Network Coding
Lecture 2

Felice Manganiello
Clemson University

ICERM
Graduate Workshop on Linear Algebra over Finite Fields & Applications

Network representation

Disclaimer: We will consider the alphabet to be $\mathbb{F}_q$.

A **network** is a 4-tuple $\mathcal{N} = (\mathcal{V}, \mathcal{E}, \mathcal{S}, \mathcal{T})$ where:

- ✓ $G = (\mathcal{V}, \mathcal{E})$ is a a finite directed acyclic multigraph with $\mathcal{V}$ is the set of vertices and $\mathcal{E}$ is the multiset of directed edges;
- ✓ $\mathcal{S} \subset \mathcal{V}$ is the set of sources;
- ✓ $\mathcal{T} \subset \mathcal{V}$ is the set of sinks (receivers, terminals);

Network representation

Disclaimer: We will consider the alphabet to be $\mathbb{F}_q$.

A **network** is a 4-tuple $\mathcal{N} = (\mathcal{V}, \mathcal{E}, \mathcal{S}, \mathcal{T})$ where:

- ✓ $G = (\mathcal{V}, \mathcal{E})$ is a a finite directed acyclic multigraph with $\mathcal{V}$ is the set of vertices and $\mathcal{E}$ is the multiset of directed edges;
- ✓ $\mathcal{S} \subset \mathcal{V}$ is the set of sources;
- ✓ $\mathcal{T} \subset \mathcal{V}$ is the set of sinks (receivers, terminals);
- ✓ $\mathcal{S} \cap \mathcal{T} = \emptyset$;
- ✓ $I(s) = \emptyset$ for $s \in \mathcal{S}$;
- ✓ $O(t) = \emptyset$ for $t \in \mathcal{T}$;

Network representation

Disclaimer: We will consider the alphabet to be $\mathbb{F}_q$.

A **network** is a 4-tuple $\mathcal{N} = (\mathcal{V}, \mathcal{E}, \mathcal{S}, \mathcal{T})$ where:

- ✓ $G = (\mathcal{V}, \mathcal{E})$ is a a finite directed acyclic multigraph with $\mathcal{V}$ is the set of vertices and $\mathcal{E}$ is the multiset of directed edges;
- ✓ $\mathcal{S} \subset \mathcal{V}$ is the set of sources;
- ✓ $\mathcal{T} \subset \mathcal{V}$ is the set of sinks (receivers, terminals);
- ✓ $\mathcal{S} \cap \mathcal{T} = \emptyset$;
- ✓ $I(s) = \emptyset$ for $s \in \mathcal{S}$;
- ✓ $O(t) = \emptyset$ for $t \in \mathcal{T}$;
- ✓ For every $v \in \mathcal{V} \setminus (\mathcal{S} \cup \mathcal{T})$ there esist a path from a source $s \in \mathcal{S}$ and a sink $t \in \mathcal{T}$ going through v.

Network representation

Disclaimer: We will consider the alphabet to be $\mathbb{F}_q$.

A **network** is a 4-tuple $\mathcal{N} = (\mathcal{V}, \mathcal{E}, \mathcal{S}, \mathcal{T})$ where:

- ✓ $G = (\mathcal{V}, \mathcal{E})$ is a a finite directed acyclic multigraph with $\mathcal{V}$ is the set of vertices and $\mathcal{E}$ is the multiset of directed edges;
- ✓ $\mathcal{S} \subset \mathcal{V}$ is the set of sources;
- ✓ $\mathcal{T} \subset \mathcal{V}$ is the set of sinks (receivers, terminals);
- ✓ $\mathcal{S} \cap \mathcal{T} = \emptyset$;
- ✓ $I(s) = \emptyset$ for $s \in \mathcal{S}$;
- ✓ $O(t) = \emptyset$ for $t \in \mathcal{T}$;
- ✓ For every $v \in \mathcal{V} \setminus (\mathcal{S} \cup \mathcal{T})$ there esist a path from a source $s \in \mathcal{S}$ and a sink $t \in \mathcal{T}$ going through v.
- ✓ $(u, v) \in \mathcal{E}$ is a perfect unit capacity channel from u to v.

