

Single-Source/Sink Network Error Correction Is as Hard as Multiple-Unicast

Wentao Huang and Tracey Ho
Department of Electrical Engineering
California Institute of Technology
Pasadena, CA
{whuang,tho}@caltech.edu

Michael Langberg
Department of Electrical Engineering
University at Buffalo, SUNY
Buffalo, NY
mikel@buffalo.edu

Joerg Kliewer
Department of ECE
New Jersey Institute of Technology
Newark, NJ
jkliewer@njit.edu

Abstract—We study the problem of communicating over a single-source single-terminal network in the presence of an adversary that may jam a single link of the network. If any one of the edges can be jammed, the capacity of such networks is well understood and follows directly from the connection between the minimum cut and maximum flow in single-source single-terminal networks. In this work we consider networks in which some edges cannot be jammed, and show that determining the network communication capacity is at least as hard as solving the multiple-unicast network coding problem for the error-free case. The latter problem is a long standing open problem.

I. INTRODUCTION

The problem of network error correction concerns reliable transmission of information in a network with point-to-point noiseless channels, in the presence of an adversary. The adversary controls a set A of channels in the network and may corrupt the information transmitted on these channels in an arbitrary way. A network error correction code, first introduced by Cai and Yeung [1], [2], is a network code that can correct adversarial errors injected into the network from a set of channels A , for all $A \in \mathcal{A}$, where \mathcal{A} is a prescribed collection of subsets of channels that characterizes the strength of the adversary. For single-source multicast, under the simplifying assumption that all channels have unit capacity and \mathcal{A} is the collection of all subsets containing Z channels, [1], [2] show that the cut-set bound is tight and characterizes the network error correction capacity, which can be achieved by a linear code. Under similar settings a variety of works, e.g., [3], [4], [5], [6], [7], [8] have proposed different efficient capacity-achieving codes and strategies.

However, under slightly more general settings such that channel capacities are non-uniform or \mathcal{A} has a more general structure, much less is known about the network error correction capacity and achievable strategies. Kim et al. [9] study a model in which channel capacities are arbitrary and show that capacity upper bounds based on cut-set approaches are generally not tight. [9] also constructs examples where linear codes are insufficient to achieve capacity. Kosut et al. [10] study a model in which the adversary controls network nodes instead of channels, which is a special case of network error correction for non-uniform \mathcal{A} in the sense that \mathcal{A} may include subsets of different sizes. In this case [10] constructs an example that linear codes are inadequate to achieve capacity.

Achievable strategies under this node adversary model are also studied in [11], [12], [13], whereas determining the capacity region remains an open problem. As opposed to the well studied and well understood setting of [1], [2], the subtlety of finding and achieving the network error correction capacity in the more general settings above motivates us to examine the fundamental complexity of the general network error correction problem.

In this paper, we show that solving the single-unicast network error correction problem with general \mathcal{A} is as hard as solving the multiple-unicast network coding problem (with no error). Specifically, we convert any unit rate k -unicast network coding problem into a corresponding network error correction problem with a single source, a single sink, and a single adversarial channel chosen from a subset of channels, such that the unit rate k -unicast is feasible with zero error if and only if the zero-error network error correction capacity is k . Under the vanishing error model, we show a similar but slightly weaker result. Specifically, in this case if the unit rate k -unicast is feasible, then a network error correction rate of k is feasible. Conversely, if a network error correction rate of k is feasible, then the unit rate k -unicast is asymptotically feasible.

Our results add to the portfolios of problems that are connected to multiple-unicast network coding, which is a long standing open problem not presently known to be in P, NP or undecidable [14], [15], [16], [17]. Previously, equivalence results have been established, e.g., between multiple-unicast network coding and multiple-multicast network coding [18], [19], index coding [20], [21], secure network coding [22], [23] and two-unicast network coding [24].

The remainder of the paper is structured as follows. In Section II, we present the models and definitions of multiple-unicast network coding and single-source single-sink network error correction. In Section III and IV, we prove the reduction from multiple-unicast to network error correction for the zero error model and vanishing error model, respectively. Finally, we conclude the paper in Section V.

II. MODELS

A. Multiple-unicast Network Coding

We model the network to be a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where the set of vertices \mathcal{V} represents network nodes and the

set of edges \mathcal{E} represents network channels. Each edge $e \in \mathcal{E}$ has a capacity c_e , which is the maximum number of bits¹ that can be transmitted on e in one transmission. An instance $\mathcal{I} = (\mathcal{G}, \mathcal{S}, \mathcal{T}, B)$ of the *multiple-unicast network coding problem*, includes a network \mathcal{G} , a set of source nodes $\mathcal{S} \subset \mathcal{V}$, a set of terminal nodes $\mathcal{T} \subset \mathcal{V}$ and an $|\mathcal{S}|$ by $|\mathcal{T}|$ requirement matrix B . The (i, j) -th entry of B equals 1 if terminal j requires the information from source i and equals 0 otherwise. We assume that B is a permutation matrix and so each source is paired with a single terminal. Let $s(t)$ be the source that is required by terminal t . Denote $[n] \triangleq \{1, \dots, [n]\}$, then each source $s \in \mathcal{S}$ is associated with a message, which is a rate R_s random variable M_s uniformly distributed over $[2^{nR_s}]$. The messages for different sources are independent. A network code of length n is defined as a set of encoding functions ϕ_e for every $e \in \mathcal{E}$ and a set of decoding functions ϕ_t for each $t \in \mathcal{T}$. For each $e = (u, v)$, the encoding function ϕ_e is a function taking as input the random variables associated with incoming edges of node u and the random variable M_u if $u \in \mathcal{S}$, and maps to values in $[2^{nc_e}]$. For each $t \in \mathcal{T}$, the decoding function ϕ_t maps all random variables associated with the incoming edges of t , to a message $\hat{M}_{s(t)}$ with values in $[2^{nR_{s(t)}}]$.

A network code $\{\phi_e, \phi_t\}_{e \in \mathcal{E}, t \in \mathcal{T}}$ is said to *satisfy* a terminal t under transmission $(m_s, s \in \mathcal{S})$ if $\hat{M}_{s(t)} = m_{s(t)}$ when $(M_s, s \in \mathcal{S}) = (m_s, s \in \mathcal{S})$. A network code is said to satisfy the multiple-unicast network coding problem \mathcal{I} with error probability ϵ if the probability that all $t \in \mathcal{T}$ are simultaneously satisfied is at least $1 - \epsilon$. The probability is taken over the joint distribution on random variables $(M_s, s \in \mathcal{S})$. Namely, the network code satisfies \mathcal{I} with error probability ϵ if

$$\Pr_{(M_s, s \in \mathcal{S})} \left\{ \bigcap_{t \in \mathcal{T}} t \text{ is satisfied under } (M_s, s \in \mathcal{S}) \right\} \geq 1 - \epsilon \quad (1)$$

For an instance \mathcal{I} of the multiple-unicast network coding problem, rate R is said to be *feasible* if $R_s = R, \forall s \in \mathcal{S}$, and for any $\epsilon > 0$, there exists a network code with sufficiently large length that satisfies \mathcal{I} with error probability at most ϵ . Rate R is said to be *feasible with zero error* if $R_s = R, \forall s \in \mathcal{S}$ and there exists a network code that satisfies \mathcal{I} with zero error probability. Rate R is said to be *asymptotically feasible* if for any $\delta > 0$, rate $(1 - \delta)R$ is feasible. The capacity of \mathcal{I} refers to the supremum over all rates R that are asymptotically feasible and the zero-error capacity of \mathcal{I} refers to the supremum over all rates R that are feasible with zero error. The given model assumes all sources transmit information at equal rate. There is no loss of generality in this assumption as a varying rate source s can be modeled by several equal rate sources collocated at s .

¹For convenience we assume that the network channels transmit binary symbols. Our results can be naturally extended to the general q -ary case.

