CSE 725

Network Coding Theory

Hung Q. Ngo

SUNY at Buffalo

Place: 242 Bell
Time: Saturdays : 2pm–5pm

Motivations

- Efficient networked data transmission is a **central** question in Computer Science
  - Many aspects remain poorly understood
  - Choosing a suitable mathematical model for data transmission is not easy
  - Current (pre-network coding) approaches: network flows, combinatorial packing

Cover and Thomas – *Elements of Information Theory*

The theory of information flow in networks does not have the same simple answers as the theory of flow of water in pipes

- **Key differences**: data can be compressed, combined with signal processing techniques and/or algebraic operations (XOR); cars on highways can't be combined (except during Buffalo's winter)

Network Information Theory

- Part of Information Theory dealing with data transmission aspects: noise, interference, correlation between data sources, etc.
- Many problems are open, too difficult
- (Ahlswede et al., 2000) opened a new door with Network Coding

(Ahlswede et al., 2000) asked

Ignoring noise and interference (which can be dealt with at the physical layer), can coding data streams together provide any benefits?

Network Coding – Canonical Example 1

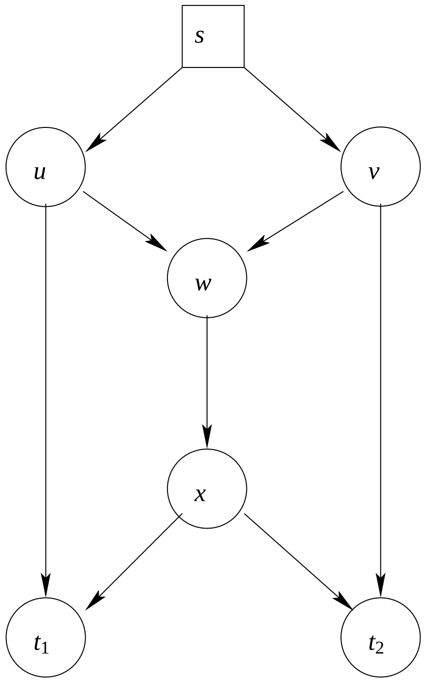

Figure: The Butterfly Network

Network Coding – Canonical Example 1

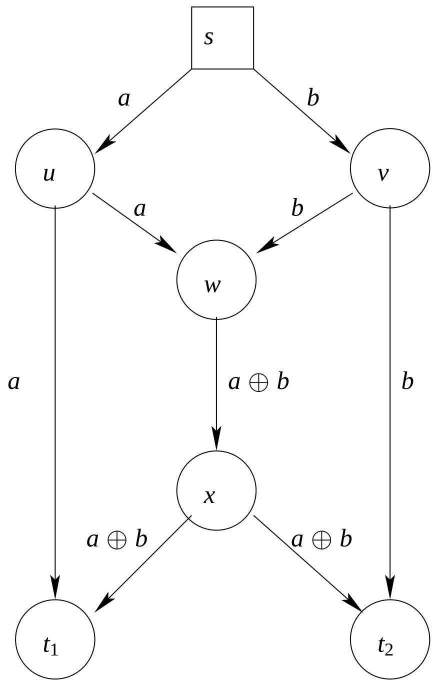

