Network Coding on Directed Acyclic Graphs

John MacLaren Walsh, Ph.D.

Multiterminal Information Theory, Spring Quarter, 2012

1 Reference

These notes are directly derived from Chapter 21 of R. W. Yeung’s Information Theory and Network Coding, Springer, 2008.

2 Motivating Example: Butterfly Network

![Butterfly Network Diagram](image)

Figure 1: The butterfly example.

We presented the butterfly example as a case where we could increase the capacity region of a network by incorporating coding between flows at intermediate network nodes. Summing the messages on link $b$ allows both $s_1$ and $s_2$ to be multicasted to both sinks $t_1, t_2$.

3 Definition of a Network Code & Coding Capacity Region

In these lectures we model a network as a directed acyclic graph $G = (V, E)$. There is a finite set of vertices or nodes $V$, and collection of edges $e \in E$ which are ordered pairs of vertices $e = (v_1, v_2)$, $v_1, v_2 \in V$. We call vertex $v_1$ the tail of edge $e = (v_1, v_2)$ and the vertex $v_2$ the head of edge $e$. A sequence of vertices $e_1, e_2, \ldots, e_k$ such that the head of edge $e_n$ is the tail of the next edge $e_{n+1}$ is called a directed path in the graph $G$. A directed path with the property that the tail of $e_1$ is the head of edge $e_k$ is called a cycle, and the graph is called acyclic if it has no cycles.

There is also a set of source nodes $S \subset V$ and sink nodes $T \subset V$ with $S \cap T = \emptyset$. Each source node $s$ is endowed with a source variable $X_s$ uniformly distributed over the set $X_s = \{1, 2, \ldots, 2^{W_s}\}$. The variables $X_s$, $s \in S$ are mutually independent, and represent messages to be sent over the network. Additionally, the edges of the graph are associated with capacity limitations $R_e$, indicating the number of bits per source time instant which can be sent over these edges.

Each sink $t \in T$ has a subset of source variables, those with indices in $\beta(t) \subset S$, that it wishes to determine. In order to make this possible, each node in the network will encode all of the messages it hears on its incoming edges (i.e. all those edges that have it as a tail) into a message to be sent on its outgoing edges (i.e. all those edges that have it as a head). These nodes do this using the functions

$$\forall i \in V \setminus (S \cup T), \ e \in \text{Out}(i) \quad k_e : \prod_{d \in \text{In}(e)} \{0, 1, \ldots, \eta_d\} \to \{0, 1, \ldots, \eta_e\}$$

where $\text{In}(i) = \{e \in E | e = (j, i) \text{ some } j \in V\}$ is the set of edges having node $i$ as their tail and $\text{Out}(i) = \{e \in E | e = (i, j) \text{ some } j \in V\}$ is the set of edges having node $i$ as their head. The source nodes encode their sources into messages.
through the functions
\[ s \in S, \ e \in \text{Out}(s), \ k_e : \mathcal{X}_s \to \{0, 1, \ldots, \eta_e\} \]

(2)
The sink nodes reproduce the source messages using the encoded messages available locally to them through the functions
\[ t \in T, \ g_t : \prod_{d \in \text{In}(t)} \{0, 1, \ldots, \eta_d\} \to \prod_{s \in \beta(t)} \mathcal{X}_s \]

(3)
The aggregate of these functions (1),(2),(3) are collectively known as a \((N, (\eta_e : e \in E), (\tau_s : s \in S))\) network code.

For this to work, all of the messages must have arrived on the incoming edges before the ones on the outgoing edges are calculated. This is enabled through the assumption of an acyclic network. For any finite directed acyclic graph, it is possible to order the nodes in a sequence such that if \(\exists e \in E, e = (i, j)\), then node \(i\) appears before node \(j\) in the sequence. By selecting such an order to perform the encoding among the nodes, every node will have all of the messages from its incoming edges before it calculates the message on the outgoing edges.