B. Single-Source Single-Sink Network Error Correction

An instance $\mathcal{I}_c = (\mathcal{G}, s, t, \mathcal{A})$ of the *single-source single-terminal network error correction problem* includes a network \mathcal{G} , a source node $s \in \mathcal{V}$, a terminal node $t \in \mathcal{V}$ and a collection of subsets of channels $\mathcal{A} \subset 2^{\mathcal{E}}$ susceptible to errors. In this problem the channels are not always reliable and an error is said to occur in a channel if the output of the channel is different from the input. More precisely, the output of a channel e is the input signal superposed by an error signal r_e , and we say there is an error on channel e if $r_e \neq 0$. For a subset $A \in \mathcal{A}$ of channels, an A -error is said to occur if an error occurs in every channel in A . For an instance \mathcal{I}_c of the single-source single-terminal network error correction problem, a network code $\{\phi_e, \phi_t\}_{e \in \mathcal{E}, t \in \mathcal{T}}$ is said to *satisfy* a terminal t under transmission m_s if $\hat{M}_s = m_s$ when $M_s = m_s$, given the occurrence of any error pattern $\mathbf{r} = (r_e, e \in \mathcal{E})$ that results in an A -error, for all $A \in \mathcal{A}$. A network code is said to satisfy problem \mathcal{I}_c with error probability ϵ if the probability that t is satisfied is at least $1 - \epsilon$. The probability is taken over the distribution on the source message M_s .

For an instance \mathcal{I}_c of the single-source single-terminal network error correction problem, rate R is said to be *feasible* if $R_s = R$ and for any $\epsilon > 0$, there exists a network code with sufficiently large length that satisfies \mathcal{I}_c with error probability at most ϵ . Rate R is said to be *feasible with zero error* if $R_s = R$ and there exists a network code that satisfies \mathcal{I}_c with zero error probability. Rate R is said to be *asymptotically feasible* if for any $\delta > 0$, rate $(1 - \delta)R$ is feasible. The capacity of \mathcal{I}_c refers to the supremum over all rates R that are asymptotically feasible and the zero-error capacity of \mathcal{I}_c refers to the supremum over all rates R that are feasible with zero error.

Throughout the paper we denote by $\mathcal{R}_{\mathcal{A}}$ the set of all possible error patterns \mathbf{r} that result in A -errors, where $A \in \mathcal{A}$.

III. REDUCTION FROM MULTIPLE-UNICAST TO NETWORK ERROR CORRECTION: ZERO ERROR CASE

In this section we reduce the multiple-unicast network coding problem (with no error) to the single-source single-terminal network error correction problem with at most a single adversarial channel. We start with the zero-error case.

Theorem 1. *Given any multiple-unicast network coding problem \mathcal{I} with source-destination pairs $\{(s_i, t_i), i = 1, \dots, k\}$, a corresponding single-source single-sink network error correction problem $\mathcal{I}_c = (\mathcal{G}, s, t, \mathcal{A})$ in which \mathcal{A} includes sets with at most one edge can be constructed as specified in Figure 1, such that the zero-error capacity of \mathcal{I}_c is k if and only if unit rate is feasible with zero error in \mathcal{I} .*

Proof: The zero-error capacity of \mathcal{I}_c is upper bounded by k , because it is the min-cut from s to t .

“ \Rightarrow ” We show that the feasibility of a zero-error rate k in \mathcal{I}_c implies the feasibility of unit zero-error rate in \mathcal{I} .

Suppose a zero-error rate k is achieved in \mathcal{I}_c by a network code with length n , and denote the source message by M , then M is uniformly distributed over $[2^{nk}]$. For any edge $e \in \mathcal{E}$,

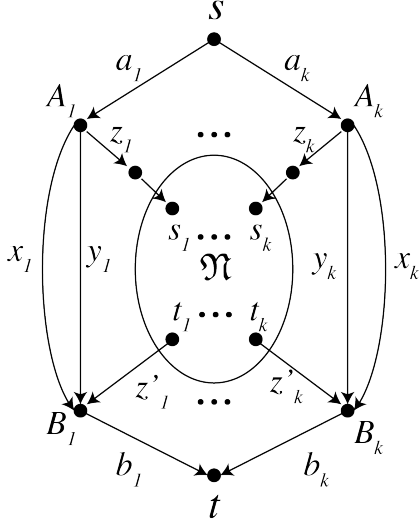


Fig. 1: In the single-source single-terminal network error correction problem \mathcal{I}_c , the source s wants to communicate with the terminal t . \mathcal{N} is a general network with point-to-point noiseless channels. All edges outside \mathcal{N} (i.e., edges for which at least one of its end-point does not belong to \mathcal{N}) have unit capacity. There is at most one error in this network, and this error can occur at any edge except $\{a_i, b_i, 1 \leq i \leq k\}$. Namely, \mathcal{A} includes all singleton sets of a single edge in the network except $\{a_i\}$ and $\{b_i\}$, $i = 1, \dots, k$. Note that there are k branches in total but only the first and the k -th branches are drawn explicitly. The multiple-unicast network coding problem \mathcal{I} is defined on the network \mathcal{N} , where the k source-destination pairs are (s_i, t_i) , $i = 1, \dots, k$, and all channels are error-free.

we denote by $e(m, \mathbf{r}) : [2^{nk}] \times \mathcal{R}_A \rightarrow [2^n]$ the signal received on edge e when the source message equals m and the error pattern \mathbf{r} occurs in the network. When the context is clear, we may denote $e(m, \mathbf{r})$ simply by e .

Let $\mathbf{b}(m, \mathbf{r}) = (b_1(m, \mathbf{r}), \dots, b_k(m, \mathbf{r}))$, then because the edges b_1, \dots, b_k form a cut-set from s to t , $\mathbf{b}(m, \mathbf{r})$ must be injective with respect to m due to the zero error decodability constraint. Formally, for two different messages $m_1 \neq m_2$, it follows from the zero error decodability constraint that $\mathbf{b}(m_1, \mathbf{r}_1) \neq \mathbf{b}(m_2, \mathbf{r}_2)$, $\forall \mathbf{r}_1, \mathbf{r}_2 \in \mathcal{R}_A$. Note that the codomain of \mathbf{b} is $[2^n]^k$, which has the same size as the set of messages $[2^{nk}]$. Therefore denote by $\mathbf{b}(m) \triangleq \mathbf{b}(m, \mathbf{0})$, then $\mathbf{b}(m)$ is a bijective function and $\mathbf{b}(m, \mathbf{r}) = \mathbf{b}(m)$, $\forall \mathbf{r} \in \mathcal{R}_A$. Similarly $\mathbf{a}(m, \mathbf{r}) = (a_1(m, \mathbf{r}), \dots, a_k(m, \mathbf{r}))$ is also a bijective function of the message, regardless of the error patterns.

For any $e \in \mathcal{E}$, denote $e(m) \triangleq e(m, \mathbf{0})$. For $i = 1, \dots, k$, we claim that for any two messages $m_1, m_2 \in [2^{nk}]$ such that $a_i(m_1) \neq a_i(m_2)$, it follows that $x_i(m_1) \neq x_i(m_2)$, $y_i(m_1) \neq y_i(m_2)$ and $z_i(m_1) \neq z_i(m_2)$. Suppose for contradiction that there exist m_1, m_2 such that $a_i(m_1) \neq a_i(m_2)$ and such that the claim is not true, i.e., $x_i(m_1) = x_i(m_2)$ or $y_i(m_1) = y_i(m_2)$ or $z_i(m_1) = z_i(m_2)$.

First consider the case that $x_i(m_1) = x_i(m_2)$. Because of the one-to-one correspondence between m and \mathbf{a} , there exists a message $m_3 \neq m_1, m_2$ and such that $\mathbf{a}(m_3) = (a_1(m_1), \dots, a_{i-1}(m_1), a_i(m_2), a_{i+1}(m_1), \dots, a_k(m_1))$. Then $x_i(m_1) = x_i(m_3)$ because by hypothesis $x_i(m_1) = x_i(m_2)$. Consider the following two scenarios. In the first scenario, m_1 is transmitted, and an error turns $y_i(m_1)$ into $y_i(m_3)$; in the second scenario, m_3 is transmitted, and an error turns $z_i(m_3)$ into $z_i(m_1)$. Then the cut-set signals $a_1, \dots, a_{i-1}, x_i, y_i, z_i, a_{i+1}, \dots, a_k$ are exactly the same in both scenarios, and so it is impossible for t to distinguish m_1 from m_3 , a contradiction to the zero error decodability constraint. Therefore $x_i(m_1) \neq x_i(m_2)$. With a similar argument it follows that $y_i(m_1) \neq y_i(m_2)$ and $z_i(m_1) \neq z_i(m_2)$, and the claim is proved.