Figure: Network Coding on the Butterfly Network

Network Coding – Canonical Example 2

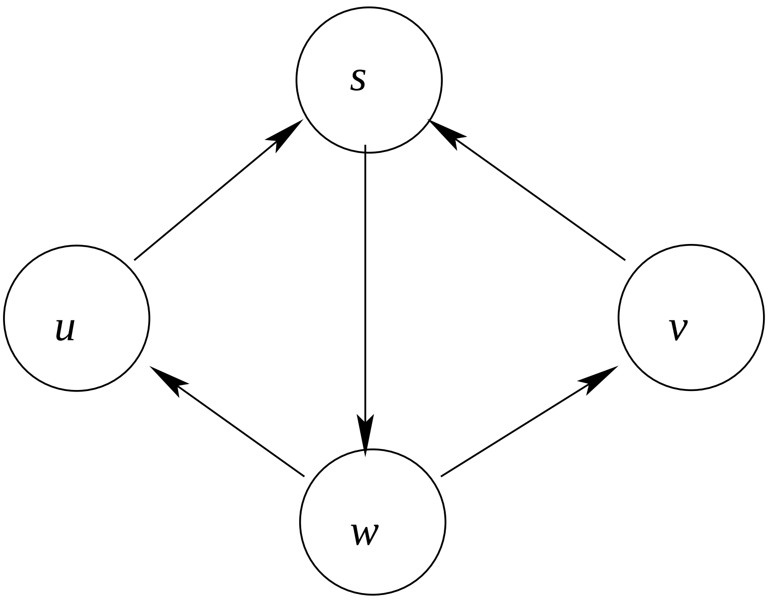

Figure: The Wheatstone Bridge

Network Coding – Canonical Example 2

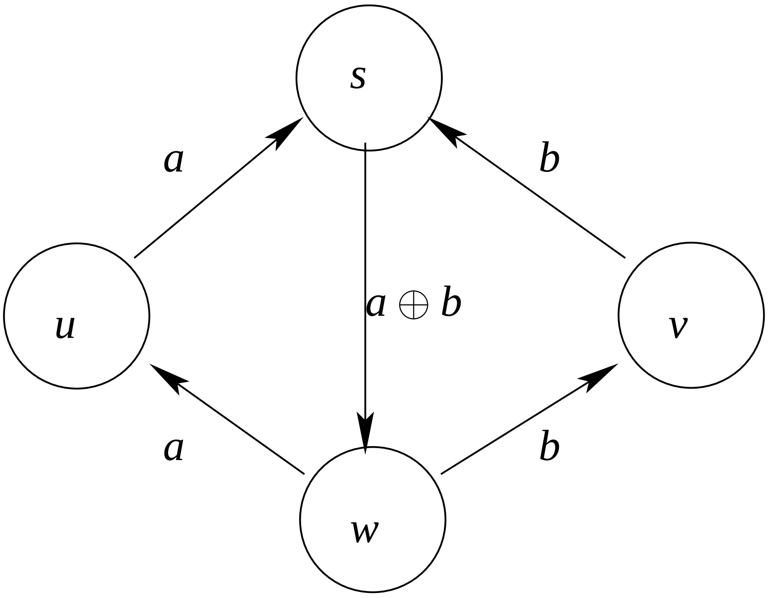

Figure: Network Coding on the Wheatstone Bridge

Network Coding – Canonical Example 3

- Two base stations B_1 and B_2 want to communicate
- They are out of each other's radio range
- There is a relay transceiver T in the middle, within range of both

How efficient can they communicate with network coding?

- B_1 sends b_1 to T
- B_2 sends b_2 to T
- T sends $b_1 \oplus b_2$ to B_1 and B_2 at the same time

Current Research on Network Coding

Network coding has interesting connections in and applications to

- Coding and information theory
- Networking, including multicasting, switching, wireless communications, content distribution
- Complexity theory, cryptography, operations research, matrix theory

This seminar

I will present materials from the perspective of coding and information theory

- Single-source, multiple-source network coding
- Cyclic, acyclic networks
- Directed, undirected networks
- Randomized, deterministic algorithms
- Centralized, distributed algorithms
- Some aspects of network coding in practice

You will present mostly applications of network coding to other areas, and/or some theoretical issues that's not address in class.

