Superposition Coding for Linear Operator Channels over Finite Fields

Shenghao Yang
Institute for Theoretical Computer Science
Institute for Interdisciplinary Information Sciences
Tsinghua University, P. R. China
Email: shyang@tsinghua.edu.cn

***Abstract*—A coding approach based on the superposition structure is proposed for linear operator channels. Under a subspace decoding rule, a lower bound on the maximum achievable rate of this coding approach is characterized. Under the subspace decoding rule, this coding approach is capacity achieving for a class of linear operator channels, and it can potentially achieve higher rates than the subspace coding approach.**

I. Introduction

Fix a finite field $\mathbb{F}$ with q elements. A linear operator channel (LOC) with input random variable $X \in \mathbb{F}^{T\times M}$ and output random variable $Y \in \mathbb{F}^{T\times N}$ is given by

$$Y = XH, \tag{1}$$

where H, called the *transfer matrix*, is a random variable over $\mathbb{F}^{M\times N}$. We assume that X and H are independent, and the transfer matrices in different channel uses are independent and follow the same distribution. The distribution of H is given a priori to both the transmitter and the receiver, but the instances of H are not known by either the transmitter or the receiver. An LOC models the communication through a network employing linear network coding [1]. In this paper, we focus on the coding problem of the noncoherent transmission of LOCs with an arbitrary distribution of H.

Existing works on coding for LOCs are mostly in two frameworks. When $T \geq M$, part of X can be used to transmit an identity matrix so that the receiver can recover the instances of H. This approach, called *channel training*, was proposed for random linear network coding [2]. Channel training can achieve at least $(1 - M/T)$ fraction of the capacity of the LOC [3], so it becomes efficient when T is much larger than M. A channel training scheme with low encoding/decoding complexity has been proposed in [4], [5] by generalizing fountain codes.

When T is not much larger than M, the overhead used to explicitly recover the instances of H is dominating. In this scenario, subspace coding has advantage over channel training. We call the vector space spanned by the row (column) vectors of a matrix the row (column) space of the matrix. Koetter and Kschischang observe that in an LOC, the column space of Y is always a subspace of the column space of X [6]. They propose a coding approach using the column subspaces for encoding and decoding and discuss coding schemes for one use of an LOC [6]. We refer to the coding schemes using the column subspaces for encoding and decoding as *(KK) subspace coding*. Thereafter, subspace coding generated a lot of research interests (see [7]–[9], et al.) and the study of subspace coding is also extended from one use to multiple uses of an LOC [3], [10], [11]. For general LOCs, however, good subspace codes are still unknown.

In this paper, instead of studying efficient subspace coding schemes, we propose a coding approach for LOCs that can potentially achieve higher rates than subspace coding does. The motivation is that since the maximum achievable rate of subspace coding is in general less than the capacity of the LOC, theoretically, there exist codes that can achieve rates higher than subspace coding. The discussion hereafter is for general values of T, M, N and q unless otherwise specified.

Our coding approach for LOCs is based on the observation that the row spaces of X and Y can also be used to transmit information. A code using this approach includes a set of cloud centers and a set of satellite codes each of which corresponds to a cloud center. A cloud center is a sequence of subspaces, and the corresponding satellite code is a set of sequences of matrices whose row spaces form a sequence identical to the cloud center. During the encoding, part of the message is first encoded to a cloud center. The remaining of the message is encoded to a codeword of the corresponding satellite code. Due to the similarity to the superposition coding for broadcast channels [12], this approach is called *Subspace-matrix superposition (Sumas) coding*.

We characterize the achievable rates of Sumas codes under a subspace decoding rule. The cloud center is first recovered using the row spaces of the received matrices, which tells which satellite code is used in encoding. Then, the corresponding satellite code is decoded using the column spaces of the received matrices by only checking the inclusion relationship between subspaces. Under the above subspace coding rule, we obtain a lower bound on the maximum achievable rate of the Sumas codes. We demonstrate a class of LOCs for which Sumas codes are capacity achieving, and we show by example

Except the work about row-space-symmetric LOCs, all the research and development work as described in the paper was conducted while the author was working in the Institute of Network Coding, The Chinese University of Hong Kong.

