of source nodes $S = \{s_1, \ldots, s_k\} \subset V$, a set of terminal nodes $T = \{t_1, \ldots, t_{|T|}\} \subset V$, a capacity vector $C = [c_e : e \in E]$, and a demand matrix $D = [d_{i,j} : (i, j) \in [k] \times [|T|]$, where for any real number $N \geq 1$, $[N] = \{1, \ldots, \lfloor N \rfloor\}$. Without loss of generality, we assume that each source $s \in S$ has no incoming edges and that each terminal $t \in T$ has no outgoing edges. Capacity vector C describes the capacity c_e for each edge $e \in E$. Binary demand matrix D specifies which sources are required at each terminal; namely, $d_{i,j} = 1$ if and only if terminal t_j requires the information originating at source s_i.
2) The block length n specifies the number of available network uses.
3) The source vector $X = (X_1, \ldots, X_k)$ describes the source random variable X_i available at each source s_i. Random variables $X_1, \ldots, X_k$ may be independent or dependent.

A solution $\mathbf{X}_n = \{X_e\}$ to the k-source coding problem $(\mathcal{I}, n, X)$ assigns a random variable X_e to each edge $e \in E$ with the following two properties:

1) **Functionality**: for $e = (u, v)$, X_e is a deterministic function f_e of the random variables $X_{e'}$ corresponding to incoming edges of node u. Equivalently, setting $In(u)$ to be the set of incoming edges of u, $Out(u)$ the set of outgoing edges from u, and setting $X_{In(u)} = \{X_{e'} \mid e' \in In(u)\}$, $X_{Out(u)} = \{X_{e'} \mid e' \in Out(u)\}$; then

$H(X_{Out(u)}|X_{In(u)}) = 0$. If e is an edge leaving a source node s, then the sources X_i originating at node s are the input to f_e.

2) **Capacity**: For each edge e it holds that $H(X_e) \leq c_e n$.

Given the acyclic structure of G, the value X_e transmitted across each edge can be defined inductively as a function of the source random variables X using the topological order of G.

A network source coding problem $(\mathcal{I}, n, X)$ is said to be *feasible* if there exists a network source code $\mathbf{X}_n$ such that all terminals can decode the information they require. Formally, for each terminal t, one defines a decoding function g_t that maps the information on the incoming edges of t to the sources demanded by t. Equivalently, one requires that for every message X_i required by t it holds that $H(X_i|X_{In(t)}) = 0$.

In this work we are interested in the comparison between source coding problems in which the source random variables are independent and those in which the source random variables are *almost* independent.

Definition 1 (δ-dependence): A set of random variables $X_1, \ldots, X_k$ is said to be δ-dependent if $\sum_i H(X_i) - H(X_1, \ldots, X_k) \leq \delta$. Independent random variables are 0-dependent.

Let $R = (R_1, \ldots, R_k)$ be a rate vector. A network instance $\mathcal{I}$ is said to be *R-feasible on independent sources* if and only if there exists a block length n and a set of independent random variables $X = (X_1, \ldots, X_k)$ with each X_i uniform over $\mathcal{X}_i = [2^{R_i n}]$ such that the k-source coding problem $(\mathcal{I}, n, X)$ is feasible. Similarly, a network instance $\mathcal{I}$ is said to be *R-feasible on δ-dependent sources* if and only if there exists a block length n and a vector of random variables $X = (X_1, \ldots, X_k)$ for which (a) $\sum_{i=1}^{k} H(X_i) - H(X) \leq \delta n$ and (b) for all i, $H(X_i) \geq R_i n$; such that the k-source coding problem $(\mathcal{I}, n, X)$ is feasible.

In what follows we address the following question:

Question 2: If $\mathcal{I}$ is $R = (R_1, \ldots, R_k)$-feasible on δ-dependent sources, what can be said about its feasibility on independent sources? For example, is it true that in this case $\mathcal{I}$ is $R_\delta = (R_1 - \delta, \ldots, R_k - \delta)$ feasible on independent sources?

While we do not resolve Question 2, we show a strong connection between this question and the *edge removal* problem studied in [12], [13], [14].

