Noncoherent Multisource Network Coding

Mahdi Jafari Siavoshani
EPFL
Email: mahdi.jafarisiavoshani@epfl.ch

Christina Fragouli
EPFL
Email: christina.fragouli@epfl.ch

Suhas Diggavi
EPFL
Email: suhas.diggavi@epfl.ch

***Abstract*— We examine the problem of multiple sources transmitting information to one or more receivers that require the information from all the sources, over a network where the network nodes perform randomized network coding. We consider the noncoherent case, where neither the sources nor the receivers have any knowledge of the intermediate nodes operations.**

We formulate a model for this problem, inspired from block-fading noncoherent MIMO communications. We prove, using information theoretic tools, that coding over subspaces is sufficient to achieve the capacity, and give bounds for the capacity. We then examine the associated combinatorial problem of code design. We extend the work by Koetter and Kschischang [3] to code constructions for the multisource case. Our constructions can also be viewed as coding for the noncoherent multiple-access finite-field channel.

I. Introduction

In network operation with network coding, a promising technique has nodes randomly combine their incoming packets over a finite field [1]. This approach is well suited to dynamically changing or large scale networks as it avoids the need for global synchronization. We consider a scenario where multiple sources transmit independent information over such a network, towards a single receiver, or towards multiple receivers all of which request all the information from the sources.

We are interested in noncoherent communication, where neither the sources nor the receivers have any knowledge of the network topology or the network nodes operations. In large dynamically changing networks, collecting network information comes at a cost, as it consumes bandwidth that could instead have been used for information transfer. Noncoherent communication allows for creating end-to-end systems completely oblivious to the network state.

In this paper, we assume that time is slotted, and at each time slot the sources insert in total m source packets X while a receiver observes n packets Y. The packets Y are related to X through a linear transformation, called transfer function, that is unknown to both the sources and the receiver and changes from timeslot to timeslot. This description bears close similarities to block-fading noncoherent MIMO communication [2]. Inspired by this, we formulate a model for noncoherent communication over networks employing random network coding operations, and examine it using both information theoretic and combinatorial code design tools.

Apart from the model formulation, our contributions include the following. We first prove that coding using subspaces as codewords is optimal from an information theory point of view in the sense that such strategies are sufficient to achieve the capacity of our channel. We then establish achievable rates for some coding strategies. For the point to point channel we find that for large fields and large coherence time, the optimal input distribution is uniform over subspaces of certain dimension.

We then proceed to the associated combinatorial problem of code design. Combinatorial codes using subspaces have been recently proposed for the single source case by Koetter and Kshischang [3], who use subspaces to design elegant error and erasure correction schemes. We extend their work to the multiple source case where a subset of the sources is active. In our case, we need to select codebooks that allow us to convey both the information messages as well as the identity of the source transmitting each information message.

The paper is organized as follows. Section II describes our model. Section III presents information theoretic results. Section IV focuses on algebraic code constructions, and finally Section V concludes the paper.

II. The noncoherent finite field channel model

Consider a network where nodes perform uniform at random network coding over a finite field $\mathbb{F}_q$. We discuss first the case of a single source and a single receiver. We assume slotted time, and a "block" time-varying channel. At timeslot l, the receiver observes

$$Y(l) = G(l)X(l), \tag{1}$$

where $X(l)$ is an $m \times T$, $G(l)$ is an $n \times m$ and $Y(l)$ is an $n \times T$ matrix defined over the finite field $\mathbb{F}_q$ (in the rest of the paper we will omit for convenience the index l). That is, at each time slot, the receiver receives n packets of length T, that depend on a set of m packets of length T sent by the source. The source packets are independent from time-slot to time-slot. This block operation of the channel, where the received packets Y depend on a different set of sent packets X, is exactly like the standard network coding model in [1].

The block length T can be interpreted as the coherence time of the channel, during which the transfer matrix G remains constant. T is finite and fixed. If T were arbitrarily large, we could send a set of "training symbols" of finite length for the receiver to learn G and then communicate using perfect channel knowledge. In fact, this is the prevailing approach for implementing network coding in practice [5]. Information packets are divided in what is called generations, the number m of source packets corresponds to the generation size, and

This work was supported by the Swiss National Science Foundation award 200021-103836/1 and by the EU project N-CRAVE, FP7 ICT-2007-215252.

the training symbols are the coding vectors appended to the information packets. As was observed in [3], this approach leads to a rate loss that becomes pronounced as T decreases.