Let \(\tilde{g}_t(X_S)\) represent the composition of all functions from the sources to the sink \(t\). We say that a collection of source rates \(\omega_s, s \in S\) are achievable if for arbitrarily small \(\epsilon > 0\), there exists a network code such that

\[ N^{-1} \log_2 \eta_e \leq R_e + \epsilon \quad e \in E \]

(4)
\[ \tau_s \geq \omega_s - \epsilon \]

(5)
\[ \mathbb{P}[\tilde{g}_t(X_S) \neq X_{\beta(t)}] \leq \epsilon \]

(6)
The set of all achievable rate vectors \(\omega\), denoted by \(\mathcal{R}\), is the network coding capacity region.

## 4 Network Coding Capacity Region

If a collection of rates were \(\omega\) achievable with zero probability of error and a block length of \(N = 1\) for a particular network code, then, after identifying the random variables \(Y_s, s \in S\) with the sources \(X_s, s \in S\), and the random variables \(U_e\) with the coded message on edge \(e\), for this code we would have the inequalities

\[ H(Y_s) \geq \omega_s \quad s \in S \]

(7)
\[ H(Y_S) = \sum_{s \in S} H(Y_s) \]

(8)
\[ H(U_{\text{Out}(s)}|Y_s) = 0 \quad s \in S \]

(9)
\[ H(U_{\text{Out}(i)}|U_{\text{In}(i)}) = 0 \quad i \in V \setminus (S \cup T) \]

(10)
\[ H(U_e) \leq R_e \]

(11)
\[ H(Y_{\beta(t)}|U_{\text{In}(t)}) = 0 \]

(12)

Indeed

- (7) reflects the fact that the sources must be uniform over a set with cardinality \(2^{N\tau_s}\) with \(\tau_s \geq \omega_s\)
- (8) reflects the requirement that the sources are independent of one another.
- (9) reflects that the message encoded by a source node is a function of the source available to it, i.e. (2)
- (10) reflects that the messages on the outgoing edges from a node are a function of the messages on its incoming edges, i.e. (1)
- (11) reflects the limitations on edge capacity (4)
- (12) indicates the zero probability of error reconstruction.

Of course, our notion of an achievable rate in the network coding capacity region \(\mathcal{R}\), was the usual Shannon lossless notion, which allows a non-zero, but finite probability of error as indicated by (6) and an arbitrarily large block length \(N\) and closure in rate space as represented by (4) and (6). Surprisingly, this lossless network coding capacity region can be written directly in terms of the inequalities (7,8,9,10,11,12) with an expression we shall define presently.

The first bit of notation is to stack subset entropies into a vector. That is, given a collection of \(M = |V| + |E|\) random variables, there are \(2^M - 1\) non-empty subsets of the random variables, and to each such subset we have an associated entropy.
We stack the entropies of these subsets into a vector \( \mathbf{h} \) of dimension \( 2^M - 1 \), and will index this vector via the subset, so that for instance \( h_A \) will represent the joint entropy of the random variables in \( A \). (The ordering for the indexing can be done, for instance, by using the integer associated with the length \( M \) binary string whose \( k \)th bit indicates whether or not \( k \) is in \( A \).)

We then consider each of the inequalities (8,9,10,11,12) as linear inequalities for this vector, defining the linear constraint sets

\[
\begin{align*}
\mathcal{L}_1 & = \left\{ \mathbf{h} \in \mathbb{R}^{2^M - 1}_+ \mid h_{Y_s} = \sum_{s \in S} h_{Y_s} \right\} \\
\mathcal{L}_2 & = \left\{ \mathbf{h} \in \mathbb{R}^{2^M - 1}_+ \mid h_{U_{\text{out}(s)} \cup Y_s} - h_{Y_s} = 0 \quad \forall s \in S \right\} \\
\mathcal{L}_3 & = \left\{ \mathbf{h} \in \mathbb{R}^{2^M - 1}_+ \mid h_{U_{\text{out}(s)} \cup U_{\text{in}(i)}} - h_{U_{\text{in}(i)}} = 0 \quad \forall i \in V \setminus (S \cup T) \right\} \\
\mathcal{L}_4 & = \left\{ \mathbf{h} \in \mathbb{R}^{2^M - 1}_+ \mid h_{U_e} \leq R_e \quad \forall e \in E \right\} \\
\mathcal{L}_5 & = \left\{ \mathbf{h} \in \mathbb{R}^{2^M - 1}_+ \mid h_{Y_{\beta(t)} \cup U_{\text{in}(t)}} - h_{U_{\text{in}(t)}} = 0 \quad \forall t \in T \right\}
\end{align*}
\]