Remark 1: It is natural to also define the notion of feasibilty with respect to *all* δ-dependent sources. Namely, a network instance $\mathcal{I}$ is said to be *strongly* R-feasible on δ-dependent sources if and only if there exists a block length n such that for *every* vector of random variables $X = (X_1, \ldots, X_k)$ for which (a) $\sum_{i=1}^{k} H(X_i) - H(X) \leq \delta n$ and (b) for all i, $H(X_i) \geq R_i n$; the k-source coding problem $(\mathcal{I}, n, X)$ is feasible.

Under this definition it is not hard to verify that any instance $\mathcal{I}$ which is *strongly* $R = (R_1, \ldots, R_k)$ feasibility for δ-dependent sources is also $R_{\delta/k} = (R_1 - \delta/k, \ldots, R_k - \delta/k)$ feasible for independent sources. To see this connection, consider a set of δ-dependent sources $X_1, \ldots, X_k$ where each X_i is equal to the pair (Y_i, Z) where for blocklength n the Y_i's are uniform in $[2^{(R_i - \delta/k)n}]$ and independent, and Z is uniform in $[2^{\delta n/k}]$ and also independent of the Y_i's. It follows that $X = X_1, \ldots, X_k$ is δ-dependent. Now, if $\mathcal{I}$ is feasible on source information X with block length n, then it is feasible with source information $X'_i = (Y_i, z)$ for any fixed value $z \in [2^{\delta n/k}]$. However, the random variables X'_i are now independent and of rate $H(X'_i) = (R_i - \delta/k)n$. We conclude that $\mathcal{I}$ is $R_{\delta/k} = (R_1 - \delta/k, \ldots, R_k - \delta/k)$ feasible on independent sources.

A. The edge removal proposition

The edge removal problem compares the rates achievable on a given instance $\mathcal{I}$ before and after an edge e of capacity c_e is removed from the network G.

Proposition 1 (Edge removal): Let $\mathcal{I} = (G, S, T, C, D)$ be a network instance. Let $e \in G$ be an edge of capacity δ. Let $\mathcal{I}^e = (G^e, S, T, C, D)$ be the network obtained by replacing G with the network G^e in which edge e is removed. Let $c > 0$ be some constant. Let $R = (R_1, \ldots, R_k)$, and let $R_{c\delta} = (R_1 - c\delta, \ldots, R_k - c\delta)$. There exists a universal constant c, such that if $\mathcal{I}$ is R-feasible on independent sources then $\mathcal{I}^e$ is $R_{c\delta}$-feasible on independent sources.

B. The source coding proposition

Addressing Questions 1 and 2 we consider the following proposition:

Proposition 2 (Source coding): Let $\mathcal{I} = (G, S, T, C, D)$ be a network instance. Let $c > 0$ be some constant. Let $\delta > 0$. Let $R = (R_1, \ldots, R_k)$, and let $R_{c\delta} = (R_1 - c\delta, \ldots, R_k - c\delta)$. There exists a universal constant c, such that if $\mathcal{I}$ is R-feasible on δ-dependent sources then $\mathcal{I}$ is $R_{c\delta}$-feasible on independent sources.

III. Main result

Our main result shows that the two propositions above are equivalent.

Theorem 1: Proposition 1 is true if and only if Proposition 2 is true.

To prove Theorem 1 we will use the following two lemmas proven in Sections V and VI respectively.

Lemma 1 (Edge removal lemma): Let $\mathcal{I} = (G, S, T, C, D)$ be a network instance. Let $e \in G$ be an edge of capacity δ. Let $\mathcal{I}^e = (G^e, S, T, B)$ be the instance obtained by replacing G with the network G^e in which edge e is removed. Let $R = (R_1, \ldots, R_k)$, and let $R_\delta = (R_1 - \delta, \ldots, R_k - \delta)$. If $\mathcal{I}$ is R-feasible on independent sources then $\mathcal{I}^e$ is R_δ-feasible on δ-dependent sources.