In our model, the transfer matrix G changes independently from timeslot to timeslot, according to the uniform at random linear combining performed by the network intermediate nodes. Although in general matrix G has some structure related to the topology of the network (see for example [6]), we will here assume that the entries of G are selected according to the uniform distribution. We argue that this is a reasonable choice as a starting point, especially for large scale dynamically changing networks, because: (i) in large networks with high probability all the elements of matrix G will be random variables (no constant elements), and (ii) the network topology changes introduce additional randomness in the matrix structure. The model given in (1) along with the modeling for G given above is clearly information stable and hence the capacity is given by

$$C = \sup_{p(x)} \frac{1}{T} I(X;Y),$$

where $p(x)$ is the input distribution. For a coding strategy that induces an input distribution $p(x)$, the achievable rate is

$$R = \frac{1}{T} I(X;Y).$$

The generalization of this model to multiple receivers is straightforward. We thus next consider the case of multiple sources, and the multiple access channel corresponding to (1). This can be expressed as

$$Y(l) = \sum_{u=1}^{N} G_u(l) X_u(l) = G_{MAC}(l) X_{MAC}(l), \qquad (2)$$

where we have N sources, each source u inserting m_u packets in the network. Thus $X_u(l)$ is an $m_u \times T$, $G_u(l)$ is an $n \times m_u$ and $Y(l)$ is an $n \times T$ matrix over $\mathbb{F}_q$. We can also collect all $G_u(l)$ in the $n \times \sum_{u=1}^{N} m_u$ matrix G_{MAC} and all $X_u(l)$ in the $\sum_{u=1}^{N} m_u \times T$ matrix $X_{MAC}(l)$. Each source u then controls m_u rows of the matrix $X_{MAC}(l)$.

Our models (1) and (2) can easily be extended to include noise. For example, introducing erasures in (1) can be modeled by randomly removing rows of matrix Y, or removing rows of matrix X. These operations correspond to making all zero a row or a column of matrix G. Additive noise can be introduced through a matrix $Z(l)$

$$Y(l) = G(l)X(l) + Z(l), \qquad (3)$$

that follows a given distribution. Constraining the rank of matrix $Z(l)$, for example to be smaller or equal to k, corresponds to the error constraints of the channel model in [3]. In the rest of this paper we will focus our attention to the noiseless case, given in (1) and (2).

III. Information Theoretic Analysis

We start from the case of a single source, and then consider the multiple access case. We assume coding over an arbitrarily large number of timeslots, i.e., instantiations of (1) and (2).

A. Single Source

We first calculate the mutual information between X and Y in (1). We will use $\langle X \rangle$ to denote the subspace spanned by the rows of the matrix X, and $\dim(\pi)$ to denote the dimension of a subspace π. The number of distinct d dimensional subspaces of a T-dimensional space $\mathbb{F}_q^T$ is equal to

$$\left[\begin{array}{c} T \\ d \end{array} \right]_q = \frac{(q^T-1)\cdots(q^{T-d+1}-1)}{(q^d-1)\cdots(q-1)}, \qquad (4)$$

where $\left[\begin{array}{c} T \\ d \end{array} \right]_q$ is called the Gaussian number.

Recall that the distribution of each entry of G is uniform and i.i.d and also G is i.i.d over different blocks. Thus, every row of G is independent of other rows, and conditioned on a sent matrix $X = x_0$, the rows of the received matrix Y are also independent from each other. The independence of rows of Y allows us to write

$$\Pr(Y = y_0 | X = x_0) = \begin{cases} q^{-n\dim(\langle x \rangle)} & \langle y_0 \rangle \subseteq \langle x_0 \rangle, \\ 0 & \text{otherwise.} \end{cases}$$

We can equivalently express this in terms of subspaces as

$$\Pr(Y = y_0 | \langle X \rangle = \pi_x) = \begin{cases} q^{-n\dim(\pi_x)} & \langle y_0 \rangle \subseteq \pi_x, \\ 0 & \text{otherwise.} \end{cases}$$

