Rotational Coding Achieves Multicast Capacity of Deterministic Wireless Networks

Mohammad A. (Amir) Khojastepour
NEC Laboratories America
4 Independence Way, Suite 200
Princeton, NJ, 08540
Email: amir@nec-labs.com

Alireza Keshavarz-Haddad
School of Electrical and Computer Engineering
Shiraz University
Zand Ave, Shiraz, Iran
Email: alireza@alumni.rice.edu

***Abstract*— In this paper we study the maximum throughput of network coding schemes for a single multicast session in wireless networks. We adapt "deterministic channel model" proposed in [1] for modeling wireless interference. We introduce a novel *rotational coding* scheme that can achieve the well-known minimum cutset bound. This coding scheme has lower encoding complexity in comparison with the existing random linear coding schemes which makes it a good candidate for practical systems. Moreover, we present a fundamental result on the rate of information that can be sent through any cutset. This can be used as criteria for analyzing the maximum throughput of linear network codes.**

I. Introduction

In recent years, network coding has become an important research topic in network information theory. It has been shown that network coding can help to improve the throughput, energy consumption, delay, robustness, and some other performance metrics of communication networks (see [2]). Network coding was first introduced in the seminal paper by Ahlswede et al. [3] in which it was proved that the maximum flow capacity of a single multicast session can be achieved using network coding in wired networks with directional links. Later, [4] and [5] show constructively that the random and deterministic linear network codes can achieve the minimum cutset bound of a single multicast session as well. Several network coding schemes for multicasting have been proposed more recently; some important literatures can be found in [6].

Recently, a deterministic approach to study wireless networks was introduced in [1]. This model incorporates both broadcast and interference challenges in the wireless network. However, by removing the randomness this model makes the challenging problem of network coding for wireless network analytically tractable. For example, the problem of maximum flow capacity of a single multicast session in wireless networks was studied in [7] where it was shown that similar to the results for wired networks [3], the minimum cutset bound can be achieved. In [7], a random network coding scheme is proposed to show the achievability of the cutset upper bound. The presented proof contains two steps: first proving the result for the layered networks (will be defined in Section II) and second extending the proof for acyclic networks (no directional loop in the network) by unfolding the network over the time and use the results for layered networks. A shorter direct proof for achievability of the random linear codes is presented in [8].

In this paper we study the maximum throughput of network coding schemes for a single multicast session in wireless networks. We adapt "deterministic channel model" of [1] for modeling the channel. We introduce a novel network coding scheme called *rotational coding* which has a different construction from the existing random coding schemes. It has much lower computational complexity for encoding in relay nodes and therefore it is more applicable for practical uses. Moreover, we present a fundamental theorem on the maximum rate of information that can be sent through any cutset of a network under linear coding schemes when using deterministic channel model. Next, we apply the theorem for rotational coding scheme and prove that our coding scheme can achieve the minimum cutset bound of a single multicast session in layered wireless networks.

This paper is organized as follows. In Section II, we describe the network model and notations. In Section III, we introduce rotational network coding scheme and prove its achievability. Finally, we conclude the paper in Section IV.

II. Network Model and Notations

We consider a wireless network as a directed graph where each node can transmit the same message into *all* its outgoing links and receives the *superposition* of the signals arrived from the incoming links. We adopt the deterministic model in [1], [7] to model the gain of the links and how the superposition is performed. Notice that in this model the nodes are *full-duplex* i.e. they can simultaneously transmit and receive data; so there is no concept of scheduling for the transmissions of network nodes.

We assume that the network contains $1+N$ nodes, where one of them is the source of a multicast session and the rest of the nodes are relay nodes or terminals (destinations) of the session. We use a universal index k for every node Φ_k where $k = 0, 1, \ldots, N$.

A. Deterministic Channel Model

In this channel model, the output signal from node Φ_k at time-slot t is considered as a column vector $\mathbf{y}_t^k = [y_{t,1}^k, y_{t,2}^k, \ldots, y_{t,q}^k]^\dagger$ of size q, where each element is a value in Galois Field $\mathbb{F}(p^n)$ for some prime number p and positive integer n. Here $\dagger$ is used to denote the matrix transpose operation. Each link from the node Φ_i to Φ_k in the network is denoted by its transfer function $\mathbf{G}_i^k$ which is a $q \times q$ matrix with the entries in $\mathbb{F}(p^n)$. The output of this link is equal to $\mathbf{G}_i^k \mathbf{y}_t^i$. The received vector or input at the node Φ_k is a column vector $\mathbf{x}_t^k = [x_{t,1}^k, x_{t,2}^k, \ldots, x_{t,q}^k]^\dagger$ which is the superposition of the outputs of the links arriving at node Φ_k defined on component-by-component basis, i.e.,

$$\mathbf{x}_t^k = \sum_{i=0}^{n} \mathbf{G}_i^k \mathbf{y}_t^i \tag{1}$$

where $\mathbf{G}_i^k$ is the transfer function when there is an outgoing link from Φ_i to Φ_k, otherwise it set to $q \times q$ matrix $\mathbf{0}$.

If we stack together the received vectors at multiple nodes $\Phi_k, k \in \mathcal{B} = \{j_1, \ldots, j_b\}$, assuming they are characterized by the output at the nodes $\Phi_k, k \in \mathcal{A} = \{i_1, \ldots, i_a\}$ (i.e. all incoming links of $\mathcal{B}$ originated in $\mathcal{A}$), then the transfer function is given by

$$\begin{bmatrix} \mathbf{x}_t^{j_1} \\ \mathbf{x}_t^{j_2} \\ \vdots \\ \mathbf{x}_t^{j_b} \end{bmatrix} = \begin{bmatrix} \mathbf{G}_{i_1}^{j_1} & \mathbf{G}_{i_2}^{j_1} & \ldots & \mathbf{G}_{i_a}^{j_1} \\ \mathbf{G}_{i_1}^{j_2} & \mathbf{G}_{i_2}^{j_2} & \ldots & \mathbf{G}_{i_a}^{j_2} \\ \vdots & & & \vdots \\ \mathbf{G}_{i_1}^{j_b} & \mathbf{G}_{i_2}^{j_b} & \ldots & \mathbf{G}_{i_a}^{j_b} \end{bmatrix} \begin{bmatrix} \mathbf{y}_t^{i_1} \\ \mathbf{y}_t^{i_2} \\ \vdots \\ \mathbf{y}_t^{i_a} \end{bmatrix}$$

We denote the above matrix as $\mathbf{G}_{\mathcal{A}}^{\mathcal{B}}$, and will use this matrix notation in several places again.