Additionally, introduce the following notation

1. \( \text{Proj}_{Y_s}(B) := \{(h_{Y_s} : s \in S) \mid h \in B\} \)
2. \( \Lambda(B) := \{h \in \mathbb{R}^{|S|}_+ \mid h \leq h', \exists h' \in B\} \)
3. \( \text{conv}(B) \) the convex hull of the set \( B \)
4. \( \bar{B} \) the closure of the set \( B \)
5. \( \Gamma^* = \left\{ h \in \mathbb{R}^{2^M - 1}_+ \mid \exists (Z_1, Z_2, \ldots, Z_M) \text{ finite discrete random variables with } h_A = H(Z_A) \forall A \subseteq \{1, \ldots, M\} \right\} \)

Yeung and his co-workers have shown that the network coding capacity region \( \mathcal{R} \) is equal to

\[
\mathcal{R}' = \Lambda \left( \text{Proj}_{Y_s} \left( \text{conv} \left( \Gamma^* \cap \mathcal{L}_{123} \right) \cap \mathcal{L}_4 \cap \mathcal{L}_5 \right) \right)
\]

### 4.1 Converse Sketch

For the converse, we must show that \( \mathcal{R} \subseteq \mathcal{R}' \). Consider an achievable rate vector \( \omega \in \mathcal{R} \) and a monotone decreasing sequence \( \epsilon_k \to 0 \) as \( k \to \infty \). Then for every \( k \) for every \( N \) sufficiently large there exists a network code such that

\[
N^{-1} \log_2 \eta_e \leq R_e + \epsilon_k \quad e \in E
\]

\[
\tau_s \geq \omega_s - \epsilon_k
\]

\[
\Pr \left[ \bar{g}_t(X_S) \neq X_{\beta(t)} \right] \leq \epsilon_k
\]

Let \( U_e \) be the message sent by the network code on edge \( e \) for all \( e \in E \), and identify \( Y_s = X_s \) as the source variables. Because the source variables are independent and because encodings on outgoing edges are a function of the messages on incoming edges to a node, the inequalities (7,8,9,10) hold among these finite discrete random variables \( Y_s, s \in S, U_e, e \in E \).

Now, Fano’s inequality states that

\[
H(Y_{\beta(t)} | U_{\text{in}(t)}) \leq 1 + \Delta^{(k)}_t \log_2 |Y_{\beta(t)}| = 1 + \Delta^{(k)}_t H(Y_{\beta(t)}) \leq 1 + \epsilon_k H(Y_{\beta(t)})
\]

While we can upper bound the entropy of \( H(Y_{\beta(t)}) \) using

\[
H(Y_{\beta(t)}) = I(Y_{\beta(t)}; U_{\text{in}(t)}) + H(Y_{\beta(t)} | U_{\text{in}(t)}) \leq H(U_{\text{in}(t)}) + H(Y_{\beta(t)} | U_{\text{in}(t)})
\]

\[
\leq H(U_{\text{in}(t)}) + 1 + \Delta^{(k)}_t H(Y_{\beta(t)}) \leq \sum_{e \in \text{in}(t)} \log_2(\eta_e + 1) + 1 + \Delta^{(k)}_t H(Y_{\beta(t)}) \leq \sum_{e \in \text{in}(t)} N(R_e + 2\epsilon_k) + \epsilon^{(k)} H(Y_{\beta(t)})
\]

