Abstract—Can we design a communication network just like a huge linear time-invariant filter? To answer this question, we generalize the celebrated mincut-maxflow theorem to linear time-invariant networks where edges are labeled with transfer functions instead of integer capacity constraints. We prove that when the transfer functions are linear time-invariant, the fundamental design limit, mincut, is achievable by a linear time-invariant scheme regardless of the topology of the network. Whereas prior works are based on layered networks, our proof has a novel way of converting an arbitrary relay network to an equivalent acyclic single-hop relay network, which we call Network Linearization. This theorem also reveals a strong connection between network coding and linear system theory.

I. INTRODUCTION

The emergence of network coding [1] changed our perspective on information flow in networks. The relay nodes could now process and mix signals rather than just routing them. This freedom to operate inside the nodes liberated us to start considering operations and stream mixing outside the nodes as well. This extension is a natural fit for wireless communication whose distinguishing features are broadcast and superposition [2].

A deterministic linear network is a noiseless network in which the nodes are connected by transfer functions rather than orthogonal links. In deterministic networks, the traditional mincut concept is generalized to the rank of the channel transfer matrix.

Avestimehr et al. [3] showed that a deterministic linear network can be a good approximation for a Gaussian network and proved that the capacity of the deterministic linear network is equal to the mincut of the network when we allow arbitrary complexity for the relay nodes. In their proof, rather than directly considering arbitrary networks, they first considered layered networks in which the only edges are from one layer to the next layer. The idea to convert an arbitrary network to a layered network is network unfolding. As we can see in Fig. 1, by introducing duplicated nodes over the time, any arbitrary network can be approximated as a layered network. Moreover, the capacity of the layered network approaches the capacity of the original network as the time expansion gets large. Since layered networks have a quite attractive and simple topology, a series of works by others [4], [5], [6] exclusively focused on it and developed algorithms that find deterministic linear schemes for layered networks.

However, what these papers are overlooking is that when we fold the unfolded network back to its physical topology a time-invariant scheme can become a time-varying scheme. The example shown in Fig. 1 shows that network-coding design based on an unfolded network can cause a significant problem even in a simple network with one source, one relay and one destination.

The source transmits $u_1, \cdots, u_6$ to the destination. The letters on the arrows of the unfolded network represent the flows of information. We can easily check that the network coding scheme shown in the figure is mincut achieving. However, when we fold it back, we can see problems for implementation. First of all, the scheme is time-varying at the relay. Thus, every node in the network has to be synchronized to a common clock. Moreover, the transmitted signal at a given time step may depend on all of its previously received signals, which may require a large memory.

It seems practically preferable to have a linear time-invariant scheme. However, whether we can achieve the capacity of an arbitrary deterministic network with a linear time-invariant scheme had remained an open question.

To answer this question, we adapt Koetter and Medard’s novel view of network coding [7]. It is worth mentioning that Kim et al. [8] also attempt a generalization of Koetter and Medard’s work. However, their results turn out to be limited to layered networks. Koetter and Medard considered a communication network as a linear time-invariant system. Then, the capacity of the network is simply the rank of the transfer function. In this way, we can convert a network coding problem to an algebra problem for a linear time-invariant system. However, the way they wrote the proofs is still based on communication network theory rather than linear system theory [9]. For example, they justify that the maximum rank of the transfer function matrix is the mincut of the network by using the Ford-Fulkerson algorithm [7, Theorem 2]. As a result, it is unclear how to extend their
results to deterministic networks with LTI-transfer-function links without the corresponding theorems in communication network theory.

In this paper, we extend Koetter and Médard’s result to such deterministic networks, i.e., we prove that we can achieve the mincut of an arbitrary deterministic network with a linear time-invariant scheme. The main idea is network linearization, which is based on linear system theory. First, we introduce internal states at the relays and represent the network as a linear equation. By algebraically converting the linear equation back to a network, we can obtain a single-hop acyclic network that exhibits an equivalent behavior to the original network.

The rest of the paper is organized as follows: We first formulate the problem and state the algebraic mincut-maxflow theorem which generalizes the celebrated mincut-maxflow theorem [10], [11]. We introduce the proof ideas and then give a formal proof.