B. Time-Frame Operations for Linear Coding Schemes

In linear coding schemes each node performs a linear operation over a time-frame of T time-slots, i.e., after receiving T vectors $\mathbf{x}_t^k$, $t = mT+1, \ldots, mT+T$ at time-frame m, the node Φ_k linearly maps qT received symbols by a matrix $\mathbf{F}_k$ of size $qT \times qT$ to find qT new symbols which are outgoing symbols. Then, puts these symbols into T column vectors $\mathbf{y}_t^k$, $t = (m+1)T+1, \ldots, (m+1)T+T$ that will be transmitted in the next T time-frame (time-frame $m+1$). The linear operation can be formulated as following:

$$\begin{bmatrix} y_{(m+1)T+1,1}^k \\ \vdots \\ y_{(m+1)T+T,1}^k \\ y_{(m+1)T+1,2}^k \\ \vdots \\ y_{(m+1)T+T,2}^k \\ \vdots \\ y_{(m+1)T+T,q}^k \end{bmatrix} = \mathbf{F}_k \begin{bmatrix} x_{mT+1,1}^k \\ \vdots \\ x_{mT+T,1}^k \\ x_{mT+1,2}^k \\ \vdots \\ x_{mT+T,2}^k \\ \vdots \\ x_{mT+T,q}^k \end{bmatrix} \tag{2}$$

Pay attention to the order of indices in (2); this order will be used later in Section III to describe our coding scheme. We denote the above vectors corresponding to outgoing symbols at time-frame $m+1$ and incoming symbols at time m of node k by $\mathbf{Y}_{(m+1)}^k$ and $\mathbf{X}_m^k$ respectively. We also define $\tilde{\mathbf{y}}_{(m+1),j}^k = [y_{(m+1)T+1,j}^k, y_{(m+1)T+2,j}^k, \ldots, y_{(m+1)T+T,j}^k]^\dagger$ and $\tilde{\mathbf{x}}_{m,j}^k = [x_{mT+1,j}^k, x_{mT+2,j}^k, \ldots, x_{mT+T,j}^k]^\dagger$ for all $j = 1, \ldots, q$.

Let $\mathbf{H}_{\mathcal{A}}^{\mathcal{B}} = \mathbf{G}_{\mathcal{A}}^{\mathcal{B}} \otimes \mathbf{I}$ where $\otimes$ is the Kronecker matrix product and $\mathbf{I}$ is a $T \times T$ identity matrix. Hence,

$$\mathbf{X}_m^k = \sum_{i=0}^{n} \mathbf{H}_i^k \mathbf{Y}_m^i \tag{3}$$

C. Notations of Layered Network Model

In this paper we assume that the wireless network is layered, i.e., the nodes are divided into $L+1$ layers namely layer 0, 1, ..., L such that the input for any node in layer l depends on the output of the nodes in layer $l-1$ and the transfer function of the links from the nodes in layer $l-1$ to the nodes in layer l. There exists exactly one node in layer 0 which is the source node of the single multicast session.

We set matrices $\mathbf{F}_k$ based on only the node index number k (independent from time-frame number m). The structure of layered network implies that at time-frame $m+l$ the nodes of layer l receive a linear

combinations of symbols which have been sent by the source at time-frame m. In other words, the symbols which are sent at different time-frames are not mixed at any node.

Since every node performs a linear transformation there will be a linear mapping between the sent symbols from the source (Φ_0) and received symbols by an arbitrary node. We denote this linear transformation that maps input symbols of Φ_0 to an arbitrary node Φ_i in layer l by $\mathbf{S}_i$. Then, for all m

$$\mathbf{X}^i_{m+l} = \mathbf{S}_i \mathbf{X}^0_m \tag{4}$$

Now, consider a node Φ_k in layer l. $\mathbf{S}_k$ can be written in terms of matrices of $\mathbf{F}_i$ and $\mathbf{H}^k_i$ of nodes in layer $l-1$ as following:

$$\mathbf{S}_k = \sum_{i \in \text{layer } l-1} \mathbf{H}^k_i \mathbf{F}_i \mathbf{S}_i \tag{5}$$

$$= \sum_{\text{paths from } i_0=0 \text{ to } i_l = k} \prod_{j=0}^{l-1} \mathbf{H}^{i_{j+1}}_{i_j} \mathbf{F}_{i_j} \tag{6}$$

(6) is obtained by repeating (5) for layers $l-2, \ldots, 1$.

Assume that the source selects its input vector from a vector space $V_0 = \mathbb{F}(p^n)^{qT}$. For an arbitrary set of nodes $\mathcal{A} = \{i_1, i_2, \ldots, i_a\}$ in layer l, we define Null($\mathcal{A}$) as the subspace of $\mathbb{F}(p^n)^{qT}$ which is mapped to $\mathbf{0}$ in all nodes of $\mathcal{A}$. In other words,

$$\text{Null}(\mathcal{A}) = \{\mathbf{X} \in \mathbb{F}(p^n)^{qT} : \forall k \in \mathcal{A},\ \mathbf{S}_k \mathbf{X} = \mathbf{0}\} \tag{7}$$

We denote the largest subspace of $\mathbb{F}(p^n)^{qT}$ orthogonal to $\mathcal{A}$ by $V_{\mathcal{A}}$ and call it *decodable space* by the set $\mathcal{A}$. It is straightforward to show that given the received vectors of the nodes of $\mathcal{A}$ we can uniquely determine which vector in $V_{\mathcal{A}}$ is sent by the source l time-frames earlier.

We denote the set theoretic difference of two sets $\mathcal{A}$ and $\mathcal{B}$, also called the relative difference of $\mathcal{A}$ from $\mathcal{B}$ by $\mathcal{A} \setminus \mathcal{B} = \mathcal{A} \cap \mathcal{B}^c$. We also use the same operator to denote the extended quotient space of two vector spaces $V_{\mathcal{A}}$ and $V_{\mathcal{A}} \cap V_{\mathcal{B}}$ as $V_{\mathcal{A}} \setminus V_{\mathcal{B}}$ which is defined as the largest subspace of $V_{\mathcal{A}}$ that is orthogonal to $V_{\mathcal{B}}$. The same operator may be applied to any two matrices, say, $\mathbf{G}_1$ and $\mathbf{G}_2$, as $\mathbf{G}_1 \setminus \mathbf{G}_2$ to denote the extended quotient space between the two vector spaces formed by the span of the column vectors of $\mathbf{G}_1$ and $\mathbf{G}_2$.