The claim above suggests that x_i, y_i and z_i , as functions of a_i , are injective. They are also surjective functions because the domain and codomain are both $[2^n]$. Hence there are one-to-one correspondences between a_i, x_i, y_i and z_i .

Next we show that for any two messages m_1, m_2 , if $b_i(m_1) \neq b_i(m_2)$, then $z'_i(m_1) \neq z'_i(m_2)$. Suppose for contradiction that there exists $m_1 \neq m_2$ such that $b_i(m_1) \neq b_i(m_2)$ and $z'_i(m_1) = z'_i(m_2)$. Then if m_1 is transmitted and an error \mathbf{r}_1 turns $x_i(m_1)$ into $x_i(m_2)$, the node B_i will receive the same signals as in the case that m_2 is transmitted and an error \mathbf{r}_2 turns $y_i(m_2)$ into $y_i(m_1)$. Therefore $b_i(m_1, \mathbf{r}_1) = b_i(m_2, \mathbf{r}_2)$. But, as shown above, because $b_i(m_1, \mathbf{r}_1) = b_i(m_1)$ and $b_i(m_2, \mathbf{r}_2) = b_i(m_2)$, it follows that $b_i(m_1) = b_i(m_2)$, a contradiction. This suggests that if $z'_i(m_1) = z'_i(m_2)$ then $b_i(m_1) = b_i(m_2)$ and therefore b_i is a function of z'_i . The function is surjective because b_i takes all 2^n possible values. Then since the domain and the codomain are both $[2^n]$, it follows that b_i must be a bijective function of z'_i . With the same argument it follows that b_i is also a bijective function of x_i .

Hence z_i is a bijection of a_i , a_i is a bijection of x_i , x_i is a bijection of b_i , and b_i is a bijection of z'_i . Therefore for all $1 \leq i \leq k$, z_i is a bijection of z'_i , and therefore unit rate is feasible with zero error in \mathcal{I} .

“ \Leftarrow ” Conversely, we show that the feasibility of the unit zero-error rate in \mathcal{I} implies the achievability of a zero-error rate of k in \mathcal{I}_c . A constructive scheme is shown in Figure 2. In \mathcal{I}_c , the source lets $M = (M_1, \dots, M_k)$, where the M_i 's are i.i.d. uniformly distributed over $[2^n]$. Let the network code be $a_i(M) = x_i(M) = y_i(M) = z_i(M) = z'_i(M) = M_i$, $i = 1, \dots, k$, and let node B_i , $i = 1, \dots, k$, perform majority decoding. It is straightforward to see that the scheme ensures that $b_i(M) = M_i$ under all error patterns in \mathcal{R}_A . Therefore rate k is feasible with zero error in \mathcal{I}_c . This rate achieves capacity since it is equal to the min-cut from s to t . ■

IV. REDUCTION FROM MULTIPLE-UNICAST TO NETWORK ERROR CORRECTION: VANISHING ERROR CASE

In this section we show that a similar but slightly weaker result holds under the vanishing error model.

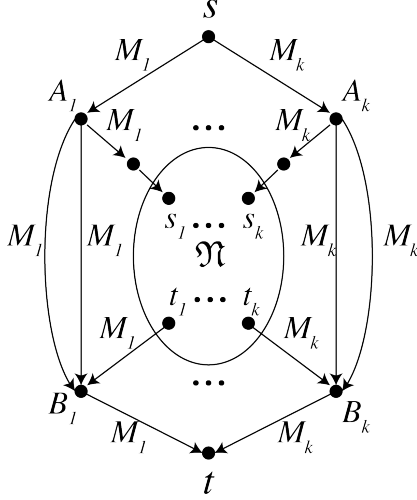


Fig. 2: A scheme to achieve zero-error rate k in \mathcal{I}_c given that unit rate is feasible with zero error in \mathcal{I} . $M = (M_1, \dots, M_k)$ and node B_i performs majority decoding.

Theorem 2. Given any multiple-unicast network coding problem \mathcal{I} with source-destination pairs $\{(s_i, t_i), i = 1, \dots, k\}$, a corresponding single-source single-sink network error correction problem $\mathcal{I}_c = (\mathcal{G}, s, t, A)$ in which A includes sets with at most a single edge can be constructed as specified in Figure 1, such that if unit rate is feasible in \mathcal{I} then rate k is feasible in \mathcal{I}_c . Conversely, if rate k is feasible in \mathcal{I}_c then unit rate is asymptotically feasible in \mathcal{I} .

We first give a simple lemma.

Lemma 1. Let X, Y, Z be three arbitrary random variables. Then

$$I(X; Z) \geq I(X; Y) + I(Y; Z) - H(Y).$$

Proof:

$$\begin{aligned} I(X; Z) &= H(Z) - H(Z|X) \\ &\geq H(Z) - H(Z, Y|X) \\ &= H(Z) - H(Y|X) - H(Z|Y, X) \\ &\geq H(Z) - H(Y|X) - H(Z|Y) \\ &= H(Y) - H(Y|X) + H(Z) - H(Z|Y) - H(Y) \\ &= I(X; Y) + I(Y; Z) - H(Y) \end{aligned}$$

In the following we prove Theorem 2.

“ \Rightarrow ” We first show the feasibility of rate k in \mathcal{I}_c implies the asymptotic feasibility of unit rate in \mathcal{I} .

We will take the following path. In \mathcal{I}_c , we apply the network code that achieves rate k with message uniformly distributed over a selected subset of $[2^{nk}]$. Then we show that for $i = 1, \dots, k$, this induces a large mutual information between random variables a_i and b_i , between b_i and z'_i , and between z_i and a_i . Hence it implies a large mutual information

between z_i and z'_i and finally we show that this implies the asymptotic feasibility of unit rate in \mathcal{I} .

Step 1: Select a subset of messages.

Suppose in \mathcal{I}_c a rate of k is achieved by a network code with length n and with a probability of error ϵ . Let M be the source message uniformly distributed over $\mathcal{M} = [2^{kn}]$ and let \hat{M} be the output of the decoder at the terminal. Partition \mathcal{M} into good and bad messages $\mathcal{M}^g + \mathcal{M}^b$ in the way that $m \in \mathcal{M}^g$ if the network code satisfies t under transmission m , i.e., the terminal decodes successfully $\hat{M} = m$ when $M = m$ for all $r \in \mathcal{R}_A$. Therefore if $m \in \mathcal{M}^b$ then there exists $r \in \mathcal{R}_A$ such that $\hat{M} \neq m$ if $M = m$ and r occurs, i.e., r results in a decoding error. By the hypothesis on the probability of error it follows that $|\mathcal{M}^b| \leq 2^{kn}\epsilon$.

For $i = 1, \dots, k$, let $x_i(m, r) : \mathcal{M} \times \mathcal{R}_A \rightarrow [2^n]$ be the signal received from channel x_i when m is transmitted by the source and the error pattern r happens. Let $x_i(m) = x_i(m, \mathbf{0})$, $\mathbf{x}(m, r) = (x_1(m, r), \dots, x_k(m, r))$ and $\mathbf{x}(m) = (x_1(m), \dots, x_k(m))$. We define functions $a_i, b_i, y_i, z_i, z'_i, \mathbf{a}, \mathbf{b}, \mathbf{y}, \mathbf{z}, \mathbf{z}'$ in a similar way.