Lemma 2 (Collocated source coding): Consider a network instance $\mathcal{I} = (G, S, T, C, D)$ in which all sources are collocated at a single node in G (i.e., each $s_i \in S$ equals the same vertex $s \in V$). Let $c > 0$ be some constant. Let $\delta > 0$. Let $R = (R_1, \ldots, R_k)$, and let $R_{c\delta} = (R_1 - c\delta, \ldots, R_k - c\delta)$. There exists a universal constant c, such that if $\mathcal{I}$ is R-feasible on δ-dependent sources then $\mathcal{I}$ is $R_{c\delta}$-feasible on independent sources.

The following corollary of Lemma 2 is also proven in Section VI:

Corollary 1 (Super source): Let $\mathcal{I} = (G, S, T, C, D)$ be a network instance. Let $c > 0$ be some (sufficiently large) constant. Let $\delta > 0$. Assume there is a vertex $s \in V$ (so-called a "super source") which has knowledge of all source information $X = (X_1, \ldots, X_k)$, and in addition has outgoing edges of capacity $c\delta$ to each and every source node in S. Let $R = (R_1, \ldots, R_k)$, and let $R_{c\delta} = (R_1 - c\delta, \ldots, R_k - c\delta)$. There exists a universal constant c, such that if $\mathcal{I}$ is R-feasible on δ-dependent sources then $\mathcal{I}$ is $R_{c\delta}$-feasible on independent sources.

IV. Proof of Theorem 1

We now present the proof of Theorem 1 using Lemma 1 and Corollary 1. Our proof will have two directions, each given in a separate subsection below.

A. Proposition 1 implies Proposition 2

In what follows, we show for the constant c in Corollary 1 that Proposition 1 is true with constant c_1 implies Proposition 2 with constant $c_2 = c + c_1$.

Proof: Let $\mathcal{I} = (G, S, T, C, D)$ be a network instance which is R-feasible on δ-dependent sources. We show that $\mathcal{I}$ is $R_{c_2\delta}$-feasible on independent sources for $c_2 = c + c_1$.

We consider 2 additional instances $\mathcal{I}_1 = (G_1, S_1, T, C, D)$ and $\mathcal{I}_2 = (G_2, S_2, T, C, D)$ similar to those considered in [9], [14]. We start by defining the network G_2; network G_1 is then obtained from network G_2 by a single edge removal.

Network G_2 is obtained from G by adding k new source nodes $s'_1, \cdots, s'_k$, a new "super-node" s, and a relay node r. For each $s_i \in G$, there is a capacity-R_i edge (s'_i, s_i) from new source s'_i to old source s_i. For each $s'_i \in G_2$, there is a capacity-R_i edge (s'_i, s) from new source s'_i to the super-node s. Let c be the constant from Corollary 1. There is a capacity-$c\delta$ edge (s, r) from the super-source s to the relay r; this edge is the network *bottleneck*. Finally, the relay r is connected to each source node s_i by an edge (r, s_i) of capacity $c\delta$. The new source set S_2 is $\{s'_1, \ldots, s'_k\}$. For $\mathcal{I}_1$, we set $S_1 = S_2$, and $G_1 = G_2$ apart from the removal of the bottleneck edge (s, r) of capacity $c\delta$.

Our assertion now follows from the following arguments. First note that $\mathcal{I}_1$ is R feasible on δ-dependent sources. This follows directly from our construction. Similarly, $\mathcal{I}_2$ is also R feasible on δ-dependent sources. Now, by Corollary 1, instance $\mathcal{I}_2$ is $R_{c\delta}$-feasible on independent sources. Using, Proposition 1, we have that $\mathcal{I}_1$ is $R_{(c+c_1)\delta}$ feasible on independent sources. (Here c_1 is the universal constant from Proposition 1.) Finally, we conclude that $\mathcal{I}$ is also $R_{(c+c_1)\delta}$ feasible on independent sources. ■

B. Proposition 2 implies Proposition 1

We now prove that Proposition 2 is true with constant c_2 implies Proposition 1 is true with constant $c_1 = 1 + c_2$.