III. Novel Rotational Coding Scheme

In this section, we introduce novel "rotational linear coding" scheme which can be employed for different network coding applications. We start by presenting the coding scheme for a generic node Φ_k which depends on some parameters that are called *keys*. Next, we explain the relationship between the coding performed at different nodes in terms of these keys which have to be chosen such that the interaction between different nodes leads to maximization of the rank of the decodable space at any subset of nodes.

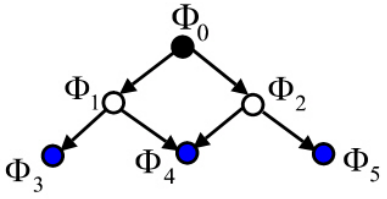

Fig. 1. The maximum throughput of a multicast session from Φ_0 to $\{\Phi_3, \Phi_4, \Phi_5\}$ cannot be achieved without setting the time-frame size T to be larger than 1.

Here we give an example to demonstrate how the linear operation is important for achieving the maximum throughput. Consider a layered network where $l_1 = 2$, and $\mathbf{G}^1_0 = \mathbf{G}^2_0 = \mathbf{I}$, also $l_2 = 3$, and $\mathbf{G}^3_1 = \mathbf{G}^4_1 = -\mathbf{G}^4_2 = -\mathbf{G}^5_2 = \mathbf{I}$ (see Fig. 1), where $\mathbf{I}$ is an identity matrix of size q. It is immediate that using, e.g., the same coding at the nodes Φ_1, Φ_2 in layer one will lead to no information reception at node Φ_4 while, in fact, it should be possible to send full rank q information to all three nodes Φ_3, Φ_4, Φ_5 simultaneously. This very same example on $\mathbb{F}(2)$ with $q = 1$ also shows that it is impossible to achieve the full rank transmission to all three nodes in layer 2, if the coding is performed in symbol level, i.e, $T = 1$, rather than over a block of T symbols. However, it is not hard to see that $T = 2$ can solve the problem.

Interestingly, we can show that for any finite number T there exists a network that needs a time-slot $T' > T$ for the linear mapping ($\mathbf{F}$) of the nodes to achieve the multicast capacity. In this network topology of Fig. 2, the transfer functions of the links from the source to all η nodes in layer 1 are equal to $\mathbf{G} = \mathbf{I}$. There are $\binom{\eta}{2}$ nodes in layer 2 where each one are correspond two exactly two nodes in layer 1. For each node in layer 2, there are two links from its two corresponding nodes in layer 1 with transfer functions equal to $\mathbf{I}$ and $-\mathbf{I}$. Clearly, if the corresponding two nodes in layer 1 have the same output then the node in layer 2 will receive $\mathbf{0}$ as superposition of received signals; i.e. it will not be able to decode the multicast messages. If we consider a network where $\eta > (p^n)^{T^2}$ and we assume that the

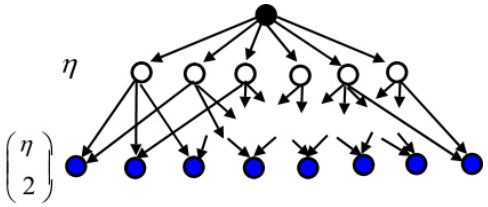

Fig. 2. An example for any finite T, there exists a network which needs $T' > T$ symbols of as a time-frame to achieve cutset bound

linear mapping of the nodes ($\mathbf{F}$) is a $T \times T$ matrix, then the linear mapping of at least two nodes in layer two will be the same and they will have equal outputs, hence, their corresponding node in layer 2 will receive zero always.

A. Rotational Coding Scheme for Layered Networks

We define *key* $f_{i,j,k}$ in layered networks to be an integer associated with the j^{th} element of the received vector and i^{th} element of the transmitted vector by node Φ_k. Thus each node has q^2 keys. The coding at node Φ_k is performed as follows. Let the rotation of a vector, say $\tilde{\mathbf{x}}^k_{m,j} = [\alpha_1, \alpha_2, \ldots, \alpha_T]^\dagger$, by the integer value s be defined as $\tilde{\mathbf{x}}^k_j\{s\} = [\alpha_{s+1}, \alpha_{s+2}, \ldots, \alpha_T, \alpha_1, \ldots, \alpha_s]^\dagger$. Using the set of keys $f_{i,j,k}$, the vector $\tilde{\mathbf{y}}^k_{(m+1),i}$ is obtained as:

$$\tilde{\mathbf{y}}^k_{(m+1),i} = \sum_{j=1}^{q} \tilde{\mathbf{x}}^k_{m,j}\{f_{i,j,k}\} \tag{8}$$

The linear operation matrix $\mathbf{F}_k$ of node Φ_k can be obtained from (8) as

$$\begin{bmatrix} \tilde{\mathbf{y}}^k_{(m+1),1} \\ \tilde{\mathbf{y}}^k_{(m+1),2} \\ \vdots \\ \tilde{\mathbf{y}}^k_{(m+1),q} \end{bmatrix} = \begin{bmatrix} \mathbf{I}\{f_{1,1,k}\} & \ldots & \mathbf{I}\{f_{1,q,k}\} \\ \mathbf{I}\{f_{2,1,k}\} & \ldots & \mathbf{I}\{f_{2,q,k}\} \\ \vdots & & \vdots \\ \mathbf{I}\{f_{q,1,k}\} & \ldots & \mathbf{I}\{f_{q,q,k}\} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{x}}^k_{m,1} \\ \tilde{\mathbf{x}}^k_{m,2} \\ \vdots \\ \tilde{\mathbf{x}}^k_{m,q} \end{bmatrix}$$

where the rotation for the matrices $\mathbf{I}\{f_{i,j,k}\}$ are defined similar to the rotation for a column vector by rotating the rows of the matrix.

One proper choice of the keys is as follows. Enumerate all possible choices of the keys and assign each pair (i,j,k) a unique number $e_{i,j,k}$ between 1 and $q^2 \times (number\ of\ nodes\ in\ the\ network)$. Define $f_{i,j,k} = (q+1)^{e_{i,j,k}}$. It can be verified that for any two disjoint subset of $(i,j,k) \in \Xi_1$ and $(i,j,k) \in \Xi_2$, we have

$$\sum_{(i,j,k)\in\Xi_1} \alpha_{i,j,k} \cdot f_{i,j,k} \neq \sum_{(i,j,k)\in\Xi_2} \alpha_{i,j,k} \cdot f_{i,j,k} \tag{9}$$

where $\alpha(.,.,.) \in \{1, 2, \ldots, q\}$. We choose T very large compare to the values of $f_{i,j,k}$. When the keys are selected as above, then the linear operation matrix $\mathbf{F}_k$ will have some important properties that will be described in the following lemmas.