Thus the conditional entropy can be calculated as

$$H(Y | \langle X \rangle = \pi_x) = n \dim(\pi_x) \log_2 q,$$

and, using the notation $P_X(\pi_x) = \Pr(\langle X \rangle = \pi_x)$, we get

$$H(Y|X) = n \log_2 q \sum_{\pi_x} \dim(\pi_x) P_X(\pi_x).$$

Let $P_Y(y_0) = \Pr(Y = y_0)$. To compute $H(Y) = -\sum_{y_0} P_Y(y_0) \log_2 P_Y(y_0)$ we can use that

$$P_Y(y_0) = \sum_{\pi_x:\ \langle y_0 \rangle \subseteq \pi_x} q^{-n\dim(\pi_x)} P_X(\pi_x).$$

We can then calculate $I(X;Y)$ as

$$I(X;Y) = H(Y) - n \log_2 q \sum_{\pi_x} \dim(\pi_x) P_X(\pi_x).$$

From the above equation, we conclude that the optimal input distribution needs to only optimize the probability distribution of subspaces, as it is only subspace properties that appear in the mutual information. This is very intuitive since, for G unknown both at the transmitter and the receiver, we can only convey the subspace that is occupied by X.

Lower Bound

We now provide a lower bound on the channel capacity that holds for large values of the finite field size q. Consider an input distribution that is uniform over all subspaces of a fixed dimension k, that is,

$$\Pr\left(\langle X \rangle = \pi, \dim(\pi) = r\right) = \begin{cases} \left[\begin{array}{c} T \\ k \end{array} \right]_q^{-1} & r = k, \\ 0 & \text{otherwise,} \end{cases}$$

where k is a number such that $1 \leq k \leq \min(m,T)$. For this distribution, the mutual information $I \triangleq I(X;Y)$ becomes

$$I = -nk\log_2 q - \sum_{d_y=0}^{\min(k,n)} \sum_{\substack{y_0: \\ \dim(\langle y_0\rangle)=d_y}} \psi(y_0)\log_2\psi(y_0),$$

where

$$\begin{aligned}\psi(y_0) &= \left(\sum_{\substack{\pi: \\ \langle y_0\rangle\subseteq\pi,\ \dim(\pi)=k}} q^{-nk} P_X(\pi)\right) \\ &= \left(q^{-nk}\begin{bmatrix} T-d_y \\ k-d_y\end{bmatrix}_q \begin{bmatrix} T \\ k\end{bmatrix}_q^{-1}\right) \triangleq \psi(d_y),\end{aligned}$$

only depends on $d_y = \dim(y_0)$. Thus for the mutual information we can write

$$I = -nk\log_2 q - \sum_{d_y=0}^{\min(k,n)} S_{d_y}\begin{bmatrix} T \\ d_y\end{bmatrix}_q \psi(d_y)\log_2\psi(d_y),$$

where S_{d_y} is the number of different $n\times T$ matrices with rows spanning a specific subspace $\pi \in \mathbb{F}_q^T$ of dimension d_y.

For q very large it holds that $\begin{bmatrix} T \\ d\end{bmatrix}_q = q^{d(T-d)}(1+\mathcal{O}(q^{-1}))$ and $S_d = q^{nd}(1+\mathcal{O}(q^{-1}))$ [7]. Then,

$$\begin{aligned}I =& -nk\log_2 q - \sum_{d_y=0}^{\min(k,n)} (1+\mathcal{O}(q^{-1}))\, q^{-(n-d_y)(k-d_y)} \\ &\times \log_2\left((1+\mathcal{O}(q^{-1}))\, q^{-nk} q^{-d_y(T-k)}\right) \\ =& \left[-nk + q^{-(n-\delta)(k-\delta)}\left[nk+\delta T-\delta k\right] + \mathcal{O}(q^{-1})\right]\log_2 q,\end{aligned}$$

where $\delta = \min(n,k)$. So we have

$$I = \left[(T-k)\times\min(n,k) + \mathcal{O}(q^{-1})\right]\log_2 q,$$

for $1 \leq k \leq \min(m,T)$. It can be easily observed that for fixed values of m, n, and $T > 1$, there exists some q_0 such that for $q > q_0$ the mutual information I will be maximized for $k = \Delta \triangleq \min(m,n,\lfloor T/2\rfloor)$. Thus

$$I_{\max} = \left[\Delta\,(T-\Delta) + \mathcal{O}(q^{-1})\right]\log_2 q,$$

and we have the following lower bound on the capacity C.

Theorem 1. *For large finite field size q, the rates up to*

$$\frac{1}{T}\times\left[\Delta\,(T-\Delta) + \mathcal{O}(q^{-1})\right]\log_2 q$$

are achievable, where $\Delta = \min(m,n,\lfloor T/2\rfloor)$.