Proof: Let $\mathcal{I} = (G, S, T, C, D)$ be a network instance. Let $e \in G$ be an edge (of capacity δ). Let $\mathcal{I}^e = (G^e, S, T, C, D)$ be the instance obtained by replacing G with the network G^e in which the edge e is removed. Assume that $\mathcal{I}$ is R feasible on independent sources. Lemma 1 implies that $\mathcal{I}^e$ is R_δ feasible on δ-dependent sources. Now, using Proposition 2, it holds that $\mathcal{I}^e$ is $R_{(1+c_2)\delta}$ feasible on independent sources. This suffices to complete our proof with $c_1 = 1 + c_2$. ■

V. Proof of Lemma 1

We start with the following definition.

Definition 2: Let m be an integer. A vector $(h_\alpha)_{\alpha \subset [m]}$ indexed by all subsets of $[m]$ is said to be *entropic* if there exist a vector of random variables $(X_1, \ldots, X_m)$ such that h_α is the joint entropy of the random variables $\{X_i \mid i \in \alpha\}$. Let Γ^*_m be the set of all entropic vectors representing m random variables.

The only property we will need from Γ^* in this work is that it is closed with respect to linear combinations over the positive integers, namely:

Fact 1 (e.g., [15], p. 366): For $\{h_i\}_{i=1}^{\ell} \subset \Gamma^*_m$ and positive integers $\{a_i\}_{i=1}^{\ell}$ it holds that $\sum a_i h_i \in \Gamma^*_m$.

We now turn to prove Lemma 1. Let $\mathcal{I} = (G, S, T, C, D)$ be a network instance. Let $R = (R_1, \ldots, R_k)$. Let $e \in G$ be an edge of capacity δ. Let $\mathcal{I}^e = (G^e, S, T, C, D)$ be the instance obtained by replacing G with the network G^e in which edge e is removed. Let $R_\delta = (R_1 - \delta, \ldots, R_k - \delta)$. Assume that $\mathcal{I}$ is R feasible on independent sources. Thus, there exists a block length n, and a code $\mathbf{X}_n$ realizing the feasibility of the k-source coding problem $(\mathcal{I}, n, X)$ with $X = \{X_1, \ldots, X_k\}$ in which each X_i is uniformly and independently distributed over $[2^{R_i n}]$. Let $\mathbf{X} = (X_1, \ldots, X_k, \mathbf{X}_n) = (X_1, \ldots, X_k, \{X_e\}_{e \in E})$.

Consider the entropic vector $\mathbf{h}$ corresponding to $\mathbf{X}$. Let Σ be the union of the supports of the random variables $(X_1, \ldots, X_k, \{X_e\}_{e \in E})$. It holds that $|\Sigma|$ is finite. This follows from the fact that in our setting the probability distribution governing $\mathbf{X}$ (and in particular the source random variables $X_1, \ldots, X_k$) has finite support of size $\Pi_{i=1}^{k} 2^{R_i n}$. In what follows, we denote the support size $\Pi_i 2^{R_i n}$ by N. Notice, that all events of the form $X_{e'} = \sigma$ for $e' \in E$ have probability which is an integer multiple of $1/N$. We will use this fact shortly in our analysis.

Let e be the edge that we are removing, and assume that $c_e = \delta$. For any value $\sigma \in \Sigma$, consider the vector of random variables drawn from the conditional distribution on $\mathbf{X}$ given $X_e = \sigma$; we denote this random variable by $\mathbf{X}^\sigma = \{(X_{e'} | X_e = \sigma)\}_{e' \in E \cup \{1,\ldots,k\}}$. Let $\mathbf{h}^\sigma$ be the entropic vector in Γ^* corresponding to $\mathbf{X}^\sigma$, and consider the convex combination $\mathbf{h}^e = \sum_\sigma \Pr[X_e = \sigma]\mathbf{h}^\sigma$. As Γ^* is not convex, the vector $\mathbf{h}^e$ is not necessarily in Γ^*. However, as noted above, there exist integers n_σ such that $\Pr[X_e = \sigma] = n_\sigma / N$. Thus, by Fact 1,

$$N \cdot \mathbf{h}^e = N \cdot \sum_\sigma \Pr[X_e = \sigma]\mathbf{h}^\sigma = \sum_\sigma n_\sigma \mathbf{h}^\sigma \in \Gamma^*$$

Let $\mathbf{X}^e = (X_1^e, X_2^e, \ldots, X_k^e, \{X_{e'}^e\}_{e' \in E})$ be the random variables corresponding to $N \cdot \mathbf{h}^e$. In what follows, we show via the code $\mathbf{X}^e$ that the problem $(\mathcal{I}^e, Nn, \{X_1^e, \ldots, X_k^e\})$ is $R_\delta = (R_1 - \delta, \ldots, R_k - \delta)$ feasible on δ-dependent sources.