Solving for an inequality for \( H(Y_{\beta(t)}) \) we get

\[
H(Y_{\beta(t)}) \leq \frac{N}{1 - \epsilon_k} \left( \sum_{e \in \text{in}(t)} R_e + \frac{2\epsilon_k}{N} \right)
\]
which when substituted back into Fano (22) gives

\[ H(Y_{\beta(t)}|U_{\text{in}(t)}) \leq N \left[ \frac{1}{N} + \frac{\epsilon_k}{1 - \epsilon_k} \left( \sum_{\tau \in \text{in}(t)} R_{\tau} + \frac{2\epsilon_k}{N} \right) \phi(t, N, \epsilon_k) \right] . \]  

(26)

Here, it is clear the function \( \phi_t(N, \epsilon_k) \) is bounded, is monotone decreasing in both \( k \) and \( N \), and approaches 0 as \( k, N \to \infty \).

Moving next to the edge capacity constraints, and the entropies of the sources we observe that

\[ H(U_e) \leq \log_2(\eta_e + 1) \leq N (R_e + 2\epsilon_k) \quad H(Y_s) \geq N(\omega_s - \epsilon_k) \]  

(27)

If we define the half spaces, reminiscent of \( L \)

\[ L_{4,\epsilon_k}^N = \left\{ h \in \mathbb{R}_{\geq 0}^{2M-1} \mid h_{U_e} \leq N(R_e - 2\epsilon_k) \forall e \in E \right\} \]  

(28)

\[ L_{5,\epsilon_k}^N = \left\{ h \in \mathbb{R}_{\geq 0}^{2M-1} \mid h_{Y_{\beta(t)}(t)\cup U_{\text{in}(t)}} - h_{U_{\text{in}(t)}} \leq N\phi_t(N, \epsilon_k) \forall t \in T \right\} \]  

(29)

we observe that the subset entropies of this network code \( h^{(k)} \) lie in the set

\[ h^{(k)} \in \Gamma_{M}^{**} \cap L_{123} \cap L_{4,\epsilon_k}^N \cap L_{5,\epsilon_k}^N, \quad h_{Y_s} \geq N(\omega_s - \epsilon_k), \quad s \in S \]  

(30)

Now, since \( 0 \in \Gamma_{M}^{**} \cap L_{123} \cap L_{4,\epsilon_k}^N \cap L_{5,\epsilon_k}^N \), and \( N^{-1}h \) can be viewed as a convex combination with \( h \) and zero, we observe that

\[ N^{-1}h^{(k)} \in \text{conv}(\Gamma_{M}^{**} \cap L_{123}) \cap L_{4,\epsilon_k} \cap L_{5,\epsilon_k}, \quad N^{-1}h_{Y_s} \geq \omega_s - \epsilon_k \]  

(31)

Defining the constraint set in (31) to be \( B^{(N,k)} \) we observe that the \( B^{(N,k)} \)'s are monotone decreasing in that

\[ B^{(N+1,k)} \subseteq B^{(N,k)} \quad B^{(N,k+1)} \subseteq B^{(N,k)} \]  

(34)

hence

\[ \lim_{N,k \to \infty} B^{(N,k)} = \bigcap_{N=1}^{\infty} \bigcap_{k=1}^{\infty} B^{(N,k)} \]  

(35)

and the latter set, since it involves the intersection of the inequalities in \( L_{4,\epsilon_k} \) and \( L_{5,\epsilon_k} \), becomes

\[ \lim_{N,k \to \infty} N^{-1}h^{(k)} \in \text{conv}(\Gamma_{M}^{**} \cap L_{123}) \cap L_{4} \cap L_{5}, \quad \lim_{N,k \to \infty} N^{-1}h_{Y_s} \geq \omega_s, \quad s \in S \]  

(36)

Rearranging this fact, we have that if \( \omega \) is achievable, then

\[ \omega \in A \left( \text{proj}_{Y_s} \left( \text{conv}(\Gamma_{M}^{**} \cap L_{123}) \cap L_{4} \cap L_{5} \right) \right) \]  

(37)

which is what we needed to prove.