Reachability as Partial Order

Definition

A vertex $v \in \mathcal{V}$ is said to be *reachable* from $u \in V$ if $v = u$ or if there is a directed path in G from u to v.

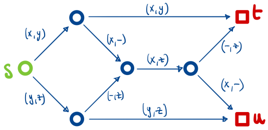
Lemma

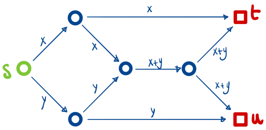
Reachability in G induces a partial order on the vertices of $\mathcal{V}$ where for $v, u \in \mathcal{V}$, $v \preceq u$ if v reachable from u.

Communication Flow

Definition

The function applied by a particular node $v \in \mathcal{V}$ as the local encoding function at v.





Definition

We refer with **single channel use** of a network $\mathcal{N}$ we mean a particular assignment of elements of $\mathbb{F}_q$ to each of the edges of $\mathcal{E}$, consistent with the particular local encoding functions implemented by the network nodes.

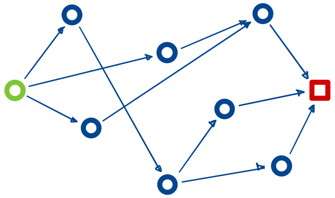
- One Source Networks
  - The Unicast Network
  - The Multicast Network
- Optimality of Linear Network Coding.
- Insufficiency of Linear Network Coding.

Unicast Network

We start by studying the networks with only one source, meaning $\mathcal{S} = \{s\}$.

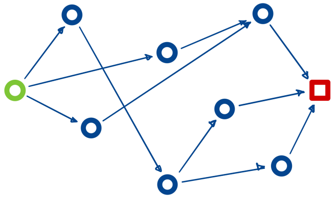
Definition

$\mathcal{N}(\mathcal{V}, \mathcal{E}, \{s\}, \{t\})$ is a **unicast network**.



Unicast Network

We start by studying the networks with only one source, meaning $\mathcal{S} = \{s\}$.



Definition

$\mathcal{N}(\mathcal{V}, \mathcal{E}, \{s\}, \{t\})$ is a **unicast network**.

Definition

Let $N = (\mathcal{V}, \mathcal{E}, \mathcal{S}, \mathcal{T})$ be a network. An *(edge) cut* between s and a terminal $t \in \mathcal{T}$ is a subset $\mathcal{E}' \subseteq \mathcal{E}$ with the property that every directed path from s to t has an edge from $\mathcal{E}'$. We denote by $\mathrm{mincut}(s, t)$ the minimum cardinality of an edge-cut between s and t.

Unicast Network Capacity

Let $N = (\mathcal{V}, \mathcal{E}, \mathcal{S}, \mathcal{T})$ be a unicast network.

Lemma (Bottleneck)

$$\rho(\mathcal{N}) \leq \mathrm{mincut}(s, t)$$

Unicast Network Capacity

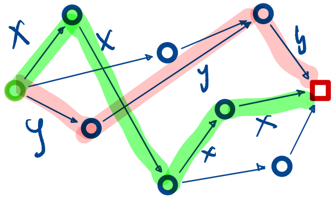
Let $N = (\mathcal{V}, \mathcal{E}, \mathcal{S}, \mathcal{T})$ be a unicast network.

Lemma (Bottleneck)

$$\rho(\mathcal{N}) \leq \text{mincut}(s, t)$$

Theorem (Menger)

For all $t \in \mathcal{T}$, the integer $\text{mincut}(s, t)$ is the maximum number of edge-disjoint directed paths in $G = (\mathcal{V}, \mathcal{E})$ connecting s and t.