978-1-4673-0223-4/12/$31.00 ©2012 IEEE

that Sumas codes can achieve rates strictly higher than the maximum achievable rate of subspace codes.

II. Preliminaries

For a matrix $\mathbf{X}$, let $\text{rk}(\mathbf{X})$ be its rank, let $\mathbf{X}^\top$ be its transpose, and let $\langle \mathbf{X} \rangle$ be the subspace spanned by the column vectors of $\mathbf{X}$. We call $\langle \mathbf{X} \rangle$ and $\langle \mathbf{X}^\top \rangle$ the column space and the row space of $\mathbf{X}$, respectively.

The vectors in $\mathbb{F}^t$ are regarded as column vectors. The *projective space* $\text{Pj}(\mathbb{F}^t)$ is the collection of all subspaces of $\mathbb{F}^t$. Let $\text{Pj}(m, \mathbb{F}^t)$ be the subset of $\text{Pj}(\mathbb{F}^t)$ that contains all the subspaces with dimension less than or equal to m. Let $\text{Fr}(\mathbb{F}^{m\times r})$ be the set of full rank matrices in $\mathbb{F}^{m\times r}$. Define

$$\chi_r^m = \begin{cases} (q^m - 1)(q^m - q)\cdots(q^m - q^{r-1}) & 0 < r \le m \\ 1 & r = 0 \end{cases} \tag{2}$$

For $r \le m$, we can count that $|\text{Fr}(\mathbb{F}^{m\times r})| = \chi_r^m$. The *Grassmannian* $\text{Gr}(r, \mathbb{F}^m)$ is the set of all r-dimensional subspaces of $\mathbb{F}^m$. The *Gaussian binomial*

$$\begin{bmatrix} m \\ r \end{bmatrix} \triangleq \frac{\chi_r^m}{\chi_r^r}$$

is the number of r-dimensional subspaces of $\mathbb{F}^m$ [13], i.e., $|\text{Gr}(r, \mathbb{F}^m)| = \left[\begin{smallmatrix} m \\ r \end{smallmatrix}\right]$.

III. LOCs from Two Points of View

To understand the coding for LOCs, we discuss the properties of LOCs from two different angles.

A. Linear Combination View

For an LOC given in (1), the column space of Y is always a subspace of the column spaces of X, i.e., $\langle Y \rangle \subset \langle X \rangle$. This point of view was first employed by Koetter and Kschischang in their approach for random linear network coding [6], in which they define a channel with subspaces as input and output and discuss the coding problem for this subspace channel.

We refer to the coding schemes of LOCs using the column subspaces for encoding and decoding as *(KK) subspace coding*. An n-block subspace code is a subset of $(\text{Pj}(\min\{T, M\}, \mathbb{F}^T))^n$. To apply a subspace code in an LOC, the subspaces in a codeword need to be converted to matrices. For $U \in \text{Pj}(\min\{T, M\}, \mathbb{F}^T)$, this conversion can be done by a transition probability $P_{X|\langle X \rangle}(\cdot|U)$. Given a transition matrix $P_{X|\langle X \rangle}$, we can define a new channel with input $\langle X \rangle$ and output $\langle Y \rangle$ that the subspace code actually applied to.

When using column subspaces for encoding and decoding, the maximum achievable rate of a memoryless LOC is

$$C_{\text{SS}} \triangleq \max_{P_{X|\langle X \rangle}} \max_{p_{\langle X \rangle}} I(\langle X \rangle; \langle Y \rangle).$$

But an optimal input distribution for subspace coding is difficult to find in general since the above maximization problem is not concave.

In general C_{SS} is less than the capacity C of a memeoryless LOC, but $C_{\text{SS}} = C$ for some special cases. For example, when H contains uniformly i.i.d. components, and when H is conditionally uniform given $\text{rk}(H)$ (also called u.g.r.), subspace coding is optimal [14]–[16]. A more general sufficient condition such that $C_{\text{SS}} = C$ given in [17] is summarized as follows.