Communications Network

- A finite directed graph $G = (V, E)$ (possibly with multiple edges between any pair of vertices)
  - Undirected graphs can also be used (later)
- A node with in-degree zero is a *source node*
- Each edge is a noiseless *channel*
- Capacity of each edge is 1 (1 “packet” per time unit)
- Assume no processing delay and no propagation delay

In the first few weeks, we will assume

- G is acyclic
- One single source s

More notations

- For each $v \in V$,
  - $\partial^-(v) = \{e \mid e = (u, v) \in E\}$
  - $\partial^+(v) = \{e \mid e = (v, w) \in E\}$
- $\partial^-(s)$ contains ω imaginary channels without originating nodes
- We want to send one single message $\mathbf{x}$ of size ω data units
  - Each data unit is an element of a *base field* $F = \mathbb{F}_q$ (In practice, $q = 2^m$ for m-bit data units.)
  - $\mathbf{x}$ is thus a vector $\mathbf{x} = (x_1, \dots, x_\omega) \in \mathbb{F}_q^\omega$
- To send $\mathbf{x}$ over the network, every channel e participates by carrying a “symbol” (i.e. data unit) $\tilde{f}_e(\mathbf{x}) \in F$

Network Code

In the single-source, acyclic case, there are two equivalent ways to define the network code

- Local encoding mapping
- Global encoding mapping

Local Encoding Mapping

Definition

Let F be a finite field, ω a positive integer. An ω-dimensional F-valued *network code* on an acyclic communications network consists of a local encoding mapping

$$\tilde{k}_e : F^{|\partial^-(v)|} \longrightarrow F$$

for each node v and each channel $e \in \partial^+(v)$.

Note:

- this definition does not explicitly give $\tilde{f}_e(\mathbf{x})$
- since the graph is acyclic, these values $\tilde{f}_e(\mathbf{x})$ can be computed recursively, however

Global Encoding Mapping

Definition

Let F be a finite field, ω a positive integer. An *ω-dimensional F-valued network code* on an acyclic communications network consists of a local encoding mapping $\tilde{k}_e : F^{|\partial^-(v)|} \to F$ and a global encoding mapping $\tilde{f}_e : F^\omega \to F$ for each channel e in the network such that:

(i) $\forall v \in V, e \in \partial^+(v)$, $\tilde{f}_e(\mathbf{x})$ is uniquely determined by $(\tilde{f}_{e'}(\mathbf{x}), e' \in \partial^-(v))$, and $\tilde{k}_e$ is the mapping defined via

$$(\tilde{f}_{e'}(\mathbf{x}), e' \in \partial^-(v)) \to \tilde{f}_e(\mathbf{x}).$$

(ii) For the ω imaginary channels e, the mappings $\tilde{f}_e$ are the projections from F^ω to the ω different coordinates, respectively

Road-map for the first few weeks

Loosely, a network coding solution is a network code allowing receivers to decode the message.

Theorem (Ahlswede et al., 2000)

For acyclic networks, there always exists a network coding solution such that the maximum throughput of single-source multicast is equal to the capacity of a minimum cut separating the source and some sink

Practical concerns: solution description may be very large

- (Lehman and Lehman, SODA 2005) showed that a doubly-exponential q (alphabet size) might be necessary for some (non-multicast) problem
- Consider $n = |V|$, $q = 2$, and some vertex v with in-degree $m = \Theta(n)$. The number of functions from $\mathbb{F}_2^m \to \mathbb{F}_2$ is doubly-exponential in m, thus there are functions requiring $2^m = 2^{\Theta(n)}$ bits to encode.

Road-map for the first few weeks

It would be nice to have linear network coding solution

Theorem (Li et al., 2003)

For acyclic network, multicast problem, there exists a linear network coding solution over some alphabet.