Unicast Network Capacity

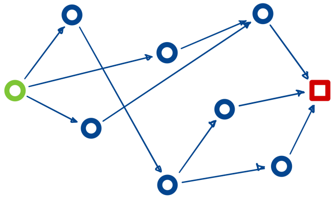
Let $N = (\mathcal{V}, \mathcal{E}, \mathcal{S}, \mathcal{T})$ be a unicast network.

Lemma (Bottleneck)

$$\rho(\mathcal{N}) \leq \mathrm{mincut}(s, t)$$

Theorem (Menger)

For all $t \in \mathcal{T}$, the integer $\mathrm{mincut}(s, t)$ is the maximum number of edge-disjoint directed paths in $G = (\mathcal{V}, \mathcal{E})$ connecting s and t.



Lemma

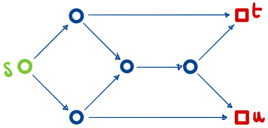
Routing maximizes $\rho(\mathcal{N})$, meaning $\rho(\mathcal{N}) = \mathrm{mincut}(s, t)$.

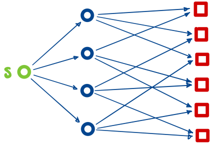
- One Source Networks
  - The Unicast Network
  - The Multicast Network
- Optimality of Linear Network Coding.
- Insufficiency of Linear Network Coding.

Multicast network [1, 2]

Definition

$\mathcal{N}(\mathcal{V}, \mathcal{E}, \{s\}, \mathcal{T})$ where $|\mathcal{T}| \geq 2$ and every $t \in \mathcal{T}$ demands s is a **multicast network**.

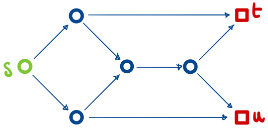


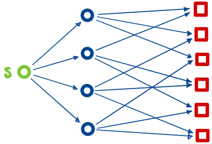


Multicast network [1, 2]

Definition

$\mathcal{N}(\mathcal{V}, \mathcal{E}, \{s\}, \mathcal{T})$ where $|\mathcal{T}| \geq 2$ and every $t \in \mathcal{T}$ demands s is a **multicast network**.





Theorem (Linear Network Multicasting Theorem)

Let $\mathcal{N} = (\mathcal{V}, \mathcal{E}, \{s\}, \mathcal{T})$, then

$$\rho(\mathcal{N}) \leq \min\{\text{mincut}(s, t) \mid t \in \mathcal{T}\}.$$

Furthermore, the bound is achievable, for sufficiently large q, with **linear network coding**.

Theorem "Ingredients"

What is Linear Network Coding?

"Network coding where nodes linearly combine their incoming messages"

What is Linear Network Coding?

"Network coding where nodes linearly combine their incoming messages"

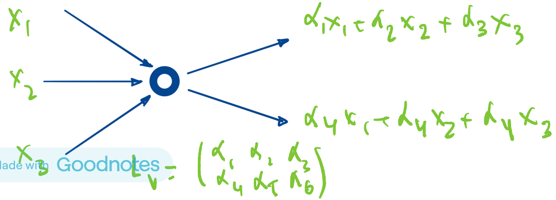
Denote by $x \in \mathbb{F}_q^{|O(s)|}$ the original message sent by the source. For $v \in \mathcal{V} \setminus (\mathcal{S} \cup \mathcal{T})$, let

$$x_{I(v)} \in \mathbb{F}_q^{|I(v)|} \text{ and } x_{O(v)} \in \mathbb{F}_q^{|O(v)|}$$