An LOC is called *rank symmetric* if there exists a function $\mu : \mathbb{Z}^+ \times \mathbb{Z}^+ \to [0\ 1]$ such that

$$P_{Y|X}(\mathbf{Y}|\mathbf{X}) = \begin{cases} \mu(\text{rk}(\mathbf{X}), \text{rk}(\mathbf{Y})) & \langle \mathbf{Y} \rangle \subseteq \langle \mathbf{X} \rangle \\ 0 & o.w. \end{cases}$$

We have that for rank-symmetric LOCs,

$$C = C_{\text{SS}} = \max_{p_{\text{rk}(X)}} [J(\text{rk}(X); \text{rk}(Y)) + I(\text{rk}(X); \text{rk}(Y))], \tag{3}$$

where

$$J(\text{rk}(X); \text{rk}(Y)) \triangleq \sum_{s \le r} p_{\text{rk}(X)\,\text{rk}(Y)}(r, s) \log \frac{\chi_s^T}{\chi_s^r} \ge 0. \tag{4}$$

(Note that an LOC with a u.g.r. transfer matrix is always rank symmetric, but when $T < M$, a rank-symmetric LOC may not have a u.g.r. transfer matrix [18].)

B. Linear Operation View

For an LOC given in (1), the transpose of the transfer matrix $H^\top$ is a (random) linear operator that maps a vector in $\mathbb{F}^M$ to a vector in $\mathbb{F}^N$. We see that $H^\top$ induces a linear operator on subspaces of $\mathbb{F}^M$ given by

$$H^\top U \triangleq \{H^\top x : x \in U\}.$$

This operation gives a new channel with input $\langle X^\top \rangle$ and output $\langle Y^\top \rangle$.

Definition 1: An LOC in (1) induces a new channel with input $\langle X^\top \rangle$, output $\langle Y^\top \rangle$, and the channel law

$$P_{\langle Y^\top \rangle | \langle X^\top \rangle}(V|U) = \Pr\{H^\top U = V\}.$$

We call this channel the *subspace core* of the LOC.

The subspace core of an LOC is unique. We can show that

$$C \ge \max_{p_{\langle X^\top \rangle}} [J(\text{rk}(X); \text{rk}(Y)) + I(\langle X^\top \rangle; \langle Y^\top \rangle)] \tag{5}$$

with equality when the LOC satisfies that for any $\langle \mathbf{Y} \rangle \subset \langle \mathbf{X} \rangle$,

$$P_{Y|X}(\mathbf{Y}|\mathbf{X}) = \frac{1}{\chi_{\text{rk}(\mathbf{Y})}^{\text{rk}(\mathbf{X})}} P_{\langle Y^\top \rangle | \langle X^\top \rangle}(\langle \mathbf{Y}^\top \rangle | \langle \mathbf{X}^\top \rangle).$$

We call an LOC with the above property a *row-space-symmetric* LOC. Note that a rank-symmetric LOC is row-space-symmetric.

In the next section, we show a coding approach that combines these two views of an LOC.

IV. Subspace-Matrix Superposition Codes

Let x^n be a vector of n components, where the ith component is x_i. For $U \in \text{Pj}(\min\{T, M\}, \mathbb{F}^M)$, let

$$\phi_T(U) \triangleq \{\mathbf{X} \in \mathbb{F}^{T\times M} : \langle \mathbf{X}^\top \rangle = U\},$$

i.e., $\phi_T(U)$ is the set of all input matrices with row space U.

Definition 2: An n-block subspace-matrix superposition (Sumas) code contains a set of cloud centers and a set of satellite codes each of which corresponds to a cloud center. The set of cloud centers $\mathcal{U}$, which is also called a cloud code, is a subset of $(\mathrm{Pj}(\min\{T,M\},\mathbb{F}^M))^n$. The satellite code corresponding to $U^n\in\mathcal{U}$ is a subset of $\phi_T(U_1)\times\cdots\times\phi_T(U_n)$, and is denoted by $\mathcal{S}(U^n)$.