Effectively, the random variables in $\mathbf{X}^e$ correspond to the variables in $\mathbf{X}$ conditioned on X_e. For any subset $\alpha \subseteq E \cup \{1, 2, \ldots, k\}$ let $X_\alpha = \{X_{e'} \mid e' \in \alpha\}$. Similarly define X^e_α. Then,

$$\begin{aligned} H(X^e_\alpha) &= N \cdot \mathbf{h}^e_\alpha = N \cdot \sum_\sigma \Pr[X_e = \sigma] h^\sigma_\alpha \\ &= N \cdot \sum_\sigma \Pr[X_e = \sigma] H(X_\alpha \mid X_e = \sigma) \\ &= NH(X_\alpha | X_e) = N(H(X_\alpha, X_e) - H(X_e)). \end{aligned}$$

We conclude that for each subset α (and in particular for $\alpha = \{i\}$ corresponding to a certain source s_i) it holds that

$$NH(X_\alpha) \geq H(X^e_\alpha) \geq N(H(X_\alpha) - \delta n).$$

This implies that $X^e_1, \ldots, X^e_k$ are δ-dependent with $H(X^e_i) \geq Nn(R_i - \delta)$. Namely, for $S = \{1, \ldots, k\}$:

$$\begin{aligned} H(X^e_S) &= NH(X_S | X_e) = N(H(X_S) - H(X_e)) \\ &= N\left(\sum_{i=1}^k H(X_i) - \delta n\right) \\ &\geq N\left(\sum_{i=1}^k H(X_i | X_e) - \delta n\right) \\ &= \sum_i H(X^e_i) - \delta N n. \end{aligned}$$

In addition, we have that $H(X^e_e) = N \cdot H(X_e | X_e) = 0$, and thus throughout we may consider the value of X^e_e to be a constant. Setting X_e to a constant corresponds to communication over the graph G^e (that does not contain the edge e at all).

We now turn to analyze the *Functionality* and *Capacity* constraints with respect to $\mathbf{X}^e$. For *Functionality*, let u be a vertex in G, and let $Out(u)$ and $In(u)$ be its set of outgoing and incoming edges. For every vertex u that is not the head of e, we have that $H(X_{Out(u),In(u)} | X_{In(u)}) = H(X_{Out(u)} | X_{In(u)}) = 0$, and

$$\begin{aligned} 0 &\leq H(X^e_{Out(u)} | X^e_{In(u)}) \\ &= H(X^e_{Out(u),In(u)}) - H(X^e_{In(u)}) \\ &= N(H(X_{Out(u),In(u)} | X_e) - H(X_{In(u)} | X_e)) \\ &= N(H(X_{Out(u)} | X_{In(u)}, X_e)) \\ &\leq N(H(X_{Out(u)} | X_{In(u)})) = 0. \end{aligned}$$

The third equality follows from the chain rule. The final inequality follows since conditioning reduces entropy.

When u is the head of e, the set $In(u)$ in G differs from the set $In^e(u)$ in G^e. Specifically, $In(u) = In^e(u) \cup \{e\}$. In this case,

$$\begin{aligned} 0 &\leq H(X^e_{Out(u),In^e(u)}) - H(X^e_{In^e(u)}) \\ &= N(H(X_{Out(u),In^e(u)} | X_e) - H(X_{In^e(u)} | X_e)) \\ &= N(H(X_{Out(u),In^e(u),e} | X_e) - H(X_{In^e(u),e} | X_e)) \\ &= N(H(X_{Out(u),In(u)} | X_e) - H(X_{In(u)} | X_e)) \\ &\leq N \cdot H(X_{Out(u)} | X_{In(u)}) = 0 \end{aligned}$$

For the *Capacity* constraints, for each edge $e' \in E$ it holds that $H(X^e_{e'}) = NH(X_{e'} | X_e) \leq NH(X_{e'}) \leq Nnc_{e'}$.