Lemma 2. There exists $\mathcal{M}^\circ \subset \mathcal{M}^g$ such that for any $m_1, m_2 \in \mathcal{M}^\circ$, $m_1 \neq m_2$, it follows that $\mathbf{a}(m_1) \neq \mathbf{a}(m_2)$, $\mathbf{b}(m_1) \neq \mathbf{b}(m_2)$ and $\mathbf{z}'(m_1) \neq \mathbf{z}'(m_2)$, and such that $|\mathcal{M}^\circ| \geq 2^{kn}(1 - \epsilon')$, where $\epsilon' = 4\epsilon$.

Proof: As \mathbf{a} and \mathbf{b} are cut-set signals, for any $m_1, m_2 \in \mathcal{M}^g$, $m_1 \neq m_2$, it follows that $\mathbf{a}(m_1) \neq \mathbf{a}(m_2)$ and $\mathbf{b}(m_1) \neq \mathbf{b}(m_2)$. Setting $\mathcal{B}^g = \{\mathbf{b}(m) : m \in \mathcal{M}^g\}$, it holds that $|\mathcal{B}^g| \geq (1 - \epsilon)2^{kn}$, and setting $\mathcal{B}^b = [2^n]^k \setminus \mathcal{B}^g$, it holds that $|\mathcal{B}^b| \leq 2^{kn}\epsilon$.

For any $m \in \mathcal{M}^g$, it is said to be a *poor* message if there exists another $m' \in \mathcal{M}^g$ such that $\mathbf{z}'(m) = \mathbf{z}'(m')$. Consider an arbitrary poor message m_1 . By definition there exists $m_2 \in \mathcal{M}^g$, $m_2 \neq m_1$, such that $\mathbf{z}'(m_1) = \mathbf{z}'(m_2)$. Since $\mathbf{b}(m_1) \neq \mathbf{b}(m_2)$, there exists j such that $b_j(m_1) \neq b_j(m_2)$. Let \mathbf{r}_1 be the error pattern that changes the signal on x_j to $x_j(m_2)$, and let \mathbf{r}_2 be the error pattern that changes the signal on y_j to $y_j(m_1)$. Then if m_1 is sent and \mathbf{r}_1 happens, node B_j will receive the same inputs as in the situation that m_2 is sent and \mathbf{r}_2 happens. Therefore $b_j(m_1, \mathbf{r}_1) = b_j(m_2, \mathbf{r}_2)$ and it follows that either $b_j(m_1, \mathbf{r}_1) \neq b_j(m_1)$ or $b_j(m_2, \mathbf{r}_2) \neq b_j(m_2)$. In the former case, the tuple of signals $(b_1(m_1, \mathbf{r}_1), \dots, b_j(m_1, \mathbf{r}_1), \dots, b_k(m_1, \mathbf{r}_1)) = (b_1(m_1), \dots, b_j(m_1, \mathbf{r}_1), \dots, b_k(m_1))$ will be decoded by the terminal to message m_1 (because by hypothesis $m_1 \in \mathcal{M}^g$, which is decodable under any error $\mathbf{r} \in \mathcal{R}_A$). It is therefore an element of \mathcal{B}^b as it does not equal $\mathbf{b}(m_1)$. Similarly, in the latter case, $(b_1(m_2, \mathbf{r}_2), \dots, b_j(m_2, \mathbf{r}_2), \dots, b_k(m_2, \mathbf{r}_2)) = (b_1(m_2), \dots, b_j(m_2, \mathbf{r}_2), \dots, b_k(m_2))$ will be decoded by the terminal to message m_2 and is an element of \mathcal{B}^b . For $\hat{\mathbf{z}} \in [2^n]^k$, let $\mathcal{M}(\hat{\mathbf{z}})$ be the set of messages $\{m \in \mathcal{M}^g : \mathbf{z}'(m) = \hat{\mathbf{z}}\}$. Then if $|\mathcal{M}(\hat{\mathbf{z}})| > 1$, by the argument above, there are at least $\lfloor \frac{|\mathcal{M}(\hat{\mathbf{z}})|}{2} \rfloor \geq \frac{|\mathcal{M}(\hat{\mathbf{z}})|}{3}$ elements of \mathcal{B}^b , such that each of them will be decoded by the terminal to some message

$m \in \mathcal{M}(\hat{z}')$. Let $\mathcal{M}^{\text{poor}}$ be the set of all poor messages, then

$$|\mathcal{M}^{\text{poor}}| = \sum_{\hat{z}': |\mathcal{M}(\hat{z}')| > 1} |\mathcal{M}(\hat{z}')| \leq 3|\mathcal{B}^b| \leq 3\epsilon \cdot 2^{kn}.$$

Let $\mathcal{M}^\circ = \mathcal{M}^g \setminus \mathcal{M}^{\text{poor}}$, then $|\mathcal{M}^\circ| = |\mathcal{M}^g| - |\mathcal{M}^{\text{poor}}| \geq (1 - \epsilon')2^{kn}$, where $\epsilon' = 4\epsilon$. This proves the assertion. \blacksquare

Let $\mathcal{A}^\circ = \{\mathbf{a}(m) | m \in \mathcal{M}^\circ\}$ and $\mathcal{A}^\times = [2^n]^k \setminus \mathcal{A}^\circ$, then $|\mathcal{A}^\circ| = |\mathcal{M}^\circ| \geq (1 - \epsilon')2^{kn}$ since by Lemma 2, $\mathbf{a}(m_1) \neq \mathbf{a}(m_2)$ for $m_1, m_2 \in \mathcal{M}^\circ, m_1 \neq m_2$. Therefore $|\mathcal{A}^\times| \leq 2^{kn}\epsilon'$. Similarly let $\mathcal{B}^\circ = \{\mathbf{b}(m) | m \in \mathcal{M}^\circ\}$ and $\mathcal{B}^\times = [2^n]^k \setminus \mathcal{B}^\circ$, then $|\mathcal{B}^\times| \leq 2^{kn}\epsilon'$. For $i = 1, \dots, k$, let $\mathcal{A}_i^\circ = \{a_i(m) | m \in \mathcal{M}^\circ\}$, then $|\mathcal{A}^\circ| \geq (1 - \epsilon')2^{kn}$ implies that $|\mathcal{A}_i^\circ| \geq (1 - \epsilon')2^n$. For $\hat{a}_i \in \mathcal{A}_i^\circ$, let $\mathcal{M}(\hat{a}_i) = \{m \in \mathcal{M}^\circ : a_i(m) = \hat{a}_i\}$ and define $N(\hat{a}_i) = |\mathcal{M}(\hat{a}_i)|$. Furthermore define $\mathcal{A}_{i,l}^\circ = \{\hat{a}_i \in \mathcal{A}_i^\circ | N(\hat{a}_i) \geq (1 - \epsilon')2^{(k-1)n}\}$.

We show that the size of $\mathcal{A}_{i,l}^\circ$ is large. Consider any $\hat{a}_i \in \mathcal{A}_i^\circ \setminus \mathcal{A}_{i,l}^\circ$, then by definition $|\{(a_1, \dots, a_k) \in \mathcal{A}^\circ : a_i = \hat{a}_i\}| < (1 - \epsilon')2^{(k-1)n}$. And because $|\{(a_1, \dots, a_k) \in [2^n]^k : a_i = \hat{a}_i\}| = 2^{(k-1)n}$, there are at least $\epsilon' \cdot 2^{(k-1)n}$ elements of \mathcal{A}^\times such that their i -th entry equals to \hat{a}_i . Therefore $|\mathcal{A}_i^\circ \setminus \mathcal{A}_{i,l}^\circ| \cdot \epsilon' \cdot 2^{(k-1)n} \leq |\mathcal{A}^\times| \leq 2^{kn}\epsilon'$, and $|\mathcal{A}_i^\circ \setminus \mathcal{A}_{i,l}^\circ| \leq 2^n/\epsilon'$. So $|\mathcal{A}_{i,l}^\circ| \geq |\mathcal{A}_i^\circ| - 2^n/\epsilon' \geq (1 - \epsilon' - 1/l)2^n$. Define $\mathcal{B}_{i,l}^\circ, \mathcal{B}_{i,l}^\times, \mathcal{Z}_{i,l}^\circ$ and $\mathcal{Z}_{i,l}^\times$ similarly, then it follows from the same argument that $|\mathcal{B}_{i,l}^\circ|, |\mathcal{Z}_{i,l}^\circ| \geq (1 - \epsilon' - 1/l)2^n$.