II. PROBLEM STATEMENT AND MAIN RESULT

An LTI point-to-point network \( N(Z) \) is a collection of a transmitter, relays, and a receiver\(^1\). There are also transfer functions connecting them. The network can be represented by a graph. Consider a digraph \((W, E)\) with a totally ordered set of vertices \( W \) and a set of edges \( E \). \( W \) is partitioned into the sets \( N_0, \cdots, N_{v+1} \), i.e., \( N_i \subseteq W \), \( N_i \cap N_j = \emptyset \) for \( i \neq j \), and \( \bigcup_{i=0}^{v+1} N_i = W \). We consider a set of vertices \( N_i \) as a node. More precisely, \( N_0 \) is a transmitter, \( N_{v+1} \) is a receiver, and the \( v \) others are relays. To emphasize this, we also denote \( N_{tx} := N_0 \) and \( N_{rx} := N_{v+1} \). For a given node \( N_i \), the elements of \( N_i \) are again partitioned into two subsets \( N_{i,in} \) and \( N_{i,out} \) which are called input and output vertices of the node \( i \). The inputs and the outputs are defined in a channel-centric perspective. \( N_{i,in} \) represents the signals going out from the node \( i \) into the channels and \( N_{i,out} \) represents the signals coming out from the channel into the node \( i \). The transmitter node does not receive signals and the receiver node does not transmit signals, so \( N_{tx,out} = \emptyset \) and \( N_{rx,in} = \emptyset \). We denote \( d_{tx} := |N_{tx,in}| \), \( d_{tx} := |N_{tx,out}| \), \( d_{rx} := |N_{rx,in}| \), and \( d_{rx} := |N_{rx,out}| \).

Let \( Z = \{z_1, \cdots, z_{|Z|}\} \) and \( K = \{k_1, \cdots, k_{|K|}\} \) be sets of variables. Given a coefficient field \( \mathbb{F} \), let \( \mathbb{F}[Z], \mathbb{F}[K], \mathbb{F}[Z, K] \) be the field of all rational functions in variables \( Z, K, Z \cup K \) with coefficients in \( \mathbb{F} \) respectively. An edge is denoted by a triplet \((w', w'', h_{w', w''}(Z, K)) \in E\) where \( w', w'' \in W \) and \( h_{w', w''}(Z, K) \in \mathbb{F}[Z, K] \). \( w' \) is called the tail of the edge and \( w'' \) is called the head of the edge.

Since a lack of physical connection between two vertices \( w' \) and \( w'' \) can be represented as \( h_{w', w''}(Z, K) = 0 \), we assume that every input vertex is connected to every output vertex, including “self-loops” connecting the vertices within individual nodes. The internode transfer functions are described by rational functions on \( Z \). Formally, for all \( i, j \in \{0, \cdots, v+1\} \) and \( w' \in N_{in,i}, w'' \in N_{out,j} \),

\[
(w', w'', h_{w', w''}(Z, K)) \in E \quad \text{and} \quad h_{w', w''}(Z, K) \in \mathbb{F}[Z].
\]

Inside each node, edges are fully connecting its output vertices to its input vertices by transfer functions in \( \mathbb{F}[K] \). These transfer functions are chosen from \( \mathbb{F}[K] \) to reflect the design freedom at the relays independent from the channel gains. Let’s call these edges internal edges. Then we can see that \( N_{tx} \) and \( N_{rx} \) do not have internal edges. Moreover, we assume the transfer functions of the internal edges of all nodes are in the form of \( k_i \in K \) and all distinct. This distinct transfer function assumption guarantees enough design freedom at the relays.

Formally, for all \( i \in \{1, \cdots, v\} \) and \( w' \in N_{out,i}, w'' \in N_{in,i} \), \((w', w'', h_{w', w''}(Z, K)) \in E \) and \( h_{w', w''}(Z, K) = k_{w', w''} \) where \( k_{w', w''} \in K \). If \((w'_1, w''_1)\) and \((w'_2, w''_2)\) are distinct internal edges, \( h_{w'_1, w''_1} \neq h_{w'_2, w''_2} \).