The encoding of a Sumas code has two steps: i) Map the first part of the message to a cloud center; ii) Pick the satellite code corresponding to the cloud center in the first step, and map the remaining part of the message to a codeword of the satellite code.

Generally, the decoding of a Sumas code has two inverse steps. We first decode the cloud code using the received matrices and recover the transmitted cloud center, which tells us which satellite code is used in encoding. Then we decode the corresponding satellite code. The achievable rates of the cloud code and the satellite codes can be characterized from a broadcast channel perspective.

The Sumas codes can be regarded as codes for a broadcast channel with input X and output $Y_1=Y$ and $Y_2=Y$, where X and Y are related by (1). Let R_1 be the rate for output Y_1 and R_2 be the rate for output Y_2. This is a degraded broadcast channel [19] and any rate pair (R_1,R_2) such that for some p_X

$$R_1\le I(X;Y|\langle X^\top\rangle)$$
$$R_2\le I(\langle X^\top\rangle;Y)$$

is achievable by the Sumas codes (where R_1 is achieved by the satellite codes and R_2 is achieved by the cloud code). Since

$$I(X;Y|\langle X^\top\rangle)+I(\langle X^\top\rangle;Y)=I(X;Y),$$

the Sumas codes achieve the capacity of the LOC.

The above discussion about the achievable rates is just for the sake of introducing the coding structure. When the discussion is not limited to any decoding rules, jointly typical decoding is usually assumed. However, we are interested in the property of Sumas codes under special decoding rules that have simple decoding algorithms. This may give insights in the coding design.

A. Subspace Decoding Rule

We study the performance of the Sumas codes under the following subspace decoding rule.

Definition 3 (Subspace Decoding Rule): Let $\mathbf{Y}_1,\ldots,\mathbf{Y}_n$ be the received matrices. We first use $(\langle\mathbf{Y}_1^\top\rangle,\ldots,\langle\mathbf{Y}_n^\top\rangle)$ to decode the cloud code. After the cloud code is decoded, we know the cloud center in the first step of encoding and hence the satellite code used in encoding. Let $\hat{U}^n$ be the cloud center recovered. To decode the satellite code, we try to find a codeword $(\mathbf{B}_1,\ldots,\mathbf{B}_n)$ in $\mathcal{S}(\hat{U}^n)$ satisfying $\langle\mathbf{Y}_i\rangle\subset\langle\mathbf{B}_i\rangle$ for $i=1,\ldots,n$. If there exist more than one such codewords, an error occurs.

In the subspace decoding rule, only the row spaces and the column spaces of the received matrices are used. The decoding of satellite codes is more specific since only the inclusion relationship between subspaces is checked.

Theorem 1: The maximum achievable rate of Sumas codes under the subspace decoding rule is at least

$$\max_{p_{\langle X^\top\rangle}}\left[J(\mathrm{rk}(X),\mathrm{rk}(Y))+I(\langle X^\top\rangle;\langle Y^\top\rangle)\right],$$

where J is defined in (4). As a function of $p_{\langle X^\top\rangle}$, $J(\mathrm{rk}(X);\mathrm{rk}(Y))+I(\langle X^\top\rangle;\langle Y^\top\rangle)$ is concave.