### 4.2 Obtaining Inner and Outer Bounds

We discussed that the capacity region presented in (18) is implicit, in that we don’t generally know all of the inequalities necessary to describe \( \Gamma_M^{**} \) (in fact, the closure of its infinite cardinality counterpart \( \Gamma_M^{**} \) is not even polyhedral for \( M \geq 4 \)). It is possible to obtain inner and outer bounds on the capacity region by substituting in inner and outer bounds for \( \Gamma_M^{**} \). We discussed a polyhedral outer bound for \( \Gamma_M^{**} \) known as the Shannon outer bound \( \Gamma_M \). One way to write the Shannon outer bound is the set of vectors obeying the properties that entropy is a non-decreasing set function that is sub-modular:

\[ \Gamma_M := \left\{ h \in \mathbb{R}_{\geq 0}^{2M-1} \mid h_A + h_B \geq h_{A\cup B} + h_{A \cap B} \quad A \subseteq A', \quad A, B \subseteq \{1, \ldots, M\} \right\} \]  

(38)

Inner bound TBD.
4.3 Achievability Sketch

We begin by proving an alternate form of the capacity region (18). Let \( D(\cdot) \) be a set operator which scales all the points in the set by numbers in between zero and one:

\[
D(A) = \{ \alpha \mathbf{h} | \mathbf{h} \in A, \alpha \in [0, 1] \}
\]

We will prove that the convex hull can be replaced by scalings in the capacity region expression, so that

\[
\Lambda \left( \text{proj}_{Y_s} \left( \text{conv}(\Gamma_M^* \cap L_{123}) \cap L_4 \cap L_5 \right) \right) = \Lambda \left( \text{proj}_{Y_s} \left( D(\Gamma_M^* \cap L_{123}) \cap L_4 \cap L_5 \right) \right). \tag{40}
\]

To do this, we will show that

\[
D(\Gamma_M^* \cap L_{123}) = \text{conv}(\Gamma_M^* \cap L_{123}) \tag{41}
\]

Consider a point \( \mathbf{h} \in D(\Gamma_M^* \cap L_{123}) \). It is the limit of some sequence \( \mathbf{h}_k \in D(\Gamma_M^* \cap L_{123}) \), where \( \mathbf{h}_k = \alpha_k \hat{\mathbf{h}}_k \) for some \( \hat{\mathbf{h}}_k \in \Gamma_M^* \cap L_{123} \) and \( \alpha_k \in [0, 1] \]. Noting that \( \mathbf{0} \in \Gamma_M^* \cap L_{123} \), we can view \( \alpha_k \hat{\mathbf{h}}_k \) as the convex combination \( (\alpha_k \hat{\mathbf{h}}_k + (1 - \alpha_k) \mathbf{0}) \in \text{conv}(\Gamma_M^* \cap L_{123}) \). This shows that \( D(\Gamma_M^* \cap L_{123}) \subseteq \text{conv}(\Gamma_M^* \cap L_{123}) \). To prove the other containment, we show that \( D(\Gamma_M^* \cap L_{123}) \) is a convex set containing \( \Gamma_M^* \cap L_{123} \). (Since the convex hull \( \text{conv}(\Gamma_M^* \cap L_{123}) \) is defined as the smallest convex set containing \( \Gamma_M^* \cap L_{123} \), the convexity of \( D(\Gamma_M^* \cap L_{123}) \) will guarantee that it contains \( \text{conv}(\Gamma_M^* \cap L_{123}) \subseteq D(\Gamma_M^* \cap L_{123}) \).