Theorem (Koetter and Médard, INFOCOM 2003)

- *Alphabet size only needs to be polynomial*
- *It takes polynomial-space (in instance size and $\log$ of alphabet size) to write down a linear solution*

Road-map for the first few weeks

More practical concerns: OK, a good solution exists, but can we find one efficiently?
YES

- (Ho et al., 2003) gave randomized algorithms
- (Sanders et al., SPAA'03) and (Jaggi et al., ISIT'03) gave deterministic algorithms

Linear encoding mappings

- When $\tilde{f}_e$ is linear,
$$\tilde{f}_e(\mathbf{x}) = \langle \mathbf{f}_e, \mathbf{x} \rangle, \quad \mathbf{f}_e \in F^{\omega}$$
- When $\tilde{k}_e$ is linear ($e = (u, v)$),
$$\tilde{k}_e(\mathbf{y}) = \langle \mathbf{k}_e, \mathbf{y} \rangle, \quad \mathbf{k}_e, \mathbf{y} \in F^{\partial^-(u)}.$$

Theorem

The local encoding mappings are linear if and only if the global encoding mappings are linear.

($\Rightarrow$) obvious by induction.
($\Leftarrow$) needs a little more setup

Global Linearity Implies Local Linearity

Definition

Strictly speaking, a map f is linear iff
$$f(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha f(\mathbf{x}) + \beta f(\mathbf{y}), \quad \forall \alpha, \beta, \mathbf{x}, \mathbf{y}.$$

Lemma (1)

$f : F^m \to F$ *is linear iff* $\exists \mathbf{f} \in F^m$ *s.t.* $f(\mathbf{x}) = \langle \mathbf{f}, \mathbf{x} \rangle$

Lemma (2)

Let S *by a subspace of* F^m*, then* $f : S \to F$ *is linear iff* $\exists \mathbf{f} \in F^m$ *s.t.* $f(\mathbf{x}) = \langle \mathbf{f}, \mathbf{x} \rangle$

Proof of Lemma (1)

- Sufficiency is immediate.
- For necessity, let $\mathbf{u}_i$ be the ith unit vector in the natural basis for F^m.
- It follows that

$$f(\mathbf{x}) = f(\sum_i x_i \mathbf{u}_i) = \sum_i f(\mathbf{u}_i) x_i.$$

Proof of Lemma (3)

Sufficiency is obvious. For necessity:

- Suppose $\dim(S) = k \leq m$. Let $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ be a basis for S. Let $\mathbf{A}$ be a $m \times k$ matrix with columns $\mathbf{u}_i$
- $\mathbf{x} \in S$ iff $\mathbf{x} = \mathbf{A}\mathbf{y}$ for some $\mathbf{y} \in F^k$
- Let $\mathbf{B}$ be $\mathbf{A}$'s left inverse, i.e. $\mathbf{B}\mathbf{A} = \mathbf{I}$
- Define $g : F^k \to F$ by $g(\mathbf{y}) = f(\mathbf{A}\mathbf{y})$, then

$$g(\alpha\mathbf{y} + \beta\mathbf{y}') = \cdots = \alpha g(\mathbf{y}) + \beta g(\mathbf{y}')$$

- Thus, $\exists \mathbf{g} \in F^k$ such that $g(\mathbf{y}) = \langle \mathbf{g}, \mathbf{y} \rangle$
- Hence,

$$f(\mathbf{x}) = f(\mathbf{A}\mathbf{y}) = g(\mathbf{y}) = \langle \mathbf{g}, \mathbf{y} \rangle = \langle \mathbf{g}, \mathbf{B}\mathbf{x} \rangle = \mathbf{g}^T\mathbf{B}\mathbf{x} = (\mathbf{B}^T\mathbf{g})^T\mathbf{x}$$
$$= \langle \mathbf{B}^T\mathbf{g}, x \rangle = \langle \mathbf{f}, \mathbf{x} \rangle$$

Global Linearity Implies Local Linearity

Consider $e = (u, v)$, $\partial^-(u) = \{e_1, \dots, e_m\}$, and the subspace

$$S = \{(\tilde{f}_{e_1}(\mathbf{x}), \dots, \tilde{f}_{e_m}(\mathbf{x})) \mid \mathbf{x} \in F^\omega\} \subseteq F^m$$