The proof of the above theorem is postponed to the next subsection. Here we demonstrate that Sumas codes can potentially achieve rate higher than KK subspace codes. We consider, as an example, the row-space-symmetric LOCs defined in the end of Section III-B. By Theorem 1 and (5), we see that Sumas codes achieve the capacity of row-space-symmetric LOCs. We claim that for row-space-symmetric LOCs,

$$C_{\mathrm{SS}}\le\max_{p_{\langle X^\top\rangle}}\left[J(\mathrm{rk}(X),\mathrm{rk}(Y))+I(\langle X^\top\rangle;\mathrm{rk}(Y))\right]\qquad(6)$$

(the proof is omitted). Since

$$I(\langle X^\top\rangle;\langle Y^\top\rangle\ge I(\langle X^\top\rangle;\mathrm{rk}(Y)),\qquad(7)$$

subspace coding may not achieve the capacity of row-space-symmetric LOCs. If the inequality in either (6) or (7) is strict, Sumas codes can potentially achieve rate strictly higher than subspace codes. We demonstrate that $I(\langle X^\top\rangle;\langle Y^\top\rangle)$ can be strictly larger than $I(\langle X^\top\rangle;\mathrm{rk}(Y))$. We can check that all LOCs with $T=1$, $M=N=2$ over the binary field $\mathbb{F}_2$ are row-space symmetric. Fix such an LOC with the transfer matrix being the identity matrix. On one hand, $I(\langle X^\top\rangle;\mathrm{rk}(Y))\le H(\mathrm{rk}(Y))\le\log_2(\min\{T,M\}+1)=1$. On the other hand, $I(\langle X^\top\rangle;\langle Y^\top\rangle)=\log_2|\mathrm{Pj}(\min\{T,M\},\mathbb{F}_2^M)|=2$.

B. Achievable Rates

We evaluate the achievable rates of Sumas codes under the subspace decoding rule.

Lemma 2: Let X be the uniform random variable with support $\phi_T(U)$, $U\in\mathrm{Pj}(\min\{M,T\},\mathbb{F}^M)$. For $V\in\mathrm{Pj}(\dim(U),\mathbb{F}^T)$,

$$\Pr\{\langle X\rangle\supset V\}=\frac{\chi_{\dim(V)}^{\dim(U)}}{\chi_{\dim(V)}^{T}}.$$

Proof: Let $r=\dim(U)$ and $s=\dim(V)$. We first show that $|\phi_T(U)|=\chi_r^T$, where the RHS is defined in (2). Fix $\mathbf{D}\in\mathrm{Fr}(\mathbb{F}^{r\times M})$ with $\langle\mathbf{D}^\top\rangle=U$. Rewrite

$$\phi_T(U)=\{\mathbf{BD}:\mathbf{B}\in\mathrm{Fr}(\mathbb{F}^{T\times r})\}.\qquad(8)$$

So $|\phi_T(U)|=|\mathrm{Fr}(\mathbb{F}^{T\times r})|=\chi_r^T$.

Let us check the distribution of $\langle X\rangle$. For $V'\in\mathrm{Gr}(r,\mathbb{F}^T)$,

$$\begin{aligned}\phi_T(U|V')&\triangleq\{\mathbf{X}\in\mathbb{F}^{T\times M}:\langle\mathbf{X}\rangle=V',\langle\mathbf{X}^\top\rangle=U\}\\&=\{\mathbf{BD}:\mathbf{B}\in\mathrm{Fr}(\mathbb{F}^{T\times r}),\langle\mathbf{B}\rangle=V'\}\qquad(9)\\&=\{\mathbf{BD}:\mathbf{B}^\top\in\phi_r(V')\},\end{aligned}$$

where (9) is obtained similar to (8). The sets $\phi_T(U|V')$, $V'\in\mathrm{Gr}(r,\mathbb{F}^T)$, give a partition of $\phi_T(U)$, and all have the same cardinality χ_r^r. Thus, $\langle X\rangle$ has a uniform distribution over $\mathrm{Gr}(r,\mathbb{F}^T)$.