At each vertex and edge, the signal is processed as follows: Each vertex \( w \in W \) adds all the signals coming from the edges whose head is \( w \) and transmits to the edges whose tail is \( w \). Each edge \( e \in E \) multiplies the signal coming from its tail with its transfer function and transmits to its head.

Denote a transfer function matrix from the input vertices of the node \( N_i \) to the output vertices of the node \( N_j \) as \( H_{i,j}(Z) \). Here, let “\( tx \)” represent 0 and “\( rx \)” represent \( v+1 \). In the same way, we denote a transfer function from a set of nodes \( A \) to a set of nodes \( B \) as \( H_{A,B}(Z) \). We also denote the transfer function matrix from the output vertices of \( N_i \) to the input vertices of \( N_j \) as \( K_{i,j}(K) \). Then, \( H_{i,j}(Z) \in \mathbb{F}[Z]^{d_{out,j} \times d_{in,i}} \) and \( K_{i,j}(K) \in \mathbb{F}[K]^{d_{in,i} \times d_{out,j}} \). For briefness, we write \( H_{i,j}(Z) \) as \( H_{ij} \) and \( K_{i,j}(K) \) as \( K_{ij} \) when it does not cause confusion.

Let \( G(Z, K_i) \) be the transfer function from the input vertices of the transmitter node to the output vertices of the receiver node. \( G(Z, K_i) \) can be written in terms of \( H_{ij} \) and \( K_{ij} \)[12].

**Theorem 1.** With the above definitions, the transfer function matrix \( G(Z, K_i) \) is given as

\[
G(Z, K_i) = [H_{1,rx} K_1 \cdots H_{v,rx} K_v] \quad \begin{bmatrix} I - [H_{1,1} K_1 \cdots H_{1,v} K_v]^{-1} [H_{tx,1} \cdots H_{tx,v}] \end{bmatrix} [H_{tx,1} \cdots H_{tx,v}] + H_{tx,rx}.
\]

**Proof:** As Fig. 3, let \( U, X_i \) and \( Y \) be vectors of signals (which are rational functions on \( Z \) and \( K_i \)) at the input vertices of the transmitter, the output vertices visible at node \( i \), and the output vertices visible at the receiver. Then, we have the following relations between \( U, X_i \) and \( Y \):

\[
X_1 = \begin{bmatrix} H_{1,1} K_1 & \cdots & H_{1,v} K_v \end{bmatrix} X_1 + \begin{bmatrix} H_{e,1} \\ \vdots \\ H_{e,v} \\ H_{tx,v} \end{bmatrix} U,
\]

\[
Y = \begin{bmatrix} H_{1,rx} K_1 & \cdots & H_{v,rx} K_v \end{bmatrix} X_1 + \begin{bmatrix} H_{tx,1} \\ \vdots \\ H_{tx,v} \end{bmatrix} U.
\]

\(^1\)The LTI networks considered here are essentially the same as the linear deterministic model studied in [3] except that the LTI network restricts the relay design to be linear time-invariant and the underlying field can be real or complex rather than a finite field.
Thus, simple algebra gives the theorem. Here, the existence of the matrix inverse follows from the invertibility of $I$ and the definition of $K_i$. Therefore, from end-to-end perspective, the point-to-point LTI network $\mathcal{N}(Z)$ can be thought as a MIMO (multiple-input multiple-output) channel whose channel matrix is $G(Z, K)$. Since the capacity of MIMO channels is closely related to the rank of the channel matrix [2], we can think of the rank of $G(Z, K)$ as the maxflow of the network.

One key fact about LTI networks is that the well-known mincut-maxflow theorem [11], [10] can be extended to them.

**Theorem 2** (Algebraic Mincut-Maxflow Theorem). With the above definitions,

$$
\text{rank } G(Z, K_i) = \min_{V: V \subseteq \{0, \ldots, v+1\}, V \ni tx, V \not\ni rx} \text{rank } H_{V, V^c}(Z).
$$

**Proof:** See Section IV.

Even though the theorem is written for variables $K$, what it tells us is that the mincut is achievable by a linear time-invariant scheme when we extend the field size enough [7, Lemma 1].