$\tilde{k}_e : S \to F$, defined by

$$\tilde{k}_e(\mathbf{y}) = \tilde{k}_e(\tilde{f}_{e_1}(\mathbf{x}), \dots, \tilde{f}_{e_m}(\mathbf{x})) = \tilde{f}_e(\mathbf{x}).$$

Thus,

$$\begin{aligned}
\tilde{k}_e(\alpha\mathbf{y}_1 + \beta\mathbf{y}_2) &= \tilde{k}_e(\alpha\tilde{f}_{e_1}(\mathbf{x}_1) + \beta\tilde{f}_{e_1}(\mathbf{x}_2), \dots, \alpha\tilde{f}_{e_m}(\mathbf{x}_1) + \beta\tilde{f}_{e_m}(\mathbf{x}_2)) \\
&= \tilde{k}_e(\tilde{f}_{e_1}(\alpha\mathbf{x}_1 + \beta\mathbf{x}_2), \dots, \tilde{f}_{e_m}(\alpha\mathbf{x}_1 + \beta\mathbf{x}_2)) \\
&= \tilde{f}_e(\alpha\mathbf{x}_1 + \beta\mathbf{x}_2) \\
&= \alpha\tilde{f}_e(\mathbf{x}_1) + \beta\tilde{f}_e(\mathbf{x}_2) \\
&= \alpha\tilde{k}_e(\mathbf{y}_1) + \beta\tilde{k}_e(\mathbf{y}_2)
\end{aligned}$$

Local description of linear network codes

Definition

An ω-dimensional F-valued *linear network code* on an acyclic communications network consists of a local encoding kernel $\mathbf{K}_v$ for every node v, where

$$\mathbf{K}_v = (k_{e,e'})_{e \in \partial^-(v), e' \in \partial^+(v)}$$

is a matrix with entries in F.

Global description of linear network codes

Definition

An ω-dimensional F-valued *linear network code* on an acyclic communications network consists of a local encoding kernel $\mathbf{K}_v$ for every node v, where

$$\mathbf{K}_v = (k_{e,e'})_{e \in \partial^-(v), e' \in \partial^+(v)}$$

is a matrix with entries in F, and a global encoding kernel $\mathbf{f}_e$ for every edge e, where $\mathbf{f}_e \in F^\omega$, such that

(i) for each $e \in \partial^+(v)$,

$$\mathbf{f}_e = \sum_{e' \in \partial^-(v)} k_{e',e} \mathbf{f}_{e'}$$

(ii) the vectors $\mathbf{f}_e$ for the ω imaginary channels form the natural basis of the vector space F^ω

Back to the Butterfly Network

Example

The network code for the butterfly network we saw is linear.

Exercise

Determine all the $\mathbf{K}_v$ and $\mathbf{f}_e$, given the solution we discussed on the butterfly network.

Exercise

Let all $k_{e,e'}$ be variables, determine the global encoding kernels $\mathbf{f}_e$ for the butterfly network.

Desirable properties of linear network codes

- Let T be the set of sinks, maxflow(s, T) is an upper bound on the information rate from s to T
- For each vertex v, let

$$\mathcal{S}_v = \text{span}\{\mathbf{f}_e \mid e \in \partial^-(v)\}.$$

v can decode iff $\dim(\mathcal{S}_v) = \omega$ ($\Rightarrow$ maxflow$(s, v) \geq \omega$)

- Whether this bound is achievable depends on the topology, ω, F, and the coding scheme
- We will define three classes of linear network codes which achieve the bound in 3 different extents: *linear multicast*, *linear broadcast*, *linear dispersion*

Linear Multicast/Broadcast/Dispersion

Definition

A linear network code is a linear multicast, linear broadcast, linear dispersion, respectively, if the following conditions hold:

(i) $\dim(\mathcal{S}_v) = \omega$ for every non-source node v with maxflow$(s, v) \geq \omega$

(ii) $\dim(\mathcal{S}_v) = \min\{\omega, \text{maxflow}(s, v)\}$ for every non-source node v

(iii) $\dim(\text{span}\{\cup_{v \in T}\mathcal{S}_v\}) = \min\{\omega, \text{maxflow}(s, T)\}$ for every collection T of non-source nodes.