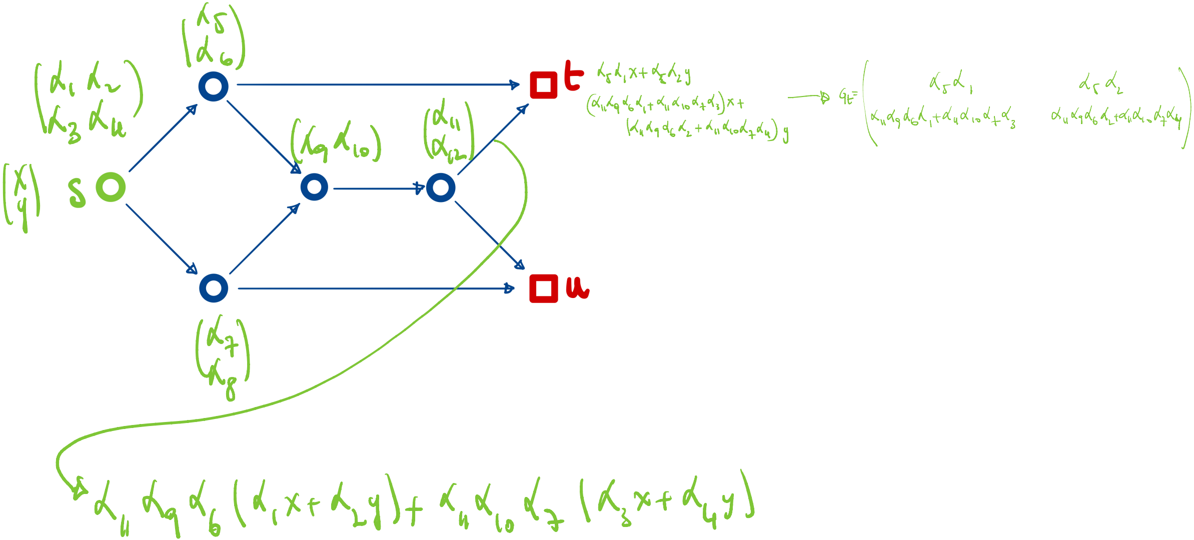
denote the vectors of messages incoming to and outgoing from the node v, respectively. Then, there exists a matrix $L_v \in \mathbb{F}_q^{|O(v)| \times |I(v)|}$ and a matrix $G_v \in \mathbb{F}_q^{|I(v)| \times |O(s)|}$ such that

$$x_{O(v)} = L_v x_{I(v)} = L_v(G_v x)$$

L_v is called the *local transfer matrix* and G_v the *global transfer matrix*.



The Butterfly Network Case



A Powerful Lemma

Lemma **(Sparse Zeros Lemma)**

Let $f \in \mathbb{F}_q[\alpha_1, \ldots, \alpha_n]$ be a polynomial, not identically zero, whose α_i-degree is at most d for all i. If $q > d$, then $f(a_1, \ldots, a_n) \neq 0$ for some $(a_1, \ldots, a_n) \in \mathbb{F}_q^n$.

Proof of the Linear Network Multicasting Theorem

Proof of the Linear Network Multicasting Theorem

Let $r = \min\{\text{mincut}(s,t) \mid t \in \mathcal{T}$, and let $x_1, x_2, \ldots, x_r \in \mathbb{F}_q$ the source messages that we wish to transmit.

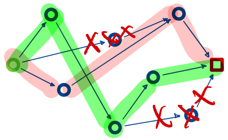
Proof of the Linear Network Multicasting Theorem

Let $r = \min\{\text{mincut}(s, t) \mid t \in \mathcal{T}$, and let $x_1, x_2, \ldots, x_r \in \mathbb{F}_q$ the source messages that we wish to transmit.

For each sink node t it is possible to find r edge-disjoint paths from s to t.

Create a **reduced network**.

Note that Note that $|I(t)| = r$ for all sink nodes $t \in \mathcal{T}$ in the reduced network, which makes each global transfer matrix G_t an $r \times r$ matrix.



Proof of the Linear Network Multicasting Theorem

Let $\alpha_1, \ldots, \alpha_\eta$ be a set of variables each corresponding to a different entry of the local transfer matrices.

Proof of the Linear Network Multicasting Theorem

Let $\alpha_1, \ldots, \alpha_\eta$ be a set of variables each corresponding to a different entry of the local transfer matrices.