We claim that

$$|\{V' \in \mathrm{Gr}(r, \mathbb{F}^T) : V \subset V'\}| = \begin{bmatrix} T-s \\ r-s \end{bmatrix}. \qquad (10)$$

Then

$$\Pr\{\langle X \rangle \supset V\} = \frac{\begin{bmatrix} T-s \\ r-s \end{bmatrix}}{|\,\mathrm{Gr}(r, \mathbb{F}^T)|} = \frac{\begin{bmatrix} T \\ r \end{bmatrix} \frac{\chi_s^r}{\chi_s^T}}{\begin{bmatrix} T \\ r \end{bmatrix}} = \frac{\chi_s^r}{\chi_s^T}.$$

We can verify (10) as follows. Fix a complementary subspace $\bar{V}$ of V such that $V \oplus \bar{V} = \mathbb{F}^T$ and $V \cap \bar{V} = \{\mathbf{0}\}$. For any $V' \in \mathrm{Gr}(r, \mathbb{F}^T)$ with $V \subset V'$, we have a unique direct sum decomposition $V' = V \oplus \tilde{V}$ such that $\tilde{V} \subset \bar{V}$. Since such V' and $\tilde{V}$ are one-to-one correspondent, the problem becomes to count the number of $(r-s)$-dimensional subspaces of the $(T-s)$-dimensional subspace $\bar{V}$. By Gaussian binomials, the number is $\begin{bmatrix} T-s \\ r-s \end{bmatrix}$. ■

We define the following sets of typical sequences according to [20, Chapter 2]. Denote by $N(a|x^n)$ the number of occurrences of a in x^n. For a random variable A with support $\mathcal{A}$, let $T^n_{[A]\delta}$ be the set of all p_A-typical sequence x^n with constant δ, i.e.,

$$\left| \frac{N(a|x^n)}{n} - p_A(a) \right| \leq \delta \ \forall a \in \mathcal{A}.$$

Further, for a random variable B with support $\mathcal{B}$ and a transition matrix $P_{B|A}$, let $\mathrm{T}^n_{[B|A]\delta}(x^n)$ be the set of $P_{B|A}$-typical sequences y^n under the condition of $x^n \in (\mathcal{A})^n$ with constant δ, i.e.,

$$\left| \frac{N(a,b|x^n,y^n)}{n} - \frac{N(a|x^n)}{n} P_{B|A}(b|a) \right| \leq \delta \ \forall a \in \mathcal{A}, b \in \mathcal{B}.$$

The delta-convention in [20] is applied so that the constant δ will be omitted from the notation.

Lemma 3: Consider a satellite code $\mathcal{S}(U^n)$ with $U^n \in T^n_{[\langle X^\top \rangle]}$. When using only column spaces of the received matrices for decoding, the maximum achievable rate of the satellite code is at least $J(\mathrm{rk}(X); \mathrm{rk}(Y))$.

Proof: We study the achievable rates of satellite codes using a random coding scheme. Assume that the code book size is 2^{nR}. For codeword $(X_1, \ldots, X_n)$, X_i is independently, uniformly at random picked in $\phi_T(U_i)$. Construct the satellite code $\mathcal{S}(U^n)$ randomly by generating 2^{nR} codewords independently.

Let $X^n \in \mathcal{S}(U^n)$ be the codeword transmitted and Y^n be the received sequence of matrices. By [20, Lemma 2.12], there exists a sequence $\epsilon_n \to 0$ such that

$$\Pr\{\mathrm{rk}(Y^n) \in \mathrm{T}^n_{[\mathrm{rk}(Y)|\langle X^\top \rangle]}(U^n)\} \geq 1 - \epsilon_n \qquad (11)$$

Here $\mathrm{rk}(Y^n) \triangleq (\mathrm{rk}(Y_1), \ldots, \mathrm{rk}(Y_n))$. Similarly, we use $\dim(U^n) = (\dim(U_1), \ldots, \dim(U_n))$.