Finally, to show that $(\mathcal{I}^e, Nn, (X^e_1, \ldots, X^e_k))$ is $R_\delta = (R_1 - \delta, \ldots, R_k - \delta)$ feasible on δ-dependent sources, we need to show for any terminal t that requires source information i that $H(X^e_i | X^e_{In(t)}) = 0$. This follows similar to the previous arguments based on the fact that $\mathbf{X}$ satisfies $H(X_i | X_{In(t)}) = 0$. Specifically,

$$\begin{aligned} 0 &\leq H(X^e_i | X^e_{In(t)}) \\ &= N(H(X_{In(t),i} | X_e) - H(X_{In(t)} | X_e)) \\ &= N(H(X_i | X_{In(t)}, X_e)) \\ &\leq N(H(X_i | X_{In(t)})) = 0 \end{aligned}$$

VI. Proof of Lemma 2 and Corollary 1

The proof that follows uses two intermediate claims proven in Section VI-A.

Claim 1: Let $\mathcal{I} = (G, S, T, C, D)$ be a network instance. Let $\delta > 0$. Let $R = (R_1, \ldots, R_k)$, and let $R_\delta = (R_1 - \delta, \ldots, R_k - \delta)$. If $\mathcal{I}$ is R-feasible on δ-dependent sources $(X_1, \ldots, X_k)$ then $\mathcal{I}$ is R_δ-feasible on 2δ-dependent sources $(\bar{X}_1, \ldots, \bar{X}_k)$ such that $\bar{X}_i$ is distributed over an alphabet $\mathcal{X}_i$ of size at most $2^{H(\bar{X}_i)(1+\delta)}$.

Using Claim 1 we may assume throughout this section that our source random variables are 2δ dependent, and that each X_i has entropy $H(X_i) \geq R_i n - \delta n$ and support $\mathcal{X}_i$ of size at most $2^{R_i n + \delta n}$. For ease of presentation, we set $H(X_i) = R_i n$ and consider $X = (X_1, \ldots, X_k)$ to be δ dependent. Thus the constants c for Lemma 2 and Corollary 1 need to be increased accordingly with respect to the constants c computed below.

Claim 2: Let $\delta \geq 0$. Let $X = (X_1, \ldots, X_k)$ be a set of random variables over alphabets $(\mathcal{X}_1, \ldots, \mathcal{X}_k)$ such that (a) $\sum_{i=1}^k H(X_i) - H(X) \leq \delta n$, (b) for each i the marginal distribution satisfies $H(X_i) = R_i n$, and (c) each X_i has support $\mathcal{X}_i$ of size at most $2^{R_i n + \delta n}$. There exist a constant c, such that for each i, there exists a partition $P_i = P_{i,1}, \ldots, P_{i,r_i}$ of $\mathcal{X}_i$ in which (a) each r_i is at least $2^{R_i n - c\delta n}$, (b) for each i, j, and j', the size of $P_{i,j}$ equals that of $P_{i,j'}$ and (c) for every $(j_1, \ldots, j_k)$ with $j_i \in [r_i]$ the product space

$$P(j_1, \ldots, j_k) = P_{1,j_1} \times P_{2,j_2} \times \ldots P_{k,j_k}$$

satisfies

$$\Pr[(X_1, \ldots, X_k) \in P(j_1, \ldots, j_k)] > 0.$$

We now prove Lemma 2 using Claim 2 above. We then prove Corollary 1. Our proof follows the line of proof appearing in [9], [14]. Let $\mathcal{I} = (G, S, T, C, D)$ be a network instance in which all sources are collocated at a single node in G. Let $R = (R_1, \ldots, R_k)$. We assume that $\mathcal{I}$ is R feasible on δ dependent sources. The following argument shows that there exists a constant c such that $\mathcal{I}$ is $R_{c\delta}$ feasible on independent sources.