Step 2: Connect a_i and b_i .

Let M° be the random variable that is uniformly distributed over \mathcal{M}° . In the following we show that if M° is the source message then $I(a_i; b_i)/n \rightarrow 1$ as $\epsilon \rightarrow 0, i = 1, \dots, k$. We start by lower bounding the entropy $H(b_i)$. Consider any $\hat{b}_i \in \mathcal{B}_{i,l}^\circ$, then

$$\Pr\{b_i = \hat{b}_i\} = \frac{N(\hat{b}_i)}{|\mathcal{M}^\circ|} \geq \frac{(1 - \epsilon')2^{(k-1)n}}{2^{kn}} = \frac{1 - \epsilon'}{2^n} \quad (2)$$

$$\Pr\{b_i = \hat{b}_i\} = \frac{N(\hat{b}_i)}{|\mathcal{M}^\circ|} \leq \frac{2^{(k-1)n}}{(1 - \epsilon')2^{kn}} = \frac{1}{2^n(1 - \epsilon')} \quad (3)$$

Therefore,

$$\begin{aligned} H(b_i) &= - \sum_{\hat{b}_i \in \mathcal{B}_{i,l}^\circ} \Pr\{\hat{b}_i\} \log(\Pr\{\hat{b}_i\}) \\ &\geq - \sum_{\hat{b}_i \in \mathcal{B}_{i,l}^\circ} \Pr\{\hat{b}_i\} \log(\Pr\{\hat{b}_i\}) \\ &\stackrel{(a)}{\geq} -|\mathcal{B}_{i,l}^\circ| \cdot \frac{1 - \epsilon'}{2^n} \cdot \log\left(\frac{1}{2^n(1 - \epsilon')}\right) \\ &\geq \left(1 - \epsilon' - \frac{1}{l}\right) 2^n \cdot \frac{1 - \epsilon'}{2^n} \cdot \log(2^n(1 - \epsilon')) \\ &= \left(1 - \epsilon' - \frac{1}{l}\right) (1 - \epsilon')(n + \log(1 - \epsilon')) \quad (4) \end{aligned}$$

where (a) is due to (2) and (3). Similarly it follows that,

$$H(a_i) \geq \left(1 - \epsilon' - \frac{1}{l}\right) (1 - \epsilon')(n + \log(1 - \epsilon')) \quad (5)$$

$$H(z'_i) \geq \left(1 - \epsilon' - \frac{1}{l}\right) (1 - \epsilon')(n + \log(1 - \epsilon')), \quad (6)$$

In the next step we upper bound $H(b_i|a_i)$. Recall that $\mathcal{M}(\hat{a}_i) = \{m \in \mathcal{M}^\circ : a_i(m) = \hat{a}_i\}$, we first prove a useful lemma.

Lemma 3. Suppose $\{b_i(m) : m \in \mathcal{M}(\hat{a}_i)\} = \{\hat{b}_i^{(1)}, \dots, \hat{b}_i^{(L)}\}$, then there exist $(L - 1)N(\hat{a}_i)$ distinct elements of \mathcal{B}^\times such that each of them will be decoded by the terminal to some message $m \in \mathcal{M}(\hat{a}_i)$.

Proof: Consider arbitrary $\hat{a}_i \in \mathcal{A}_i^\circ$, by hypothesis there exist L messages $m_1, \dots, m_L \in \mathcal{M}(\hat{a}_i)$ such that $b_i(m_j) = \hat{b}_i^{(j)}$, $j = 1, \dots, L$. For $j = 1, \dots, L$, let \mathbf{r}_j be the error pattern that changes the signal on z'_i to be $z'_i(m_j)$. Then if an arbitrary message $m_0 \in \mathcal{M}(\hat{a}_i)$ is transmitted by the source and \mathbf{r}_j happens, the node B_i will receive the same inputs as in the situation that m_j is sent and no error happens. Therefore $b_i(m_0, \mathbf{r}_j) = \hat{b}_i^{(j)}$, and so $\mathbf{b}(m_0, \mathbf{r}_j)$ takes L distinct values for $j = 1, \dots, L$. Since $m_0 \in \mathcal{M}^\circ$ is decodable under any error pattern $\mathbf{r} \in \mathcal{R}_{\mathcal{A}}$, it follows that $\mathbf{b}(m_0, \mathbf{r}_j)$ will be decoded by the terminal to m_0 for all $j = 1, \dots, L$. Among these L values, i.e., $\{\mathbf{b}(m_0, \mathbf{r}_j), j = 1, \dots, L\}$, only one is equal to $\mathbf{b}(m_0)$, and the remaining $L - 1$ of them are elements of \mathcal{B}^\times . Sum over all $m_0 \in \mathcal{M}(\hat{a}_i)$ and the assertion is proved. \blacksquare

Partition $\mathcal{A}_{i,l}^\circ$ into $\mathcal{A}_{i,L=1}^\circ + \mathcal{A}_{i,L>1}^\circ$, such that every element of $\mathcal{A}_{i,L=1}^\circ$ has a corresponding $L = 1$ as defined in Lemma 3. Then it follows from Lemma 3 that:

$$|\mathcal{A}_{i,L>1}^\circ| \cdot (1 - \epsilon') \cdot 2^{(k-1)n} \leq |\mathcal{B}^\times| \leq 2^{kn}\epsilon'$$

and so

$$|\mathcal{A}_{i,L>1}^\circ| \leq \frac{\epsilon' \cdot 2^n}{1 - \epsilon'}. \quad (7)$$

We are ready to upper bound $H(b_i|a_i)$.

$$\begin{aligned} H(b_i|a_i) &= - \sum_{\hat{a}_i \in \mathcal{A}_i^\circ} \Pr\{\hat{a}_i\} \sum_{\hat{b}_i \in \mathcal{B}_i^\circ} \Pr\{\hat{b}_i|\hat{a}_i\} \log \Pr\{\hat{b}_i|\hat{a}_i\} \\ &\leq I_1 + I_2 + I_3, \quad (8) \end{aligned}$$

where

$$\begin{aligned} I_1 &= - \sum_{\hat{a}_i \in \mathcal{A}_i^\circ \setminus \mathcal{A}_{i,l}^\circ} \Pr\{\hat{a}_i\} \sum_{\hat{b}_i \in \mathcal{B}_i^\circ} \Pr\{\hat{b}_i|\hat{a}_i\} \log \Pr\{\hat{b}_i|\hat{a}_i\} \\ I_2 &= - \sum_{\hat{a}_i \in \mathcal{A}_{i,L>1}^\circ} \Pr\{\hat{a}_i\} \sum_{\hat{b}_i \in \mathcal{B}_i^\circ} \Pr\{\hat{b}_i|\hat{a}_i\} \log \Pr\{\hat{b}_i|\hat{a}_i\} \\ I_3 &= - \sum_{\hat{a}_i \in \mathcal{A}_{i,L=1}^\circ} \Pr\{\hat{a}_i\} \sum_{\hat{b}_i \in \mathcal{B}_i^\circ} \Pr\{\hat{b}_i|\hat{a}_i\} \log \Pr\{\hat{b}_i|\hat{a}_i\} \end{aligned}$$

We now bound I_1, I_2 and I_3 respectively.