- linear dispersion $\Rightarrow$ linear broadcast $\Rightarrow$ linear multicast
- linear multicast $\not\Rightarrow$ linear broadcast $\not\Rightarrow$ linear dispersion

Generic linear network codes

Definition

An ω-dimensional F-valued linear network code on an acyclic communications network is said to be generic if:

- Let $\{e_1, \dots, e_m\}$ be any set of channels, where $e_i \in \partial^+(v_i)$. (The v_i are not necessarily distinct.) Then, the vectors $\mathbf{f}_{e_i}$ are linearly independent (thus $m \le \omega$) *provided that*

$$\text{for any } i \in [m], \quad \mathcal{S}_{v_i} \not\subseteq \text{span}\{\mathbf{f}_{e_j} \mid j \neq i\}.$$

- In a sense, this is saying that every collection of global encoding kernels that can possibly be independent must be independent
- generic linear network code $\Rightarrow$ linear dispersion (will prove later)
- linear dispersion $\not\Rightarrow$ generic linear network code

Existence

Theorem (Existence of generic linear network code)

Given a positive integer ω and an acyclic network, there exists an ω-dimensional F-valued generic linear network code if $|F|$ is sufficiently large.

Theorem

Every generic linear network code is a linear dispersion

Existence

Corollary (Existence of linear dispersion)

Given a positive integer ω and an acyclic network, there exists an ω-dimensional F-valued linear dispersion if $|F|$ is sufficiently large.

Corollary (Existence of linear broadcast)

Given a positive integer ω and an acyclic network, there exists an ω-dimensional F-valued linear broadcast if $|F|$ is sufficiently large.

Corollary (Existence of linear multicast)

Given a positive integer ω and an acyclic network, there exists an ω-dimensional F-valued linear multicast if $|F|$ is sufficiently large.

Existence of a linear multicast – Koetter-Médard's Method

- The local encoding kernels $k_{e,e'}$ are variables whose values are to be determined so that the code is a linear multicast
- For every path $P = (e_1, \dots, e_m)$, define its "weight" to be

$$w_P = k_{e_1,e_2} \cdots k_{e_{m-1},e_m}.$$

- For every channel e,

$$\mathbf{f}_e = \sum_{e' \in \partial^-(s)} \left(\sum_{P:\ \text{a path from } e' \text{ to } e} w_P \right) \mathbf{f}_{e'}$$

- Thus, every component of every vector $\mathbf{f}_e$ is a polynomial in the ring $F[\{k_{e,e'} \mid (e,e') \text{ are adjacent }\}]$.

Existence of a linear multicast – Koetter-Médard's Method

- Let t be a node with maxflow$(s,t) \geq \omega$. Then, there exists ω edge-disjoint paths from $\partial^-(s)$ to $\partial^-(t)$. (Menger theorem)
- Let L_t be the $\omega \times \omega$ matrix formed by putting together the vectors $\mathbf{f}_e$, $e \in \partial^-(t)$ and e belongs to one of these paths. (This is a matrix of variables.)
- **Want:** find local encoding kernels such that all L_t have full rank.

Theorem

If F is sufficiently large, then there are local encoding kernels such that L_t have full rank for all t with maxflow$(s,t) \geq \omega$.

Proof of Koetter-Médard Theorem

- L_t has full rank iff $\det(L_t) \neq 0$
- $\det(L_t)$ is a polynomial in $F[k_{e,e'}]$
- Just need to find local encoding kernels such that the polynomial $p(\cdot) = \prod_t \det(L_t)$ is not zero.