Let n be a block length such that there exists random source variables $X = (X_1, \ldots, X_k)$ with $H(X_i) = R_i n$ and $H(X) \geq \sum_i R_i n - \delta n$ such that the corresponding communication problem is feasible. The general idea is conceptually simple. As $X_1, \ldots, X_k$ satisfy the condition in Claim 2, we may define the corresponding partitions $P_{i,j}$, product sets $P(j_1, \ldots, j_k)$, and values r_i appearing in the claim. Consider the random variables $\bar{X}_1, \ldots, \bar{X}_k$ where each $\bar{X}_i$ is uniformly distributed over the set $[r_i]$. As r_i is at least $2^{R_i n - c\delta n}$, $H(\bar{X}) \geq \sum_i R_i n - ck\delta n$. As

$$\Pr[(X_1, \ldots, X_k) \in P(j_1, \ldots, j_k)] > 0$$

for all $j_i \in [r_i]$, we can identify (at least) a single k-tuple $(x_{1,j_1}, \ldots, x_{k,j_k}) \in P(j_1, \ldots, j_k)$ for which the communication of $(x_1, \ldots, x_k)$ over $\mathcal{I}$ is successful (yields correct decoding). This enables the following communication scheme over $\mathcal{I}$ with $\bar{X}$ as input. For every k-tuple $j_1, \ldots, j_k \in [r_1] \times [r_2] \times \cdots \times [r_k]$, the (single) source node *computes* the corresponding k-tuple $(x_{1,j_1}, \ldots, x_{k,j_k}) \in P(j_1, \ldots, j_k)$ and continues the communication as if it were communicating with the original source information X. Since communication is successful, in this case every terminal that decodes some x_{i,j_i} can deduce j_i simply by finding the element of $P_i = P_{i,1}, \ldots, P_{i,r_i}$ that contains x_{i,j_i}. Lemma 2 follows.

The proof of Corollary 1 is very similar and also follows ideas of [9], [14]. As r_i of Claim 2 is (at least) of size $2^{R_i n - c\delta n}$ and $|\mathcal{X}_i| \leq 2^{R_i n + \delta n}$, the size of $P_{i,j}$ is (at most) $2^{(c+1)\delta n}$. We conclude that a super source that has access to $\bar{X}_1, \ldots, \bar{X}_k$ and has rate $(c+1)\delta$ edges to every source node s_i in G, when given a k-tuple $j_1, \ldots, j_k \in [r_1] \times [r_2] \times \cdots \times [r_k]$, can send the *location* of x_{i,j_i} in P_{i,j_i}. This allows source s_i, which has j_i, to deduce x_{i,j_i} based on the information it receives from the super source, source s_i thus continues the protocol described in the proof to Lemma 2, sending x_{i,j_i} as defined above.

A. Proof of Claims 1 and 2

1) Proof of Claim 1: The proof is based on standard typicality arguments and is omitted due to space limitations.

2) Proof of Claim 2: For simplicity of presentation, we present our proof for the case $k = 2$. A similar proof can be given for general k. We first note that the size of the support of $X = (X_1, X_2)$ is at least $2^{H(X)} \geq 2^{(R_1+R_2)n - \delta n}$, and that of X_i is in the range $[2^{R_i n}, 2^{R_i n + \delta n}]$. Here, a pair (x_1, x_2) is in the support of (X_1, X_2) if it has positive probability. Thus for an *average* x_1 there are at least $2^{R_2 n - 2\delta n}$ values x_2 for which $p(x_1, x_2) > 0$. Consider the set $A_1 \subseteq \mathcal{X}_1$ such that there are at least $2^{R_2 n - 2\delta n - 1}$ values x_2 for which $p(x_1, x_2) > 0$. It holds that the support Γ of (X_1, X_2) when X_1 is restricted to A_1 is at least of size $|A_1| 2^{R_2 n - 2\delta n - 1}$. In addition, the size of A_1 is at least $2^{R_1 n - 2\delta n - 1}$. The latter follows from the fact that

$$|A_1| 2^{R_2 n + \delta n} + (2^{R_1 n + \delta n} - |A_1|) 2^{R_2 n - 2\delta n - 1} \geq 2^{(R_1+R_2)n - \delta n}.$$