The entries of $G_t(\alpha_1, \ldots, \alpha_\eta)$ are polynomials in $\mathbb{F}_q[\alpha_1, \ldots, \alpha_\eta]$.

Proof of the Linear Network Multicasting Theorem

Let $\alpha_1, \ldots, \alpha_\eta$ be a set of variables each corresponding to a different entry of the local transfer matrices.

The entries of $G_t(\alpha_1, \ldots, \alpha_\eta)$ are polynomials in $\mathbb{F}_q[\alpha_1, \ldots, \alpha_\eta]$.

We have that $\det G_t(\alpha_1, \ldots, \alpha_\eta) \in \mathbb{F}_q[\alpha_1, \ldots, \alpha_\eta]$ and

$$G_t(\alpha_1, \ldots, \alpha_\eta) \text{ is invertible} \iff \det G_t(\alpha_1, \ldots, \alpha_\eta) \neq 0.$$

Proof of the Linear Network Multicasting Theorem

Let $\alpha_1, \ldots, \alpha_\eta$ be a set of variables each corresponding to a different entry of the local transfer matrices.

The entries of $G_t(\alpha_1, \ldots, \alpha_\eta)$ are polynomials in $\mathbb{F}_q[\alpha_1, \ldots, \alpha_\eta]$.

We have that $\det G_t(\alpha_1, \ldots, \alpha_\eta) \in \mathbb{F}_q[\alpha_1, \ldots, \alpha_\eta]$ and

$$G_t(\alpha_1, \ldots, \alpha_\eta) \text{ is invertible } \iff \det G_t(\alpha_1, \ldots, \alpha_\eta) \neq 0.$$

$\det G_t(\alpha_1, \ldots, \alpha_\eta)$ is not identically zero because if we use routing then $G_t = I_t$.

Proof of the Linear Network Multicasting Theorem

Because the network is acyclic, we note that any variable α_i may in generally appear with degree at most one in any of the entries of G_t

Proof of the Linear Network Multicasting Theorem

Because the network is acyclic, we note that any variable α_i may in generally appear with degree at most one in any of the entries of G_t

This implies that $\deg_{\alpha_i} \det G_t(\alpha_1, \ldots, \alpha_\eta) \leq r$ for each variable.

Proof of the Linear Network Multicasting Theorem

Because the network is acyclic, we note that any variable α_i may in generally appear with degree at most one in any of the entries of G_t

This implies that $\deg_{\alpha_i} \det G_t(\alpha_1, \dots, \alpha_\eta) \leq r$ for each variable.

In order for all the global transfer matrices to be invertible simultaneously, we need that

$$\prod_{t \in \mathcal{T}} \det G_t(\alpha_1, \dots, \alpha_\eta) \neq 0$$

Proof of the Linear Network Multicasting Theorem

Because the network is acyclic, we note that any variable α_i may in generally appear with degree at most one in any of the entries of G_t

This implies that $\deg_{\alpha_i} \det G_t(\alpha_1, \ldots, \alpha_\eta) \leq r$ for each variable.

In order for all the global transfer matrices to be invertible simultaneously, we need that

$$\prod_{t \in \mathcal{T}} \det G_t(\alpha_1, \ldots, \alpha_\eta) \neq 0$$

Since

$$\deg_{\alpha_i} \prod_{t \in \mathcal{T}} \det G_t(\alpha_1, \ldots, \alpha_\eta) \leq r|\mathcal{T}|,$$

then if $q \geq r|\mathcal{T}|$ by the Sparse Zero Lemma, there exists a $(a_1, \ldots, a_\eta) \in \mathbb{F}_q^\eta$ such that $\prod_{t \in \mathcal{T}} \det G_t(a_1, \ldots, a_\eta) \neq 0$.

- One Source Networks
  - The Unicast Network
  - The Multicast Network

- Optimality of Linear Network Coding.