Lemma (1)

The polynomial $\det(L_t)$ is not a zero polynomial. Thus, $p(\cdot)$ is not identically zero.

Lemma (2)

Let $p(z_1, \ldots, z_n)$ be a non-zero polynomial over F. If $|F|$ is greater than the maximum degree of any variable z_i, then there are values $a_1, \ldots, a_n \in F$ such that $p(a_1, \ldots, a_n) \neq 0$.

Proof of Lemma 1

- $\det(L_t)$ is a polynomial in $k_{e,e'}$
- assign $k_{e,e'} = 1$ if (e, e') are consecutive on a path of the ω edge-disjoint paths from s to t
- assign $k_{e,e'} = 0$ for all other pairs (e, e')
- then, $\det(L_t) = 1$ for this set of values of $k_{e,e'}$
- thus, as a polynomial, $\det(L_t)$ is not identically zero

Proof of Lemma 2

Lemma (2)

Let $p(z_1, \dots, z_n)$ be a non-zero polynomial over F. If $|F|$ is greater than the maximum degree of any variable z_i, then there are values $a_1, \dots, a_n \in F$ such that $p(a_1, \dots, a_n) \neq 0$.

Proof.

Induction on n. □

Further Notes on Koetter-Médard's Method

- Works for linear broadcast also
- Implicitly gives a (recursive) algorithm to construct linear multicasts
- However, it is not clear if there exists efficient algorithms to construct linear multicasts using this method
- Jaggi, Sanders, et al. (IEEE Trans. Info. Theory, 2005) gave a deterministic polynomial time algorithm to construct linear multicast (we will discuss later)
- Tracy Ho et al. (IEEE Trans. Info. Theory, 2006) gave randomized and distributed algorithm

Construction of Generic Linear Network Code

Assume: $|F| \geq \binom{M+\omega-1}{\omega-1}$, M is the number of channels

1: $\{\mathbf{f}_e \mid e \in \partial^-(s)\}$ is the natural basis for F^ω
2: $\mathbf{f}_e \leftarrow \mathbf{0}$, for each $e \notin \partial^-(s)$
3: **for** each node u in breath-first order **do**
4:     **for** each $e \in \partial^+(u)$ **do**
5:         Choose $\mathbf{w} \in \mathcal{S}_u$ such that $w \notin \text{span}\{\mathbf{f}_{e'} : e' \in S'\}$,
        for any set S' of $\omega - 1$ channels (not including e) for which
        $\mathcal{S}_u \not\subseteq \text{span}\{\mathbf{f}_{e'} : e' \in S'\}$
6:         $\mathbf{f}_e \leftarrow \mathbf{w}$
7:     **end for**
8: **end for**

Correctness of the Construction

- $\exists S'$ as in line 5, and $\exists w$ as in line 5
- Consider $\{e_1, \dots, e_m\}$, $e_i \in \partial^+(v_i)$, $\mathcal{S}_{v_j} \not\subseteq \text{span}\{\mathbf{f}_{e_i} \mid j \neq i\}$, $\forall j$
- WLOG, assume e_m is considered after other e_i
- We will induct that $\{\mathbf{f}_{e_1}, \dots, \mathbf{f}_{e_m}\}$ are independent
- We know $\{\mathbf{f}_{e_1}, \dots, \mathbf{f}_{e_{m-1}}\}$ are independent by induction hypothesis
- $\mathcal{S}_{v_m} \not\subseteq \text{span}\{\mathbf{f}_{e_1}, \dots, \mathbf{f}_{e_{m-1}}\} \Rightarrow m - 1 < \omega \Rightarrow m \leq \omega$
- If $m = \omega$, then $\mathbf{f}_{e_m} = w \notin \text{span}\{\mathbf{f}_{e_1}, \dots, \mathbf{f}_{e_{m-1}}\}$
- If $m < \omega$, let $R = \{\mathbf{f}_{e_1}, \dots, \mathbf{f}_{e_{m-1}}\}$, $|R| \leq \omega - 2$.
- $\exists$ imaginary channels e', e'' such that $R \cup \{\mathbf{f}_{e'}, \mathbf{f}_{e''}\}$ are independent
- Either $\mathcal{S}_{v_m} \not\subseteq R' = R \cup \{\mathbf{f}_{e'}\}$ or $\mathcal{S}_{v_m} \not\subseteq R'' = R \cup \{\mathbf{f}_{e''}\}$
- Replace R by R' or R'', until $|R| = \omega - 1$