Consider two points \( \hat{\mathbf{h}}_1, \hat{\mathbf{h}}_2 \in D(\Gamma_M^* \cap L_{123}) \), and select any \( \lambda \in [0, 1] \). These points are limits of the sequences \( \mathbf{h}^{(k)}_1 = \alpha_1^{(k)} \hat{\mathbf{h}}^{(k)}_1 \) and \( \mathbf{h}^{(k)}_2 = \alpha_2^{(k)} \hat{\mathbf{h}}^{(k)}_2 \) with \( \alpha_1^{(k)}, \alpha_2^{(k)} \in (0, 1) \) and \( \hat{\mathbf{h}}^{(k)}_1, \hat{\mathbf{h}}^{(k)}_2 \in \Gamma_M^* \cap L_{123} \). Select a sequence of positive integers \( n_1^k, n_2^k \in \mathbb{N} \) with \( n_1^k, n_2^k \to \infty \) as \( k \to \infty \) and with

\[
\frac{n_1^k \alpha_1^k}{n_2^k \alpha_2^k} = \frac{\lambda}{1 - \lambda} \tag{42}
\]

Letting the collection of random variables \( (Z_1, \ldots, Z_2^2) \) and \( (Z_1^2, \ldots, Z_m^2) \) be random variables obtaining any \( \hat{\mathbf{h}}_1, \hat{\mathbf{h}}_2 \in \Gamma_M^* \cap L_{123} \) respectively, we observe that the collection of random variables \( Z_1, \ldots, Z_m \) defined via

\[
Z_m = (\mathbf{Z}_{m,n_1}, \ldots, \mathbf{Z}_{m,n_2}), \quad \forall m \in \{1, \ldots, M\}
\]

are associated with the entropies \( n_1 \hat{\mathbf{h}}_1 + n_2 \hat{\mathbf{h}}_2 \in \Gamma_M^* \cap L_{123} \), hence \( (n_1^k \hat{\mathbf{h}}^{(k)}_1 + n_2^k \hat{\mathbf{h}}^{(k)}_2) \in \Gamma_M^* \cap L_{123} \). Additionally for \( k \) sufficiently large we have

\[
\frac{\alpha_1^{(k)} \alpha_2^{(k)}}{n_1^k \alpha_2^{(k)} + n_2^k \alpha_1^{(k)}} \leq 1 \tag{44}
\]

This then shows that

\[
\frac{\alpha_1^{(k)} \alpha_2^{(k)}}{n_1^k \alpha_2^{(k)} + n_2^k \alpha_1^{(k)}} (n_1^k \hat{\mathbf{h}}^{(k)}_1 + n_2^k \hat{\mathbf{h}}^{(k)}_2) \in \Gamma_M^* \cap L_{123} \tag{45}
\]

for all \( k \) sufficiently large, which then implies that

\[
\lim_{k \to \infty} \left( \frac{\alpha_1^{(k)} \alpha_2^{(k)}}{n_1^k \alpha_2^{(k)} + n_2^k \alpha_1^{(k)}} (n_1^k \hat{\mathbf{h}}^{(k)}_1 + n_2^k \hat{\mathbf{h}}^{(k)}_2) \right) \in \Gamma_M^* \cap L_{123} \tag{46}
\]

However, rearranging the terms inside the limit we have

\[
\lim_{k \to \infty} \left( \frac{\alpha_1^{(k)} \alpha_2^{(k)}}{n_1^k \alpha_2^{(k)} + n_2^k \alpha_1^{(k)}} (n_1^k \hat{\mathbf{h}}^{(k)}_1 + n_2^k \hat{\mathbf{h}}^{(k)}_2) \right) = \lim_{k \to \infty} \frac{n_1^k \alpha_1^{(k)}}{n_1^k \alpha_2^{(k)} + n_2^k \alpha_1^{(k)}} \alpha_1^{(k)} \hat{\mathbf{h}}^{(k)}_1 + \lim_{k \to \infty} \frac{n_2^k \alpha_2^{(k)}}{n_1^k \alpha_2^{(k)} + n_2^k \alpha_1^{(k)}} \alpha_2^{(k)} \hat{\mathbf{h}}^{(k)}_2 = \lambda \mathbf{h}_1 + (1 - \lambda) \mathbf{h}_2 \tag{47}
\]