Lemma 1: Let $\mathbf{E}$ *be a* $rT \times rT$, $r \leq q$ *submatrix of* $\mathbf{F}_k$ *by selecting any arbitrary* r *block columns and* r *block rows from the block matrix* $\mathbf{F}_k$*. Then, the rank of* $\mathbf{E}$ *is* $rT - o(T)$*, i.e., the rank of* $\mathbf{E}$ *is "almost"* rT *as* $T \longrightarrow \infty$*.*

Proof of Lemma 1: For each matrix $\mathbf{I}\{f_{i,j,k}\}$ exactly $f_{i,j,k}$ rows wrap around. Thus, at most $r \cdot \max_{i,j,k} f_{i,j,k}$ columns of $\mathbf{E}$ contains the wrap around rows. It can be directly shown that all other rows, i.e., $rT - r \cdot \max_{i,j,k} f_{i,j,k}$ rows, are linearly independent which means that the rank is at least $rT - o(T)$.

Let us throw away $M = \max_{i,j,k} f_{i,j,k}$ last rows of each block. Thus each $T \times T$ block of the matrix $\mathbf{E}$ is now converted to a block of size $(T-M) \times T$ and the new matrix $\tilde{\mathbf{E}}$ is now of the size $(T-M)r \times Tr$. Please note that any row of $\tilde{\mathbf{E}}$ has a nonzero element in each of the r sub-blocks which contains this row. However, this property does not hold for any column. We say a column *coincide* with a row if their common element of the matrix is nonzero.

We prove that all the rows of $\tilde{\mathbf{E}}$ are linearly independent in vectors space $\mathbb{F}(p^n)^{rT}$. Suppose not, thus, there exist some row indices $i \in \mathcal{I}$ such that

$$\sum_{i\in\mathcal{I}} \alpha_i \mathbf{v}_i = 0. \tag{10}$$

where $\alpha_i \in \mathbb{F}(p^n) \setminus \{0\}$ and $\mathbf{v}_i$'s are row vectors of $\tilde{\mathbf{E}}$.

We look at the block column j, $j = 1, 2, \ldots, r$ of $\tilde{\mathbf{E}}$ where each consists of T columns. From the T columns in each block column j, we pick the column $\mathbf{w}_j$ with the largest index (i.e., the column which is to the right of any other column) coincide with one of $\mathbf{v}_i, i \in \mathcal{I}$. We note that such selection is always possible, because, there is at least one row $\mathbf{v}_i$ in (10) and it coincides with at least one column in each block column j. We also consider the set of all rows $\mathbf{v}_i$ which coincide with one of $\mathbf{w}_j$, $j = 1, \ldots, r$ and we denote the set of indices of such rows $\tilde{\mathcal{I}}$, $\tilde{\mathcal{I}} \subseteq \mathcal{I}$.

Now, we note that $|\tilde{\mathcal{I}}| \leq r$, because there is at most one row in each block row j, $j = 1, 2, \ldots, r$. If there are more than one row in, say block row j, then they both coincide with two different columns in one block column which is in contradiction with the construction in which only one column with the largest index is chosen within each block column.

We construct a graph where each node corresponds to a row $\mathbf{v}_i$, $i \in \tilde{\mathcal{I}}$, and two nodes are connected by an edge if there exists a column $\mathbf{w}_j$, $j = 1, 2, \ldots, r$ which coincides with both nodes. We note that each column $\mathbf{w}_j$ have at least two nonzero elements which means the corresponding edge connects a pair of the nodes. Also, we argue that two different columns $\mathbf{w}_j$ and $\mathbf{w}_{j'}$ connect two different pairs of nodes. Suppose not. Thus, there exist two column $\mathbf{w}_j$ and $\mathbf{w}_{j'}$ such that they both connect one pair of nodes corresponding to the rows $\mathbf{v}_i$ and $\mathbf{v}_{i'}$. Then, we have

$$i + f_{i,j} = i' + f_{i',j}, \tag{11}$$

$$i + f_{i,j'} = i' + f_{i',j'} \tag{12}$$

which means $f_{i,j} - f_{i',j} - f_{i,j'} + f_{i',j'} = 0$. This is contradicting the construction of $f_{.,.,.}$. In other words, we have shown that there cannot be a loop of length two in the graph.

Thus, there will be at least r distinct edges between the nodes of the graph. Since $|\tilde{\mathcal{I}}| \leq r$ there exists a loop in the graph, say from $\mathbf{v}_{i_1}$ to $\mathbf{v}_{i_2}$ to ... $\mathbf{v}_{i_e}$, which are connected with the edges corresponding to the columns $\mathbf{w}_{j_1}$, $\mathbf{w}_{j_2}$, ..., $\mathbf{w}_{j_e}$, respectively. We have

$$\begin{aligned} i_1 - i_2 &= f_{i_2,j_1} - f_{i_1,j_1}, \\ i_2 - i_3 &= f_{i_3,j_2} - f_{i_2,j_2}, \\ &\vdots \\ i_{e-1} - i_e &= f_{i_e,j_{e-1}} - f_{i_{e-1},j_{e-1}}, \\ i_e - i_1 &= f_{i_1,j_e} - f_{i_e,j_e}. \end{aligned} \tag{13}$$

Thus, we have $f_{i_1,j_e} + \sum_{\pi=1}^{e-1} f_{i_{\pi+1},j_\pi} = \sum_{\pi=1}^{e} f_{i_\pi,j_\pi}$ which is in contradiction with the construction of $f_{.,.,.}$. Therefore, we conclude that (10) cannot hold. This complete the proof of the lemma. ■

The consequence of Lemma 1 is that by observing a restricted set (only some components of the space) from the output of the node and knowing the space of the input of the node, it is possible to *almost uniquely* determine the input. In Lemma 2 more generally we prove that $\mathbf{F}$ can transfer the maximum amount of information from input space of the node to output links.

Lemma 2: Consider an arbitrary link from node Φ_k to Φ_j. Then,

$$|\mathbf{H}_k^j \mathbf{F}_k \mathbf{S}_k| = \min(T|\mathbf{G}_k^j|, |\mathbf{S}_k|) - o(T) \tag{14}$$

Here, $|.|$ denotes the matrix rank.