$$\begin{aligned} I_1 &\leq \sum_{\hat{a}_i \in \mathcal{A}_i^\circ \setminus \mathcal{A}_{i,l}^\circ} \Pr\{\hat{a}_i\} \log |2^n| \\ &= n \sum_{\hat{a}_i \in \mathcal{A}_i^\circ \setminus \mathcal{A}_{i,l}^\circ} \Pr\{\hat{a}_i\} \\ &\leq n \frac{|\mathcal{M}^\circ| - \sum_{\hat{a}_i \in \mathcal{A}_{i,l}^\circ} N(\hat{a}_i)}{|\mathcal{M}^\circ|} \end{aligned}$$

$$\begin{aligned}
&\leq n \left(1 - \frac{|\mathcal{A}_{i,l}^\circ| \cdot (1 - l\epsilon') 2^{(k-1)n}}{|\mathcal{M}^\circ|} \right) \\
&\leq n \left(1 - \frac{(1 - 1/l - \epsilon') 2^n (1 - l\epsilon') 2^{(k-1)n}}{2^{kn}} \right) \\
&< (1/l + l\epsilon')n.
\end{aligned}$$

$$\begin{aligned}
I_2 &\leq \sum_{\hat{a}_i \in \mathcal{A}_{i,L>1}^\circ} \Pr\{\hat{a}_i\} \log |2^n| \\
&= n \sum_{\hat{a}_i \in \mathcal{A}_{i,L>1}^\circ} \Pr\{\hat{a}_i\} \\
&\leq n \frac{|\mathcal{A}_{i,L>1}^\circ| 2^{(k-1)n}}{2^{kn}(1 - \epsilon')} \\
&\stackrel{(b)}{\leq} \frac{\epsilon'}{(1 - \epsilon')(1 - l\epsilon')} n,
\end{aligned} \tag{10}$$

where (b) follows from (7). Finally, by definition if $\hat{a}_i \in \mathcal{A}_{i,L=1}^\circ$, then there is a unique $\hat{b}_i \in \mathcal{B}_i^\circ$ such that for all messages $m \in \mathcal{M}(\hat{a}_i)$, it follows that $b_i(m) = \hat{b}_i$. Therefore $I_3 = 0$. Substituting (10) and (9) into (8) we have

$$H(b_i|a_i) < \left(\frac{1}{l} + l\epsilon' \right) n + \frac{\epsilon'}{(1 - \epsilon')(1 - l\epsilon')} n.$$

Together with (4), it follows

$$\begin{aligned}
I(a_i; b_i) &> \left(1 - \epsilon' - \frac{1}{l} \right) (1 - l\epsilon')(n + \log(1 - \epsilon')) \\
&\quad - \left(\frac{1}{l} + l\epsilon' \right) n - \frac{\epsilon'}{(1 - \epsilon')(1 - l\epsilon')} n.
\end{aligned} \tag{11}$$

Step 3: Connect z'_i and b_i .

Next we show that $I(b_i; z'_i)/n \rightarrow 1$, as $\epsilon \rightarrow 0$, $n \rightarrow \infty$, for $i = 1, \dots, k$, by upper bounding $H(z'_i|b_i)$. We first make some useful observations.

Lemma 4. *Let $\mathcal{M}(z'_i) = \{m \in \mathcal{M}^\circ : z'_i(m) = z'_i\}$, then for any $m_1, m_2 \in \mathcal{M}(z'_i)$ such that $b_i(m_1) \neq b_i(m_2)$, there exists an element of \mathcal{B}^\times that will be decoded by the terminal to either m_1 or m_2 .*

Proof: Consider any $m_1, m_2 \in \mathcal{M}(z'_i)$ such that $b_i(m_1) \neq b_i(m_2)$. Let \mathbf{r}_1 be the error pattern that changes the signal on x_i to be $x_i(m_2)$, and let \mathbf{r}_2 be the error pattern that changes the signal on y_i to be $y_i(m_1)$. Then if m_1 is transmitted by the source and \mathbf{r}_1 happens, the node B_i will receive the same inputs as in the situation that m_2 is transmitted and \mathbf{r}_2 happens. Therefore $b_i(m_1, \mathbf{r}_1) = b_i(m_2, \mathbf{r}_2)$, and so either $b_i(m_1, \mathbf{r}_1) \neq b_i(m_1)$ or $b_i(m_2, \mathbf{r}_2) \neq b_i(m_2)$ because by hypothesis $b_i(m_1) \neq b_i(m_2)$. Consider the first case that $b_i(m_1, \mathbf{r}_1) \neq b_i(m_1)$, then the tuple of signals $(b_1(m_1, \mathbf{r}_1), \dots, b_k(m_1, \mathbf{r}_1)) = (b_1(m_1), \dots, b_i(m_1, \mathbf{r}_1), \dots, b_k(m_1))$ will be decoded by the terminal to message m_1 because by hypothesis $m_1 \in \mathcal{M}^\circ$ which is decodable under any error pattern $\mathbf{r} \in \mathcal{R}_A$. Therefore it is an element of \mathcal{B}^\times since it does not equals $\mathbf{b}(m_1)$. Similarly in the latter case, $b_1(m_2, \mathbf{r}_2) \neq b_1(m_2)$, then $(b_2(m_2, \mathbf{r}_2), \dots, b_k(m_2, \mathbf{r}_2)) =$

$(b_2(m_2), \dots, b_i(m_2, \mathbf{r}_2), \dots, b_k(m_2))$ is an element of \mathcal{B}^\times and will be decoded by the terminal to m_2 . Therefore in either case we find an element of \mathcal{B}^\times that will be decoded to either m_1 or m_2 . ■

Lemma 5. *Let $\mathcal{M}(z'_i, \hat{b}_i) = \{m \in \mathcal{M}^\circ : z'_i(m) = z'_i, b_i(m) = \hat{b}_i\}$, $\hat{b}_{i,z'_i} = \arg \max_{\hat{b}_i \in \mathcal{B}_i^\circ} |\mathcal{M}(z'_i, \hat{b}_i)|$ and $N(z'_i) = |\mathcal{M}(z'_i)|$, then there are at least $\frac{1}{2}(N(z'_i) - |\mathcal{M}(z'_i, \hat{b}_{i,z'_i})|)$ distinct elements of \mathcal{B}^\times that will be decoded by the terminal to some messages in $\mathcal{M}(z'_i)$.*

Proof: Let $\mathcal{W} := \mathcal{M}(z'_i)$, and we describe an iterative procedure as follow. Pick arbitrary $m_1, m_2 \in \mathcal{W}$ such that $b_i(m_1) \neq b_i(m_2)$, and then delete them from \mathcal{W} and repeat until there does not exist such m_1, m_2 . By Lemma 4, each pair of elements deleted from \mathcal{W} will generate a distinct element of \mathcal{B}^\times , which will be decoded by the terminal to either m_1 or m_2 . After the iterative procedure terminates, it follows that $|\mathcal{W}| \leq |\mathcal{M}(z'_i, \hat{b}_{i,z'_i})|$, because otherwise there must exist $m_1, m_2 \in \mathcal{W}$ such that $b_i(m_1) \neq b_i(m_2)$. Therefore at least $N(z'_i) - |\mathcal{M}(z'_i, \hat{b}_{i,z'_i})|$ elements are deleted and the lemma is proved. ■

We are now ready to upper bound $H(z'_i|b_i)$. Recall that $\mathcal{Z}'_i = \{z'_i(m) : m \in \mathcal{M}^\circ\}$. We have,

$$\begin{aligned}
H(b_i|z'_i) &= - \sum_{z'_i \in \mathcal{Z}'_i} \Pr\{z'_i\} \sum_{\hat{b}_i \in \mathcal{B}_i^\circ} \Pr\{\hat{b}_i|z'_i\} \log \Pr\{\hat{b}_i|z'_i\} \\
&= I_4 + I_5,
\end{aligned} \tag{12}$$

where

$$\begin{aligned}
I_4 &= - \sum_{z'_i \in \mathcal{Z}'_i} \Pr\{z'_i\} \Pr\{\hat{b}_{i,z'_i}|z'_i\} \log \Pr\{\hat{b}_{i,z'_i}|z'_i\} \\
&< \sum_{z'_i \in \mathcal{Z}'_i} \Pr\{z'_i\} \leq 1
\end{aligned} \tag{13}$$