For $\mathbf{Y}^n$ with $\mathrm{rk}(\mathbf{Y}^n) \in \mathrm{T}^n_{[\mathrm{rk}(Y)|\langle X^\top \rangle]}(U^n)$,

$$\begin{aligned}
& N(r, s|\dim(U^n), \mathrm{rk}(\mathbf{Y}^n)) \\
&= \sum_{V:\dim(V)=r} N(V, s|U^n, \mathrm{rk}(\mathbf{Y}^n)) \\
&\geq \sum_{V:\dim(V)=r} n\left(P_{\mathrm{rk}(Y)|\langle X^\top \rangle}(s|V) \frac{N(V|U^n)}{n} - \delta_n \right) && (12) \\
&\geq \sum_{V:\dim(V)=r} n\left(P_{\mathrm{rk}(Y)|\langle X^\top \rangle}(s|V) p_{\langle X^\top \rangle}(V) - \delta'_n \right) && (13) \\
&= n(p_{\mathrm{rk}(X)\,\mathrm{rk}(Y)}(r,s) - \delta''_n)
\end{aligned}$$

where (12) and (13) follow from $\mathrm{rk}(\mathbf{Y}^n) \in \mathrm{T}^n_{[\mathrm{rk}(Y)|\langle X^\top \rangle]}(U^n)$ and $U^n \in T^n_{[\langle X^\top \rangle]}$, respectively; and $\delta''_n \to 0$.

Let $\tilde{X}^n$ be a codeword not the same as X^n. For $\mathbf{Y}^n$ with $\mathrm{rk}(\mathbf{Y}^n) \in \mathrm{T}^n_{[\mathrm{rk}(Y)|\langle X^\top \rangle]}(U^n)$,

$$\begin{aligned}
\Pr\{\langle \tilde{X}_i \rangle \supset \langle \mathbf{Y}_i \rangle, \forall i\} &= \prod_{i=1}^{n} \Pr\{\langle \tilde{X}_i \rangle \supset \langle \mathbf{Y}_i \rangle\} \\
&= \prod_{i=1}^{n} \frac{\chi_{\mathrm{rk}(\mathbf{Y}_i)}^{\dim(U_i)}}{\chi_{\mathrm{rk}(\mathbf{Y}_i)}^{T}} && (14) \\
&= \prod_{r,s} \left(\frac{\chi_s^r}{\chi_s^T} \right)^{N(r,s|\dim(U^n),\mathrm{rk}(\mathbf{Y}^n))} \\
&\leq \prod_{r,s} \left(\frac{\chi_s^r}{\chi_s^T} \right)^{n(p_{\mathrm{rk}\,X\,\mathrm{rk}(Y)}(r,s)-\delta''_n)} && ,(15)
\end{aligned}$$

where (14) follows from Lemma 2.

The probability of decoding error using the decoding method described for satellite codes is

$$\begin{aligned}
& \Pr\{\exists \tilde{X}^n \neq X^n \text{ s.t } \langle \tilde{X}_i \rangle \supset \langle Y_i \rangle, \forall i\} \\
&= \sum_{\mathbf{Y}^n} \Pr\{\exists \tilde{X}^n \neq X^n \text{ s.t } \langle \tilde{X}_i \rangle \supset \langle \mathbf{Y}_i \rangle, \forall i\} p_{Y^n}(\mathbf{Y}^n) && (16) \\
&< \sum_{\tilde{X}^n \neq X^n} \sum_{\mathbf{Y}^n \in \mathrm{T}^n_{[\mathrm{rk}(Y)|\langle X^\top \rangle]}(U^n)} \Pr\{\langle \tilde{X}_i \rangle \supset \langle \mathbf{Y}_i \rangle, \forall i\} p_{Y^n}(\mathbf{Y}^n) \\
&\quad + \sum_{\mathbf{Y}^n \notin \mathrm{T}^n_{[\mathrm{rk}(Y)|\langle X^\top \rangle]}(U^n)} p_{Y^n}(\mathbf{Y}^n) && (17) \\
&\leq 2^{nR} \prod_{r,s} \left(\frac{\chi_s^r}{\chi_s^T} \right)^{n(p_{\mathrm{rk}(X)\,\mathrm{rk}(Y)}(r,s)-\delta'')} + \epsilon_n && (18) \\
&\leq 2^{n(R - J(\mathrm{rk}(X),\mathrm{rk}(Y)) + \delta'''_n)}
\end{aligned}$$

where (16) follows that Y^n and $\tilde{X}^n$ are independent when $\tilde{X}^n \neq X^n$; (17) follows from the union bound; and (18) follows from (11) and (15). Then following the typical random coding argument, we know that $J(\mathrm{rk}(X), \mathrm{rk}(Y))$ is achievable by the satellite code. ■