For the special case where $n = 1$, we get that

$$\begin{aligned}I_{\max}(X;Y) =& (T-k)\log_2 q - q^{-k}\log_2 k \\ &- (1-q^{-k})\log_2\left(\frac{1-q^{-k}}{1-q^{-T}}\right).\end{aligned}$$

For large q the mutual information behaves as

$$I(X;Y) \approx T-k, \tag{5}$$

which is maximized for $k = 1$. Thus, selecting matrices X that span one-dimensional subspaces leads to an achievable rate of $1-1/T$ which is close to the trivial upper bound of 1.

Capacity when $T > n + \min(m,n)$.

Consider the input distribution where all subspaces of the same dimension are chosen with the same probability. That is,

$$\Pr(\langle X\rangle = \pi, \dim(\pi) = r) = \alpha_r \begin{bmatrix} T \\ r\end{bmatrix}_q^{-1}, \tag{6}$$

where $\sum_{r=0}^{\min(m,T)} \alpha_r = 1$.

The proof of the following theorem is provided in [7].

Theorem 2. *If $T > n + \min(m,n)$, there exist a number q_0 such that for $q > q_0$, the optimal input distribution is as in (6) with $\alpha_{\min(m,n)} = 1$. Then,*

$$C = \left[\min(m,n)(T-\min(m,n)) + \mathcal{O}(q^{-1})\right]\log_2 q.$$

It is worth noting that the above equation coincides with the lower bound in Theorem 1, since under the assumptions of Theorem 2 we have that $\min(m,n,\lfloor T/2\rfloor) = \min(m,n)$.

B. Multiple Sources

For simplicity we consider the case of two sources X_1 and X_2. The well known rate region for the MAC channel is given by the union of rate pairs satisfying [4]

$$\begin{aligned}R_1 &\leq I(X_1;Y|X_2) = H(Y|X_2) - H(Y|X_1,X_2), \\ R_2 &\leq I(X_2;Y|X_1) = H(Y|X_1) - H(Y|X_1,X_2), \\ R_1+R_2 &\leq I(X_1,X_2;Y) = H(Y) - H(Y|X_1,X_2),\end{aligned}$$

for a given channel probability $\Pr(Y|X_1,X_2)$ and $P(X_1,X_2) = \Pr(X_1)\Pr(X_2)$.

For the channel described by (2), the conditional probability of Y given $X_1 = x_1$ and $X_2 = x_2$, can be written as follows

$$\begin{aligned}&\Pr(Y = y_0|\,\langle X_1\rangle = \pi_1, \langle X_2\rangle = \pi_2) = \\ &= \begin{cases} q^{-n\dim(\pi_1+\pi_2)} & \langle y_0\rangle \subseteq \pi_1+\pi_2, \\ 0 & \text{otherwise.}\end{cases}\end{aligned}$$

As can be observed from the above equation, the probability of receiving some matrix y_0 only depends on the subspaces π_1 and π_2 and not on the exact matrices x_1 and x_2 sent by the sources. So we can again conclude that considering only input distributions over subspaces is sufficient to describe the multiple access region for our channel.

For convenience, in the following we will use the notation:

$$\begin{aligned}P_Y(y_0) &\triangleq \Pr(Y = y_0), \\ P_X(\pi) &\triangleq \Pr(\langle X\rangle = \pi), \\ P_{Y|X}(y_0|\pi) &\triangleq \Pr(Y = y_0|\,\langle X\rangle = \pi), \\ P_{Y|X_1X_2}(y_0|\pi_1,\pi_2) &\triangleq \Pr(Y = y_0|\,\langle X_1\rangle = \pi_1, \langle X_2\rangle = \pi_2).\end{aligned}$$

We now write the expressions for the entropies in the rate region inequalities. Clearly,

$$H(Y|\,\langle X_1\rangle = \pi_1, \langle X_2\rangle = \pi_2) = n\dim(\pi_1+\pi_2)\log_2 q,$$

and thus for the conditional entropy we have

$$H(Y|X_1,X_2) = n\log_2 q \sum_{\pi_1,\pi_2} \dim(\pi_1+\pi_2) P_{X_1}(\pi_1) P_{X_2}(\pi_2).$$