and

$$\begin{aligned}
I_5 &= - \sum_{z'_i \in \mathcal{Z}'_i} \Pr\{z'_i\} \sum_{\hat{b}_i \neq \hat{b}_{i,z'_i}} \Pr\{\hat{b}_i|z'_i\} \log \Pr\{\hat{b}_i|z'_i\} \\
&\leq \sum_{z'_i \in \mathcal{Z}'_i} \Pr\{z'_i\} \sum_{\hat{b}_i \neq \hat{b}_{i,z'_i}} \Pr\{\hat{b}_i|z'_i\} \log |2^{kn}| \\
&= kn \sum_{z'_i \in \mathcal{Z}'_i} \sum_{\hat{b}_i \neq \hat{b}_{i,z'_i}} \Pr\{z'_i, \hat{b}_i\} \\
&= kn \sum_{z'_i \in \mathcal{Z}'_i} \frac{N(z'_i) - |\mathcal{M}(z'_i, \hat{b}_{i,z'_i})|}{|\mathcal{M}^\circ|} \\
&\leq \frac{kn}{2^{kn}(1 - \epsilon')} \sum_{z'_i \in \mathcal{Z}'_i} (N(z'_i) - |\mathcal{M}(z'_i, \hat{b}_{i,z'_i})|) \\
&\stackrel{(c)}{\leq} \frac{kn \cdot 2|\mathcal{B}^\times|}{2^{kn}(1 - \epsilon')} \\
&= \frac{kn \cdot 2 \cdot \epsilon' \cdot 2^{kn}}{2^{kn}(1 - \epsilon')} = \frac{2k\epsilon'}{1 - \epsilon'} n,
\end{aligned} \tag{14}$$

where (c) follows from Lemma 5. Substituting (13) and (14) to (12) we have

$$H(b_i|z'_i) \leq 1 + \frac{2k\epsilon'}{1-\epsilon'}n. \quad (15)$$

Step 4: Connect a_i and z_i .

Finally we discuss the connection between a_i and z_i . We may assume that $z_i = a_i$ without loss of generality in the following sense. For every network code that achieves rate k with error probability ϵ in \mathcal{I}_c , we can modify it slightly to obtain a new code such that $z_i = a_i$, and such that the code on all other edges and at the terminal are the same as the original code. This modification is feasible because the encoding function z_i of the original code is a function of a_i , and so if we let $z_i = a_i$, then the node s_i , $i = 1, \dots, k$, always has enough information to reproduce the original network code. Since from the perspective of the terminal, the modified code is the same as the original code, it also achieves rate k with error probability ϵ in \mathcal{I}_c . Hence,

$$I(a_i; z_i) = H(a_i). \quad (16)$$

Step 5: Connect z_i and z'_i .

By (11), (15), (16) and Lemma 1, we have

$$\begin{aligned} I(z_i; z'_i) &\geq I(a_i; z_i) + I(a_i; b_i) + I(b_i; z'_i) - H(a_i) - H(b_i) \\ &= I(a_i; b_i) + I(b_i; z'_i) - H(b_i) \\ &= I(a_i; b_i) - H(b_i|z'_i) \\ &> \left(1 - \epsilon' - \frac{1}{l}\right) (1 - l\epsilon')(n + \log(1 - \epsilon')) \\ &\quad - \left(\frac{1}{l} + l\epsilon'\right) n - \frac{\epsilon'}{(1 - \epsilon')(1 - l\epsilon')}n \\ &\quad - 1 - \frac{2k\epsilon'}{1 - \epsilon'}n. \end{aligned} \quad (17)$$

Recall that (17) is obtained under the assumption that the source message is uniformly distributed over \mathcal{M}° , and hence \mathbf{a} is uniformly distributed over \mathcal{A}° . Therefore the random variables $\{a_1(M^\circ), \dots, a_k(M^\circ)\}$ are not independent. Now we consider the case that \mathbf{a} is uniformly distributed over $[2^n]^k$ and the network code is the same as before. Specifically, the decoding function and the encoding functions at all edges except a_1, \dots, a_k are the same as the network code that achieves rate k with error probability ϵ in \mathcal{I}_c . We are interested in the mutual information between z_i and z'_i under this setting where the random variables $\{a_1, \dots, a_k\}$ are independent. Let $\mathbf{a} \leftarrow [2^n]^k$ denote that the distribution of \mathbf{a} is uniform over $[2^n]^k$,

$$\begin{aligned} I_{\mathbf{a} \leftarrow [2^n]^k}(z_i; z'_i) &= \Pr_{\mathbf{a} \leftarrow [2^n]^k} \{\mathbf{a} \in \mathcal{A}^\circ\} I_{\mathbf{a} \leftarrow [2^n]^k}(z_i; z'_i | \mathbf{a} \in \mathcal{A}^\circ) \\ &\quad + \Pr_{\mathbf{a} \leftarrow [2^n]^k} \{\mathbf{a} \in \mathcal{A}^\times\} I_{\mathbf{a} \leftarrow [2^n]^k}(z_i; z'_i | \mathbf{a} \in \mathcal{A}^\times). \end{aligned} \quad (18)$$

Note that

$$I_{\mathbf{a} \leftarrow [2^n]^k}(z_i; z'_i | \mathbf{a} \in \mathcal{A}^\circ) = I_{\mathbf{a} \leftarrow \mathcal{A}^\circ}(z_i; z'_i) = I(z_i; z'_i), \quad (19)$$

which is exactly the result we computed in (17). And

$$\Pr_{\mathbf{a} \leftarrow [2^n]^k} \{\mathbf{a} \in \mathcal{A}^\circ\} = |\mathcal{A}^\circ|/2^{kn} \geq (1 - \epsilon'). \quad (20)$$

Therefore by (19) and (20), $I_{\mathbf{a} \leftarrow [2^n]^k}(z_i; z'_i) \geq (1 - \epsilon')I(z_i; z'_i)$. Then by (17), for any $\epsilon_1 > 0$, by first choosing a sufficiently large l and then a sufficiently small ϵ and a sufficiently large n , it follows that $I_{\mathbf{a} \leftarrow [2^n]^k}(z_i; z'_i)/n > 1 - \epsilon_1$, $i = 1, \dots, k$. Hence unit rate is asymptotically feasible in \mathcal{I} by the channel coding theorem [25]. This completes the proof of the first part of the theorem.

“ \Leftarrow ”. Conversely, we show that if unit rate is feasible in \mathcal{I} , then rate k is feasible in \mathcal{I}_c . Again we use the constructive scheme in Figure 2. In \mathcal{I}_c , the source lets $M = (M_1, \dots, M_k)$, where the M_i 's are i.i.d. uniformly distributed over $[2^n]$. Let the network code be $a_i(M) = x_i(M) = y_i(M) = z_i(M) = z'_i(M) = M_i$, $i = 1, \dots, k$, and let node B_i , $i = 1, \dots, k$ performs majority decoding. The terminal t will not decode an error as long as the multiple-unicast instance \mathcal{I} does not commit an error. This happens with probability at least $1 - \epsilon$, which implies the feasibility of rate k in \mathcal{I}_c .

V. CONCLUSION

We summarize the results of this paper in Fig. 3 which expresses the possible connections between the feasibility of a general multiple-unicast network coding instance \mathcal{I} and its reduced single-unicast network error correction instance \mathcal{I}_c . Our results present an equivalence between multiple-unicast and network error correction under zero-error communication. Namely, determining the feasibility of zero-error unit rate in \mathcal{I} is equivalent to determining the feasibility of zero-error rate k in \mathcal{I}_c . This is expressed in Fig. 3 by the fact that the two states in the first row are connected by a single solid edge and there are no other edges connected to these two states.

For the vanishing error model, however, the implication of our results are more involved. In our reduction there is a slight slackness, which gives rise to the two dashed lines in Fig. 3. For example, consider the case that unit rate is feasible (but not with zero error) in \mathcal{I} . Then it follows that rate k is feasible (but not with zero error) in \mathcal{I}_c . This fact is expressed by the sole solid line leaving the state that unit rate is feasible (but not with zero error) in \mathcal{I} . However, if rate k is feasible (but not with zero error) in \mathcal{I}_c , then there are potentially two possibilities: (a) that unit rate is feasible (but not with zero error) in \mathcal{I} ; and (b) that unit rate is asymptotically feasible (but not exactly feasible) in \mathcal{I} . Option (a) is represented by a solid line, as indeed we have observed instance \mathcal{I} with a corresponding \mathcal{I}_c that fits this setting. Option (b) is represented by a dashed line, as on one the one hand it has not been ruled out by our analysis, but on the other hand we are not aware of how to construct instances \mathcal{I} with a corresponding \mathcal{I}_c that fit this setting.