- Insufficiency of Linear Network Coding.

Bound on the Size of the Field

Theorem (Linear Network Multicasting Theorem)

Let $\mathcal{N} = (\mathcal{V}, \mathcal{E}, \{s\}, \mathcal{T})$, then

$$\rho(\mathcal{N}) \leq \min\{\text{mincut}(s, t) \mid t \in \mathcal{T}\}.$$

Furthermore, the bound is achievable, for sufficiently large q, with linear network coding.

Bound on the Size of the Field

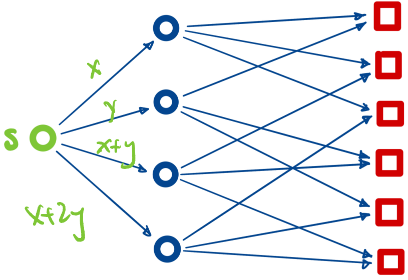
Theorem (Linear Network Multicasting Theorem)

Let $\mathcal{N} = (\mathcal{V}, \mathcal{E}, \{s\}, \mathcal{T})$, then

$$\rho(\mathcal{N}) \leq \min\{\text{mincut}(s,t) \mid t \in \mathcal{T}\}.$$

Furthermore, the bound is achievable for $q \geq |\mathcal{T}|$ with linear network coding.

Network with Connection to Coding Theory



$\begin{pmatrix} 1 \\ 1 & 2 \end{pmatrix}$

x, y

$$\begin{pmatrix} 1 & 0 & 1 & 1 & \cdots \\ 0 & 1 & 1 & 2 & \end{pmatrix} \leftarrow \text{RS code}$$

Possible Questions

- What is the cost of finding a network code?
- What is the chance for a random linear network coding to achieve capacity?

Possible Questions

- What is the cost of finding a network code?

  A linear information flow algorithm can be efficiently implemented, meaning that its running time scales as a polynomial in $|\mathcal{E}|$, $|\mathcal{T}|$, and $\log(q)$.

- What is the chance for a random linear network coding to achieve capacity?

Possible Questions

- What is the cost of finding a network code?

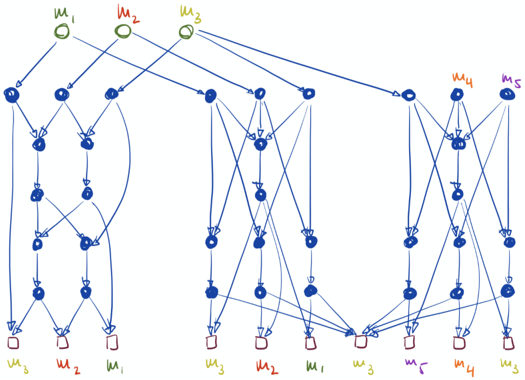
  A linear information flow algorithm can be efficiently implemented, meaning that its running time scales as a polynomial in $|\mathcal{E}|$, $|\mathcal{T}|$, and $\log(q)$.

- What is the chance for a random linear network coding to achieve capacity?

  the success probability is larger than or equal to $(1 - |\mathcal{T}|/q)^{\eta'}$

- One Source Networks
  - The Unicast Network
  - The Multicast Network
- Optimality of Linear Network Coding.
- Insufficiency of Linear Network Coding.

Insufficiency of LNC [3]



Theorem

A solvable network exists that has no linear solution over any finite field or vector dimension.

References

- S. . R. Li, R. W. Yeung, and Ning Cai, "Linear network coding," *IEEE Transactions on Information Theory*, vol. 49, pp. 371–381, Feb 2003.
- R. Koetter and M. Medard, "An algebraic approach to network coding," *IEEE/ACM Transactions on Networking*, vol. 11, pp. 782–795, Oct 2003.
- R. Dougherty, C. Freiling, and K. Zeger, "Insufficiency of linear coding in network information flow," *IEEE Transactions on Information Theory*, vol. 51, pp. 2745–2759, Aug 2005.