Proof of Lemma 2: To ease the proof we write $\mathbf{H}_k^j$ and $\mathbf{S}_k$ in new forms: $\mathbf{H}_k^j = \mathbf{C}_1\mathbf{D}_1$ where $\mathbf{C}_1$ is a full rank matrix and $\mathbf{D}_1$ is a reduced row echelon matrix (in every column there is at most one non-zero element), $\mathbf{S}_k = \mathbf{D}_2\mathbf{C}_2$ where $\mathbf{C}_2$ is a full rank matrix and $\mathbf{D}_2$ is reduced column echelon matrix (in every row there is at most one non-zero element).

We note that $\mathbf{H}_k^j$ contains q^2 blocks where the block rs is equal to $g_{rs}\mathbf{I}$ that g_{rs} is rs component in $\mathbf{G}_k^j$. Consider only the rows corresponding to the first element of each block and perform elementary row operation (this is same as doing row operation on matrix $\mathbf{G}_k^j$), we can build a new matrix (without losing the rank) which has at most one non-zero element on the columns of the first element of the blocks. If we repeat this operation for all corresponding elements of the blocks, we will obtain a matrix of blocks where in each block column contains at most one block equal to a factor of $\mathbf{I}$ and the rest of its elements are equal to zero. We denote this matrix by $\mathbf{D}_1$, and the full rank matrix which is equivalent to all row operations by $\mathbf{C}_1$). It is straightforward to show that $|\mathbf{H}_k^j| = T|\mathbf{G}_k^j| = |\mathbf{D}_2|$.

Next, we study the format of $\mathbf{S}_k$ matrix based on the structure of matrices $\mathbf{F}_i$ for different nodes and then we compute matrices $\mathbf{D}_2$ and $\mathbf{C}_2$. It can be easily shown using (5) that $\mathbf{S}_k$ has the following form: It has q^2 blocks. Each block is a linear of ρq^l rotation matrices where ρ is the number of different paths from the source to ϕ_k and l is the length of these paths. The size of rotations can be obtained as $\sum_{\pi=0}^{l-1} f(e_\pi, e_{\pi+1}, i_\pi)$ where $\Phi_0 = i_0, i_1, \ldots, i_{l-1}, i_l = \Phi_k$ is a path from the source to the node ϕ_k, and e_π's are chosen arbitrarily from $\{1, \ldots, q\}$.

Based on the construction of the keys $f(.,.,.)$ we can easily show that the rotation matrices will not cancel out each other.

Now we perform elementary column operation on matrix $\mathbf{S}_k$. We consider the column of the first non-zero element on the first row of $\mathbf{S}_k$. We add multiples of this column to all columns which have non-zero elements on the first row to cancel out them. We repeat this operation for row $2, \ldots, T$. Then first block row as a block in the form of $\mathbf{I}\{s_1\}$ and the rest of its elements are zero. We perform the same operations on other block rows, then $\mathbf{S}_k$ will be converted to a matrix which has at most one non-zero block on each block row that is in the form of $\mathbf{I}\{s\}$. We call this matrix $\mathbf{D}_2$, and the matrix correspond to column operations $\mathbf{C}_2$.

Since $\mathbf{C}_1$ and $\mathbf{C}_2$ are full rank matrices, we have $|\mathbf{H}_k^j \mathbf{F}_k \mathbf{S}_k| = |\mathbf{C}_1\mathbf{D}_1\mathbf{F}_k\mathbf{D}_2\mathbf{C}_2| = |\mathbf{D}_1\mathbf{F}_k\mathbf{D}_2|$. We set

$\mathbf{F} = \mathbf{D}_1\mathbf{F}_k\mathbf{D}_2$. Note that $\mathbf{D}_1\mathbf{F}_k$ has exactly $|\mathbf{G}_k^j|$ non-zero block rows, which some of its block row might be a linear combination of different block rows of $\mathbf{F}_k$. Now, when it is multiplied by $\mathbf{D}_2$, there will be $|\mathbf{D}_2|/T$ non-zero block columns in it. Also each block row of $\mathbf{D}_1\mathbf{F}_k$ will rotate equal to the rotation size of its corresponding non-zero block in matrix $\mathbf{D}_2$, some blocks might be the sum of rotations of different block columns of $\mathbf{D}_1\mathbf{F}_k$.

If we through out the zero block rows and zero block columns of $\mathbf{F}$, a $|\mathbf{D}_1| \times |\mathbf{D}_2|$ block matrix is obtained which we denote it by $\hat{\mathbf{F}}$. Next, we prove that every square submatrix of $\hat{\mathbf{F}}$ is (almost) full rank, therefore we have

$$|\mathbf{H}_k^j\mathbf{F}_k\mathbf{S}_k| = |\mathbf{F}| = |\hat{\mathbf{F}}| = \min(|\mathbf{D}_1|, |\mathbf{D}_2|) - o(T) \tag{15}$$

For proving the above statement, we choose a square submtarix $\mathbf{E}$ of $\hat{\mathbf{F}}$. Note that the shift size of different rotations inside $\mathbf{E}$ is equal to $f(r,s,k) + \sum_{\pi=1}^{l} f(e_\pi, e_{\pi+1}, i_\pi)$. If we follow the proof of Lemma 1 and use (9) we can show that almost all rows of $\mathbf{E}$ are independent. ■

B. Fundamental Theorem of Rank Inheritance in Layered Networks

Here we explain some important theorems on the maximum rate of information which can be sent from a source to any set of nodes in a layered network.

We start from a simple case when the network has only two layers, where the first layer is the source node and the second layer contains the rest of the nodes. This corresponds to a single hop broadcast network. Theorem 1 provides a fundamental bound on the dimension of the space seen by any set of nodes in this network.

From the following proof, it can be seen that a unique coding scheme can simultaneously achieve the rank in (16) for any arbitrary subset of the nodes. This is the reason that we use the terminology of the inherited rank for the nodes in the second layer from the rank of the nodes in layer one. Obviously, the inherited rank depends on the network topology, i.e., the channel matrices, and also the coding scheme. Nonetheless, rotational coding scheme achieves the maximum possible rank for any arbitrary network topology.

Here, we prove the lemma using rotational codes which create the codes independent of the transfer functions (i.e., forward channel knowledge) $\mathbf{G}_1^k$. As a consequence, if a node Φ_1 broadcasts to multiple destinations, the inherited rank for each destination follows (16), simultaneously.

Theorem 1 (Broadcast Theorem): Assume a layered network. Consider a single node Φ_i (i.e., $\mathcal{A} = \{i\}$) broadcasting information to a set of nodes $\{\Phi_k\}_{k\in\mathcal{B}}, \mathcal{B} = \{j_1, \ldots, j_b\}$. The maximum dimension of the decodable space of $\mathcal{B}$ is

$$|V_\mathcal{B}| = \min(|V_i|, |\mathbf{H}_\mathcal{A}^\mathcal{B}|) \tag{16}$$

and this can be (almost) achieved using rotational coding scheme.