is in \( \Gamma_M^* \cap L_{123} \), proving that it is convex. This establishes

We must now prove that any vector \( \mathbf{w} \) in the alternate rate region representation

\[
\mathbf{w} \in \Lambda \left( \text{proj}_{Y_s} \left( D(\Gamma_M^* \cap L_{123}) \cap L_4 \cap L_5 \right) \right) \tag{48}
\]

is achievable. Since we can discard any excess rate we wish not to use, this amounts to showing that any rate vector \( \mathbf{w} \in \text{proj}_{Y_s} \left( D(\Gamma_M^* \cap L_{123}) \cap L_4 \cap L_5 \right) \) is achievable. Let \( \mathbf{h} \in \text{proj}_{Y_s} \left( D(\Gamma_M^* \cap L_{123}) \cap L_4 \cap L_5 \right) \) be a vector such that \( \mathbf{w} = \text{proj}_{Y_s} \mathbf{h} \). Since \( \mathbf{h} \in D(\Gamma_M^* \cap L_{123}) \cap L_4 \cap L_5 \) it is the limit of some sequence \( \mathbf{h}_k \in D(\Gamma_M^* \cap L_{123}) \) which is of the form \( \alpha_k \hat{\mathbf{h}}_k \) with
\( \hat{h}_k \in \Gamma_M \cap \mathcal{L}_{123} \). Let \((Y_s^k : s \in S), (U_e^k : e \in E)\) be the random variables associated with \( \hat{h}_k \), since the entropies for these variables are in \( \mathcal{L}_{123} \) they obey the inequalities

\[
H(Y_s^k) = \sum_{s \in S} H(Y_s^k) \tag{49}
\]

\[
H(U_{Out(s)}^k | Y_s^k) = 0 \quad \forall s \in S \tag{50}
\]

\[
H(U_{Out(i)}^k | U_{In(i)}) = 0 \quad \forall i \in V \setminus (S \cup T) \tag{51}
\]

The equalities (50) and (51) show that we may think of the random variables \( U_e^k = f_{s,e,k}(Y_s^k) \) as a deterministic function \( f_{s,e,k} \) of the source random variable \( Y_s^k \) for each \( e \in Out(s) \) and \( s \in S \) and the random variables \( U_e^k = f_{e,k}(U_{In(i)}^k) \) as a deterministic function of the random variables \( U_{In(i)}^k \) for each \( e \in Out(i) \) for each \( i \in V \setminus (S \cup T) \).

Additionally, since the limit of the scaled entropies \( h_k \) is in \( \mathcal{L}_4 \cap \mathcal{L}_5 \) and has \( \text{proj}_{Y_S} \lim_{k \to \infty} h_k = \omega \), we have

\[
\alpha_k H(Y_s^k) = \omega_s^k \quad \forall s \in S \tag{52}
\]

\[
\alpha_k H(U_e^k) \leq R_e + \mu_k \quad \forall e \in E \tag{53}
\]

\[
\alpha_k H(Y_{\beta(t)}^k | U_{In(i)}) \leq \gamma_k \quad \forall t \in T \tag{54}
\]

where \( \gamma_k \to 0 \) and \( \mu_k \to 0 \) as \( k \to \infty \) while \( \omega_s^k \to \omega_s \).

Let \( \tilde{N}_k = \lfloor \alpha N_k \rfloor \). For each source \( s \), generate a \( 2^{N_k \tau_s^k} \times \tilde{N}_k \) dimensional matrix by sampling its elements i.i.d. according to the distribution \( p_{Y_s^k} \), let the \( j \)th row be denoted by \( Y_{s,e,k}^{\tilde{N}_k}(j), j \in \{1, \ldots, 2^{N_k \tau_s^k}\} \). For each edge, enumerate all of the length \( \tilde{N}_k \) typical sequences \( T_{e,k}^{\tilde{N}_k}(U_e^k) \) as \( U_{s,e,k}^{\tilde{N}_k}(1), \ldots, U_{s,e,k}^{\tilde{N}_k}(\eta_e^k) \). Due to the bound on the cardinality of the typical set, for such an enumeration