To compute $H(Y|X_1)$ and $H(Y|X_2)$ we will use the probability $P_{Y|X_1X_2}$. Since

$$P_{Y|X_1}(y_0|\pi_1) = \sum_{\pi_2:\ \langle y_0\rangle \subseteq \pi_1+\pi_2} P_{X_2}(\pi_2) q^{-n\dim(\pi_1+\pi_2)},$$

we can write

$$H(Y|\langle X_1\rangle = \pi_1) = -\sum_{y_0} P_{Y|X_1}(y_0|\pi_1)\log_2\left(P_{Y|X_1}(y_0|\pi_1)\right).$$

Therefore, for $H(Y|X_1)$ we have

$$H(Y|X_1) = \sum_{\pi_1} H(Y|\langle X_1\rangle = \pi_1) P_{X_1}(\pi_1)$$

$$= -\sum_{\pi_1} P_{X_1}(\pi_1) \sum_{y_0} \left(\sum_{\substack{\pi_2: \\ \langle y_0\rangle \subseteq \pi_1+\pi_2}} P_{X_2}(\pi_2) q^{-n\dim(\pi_1+\pi_2)} \right)$$

$$\times \log_2 \left(\sum_{\substack{\pi_2': \\ \langle y_0\rangle \subseteq \pi_1+\pi_2'}} P_{X_2}(\pi_2') q^{-n\dim(\pi_1+\pi_2')} \right).$$

A similar expression holds for $H(Y|X_2)$.

Finally, to calculate $H(Y) = -\sum_{y_0} P_Y(y_0)\log_2 P_Y(y_0)$, we can use the expression

$$P_Y(y_0) = \sum_{\pi_1,\pi_2:\ \langle y_0\rangle \subseteq \pi_1+\pi_2} q^{-n\dim(\pi_1+\pi_2)} P_{X_1}(\pi_1) P_{X_2}(\pi_2).$$

Achievable Region

Again for simplicity we assume two sources. Consider the input distribution

$$\Pr(\langle X_i\rangle = \pi_i, \dim(\pi_i) = r_i) = \begin{cases} \begin{bmatrix} T \\ k_i \end{bmatrix}_q^{-1} & r_i = k_i, \\ 0 & \text{otherwise,} \end{cases}$$

for $i \in \{1,2\}$ where k_i are some fixed values satisfying $0 \le k_i \le \min(m_i, T)$. Substituting this distribution in the entropy expressions we have previously calculated, we get the achievable region

$$R_i \le \frac{1}{T}\left[T\min(n,k_i) + k\max(n-k_i,0) - n\min(k,T)\right],$$

$$R_1+R_2 \le \frac{1}{T}\left[[T-\min(n,k,T)]\min(n,k,T)\right] + \frac{1}{T}\left[n\left[\min(n,k,T)-\min(k,T)\right]\right],$$

where $k = k_1+k_2$, $0 \le k_i \le \min(m_i,T)$, and $i \in \{1,2\}$. The complete proof is provided in [7].

Maximizing the above equations over different values of k_1 and k_2, we obtain the following theorem.

Theorem 3. *For the case of two sources, an achievable region is given by*

$$R_i \le \frac{1}{T}\left[\Delta_i\left(T-\Delta_i\right)\right],$$

$$R_1+R_2 \le \frac{1}{T}\left[\Delta\left(T-\Delta\right)\right],$$

where $m = m_1+m_2$, $\Delta_i = \min(m_i, n, \lfloor T/2\rfloor)$, $\Delta = \min(m, n, \lfloor T/2\rfloor)$, *and* $i \in \{1,2\}$.

IV. Combinatorial Code Constructions

In this section, we are interested in subspace combinatorial codes of fixed dimension. Subspace codes use subspace as codewords. That is, to transmit a specific value the source inserts in the network the basis element of the corresponding codeword subspace.

The idea of using subspace codes for network coding and code constructions for the single source model in (1) were proposed in [3]. In Section IV-B we consider instead the multiple sources case as in (2). We also address in Section IV-A the restricted access problem, where at each timeslot an unknown subset of at most M out of the N sources is active. Each source has 2^m messages to convey. We want, by observing n packets at the receiver, to determine which M sources are active and what are the information messages they convey. Our proposed constructions bring together and build on ideas from [8], [9] and [10].

A. Construction 1

In [8] a "lifting" construction is proposed that allows to construct a subspace code by starting from any rank metric code with codewords $m\times T$ matrices X and appending to each such matrix the same $m\times m$ identity matrix. The codewords of the resulting subspace code are the subspaces spanned by the rows of the matrices $[I \quad X]$, for all possible matrices X in the original code.