Proof of Theorem 1: It is clear that $|V_i|$ is an upper bound for $|V_\mathcal{B}|$, because the rank the dimension of output space of a linear mapping cannot be larger than the dimension of input space. On the other hand, $|\mathbf{H}_\mathcal{A}^\mathcal{B}|$ represents the rank of output space when the whole vector space $\mathcal{F}(p^n)^q$ is the input, i.e. $|\mathbf{H}_\mathcal{A}^\mathcal{B}|$ is an upper bound on $|V_\mathcal{B}|$ as well.

We can take out an square $q \times q$ submatrix of $\mathbf{G}$ from $\mathbf{G}_\mathcal{A}^\mathcal{B}$ such that $|\mathbf{G}| = |\mathbf{G}_\mathcal{A}^\mathcal{B}|$. We construct the corresponding matrix $\mathbf{H} = \mathbf{G} \otimes \mathbf{I}$, thus $|\mathbf{H}| = |\mathbf{H}_\mathcal{A}^\mathcal{B}|$. Based on the result of Lemma 2, if we use rotational coding at $\mathcal{A}$ then $|\mathbf{HFS}_i| = \min(|\mathbf{S}_i|, |\mathbf{H}|) - o(T)$, where here $|\mathbf{S}_i| = |V_i|$ and $|\mathbf{H}| \geq |V_\mathcal{B}|$. These inequalities will conclude the proof. ■

Theorem 2 gives a fundamental bound the dimension of decodable space which a set of nodes in layer $l+1$ can obtain from another set of nodes in layer l. This is a generalization of Theorem 1 which the first layer has only one node.

For example, consider two nodes Φ_1 and Φ_2 in layer one and two nodes Φ_3 and Φ_4 in layer two. We have

$$|V_3| = \min(|V_{1,2}|, |V_1| + |\mathbf{H}_2^3|, |V_2| + |\mathbf{H}_1^3|, |\mathbf{H}_{1,2}^3|), \tag{17}$$

$$|V_4| = \min(|V_{1,2}|, |V_1| + |\mathbf{H}_2^4|, |V_2| + |\mathbf{H}_1^4|, |\mathbf{H}_{1,2}^4|), \tag{18}$$

$$|V_{3,4}| = \min(|V_{1,2}|, |V_1| + |\mathbf{H}_2^{3,4}|, |V_2| + |\mathbf{H}_1^{3,4}|, |\mathbf{H}_{1,2}^{3,4}|). \tag{19}$$

Theorem 2 (Rank Inheritance Theorem): Assume a layered network. Consider a subset of transmitting nodes $\{\Phi_k\}_{k\in\mathcal{A}}, \mathcal{A} = \{i_1, \ldots, i_a\}$ sending information to the subset $\{\Phi_k\}_{k\in\mathcal{B}}, \mathcal{B} = \{j_1, \ldots, j_b\}$ in the receiving nodes. The maximum rank that space $V_\mathcal{B}$ inherits from the space $V_\mathcal{A}$ is

$$|V_\mathcal{B}| = \min_{\Theta\subseteq\mathcal{A}} (|V_\Theta| + |\mathbf{H}_{\mathcal{A}\setminus\Theta}^\mathcal{B}|) \tag{20}$$

and this can be (almost) achieved using rotational coding scheme.

Proof of Theorem 2: The main idea of the proof is as follows. We carefully partition the nodes with the index in $\mathcal{A}$ into two subset, say $\mathcal{C}$ and $\mathcal{A} \setminus \mathcal{C}$ which satisfies certain conditions. Next for each node in either partition, we redefine the space of the input signals that are going to be decoded. Also, based on the above partitioning, we determine a subspace of the space of the signals received at $\mathcal{B}$ which is used for decoding of any subset of the nodes in either of the partitions. The partitioning is chosen such that the condition of the Theorem holds for each partition individually with these redefinitions. Eventually, the inductive technique boils down to prove the case $|\mathcal{A}| = 1$ where no further partitioning is possible, which is proved in Theorem 1.

Now assume that the theorem is proved up to $|\mathcal{A}| = \eta$ and we want to prove it for $|\mathcal{A}| = \eta + 1$. Let $\mathcal{C}$ be the optimum value of Θ in (20). For all $\Omega \supseteq \mathcal{C}$ we have

$$\begin{aligned} |V_{\mathcal{C}}| + |\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\mathcal{C}}| &\leq |V_\Omega| + |\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\Omega}| && (21)\\ &\leq |V_{\mathcal{C}}| + |V_\Omega \setminus V_{\mathcal{C}}| + |\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\Omega}| && (22)\end{aligned}$$

Thus,

$$\begin{aligned} |\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\mathcal{C}}| &\leq |V_\Omega \setminus V_{\mathcal{C}}| + |\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\Omega}| && (23)\\ &\leq |V_{\Omega\setminus\mathcal{C}} \setminus V_{\mathcal{C}}| + |\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\Omega}| && (24)\\ &\leq |V_{\Omega\setminus\mathcal{C}} \setminus V_{\mathcal{C}}| + |\mathbf{H}^{\mathcal{B}}_{(\mathcal{A}\setminus\mathcal{C})\setminus(\Omega\setminus\mathcal{C})}|. && (25)\end{aligned}$$

Similarly, for all $\Omega \subseteq \mathcal{C}$ we have

$$\begin{aligned} |V_{\mathcal{C}}| + |\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\mathcal{C}}| &\leq |V_\Omega| + |\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\Omega}| && (26)\\ &\leq |V_\Omega| + |\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\mathcal{C}}| + |\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\Omega} \setminus \mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\mathcal{C}}| && (27)\end{aligned}$$

Thus,

$$\begin{aligned} |V_{\mathcal{C}}| &\leq |V_\Omega| + |\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\Omega} \setminus \mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\mathcal{C}}| && (28)\\ &\leq |V_\Omega| + |\mathbf{H}^{\mathcal{B}}_{(\mathcal{A}\setminus\Omega)\setminus(\mathcal{A}\setminus\mathcal{C})} \setminus \mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\mathcal{C}}| && (29)\\ &\leq |V_\Omega| + |\mathbf{H}^{\mathcal{B}}_{\mathcal{C}\setminus\Omega} \setminus \mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\mathcal{C}}|. && (30)\end{aligned}$$

We note that once $\mathcal{C}$ is given, the vector spaces $V_{\mathcal{C}}$ and $\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\mathcal{C}}$ are fixed and well-defined. Thus, corresponding to every subset of $\mathcal{A} \setminus \mathcal{C}$, we can consider a sub-space of the vector space $V_{\mathcal{C}}$ defined as $V_\Omega \setminus V_{\mathcal{C}}$. Also, corresponding to each subset of $\mathcal{C}$, we can consider a subspace of the input of the nodes in $\mathcal{B}$, i.e., span($\mathbf{H}^{\mathcal{B}}_\Omega$), defined as span($\mathbf{H}^{\mathcal{B}}_\Omega \setminus \mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\mathcal{C}}$), where span($\mathbf{H}$) denotes the span of the column vectors of $\mathbf{H}$.