\[
\eta_e^k \leq 2^{N_k (H(U_e^k) + \epsilon c_e,k)} \tag{55}
\]

so that

\[
\frac{1}{\tilde{N}_k} \log_2 \eta_e^k \leq \alpha_k H(U_e^k) + \epsilon c_{e,k} \leq R_e + \mu_k + \epsilon c_{e,k} \tag{56}
\]

The encoder at source node \( s \) selects at random one of the \( 2^{N_k \tau_s^k} \) rows in its matrix, then calculates the deterministic function \( f_{s,k}(Y_{s,k}^k(j)) \) operating elementwise on each of the \( \tilde{N}_k \) positions in the vector. Provided we select

\[
\epsilon_{e,k} \geq |Y_{s,k}^k| \quad \forall e \in Out(s) \tag{57}
\]

if \( Y_{s,k}^k(j) \in T_{e,k}^{\tilde{N}_k}(Y_{s,k}^k) \), then the result of these deterministic function will be in \( T_{e,k}^{\tilde{N}_k}(U_e^k) \), and together the outgoing messages will all be jointly typical, i.e. in \( T_{e,k}^{\tilde{N}_k}(U_{Out(s)}^k, Y_{s,k}^k) \). The messages sent are the associated typical sequence index from \( \{1, \ldots, \eta_e^k\} \) from the deterministic function, or 0 if the input was not typical. Via the Markov lemma if we take \( N \) sufficiently large, then we observe that if \( Y_{s,k}^k(j) \in T_{e,k}^{\tilde{N}_k}(Y_{s,k}^k) \) for each \( s \in S \), then all of the messages outgoing from these source are together jointly typical. Proceeding via the order defined by the directed acyclic graph for the operation of the encoders (such that all incoming message are available before an outgoing message is calculated), we observe that provided all incoming messages are jointly typical, the outgoing messages (calculated with the deterministic functions \( f_{e,k} \) operating on the typical sequences associated with the incoming indices) will be jointly typical themselves, and jointly typical with everything computed so far. Thus, provided that each of the selected messages \( Y_{s,k}^k(j) \) are typical, they will be jointly typical with the sequences associated with each of the incoming messages at a sink.

The sink operates by looking in the rows of the codebooks for the sources \( \beta(t) \) for a collection of codewords that are jointly typical together with the incoming messages \( U_{In(t)}^s \). By the logic above, there will be at least one collection (corresponding to the correct decoding) of such sequences provided that \( Y_{s,k}^s(j) \) are typical. If there is more than one such collection, then an error is declared. This error event is then the union, over all subsets \( A \) of \( \beta(t) \), of the events \( \mathcal{E}_A \) for which \( U_{In(t)}^s \), \( Y_{s,k}^s(j_A) \), \( Y_{A'}^s(j_{A'}) \) are jointly typical for some \( j_A \neq j_{A'} \) for each \( s \in A \). These events (associated with the independent codewords \( Y_{s,k}^s(j_A) \), \( s \in A \) winding up in the jointly typical set with \( Y_{s,k}^s(j) \), \( s \in A' \) and \( U_{In(t)}^s \), have probabilities bounded by

\[
P(\mathcal{E}_A) \leq 2^{N} \sum_{s \in A} \tau_s^k 2^{-N \tilde{N}(Y_{A}(Y_{A'},U_{In(t)})) - \epsilon c} \tag{58}
\]

These will go to zero exponentially as \( N_k \to \infty \) provided that we select \( \tau_s^k \) slightly less than \( \alpha_k H(Y_s^k) = \omega_s \). This, together with (49,50,51,56), shows that the \( \omega_s^k \) are achievable (according to the definition (4,5,6) ) for sufficiently large \( k \) and \( N_k \).