Generic LNC $\Rightarrow$ Linear Dispersion

For any $T \not\ni s$, let

$$\mathcal{S}_T := \text{span}\{\cup_{u \in T} \mathcal{S}_u\} = \text{span}[\bar{T}, T]$$

Here, $[\bar{T}, T] = \{(u, v) \in E \mid u \in \bar{T}, v \in T\}$ Want: for any such T, $\dim(\mathcal{S}_T) = \min\{\omega, \text{maxflow}(s, T)\}$. We have

- $\dim(\mathcal{S}_T) \leq \min\{\omega, \text{maxflow}(s, T)\}$, $\forall T$
- If $\dim(\mathcal{S}_T) = \omega$, we are done.
- Suppose $\dim(\mathcal{S}_T) < \omega$, will show $\exists W \supseteq T$ s.t.

$$s \in \bar{W}, \dim(\mathcal{S}_T) = |[\bar{W}, W]|$$

Generic LNC $\Rightarrow$ Linear Dispersion

We will show by induction that

$$\forall T \not\ni s, \exists W \supseteq T, s \notin W, \text{ s.t. } \dim(\mathcal{S}_T) = |[\bar{W}, W]|.$$

Induct on the number of non-source vertices not in T.

- Base case. Suppose $T = V(G) - \{s\}$. Let $\{e_1, \dots, e_m\} = \partial^+(s)$. Apply definition of Generic LNC to $\{e_1, \dots, e_m\}$, then,
  - either $m < \omega$ and $\mathbf{f}_{e_1}, \dots, \mathbf{f}_{e_m}$ are independent, $\Rightarrow W = T$ works.
  - or $m \geq \omega$ and every ω subset of $\mathbf{f}_{e_1}, \dots, \mathbf{f}_{e_m}$ are independent, $\Rightarrow$ $W = T \cup \{s\}$ works.

Generic LNC $\Rightarrow$ Linear Dispersion

- Suppose $T \subset V(G) - \{s\}$.
- If $\exists u \in U = V(G) - \{s\} \cup T$ s.t. $\mathcal{S}_u \subset \mathcal{S}_T$, then $\omega > \dim(\mathcal{S}_T) = \dim(\mathcal{S}_{T\cup\{u\}})$. By induction hypothesis, $\exists W \supseteq T \cup \{u\}$ such that

$$|[\bar{W}, W]| = \dim(\mathcal{S}_{T\cup\{u\}}) = \dim(\mathcal{S}_T).$$

- Now, assume $\forall u \in U$, there is $e \in \partial^-(u)$ such that $\mathbf{f}_e \notin \mathcal{S}_T$.
  - Let $\{e_1, \dots, e_m\} = [\bar{T}, T]$, and $e_i \in \partial^+(u_i)$.
  - Since $u_i \in U$, $\forall i$,

$$\mathcal{S}_{u_i} \not\subseteq \mathcal{S}_T = \text{span}\{\mathbf{f}_{e_1}, \dots, \mathbf{f}_{e_m}\}$$

  - Thus, for all i

$$\mathcal{S}_{u_i} \not\subseteq \text{span}\{\mathbf{f}_{e_j} \mid j \neq i\}$$

  - By definition of generic LNC, $\mathbf{f}_{e_1}, \dots, \mathbf{f}_{e_m}$ are independent, $\Rightarrow W = T$ works!