### III. IDEAS FOR NETWORK LINEARIZATION

Network linearization is the counterpart of the following linear system theory fact: Every strictly causal linear time-invariant system with an input $u[n]$ and an output $y[n]$ can be written in state-space form, i.e., can be realized as a linear equation:

$$
x[n+1] = Ax[n] + Bu[n]
$$

$$
y[n] = Cx[n]
$$

by introducing proper internal states $x[n]$. Similarly, network linearization tells that every LTI network with an arbitrary topology can be converted to an acyclic single-hop relay network by introducing proper internal states.

First, we illustrate two ideas for network linearization.

1. **Internal States:** Consider the two-hop relay network shown in the top figure of Fig. 2. The transfer function from $U$ to $Y$ is $K_2 K_1$, which is not linear in $K_1, K_2$. To write the transfer function in a linear form, we introduce an internal state $X$ at the output of the second node. Then, the transfer function matrix from $X, U$ to $Y, X$ is

$$
\begin{bmatrix}
  Y \\
  X
\end{bmatrix} =
\begin{bmatrix}
  K_2 & 0 \\
  0 & K_1
\end{bmatrix}
\begin{bmatrix}
  X \\
  U
\end{bmatrix},
$$

which is linear in $K_1, K_2$. Moreover, since

$$
\begin{bmatrix}
  K_2 & 0 \\
  0 & K_1
\end{bmatrix} =
\begin{bmatrix}
  0 & 1 \\
  1 & 0
\end{bmatrix} K_1 [0 1] +
\begin{bmatrix}
  1 & 0
\end{bmatrix} K_2 [1 0],
$$

it corresponds to the transfer function of the acyclic single-hop relay network shown in the bottom figure of Fig. 2.

2. **Circulation Arc:** Even if the transfer function can be written in a linear form by introducing internal states, there has to be a relationship between the rank of the original transfer function and the rank of the linearized transfer function. To make this connection, we borrow the circulation arc idea from an integer programming context [13, p.86]. The problem that they had was that when they tried to write the maxflow problem in linear programming form, the flow conservation law did not hold at the source and the destination. The flow at the source is negative and the flow at the destination is positive. To patch this, they introduced a circulation arc with infinite capacity from the destination to the source. Since the amount of the negative flow at the source is the same as the amount of the positive flow at the destination, the flow conservation law can be recovered. Moreover, the flow across the network can be easily measured by measuring the flow in the circulation arc.

To apply this idea to LTI networks, we use an underdetermined system. Let’s consider $x = x + K_{tx} G(Z, K_i) K_{tx}$ with unknown vector $x$. Here, $K_{tx} G(Z, K_i) K_{tx}$ is a transfer function with a preprocessing matrix $K_{tx}$ and a postprocessing matrix $K_{tx}$. If the rank of $K_{tx} G(Z, K_i) K_{tx}$ is smaller than the dimension of $x$, the equation is underdetermined. Otherwise, it is not. Thus, we can see that the rank of the transfer function can be measured by the underdeterminedness of the system.

### IV. NETWORK LINEARIZATION AND A PROOF OF ALGEBRAIC MINCUT-MAXFLOW THEOREM

Now, we will combine these ideas for network linearization. We first introduce the circulation arc. As shown in Fig. 3, an auxiliary node $N_{ax}$ with $d_{ax}$ input vertices and $d_{ax}$ output vertices is added to the original network. We also introduce $d_{ax}$ input vertices at the receiver node and $d_{ax}$ output vertices at the transmitter node. Following the previous notations, we put “$ax$” represent $v + 2$, and use the notation $H_{i,j}$ and $K_i$ as before. Let $H_{rx, ax} = H_{ax, tx} = H_{ax, ax} = K_{ax} = I_{d_{ax}}$. $K_{tx} \in \mathbb{F}[K]^{d_{tx} \times d_{ax}}$ and $K_{rx} \in \mathbb{F}[K]^{d_{ax} \times d_{tx}}$ are selected as those of the relay nodes, i.e., the transfer functions are in the form of $k_i \in K$ and they are all distinct.