Appending the identity matrix ensures that
(P1) all subspaces have dimension equal to m, and that
(P2) the distance properties of the rank metric code are inherited by the subspace code [8].
Note that instead of the identity matrix we can equivalently use any $m\times m$ full rank matrix. In the rest of this section for simplicity we will just assume that we are not interested in error correction capabilities, and thus X takes all possible values. We will also restrict our attention to binary codes, and to the symmetric case, where every source has 2^m messages to transmit, and the receiver receives $n = mM$ packets.

To build codes for the MAC case with M active users, we are going to use a lifting construction similar to [8] where now we append a different lifting matrix

$$L_i = [G_i \quad A_i],$$

to create the codewords of every source. That is, in the case of uncoded transmission, each source will transmit the subspaces spanned by the rows of the matrices $< [G_i \quad A_i \quad X] >$ where X is an arbitrary matrix.

The N matrices A_i serve to identify the set of active users. Each matrix A_i has m identical rows, each row repeating the same vector a_i. The N vectors a_i have the property that any $M+1$ of them are linearly independent, and thus, any M of them span a different M-dimensional subspace that uniquely identifies the set of active sources. For the vectors a_i we can simply use the columns of a parity check matrix corresponding to an error correction code with minimum distance $M+1$.

The N matrices G_i have the property that any M of them span the mM dimensional space. Such matrices can be constructed exactly as in the construction of (M, m) separable codes described in [10]. These matrices serve the same role as the identity matrix in the original construction, that is, ensure that (P1) and (P2) hold. Note that the size of the lifting part of the matrices depends on the number of active users M.

B. Construction 2

In this section we mainly extend the construction introduced in [3], [9], and we will use the terminology therein.

Let $\mathbb{F}_q$ be a finite field and let $\mathbb{F} = \mathbb{F}_{q^K}$ be an extension of $\mathbb{F}_q$. Let $L(x)$ be a linearized polynomial as described in [9]. We will refer to the degree of its conventional q-associate as the associate degree of $L(x)$.

We may regard $\mathbb{F}$ as vector space of dimension K over $\mathbb{F}_q$. Let each $A_i = \{\alpha_1^{(i)}, \ldots, \alpha_{l_i}^{(i)}\} \subseteq \mathbb{F}$, $i = 1, \ldots, N$, be a set of linearly independent elements in this vector space such that $\langle A_1 \rangle + \cdots + \langle A_N \rangle$ span an $l = \sum_{i=1}^{N} l_i$ dimensional vectors space over $\mathbb{F}_q$, which means $\langle A_i \rangle \cap \langle A_j \rangle = \{\overrightarrow{0}\}$ for all $i \neq j$. Clearly we have $l \leq K$. Let us define the space W_i as follows

$$W_i = \langle A_i \rangle \oplus \mathbb{F} = \{(\alpha, \beta) : \alpha \in \langle A_i \rangle, \beta \in \mathbb{F}\},$$

where subspaces W_i are disjoint because $\langle A_i \rangle$ are disjoint. We will operate on the space

$$\begin{aligned} W &= W_1 + \cdots + W_N \\ &= (\langle A_1 \rangle + \cdots + \langle A_N \rangle) \oplus \mathbb{F} \\ &= \{(\alpha, \beta) : \alpha \in \left\langle \cup_{i=1}^{N} A_i \right\rangle, \beta \in \mathbb{F}\}, \end{aligned}$$

which is a vector space of dimension $l + K$ over $\mathbb{F}_q$.

Let $u^{(i)} = (u_0^{(i)}, \ldots, u_{m_i-1}^{(i)}) \in \mathbb{F}^{m_i}$ denote a block of message symbols in source i, consisting of m_i symbols over $\mathbb{F}$ or, equivalently, Km_i symbols over $\mathbb{F}_q$. Let $\mathbb{F}^{m_i}[x]$ denote the set of linearized polynomials over $\mathbb{F}$ of associate degree at most $m_i - 1$. Let $f^{(i)}(x) \in \mathbb{F}^{m_i}[x]$, defined as

$$f^{(i)}(x) = \sum_{j=0}^{m_i-1} u_j^{(i)} x^{[j]},$$

be the linearized polynomial with coefficients corresponding to $u^{(i)}$. Finally, let each source i compute $\beta_j^{(i)} = f^{(i)}(\alpha_j^{(i)})$. Each pair $(\alpha_j^{(i)}, \beta_j^{(i)})$, $j = 1, \ldots, l_i$, may be regarded as a vector in W_i. Since $\{\alpha_1^{(i)}, \ldots, \alpha_{l_i}^{(i)}\}$ is a linearly independent set, so is $\{(\alpha_1^{(i)}, \beta_1^{(i)}), \ldots, (\alpha_{l_i}^{(i)}, \beta_{l_i}^{(i)})\}$, hence this set spans an l_i-dimensional subspace V_i of W_i.