Now, let us examine (25) more closely. Consider only the subset $\mathcal{C}' = \mathcal{A} \setminus \mathcal{C}$ of the transmitting nodes and note that $\Omega' = \Omega \setminus \mathcal{C}$ is an arbitrary subset of $\mathcal{C}'$ under the condition of (25), i.e., $\Omega \supseteq \mathcal{C}$. We note that for any subset Ω' of $\mathcal{C}'$, the rank of the corresponding input space $V_{\Omega\setminus\mathcal{C}} \setminus V_{\mathcal{C}}$ and the rank of the corresponding transfer function $\mathbf{H}^{\mathcal{B}}_{(\mathcal{A}\setminus\mathcal{C})\setminus(\Omega\setminus\mathcal{C})}$ satisfy (25) which means that we can use the assumption of the induction over this partition, i.e., $\mathcal{A} \setminus \mathcal{C}$, to decode for the spaces $V_{\Omega\setminus\mathcal{C}} \setminus V_{\mathcal{C}}$ by looking at the projection of the received vectors of nodes in $\mathcal{B}$ in the space $\mathbf{H}^{\mathcal{B}}_{(\mathcal{A}\setminus\mathcal{C})\setminus(\Omega\setminus\mathcal{C})}$.

Now, we are left with an orthogonal space of the input vectors, namely, $V_{\mathcal{C}}$ and an orthogonal space of the received vectors of the nodes in $\mathcal{B}$, i.e., $\mathbf{H}^{\mathcal{B}}_{\mathcal{C}\setminus\Omega} \setminus \mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\mathcal{C}}$. A careful review of (30) also reveals that for all $\Omega \subseteq \mathcal{C}$ a similar property holds which means we can use the assumption of the induction on this subset as well.

Assuming that the optimal partitioning solution $\Theta = \mathcal{C}$ for (20) is not one of the two special cases, namely $\mathcal{C} = \varnothing$ and $\mathcal{C} = \mathcal{A}$, we have $1 \leq |\mathcal{C}| \leq \eta$ and $1 \leq |\mathcal{A} \setminus \mathcal{C}| \leq \eta$. Thus, by using the assumption of the induction the rank $|V_{\mathcal{C}}|$ is achievable for the set $\mathcal{C}$ and the rank $|\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\mathcal{C}}|$ is achievable for the set $\mathcal{A} \setminus \mathcal{C}$. Besides, the decoded spaces belong to two independent vector space, therefore, the total rank $|V_{\mathcal{C}}| + |\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\mathcal{C}}|$ is achievable.

Now, we treat the first special case where we assume that the optimum Θ in (20), i.e., $\mathcal{C}$, is the empty set. Thus, for all Ω, we have

$$|\mathbf{H}^{\mathcal{B}}_{\mathcal{A}}| \leq |V_\Omega| + |\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\Omega}| \quad (31)$$

Put aside the case $\Theta = \varnothing$ and find the next minimum among the remaining cases of Θ, i.e., $\varnothing \subset \Theta \subseteq \mathcal{A}$. Denote this second minimum by $\mathcal{D}$ and assume $\mathcal{D} \neq \mathcal{A}$. Therefore, the steps to derive (25) hold here for $\mathcal{D}$ instead of $\mathcal{C}$. Similarly, the steps to derive (30) hold for all $\Omega \subseteq \mathcal{D}$, except for $\Omega = \emptyset$ for which the inequality is reversed and we have (31) computed at $\Omega = \mathcal{D}$. Thus

$$|\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\mathcal{D}}| \leq |V_{\Omega\setminus\mathcal{D}} \setminus V_{\mathcal{D}}| + |\mathbf{H}^{\mathcal{B}}_{(\mathcal{A}\setminus\mathcal{D})\setminus(\Omega\setminus\mathcal{D})}|, \ \Omega \supseteq \mathcal{D} \quad (32)$$

$$|V_{\mathcal{D}}| \leq |V_\Omega| + |\mathbf{H}^{\mathcal{B}}_{\mathcal{D}\setminus\Omega} \setminus \mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\mathcal{D}}|, \quad \varnothing \subset \Omega \subseteq \mathcal{D} \quad (33)$$

$$|\mathbf{H}^{\mathcal{B}}_{\mathcal{A}}| \leq |V_{\mathcal{D}}| + |\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\mathcal{D}}|. \quad (34)$$

From (34), we have $|\mathbf{H}^{\mathcal{B}}_{\mathcal{D}} \setminus \mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\mathcal{D}}| = |\mathbf{H}^{\mathcal{B}}_{\mathcal{A}} \setminus \mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\mathcal{D}}| = |\mathbf{H}^{\mathcal{B}}_{\mathcal{A}}| - |\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\mathcal{D}}| < |V_{\mathcal{D}}|$ and combining with (33) we have

$$|\mathbf{H}^{\mathcal{B}}_{\mathcal{D}} \setminus \mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\mathcal{D}}| \leq |V_\Omega| + |\mathbf{H}^{\mathcal{B}}_{\mathcal{D}\setminus\Omega} \setminus \mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\mathcal{D}}|, \ \varnothing \subset \Omega \subseteq \mathcal{D} \quad (35)$$

Now the induction can be performed using (32) and (35) to prove that the ranks $|\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\mathcal{D}}|$ and $|\mathbf{H}^{\mathcal{B}}_{\mathcal{D}} \setminus \mathbf{H}^{\mathcal{B}}_{\mathcal{A}\setminus\mathcal{D}}|$

are achievable using two subset of nodes $\mathcal{A}\backslash\mathcal{D}$ and $\mathcal{D}$ in layer one and decoding for the vectors in the subspaces $V_{\mathcal{A}\backslash\mathcal{D}}\backslash V_{\mathcal{D}}$ and $V_{\mathcal{D}}$ at nodes $\mathcal{B}$ in layer two by projecting the received signal at $\mathcal{B}$ into two subspace $\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\backslash\mathcal{D}}$ and $\mathbf{H}^{\mathcal{B}}_{\mathcal{D}}\setminus\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\backslash\mathcal{D}}$, respectively. Thus, the rank of space seen at $\mathcal{B}$ is equal to $|\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\backslash\mathcal{D}}|+|\mathbf{H}^{\mathcal{B}}_{\mathcal{D}}\setminus\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\backslash\mathcal{D}}|=|\mathbf{H}^{\mathcal{B}}_{\mathcal{A}}|$. Please note that if the solution Θ for second minimum, i.e., $\mathcal{D}$ is equal to $\mathcal{A}$, then we have to find the third minimum solution, say, $\mathcal{E}$. Using the same line of argument it can be shown that (32) and (35) are still valid where $\mathcal{D}$ is replaced with $\mathcal{E}$.