Now, we introduce the internal states. As shown in Fig. 3, let $X_{ax}, X_i$, and $Y$ be the vectors of the signals of the output vertices seen at the auxiliary node, the node $i$, and the receiver respectively.

From the system diagram, Fig. 3, we can see the following.
We will prove the equivalence between the original network \( N(Z) \) and the linearized network \( \mathcal{N}_{\text{lin}}(Z) \).

Let \( d := \dim Y + \sum_{1 \leq i \leq \ell} \dim X_i = d_{\text{tx}} + \sum_{1 \leq i \leq \ell} d_{\text{out}}. \) Then, the maxflow of \( \mathcal{N}_{\text{lin}}(Z) \) is the same as the maxflow of \( N(Z) \) by an offset \( d \).

**Lemma 1** (Maxflow Equivalence Lemma). Given the above notations,

\[
\text{rank}(K_{rx}G(Z, K_i)K_{tx}) + d = \text{rank} G_{\text{lin}}(Z, K_i).
\]

**Proof:** \( K_{rx}G(Z, K_i)K_{tx} \) is the Schur’s complement of \( G_{\text{lin}}(Z, K_i) \). Thus, we can use the fact that

\[
\text{rank} \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \text{rank} D + \text{rank}(A - BD^{-1}C)
\]

where \( D \) is invertible. See [14] for the details.

The mincut of \( \mathcal{N}_{\text{lin}}(Z) \) is also the same as the mincut of \( N(Z) \) by an offset \( d \).

**Lemma 2** (Mincut Equivalence Lemma). Given the above notations,

\[
\begin{aligned}
&\min_{V \subseteq \{0, \ldots, v+1\}, V \ni tx, V \not\ni rx} \text{rank} H_{V, V^c}, \text{rank} K_{tx}, \text{rank} K_{rx} + d = \\
&\min_{V \subseteq \{-1, \ldots, v+2\}, V \ni tx^*, V \not\ni rx^*} \text{rank} H_{V, V^c}.
\end{aligned}
\]

**Proof:** Let \( V \) be a cut of \( \mathcal{N}_{\text{lin}}(Z) \), i.e. \( V \subseteq \{-1, \ldots, v+2\}, V \ni tx^* \) and \( V \not\ni rx^* \). We will divide \( V \) into three cases.

(i) When \( tx \in V^c \),

Since \( \text{rank} \begin{bmatrix} A \\ C_{tx} \end{bmatrix} = \dim X_{ax} + d \), we have

\[
\begin{aligned}
&\min_{V \subseteq \{-1, \ldots, v+2\}, V \ni tx^*, V \not\ni rx^*, V^c \ni tx} \text{rank} H_{V, V^c} = \dim X_{ax} + d.
\end{aligned}
\]

(ii) When \( rx \in V \),

Since \( \text{rank} \begin{bmatrix} A & B_{tx} \end{bmatrix} = \dim X_{ax} + d \), we have

\[
\begin{aligned}
&\min_{V \subseteq \{-1, \ldots, v+2\}, V \ni tx^*, V \not\ni rx^*, V \ni rx} \text{rank} H_{V, V^c} = \dim X_{ax} + d.
\end{aligned}
\]
(iii) When $tx \in V$ and $rx \in V^c$. Let $V' := V \setminus \{tx\}$ and $V'' := V^c \setminus \{rx\}$. Then, we can show

$$\text{rank } H^{\text{lin}}_{V',V'} = \text{rank } H_{V',V''} + d.$$ 

Finally, by (i),(ii) and (iii) we can prove the lemma.

$$\min_{V \subseteq \{-1, \ldots, v+2\}} \text{rank } H^{\text{lin}}_{V',V'}
= \min_{V \subseteq \{-1, \ldots, v+2\}} \min_{V \ni tx, V \not
= \min_{V \subseteq \{-1, \ldots, v+2\}} \min_{V \ni tx, V \not
rank K_{tx}, \text{rank } K_{rx}} + d$$

See [14] for the details.