We know that W_i are disjoint so are V_i for different sources. Next we will use the Lemma 13 stated in [3] to show that V_i are different for two different message blocks. We denote the map that takes the message polynomial $f^{(i)}(x) \in \mathbb{F}^{m_i}[x]$ to the linear space $V_i \in \mathcal{P}(W, l_i)$ as Γ_{A_i}.

Lemma IV.1. *If $l_i \geq m_i$ then the map $\Gamma_{A_i} : \mathbb{F}^{m_i} \to \mathcal{P}(W_i, l_i)$ is injective.*

Proof. Refer to [3]. □

So provided the conditions posed by Lemma IV.1 are satisfied, we can construct our code by assigning to each source the codebook

$$\mathcal{X}_i = \left\{ \pi_j^i(u^{(i)}) : \text{for all } u^{(i)} \in \mathbb{F}^{m_i} \right\},$$

where $\pi_j^i(u^{(i)}) = \left\langle \{(\alpha_1^{(i)}, \beta_1^{(i)}), \ldots, (\alpha_{l_i}^{(i)}, \beta_{l_i}^{(i)})\} \right\rangle$ and $\beta_j^{(i)} = f^{(i)}(\alpha_j^{(i)})$.

From the above code construction we observe that we have $|\mathcal{X}_i| = q^{Km_i}$ for each user i, which results in a code with size $\log_q(|C|) = K \sum_{i=1}^{N} m_i$. The length of code is $l + K$ and the dimension of code at the receiver is l. So for the rate of the code we can write

$$R_{\text{receiver}} = \frac{mK}{l(l+K)},$$

where $m = \sum_{i=1}^{N} m_i$. This rate approaches 1 assuming that K is large with respect to l and assuming $m \approx l$.

V. Conclusions

In this paper we formulated a new model of network communication of multiple sources via a proposed noncoherent block channel modeling. We showed the optimality of the use of subspaces, characterized achievable rates for some coding strategies, and discussed combinatorial code designs.

Acknowledgments: The authors would like to thank S. Saeedi and S. Mohajer for many useful discussions.

References

[1] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow", *IEEE Trans. Inform. Theory*, vol. 46, iss. 4, pp. 1204–1216, 2000.

[2] L. Zheng and D. N. C. Tse, "Communication on the grassman manifold: a geometric approach to the non-coherent multiple antenna channel", *IEEE Trans. Inform. Theory*, vol. 48, iss. 2, pp. 359–383, February 2002.

[3] R. Koetter and F. Kschischang, "Coding for errors and erasures in random network coding", *IEEE International Symposium on Information Theory (ISIT)*, June 2007.

[4] T. Cover and J. Thomas, "Elements of Information Theory", Wiley & Sons, New York, Second edition, 2006.

[5] P. A. Chou, Y. Wu, and K. Jain, "Practical network coding," *Allerton Conference on Communication, Control, and Computing*, Monticello, IL, October 2003.

[6] M. Jafari Siavoshani, C. Fragouli and S. Diggavi, "Passive topology discovery for network coded systems", *Information Theory Workshop (ITW)*, Bergen, Norway, pp 17-21, July 2007.

[7] M. Jafari Siavoshani, C. Fragouli and S. Diggavi, "Noncoherent Multisource Network Coding", *EPFL Technical report*, May 2008.

[8] D. Silva, F. Kschischang, and R. Koetter, "A rank metric approach to error control in random network coding", submitted for publication.

[9] D. Silva and F. R. Kschischang, "Using rank-metric codes for error correction in random network coding," *IEEE International Symposium on Information Theory*, Nice, France, June 2007.

[10] G. Poltyrev and J. Snyders, "Linear codes for the sum mod-2 multiple access channel with restricted access", *IEEE Trans. Inform. Theory*, vol. 41, iss. 2, May 1995.