Thus, we can use the same argument for the partitioning $\Theta=\mathcal{D}$ since $1\leq|\mathcal{D}|\leq\eta$ and $1\leq|\mathcal{A}\backslash\mathcal{D}|\leq\eta$. Thus, by using the assumption of the induction the rank $|\mathbf{H}^{\mathcal{B}}_{\mathcal{D}}\setminus\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\backslash\mathcal{D}}|$ is achievable for the set $\mathcal{D}$ and the rank $|\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\backslash\mathcal{D}}|$ is achievable for the set $\mathcal{A}\setminus\mathcal{D}$. Besides, the decoded spaces $\mathbf{H}^{\mathcal{B}}_{\mathcal{D}}\setminus\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\backslash\mathcal{D}}$ and $\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\backslash\mathcal{D}}$ are orthogonal vector spaces, therefore, the total rank $|\mathbf{H}^{\mathcal{B}}_{\mathcal{D}}\setminus\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\backslash\mathcal{D}}|+|\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\backslash\mathcal{D}}|=|\mathbf{H}^{\mathcal{B}}_{\mathcal{A}}|$ is achievable.

The other special case is when the solution for the optimum Θ in (20), i.e., $\mathcal{C}$, is the set of all the nodes, i.e., $\mathcal{C}=\mathcal{A}$. Thus, for all Ω, we have

$$|V_{\mathcal{A}}| \quad\leq\quad |V_{\Omega}|+|\mathbf{H}^{\mathcal{B}}_{\mathcal{A}\backslash\Omega}| \tag{36}$$

The proof for this case is very similar to the case where $\mathcal{C}=\varnothing$ and the partition is found by considering a secondary (auxiliary) partition by removing the solution $\Theta=\mathcal{A}$ from the space of possible Θ. Also, if the secondary partition corresponds to $\Theta=\varnothing$, we seek the tertiary partition as before. This completes the proof of the theorem. ■

We denote the set of cutsets between the source and a terminal node k_i by $\Lambda_{k_i}=\{\Omega:\Omega\subset\{0,1,\ldots,N\},0\in\Omega,N\in\Omega^c\}$ and define minimum cutset rank between the source and the destination as

$$r_{\min}=\min_{k_i\in\text{terminals}}\ \min_{\Omega\in\Lambda_{k_i}}\{|\mathbf{G}^{\Omega^c}_{\Omega}|\} \tag{37}$$

If the multicast session contains more than one terminal, then we define $r_{\min}$ as the minimum value among the minimum cutset ranks of different terminals.

Theorem 3 (Achievability Theorem): Assume a single multicast session in a layered wireless network and define $r_{\min}$ as (37). The rate of $R=n\log_2(p)r_{\min}(1-o(T)/T)$ can be achieved using rotational coding scheme. Clearly, $R\simeq r_{\min}$ when $T\to\infty$.

Proof of Theorem 3: Assume that $\mathcal{D}=\{k_1,\ldots,k_d\}$ is the set of terminals. We employ rotational coding scheme in the network. Let $\mathbf{U}$ be a $qT\times qT$ matrix where the first $r_{\min}T$ components of its diagonal are 1 and rest of its components are 0. Also, let $V_0=\{\mathbf{X}\in\mathbb{F}(p^n)^{qT}:\mathbf{X}=\mathbf{U}\mathbf{X}\}$.

Consider an arbitrary terminal k_j. Using a similar method to proof of Lemma 2 we can show that $|\mathbf{S}_{k_j}\mathbf{U}|=r_{\min}T-o(T)$. That means almost all elements of V_0 are in deacodable space of k_j. Considering all terminals, we conclude that there is a subspace of $\overline{V}_0\subseteq V_0$ with rank of at least $r_{\min}T-o(dT)$ which is in deacodable space of all terminals. By sending the elements of $\overline{V}_0$ from the source we can achieve the rate of $R=\log_2(p^n)(r_{\min}T-o(T))/T$ in all terminals. ■

IV. Conclusion

In this paper we introduced rotational coding scheme for wireless networks under "deterministic channel model". We proved that this coding scheme can achieve the minimum cutset bound for a single multicast session in layered wireless networks. It has much lower computational complexity compared to the existing coding schemes, because, it only uses addition and rotation operations for encoding the information at each node. Moreover, we presented a fundamental theorem on the rate of information that can be sent through any cutset. This result is applicable for analyzing the throughput of linear network codes.

Acknowledgment

A. Keshavarz-Haddad would like to thank National Foundation of Elites of Iran for partial financial support.

References

[1] S. Avestimehr, S. Diggavi, and D. Tse, "A deterministic approach to wireless relay networks," in *ISIT*, 2007.

[2] C. Fragouli and E. Soljanin, "Network coding applications," *Found. Trends Netw.*, vol. 2, no. 2, pp. 135–269, 2007.

[3] R. Ahlswede, N. Cai, S. Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204–1216, 2000.

[4] S. Y. R. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Transactions on Information Theory*, vol. 49, no. 2, pp. 371–381, 2003.

[5] R. Koetter and M. Medard, "An algebraic approach to network coding," *IEEE/ACM Transactions on Network*, vol. 11, no. 5, pp. 782–795, 2003.

[6] C. Fragouli and E. Soljanin, "Network coding fundamentals," *Found. Trends Netw.*, vol. 2, no. 1, pp. 1–133, 2007.

[7] S. Avestimehr, S. Diggavi, and D. Tse, "Wireless network information flow," in *Allerton*, 2007.

[8] M. A. Khojastepour and A. Keshavarz-Haddad, "On capacity of deterministic relay networks," in *Allerton*, 2008.