All in all, under the vanishing error model, the two dashed lines in Fig. 3 do not allow us to directly determine the feasibility of \mathcal{I} based on the feasibility of \mathcal{I}_c . Nevertheless, we may consider the following problem on \mathcal{I} which can be

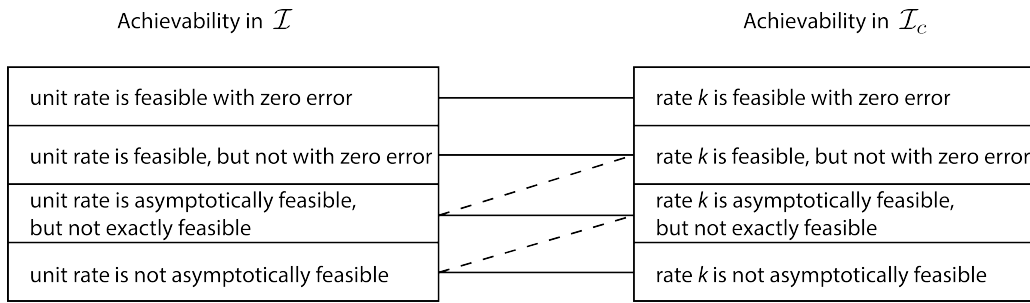


Fig. 3: Equivalence between multiple-unicast and network error correction under the construction of Fig. 1. In our reduction, given an instance \mathcal{I} of the multiple-unicast network coding problem, we construct an instance \mathcal{I}_c of the single-unicast network error correction problem. This figure expresses our current understanding of the relation between achievability in \mathcal{I} and \mathcal{I}_c . A solid line between two states means that there exist instances of \mathcal{I} and \mathcal{I}_c with the corresponding achievability. If there is no line between two states, there do not exist instances of \mathcal{I} and \mathcal{I}_c with the corresponding achievability. A dashed line between two states means that whether there is a solid line between the two states or not is still an open problem.

solved by the study of \mathcal{I}_c using our results: Given an instance \mathcal{I} , partially determine between the three possible settings in the following manner: if unit rate is feasible (but not with zero error) in \mathcal{I} , answer *yes*; if unit rate is not asymptotically feasible in \mathcal{I} , answer *no*; if unit rate is asymptotically feasible but not exactly feasible in \mathcal{I} , then any answer is considered correct. By our results, answering *yes* if and only if rate k is feasible (but not with zero error) in \mathcal{I}_c solves the problem above. Whether the partial distinction problem above on \mathcal{I} is a *difficult* one (compared to the standard feasibility of multiple-unicast network coding) is yet to be established.

Finally we note that our reduction and analysis in the paper consider the case that the connections between source destination pairs in \mathcal{I} have unit rate. Our reduction can be adapted to connections with different rates.

ACKNOWLEDGEMENT

This work has been supported in part by NSF grant CCF-1440014, CCF-1440001, CCF-1439465, and CCF-1321129.

REFERENCES

- [1] R. W. Yeung and N. Cai, "Network Error Correction, I: Basic Concepts and Upper Bounds," *Communications in Information & Systems*, vol. 6, no. 1, pp. 19–35, 2006.
- [2] —, "Network Error Correction, II: Lower Bounds," *Communications in Information & Systems*, vol. 6, no. 1, pp. 37–54, 2006.
- [3] R. Koetter and F. R. Kschischang, "Coding for Errors and Erasures in Random Network Coding," *IEEE Transactions on Information Theory*, vol. 54, no. 8, pp. 3579–3591, 2008.
- [4] S. Jaggi, M. Langberg, S. Katti, T. Ho, D. Katabi, M. Medard, and M. Effros, "Resilient Network Coding in the Presence of Byzantine Adversaries," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2596–2603, 2008.
- [5] D. Silva, F. R. Kschischang, and R. Kotter, "A Rank-Metric Approach to Error Control in Random Network Coding," *IEEE Transactions on Information Theory*, vol. 54, no. 9, pp. 3951–3967, 2008.
- [6] Z. Zhang, "Linear network error correction codes in packet networks," *IEEE Transactions on Information Theory*, vol. 54, no. 1, pp. 209–218, Jan. 2008.
- [7] L. Nutman and M. Langberg, "Adversarial Models and Resilient Schemes for Network Coding," in *IEEE International Symposium on Information Theory*, 2008, pp. 171–175.
- [8] W. Huang, T. Ho, H. Yao, and S. Jaggi, "Rateless resilient network coding against byzantine adversaries," in *IEEE INFOCOM mini conference*. IEEE, 2013, pp. 265–269.
- [9] S. Kim, T. Ho, M. Effros, and A. S. Avestimehr, "Network Error Correction With Unequal Link Capacities," *IEEE Transactions on Information Theory*, vol. 57, no. 2, pp. 1144–1164, Feb. 2011.
- [10] O. Kosut, L. Tong, and D. Tse, "Nonlinear Network Coding is Necessary to Combat General Byzantine Attacks," in *Allerton Conference on Communication, Control, and Computing*, 2009, pp. 593–599.
- [11] D. Wang, D. Silva, and F. R. Kschischang, "Robust Network Coding in the Presence of Untrusted Nodes," *IEEE Transactions on Information Theory*, vol. 56, no. 9, pp. 4532–4538, 2010.
- [12] O. Kosut, L. Tong, and D. N. C. Tse, "Polytope codes against adversaries in networks," in *IEEE International Symposium on Information Theory*, 2010, pp. 2423–2427.
- [13] P. H. Che, M. Chen, T. Ho, S. Jaggi, and M. Langberg, "Routing for Security in Networks with Adversarial Nodes," in *IEEE NetCod*, 2013, pp. 1–6.
- [14] A. R. Lehman, "Network Coding," Ph.D. dissertation, Massachusetts Institute of Technology, 2005.
- [15] M. Langberg and M. Medard, "On the multiple unicast network coding conjecture," in *Allerton Conference on Communication, Control, and Computing*, 2009, pp. 222–227.
- [16] R. Dougherty, C. Freiling, and K. Zeger, "Insufficiency of Linear Coding in Network Information Flow," *IEEE Transactions on Information Theory*, vol. 51, no. 8, pp. 2745–2759, 2005.
- [17] —, "Networks, Matroids, and Non-Shannon Information Inequalities," *IEEE Transactions on Information Theory*, vol. 53, no. 6, pp. 1949–1969, 2007.
- [18] R. Dougherty and K. Zeger, "Nonreversibility and Equivalent Constructions of Multiple-Unicast Networks," *IEEE Transactions on Information Theory*, vol. 52, no. 11, pp. 5067–5077, 2006.
- [19] M. F. Wong, M. Langberg, and M. Effros, "On a capacity equivalence between multiple multicast and multiple unicast," in *Allerton Conference on Communication, Control, and Computing*, 2013, pp. 1537–1544.
- [20] S. El Rouayheb, A. Sprintson, and C. Georghiadis, "On the Index Coding Problem and Its Relation to Network Coding and Matroid Theory," *IEEE Transactions on Information Theory*, vol. 56, no. 7, pp. 3187–3195, 2010.
- [21] M. Effros, S. El Rouayheb, and M. Langberg, "An Equivalence between Network Coding and Index Coding," in *IEEE International Symposium on Information Theory*, 2013, pp. 967–971.
- [22] T. Chan and A. Grant, "Mission impossible: Computing the network coding capacity region," in *IEEE International Symposium on Information Theory*, Jan. 2008, pp. 320–324.
- [23] W. Huang, T. Ho, M. Langberg, and J. Kliever, "On secure network coding with uniform wiretap sets," in *IEEE NetCod*, 2013, pp. 1–6.
- [24] S. Kamath, D. N. C. Tse, and C.-C. Wang, "Two-unicast is hard," in *2014 IEEE International Symposium on Information Theory (ISIT)*, 2014, pp. 2147–2151.
- [25] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. John Wiley & Sons, Jul. 2006.