Moreover, it is known that the algebraic mincut-maxflow theorem holds for the linearized network $N_{\text{lin}}(Z, K)$ [15, Theorem 4.1].

**Theorem 3 (Algebraic Mincut-Maxflow Theorem for Linearized Network [15]).** Given the above notations,

$$\text{rank } G_{\text{lin}}(Z, K) = \min_{V \subseteq \{-1, \ldots, v+2\}, V \ni tx, V \not \ni tx'} \text{rank } H^{\text{lin}}_{V',V'}$$

**Proof:** See [15], or [14] for a cleaner proof.

Now, we can prove the algebraic mincut-maxflow theorem for a general LTI network.

**Proof of Theorem 2:** Since we can arbitrarily choose $d_{ax}$, let $d_{ax} \geq \max\{d_{tx}, d_{rx}\}$. Then,

$$\text{rank } G(Z, K_i) = \text{rank } (K_{rx} G(Z, K_i) K_{tx})$$

$$= \text{rank } G_{\text{lin}}(Z, K_i) - d$$

$$= \min_{V \subseteq \{-1, \ldots, v+2\}, V \ni tx', V \not \ni tx'} \text{rank } H^{\text{lin}}_{V',V'} - d$$

$$= \min_{V \subseteq \{-1, \ldots, v+2\}, V \ni tx', V \not \ni tx'} \text{rank } H_{V',V''}$$

$$= \min_{V \subseteq \{-1, \ldots, v+2\}, V \ni tx', V \not \ni tx'} \text{rank } H_{V',V''}$$

$$= \min_{V \subseteq \{-1, \ldots, v+2\}, V \ni tx', V \not \ni tx'} \text{rank } H_{V',V''}$$

(1) is due to the following fact: If we select $K_{tx}$ as a $0 - 1$ matrix that chooses independent rows of $G(Z, K_i)$ and $K_{tx}$ as a $0 - 1$ matrix that chooses independent columns of $K_{rx} G(Z, K_i)$, (1) holds.

(2), (3) and (4) follow from Lemma 1, Theorem 3 and Lemma 2 respectively.

(5) follows from that the mincut of $N(Z)$ is not greater than $\min\{d_{tx}, d_{rx}\}$, rank $K_{tx} = d_{tx}$ and rank $K_{rx} = d_{rx}$. ■

V. DISCUSSION

We proved that the algebraic mincut-maxflow theorem holds in LTI networks. Moreover, Lemma 1 of [7] tells that even if the underlying field $F$ is finite, we can always find constant matrices in its extended field that achieve the maximum rank. Therefore, we can conclude that there exists a linear time-invariant scheme that achieves the mincut of the network.

In the proof, we also found a novel way to convert an arbitrary LTI network to an equivalent acyclic single-hop LTI network. Moreover, Lemma 1 can be extended to constant matrices $K_i \in \mathbb{F}[1]^{d_{in}, d_{out}}$ as long as $[I - H_{1,1} K_1 \cdots - H_{1,v} K_v]$

$$\begin{bmatrix}
    I - H_{1,1} K_1 & \cdots & - H_{1,v} K_v
    \vdots & \ddots & \vdots
    - H_{v,1} K_1 & \cdots & I - H_{v,v} K_v
  \end{bmatrix}$$

is invertible, which is just an existence condition of the transfer function. Therefore, we can design a network-coding scheme in the linearized LTI network, and then immediately apply it to the original LTI network.

It is also worthwhile to notice that the difference between the linearized LTI network and a 1-layered network is the existence of the direct link from the transmitter to the receiver.

We also want to emphasize the generality of the algebraic mincut-maxflow theorem. Since it holds for all linear time-invariant systems, the application of the theorem does not have to be restricted to the design of communication systems. For example, by taking the underlying field $\mathbb{F}$ to be the complex or real field, we can apply the theorem to filter design problems in signal processing or linear control system design problems in decentralized control theory. In fact, a strong connection between communication and control problems can be made by the algebraic mincut-maxflow theorem [14].

REFERENCES