Let $c' \geq 1$ be a constant to be determined later in the proof. Now, consider a random partition P_1 of A_1 into $r_1 = |A_1|/2^{c'\delta n}$ sets $P_{1,j}$ each of size $2^{c'\delta n}$. Consider also a random partition P_2 of $\mathcal{X}_2$ into $r_2 = |\mathcal{X}_2|/2^{c'\delta n}$ sets $P_{2,j}$ each of size $2^{c'\delta n}$. Here we assume that $|A_1|$ and $|\mathcal{X}_2|$ are divisible by $2^{c'\delta n}$, minor modifications in the proof are needed otherwise. For any j_1, j_2 we compute the probability that $P_{1,j_1} \times P_{1,j_2}$ intersects Γ defined above. Let $x_1 \in P_{1,j_1}$. As x_1 is in A_1, the probability that there exists an x_2 such that $(x_1, x_2) \in \Gamma$ is at least (for sufficiently large n)

$$1 - \frac{\binom{2^{R_2 n} - 2^{R_2 n - 2\delta n - 1}}{2^{c'\delta n}}}{\binom{2^{R_2 n}}{2^{c'\delta n}}} \simeq 1 - e^{-2^{(c'-2)\delta n}}$$

Using the union bound, this implies that there exist partitions P_1 and P_2 for which each and every "cell" $P_{1,j_1} \times P_{2,j_2}$ intersect Γ. Dividing the remaining elements x_1 among the sets P_{1,j_1} evenly, and noticing that $r_1 = |A_1|/2^{c'\delta n} \geq 2^{R_1 n - (c'+2)\delta n - 1}$ suffices to prove our assertion for $c' = 3$ while the the constant c in the claim equals $c' + 2 = 5$.

REFERENCES

[1] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung. Network Information Flow. *IEEE Transactions on Information Theory*, 46(4):1204–1216, 2000.

[2] S.-Y. R. Li, R. W. Yeung, and N. Cai. Linear Network Coding. *IEEE Transactions on Information Theory*, 49(2):371 – 381, 2003.

[3] R. Koetter and M. Medard. An Algebraic Approach to Network Coding. *IEEE/ACM Transactions on Networking*, 11(5):782 – 795, 2003.

[4] S. Jaggi, P. Sanders, P. A. Chou, M. Effros, S. Egner, K. Jain, and L. Tolhuizen. Polynomial Time Algorithms for Multicast Network Code Construction. *IEEE Transactions on Information Theory*, 51(6):1973–1982, June 2005.

[5] 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*, 52(10):4413–4430, 2006.

[6] R. Yeung, S. R. Li, N. Cai, and Z. Zhang. Network Coding Theory. *Foundations and Trends in Communications and Information Theory, Now Publishers Inc.*, 2(4,5):241–381, 2005.

[7] X. Yan, R. Yeung, and Z. Zhang. The capacity region for multi-source multi-sink network coding. In *International Symposium on Information Theory*, pages 116 – 120, 2007.

[8] T. Chan and A. Grant. Dualities Between Entropy Functions and Network Codes. *IEEE Transactions on Information Theory*, 54(10):4470–4487, 2008.

[9] T. Chan and A. Grant. On capacity regions of non-multicast networks. In *International Symposium on Information Theory*, pages 2378 – 2382, 2010.

[10] H. Yao and E. Verbin. Network coding is highly non-approximable. *In the Proceeding of the 47th Annual Allerton Conference on Communication, Control, and Computing*, 2009.

[11] M. Langberg and A. Sprintson. On the Hardness of Approximating the Network Coding Capacity. *IEEE Transactions on Information Theory*, 57(2):1008–1014, 2011.

[12] T. Ho, M. Effros, and S. Jalali. On equivalences between network topologies. In *Forty-Eighth Annual Allerton Conference on Communication, Control, and Computing*, 2010.

[13] S. Jalali, M. Effros, and T. Ho. On the impact of a single edge on the network coding capacity. In *Information Theory and Applications Workshop (ITA)*, 2011.

[14] M. Langberg and M. Effros. Network coding: Is zero error always possible? In *In proceedings of Forty-Ninth Annual Allerton Conference on Communication, Control, and Computing*, pages 1478–1485, 2011.

[15] R. W. Yeung. Information Theory and Network Coding. *Springer*, 2010.