Proof of Theorem 1: The cloud code of a Sumas code is used for the subspace core of an LOC. By the achievability of the Channel Coding Theorem (see [20, Corollary 6.3]), for every distribution of $\langle X^\top \rangle$, there exists a sequence of cloud codes $\mathcal{U} \subset T^n_{[\langle X^\top \rangle]}$ achieving the rate $I(\langle X^\top \rangle; \langle Y^\top \rangle)$. By Lemma 3, the maximum achievable rate of the satellite codes is at least $J(\mathrm{rk}(X), \mathrm{rk}(Y))$. So the first part of the theorem is proved.

By writing

$$p_{\mathrm{rk}(X)\,\mathrm{rk}(Y)}(r, s) = \sum_{U} P_{\mathrm{rk}(Y)|\langle X^\top\rangle}(s|U) p_{\langle X^\top\rangle}(U),$$

we see that $J(\mathrm{rk}(X), \mathrm{rk}(Y))$ is a linear function of $p_{\langle X^\top\rangle}$. Together with the fact that $I(\langle X^\top\rangle; \langle Y^\top\rangle)$ is a concave function of $p_{\langle X^\top\rangle}$, the proof is completed. ■

V. Rank-Matrix Superposition Codes

Sumas codes may not always achieve rate higher than KK subspace coding. For example, subspace coding is optimal for rank-symmetric LOCs (see Section III-A). The optimality of subspace coding is based on the argument that $I(X; Y) = I(\langle X\rangle; \langle Y\rangle)$ [14]–[17], and as far as we know, no particular coding schemes have been proposed to achieve the capacity of rank-symmetric LOCs (cf. (3)).

We introduce a special class of Sumas codes, which can achieve the capacity of rank-symmetric LOCs. The superposition coding approach can potentially design codes with efficient encoding/decoding algorithms.

For $r = 0, 1, \ldots, \min\{M, T\}$, let $U(r)$ be an r-dimensional subspace of $\mathbb{F}^M$. An n-block *rank-matrix superposition (Ramas) code* with respect to $\{U(r)\}$ is an n-block Sumas code with the cloud code being a subset of $\{U(r)\}^n$. Denote by $[m]$ the set $\{0, 1, \ldots, m\}$. We can alternatively define a Ramas code code as follows.

Definition 4: An n-block rank-matrix superposition (Ramas) code with respect to $\{U(r)\}$ contains a cloud code $\mathcal{R} \subset [\min\{T, M\}]^n$ and a set of satellite codes, each of which corresponds to a codeword in the cloud code. The satellite code corresponding to $r^n \in \mathcal{R}$, denoted by $\mathcal{S}(r^n)$, is a subset of $\phi_T(U(r_1)) \times \cdots \times \phi_T(U(r_n))$.

Note that $\{U(r)\}$ in the above definition is a subset of $\mathrm{Pj}(\min\{T, M\}, \mathbb{F}^M)$ with all the elements having different ranks. The encoding of the cloud code is equivalent to mapping the message to the ranks of the input matrices. The constraints on the satellite codes given by $\{U(r)\}$ makes Ramas codes different from the existing designs of KK subspace codes. We are interested in the performance of Ramas codes under a modified subspace decoding rule where the cloud code is decoded using the ranks of the received matrices. The proof of the following theorem is omitted.

Theorem 4: The maximum achievable rates of Ramas codes under the modified subspace decoding rule is at least

$$\max_{p_{\langle X^\top\rangle}} [J(\mathrm{rk}(X); \mathrm{rk}(Y)) + I(\mathrm{rk}(X); \mathrm{rk}(Y))].$$

VI. Concluding Remarks

We propose codes with refined coding structures that have not been unveiled under the subspace coding framework. The design of a Sumas (Ramas) code depends on the distribution of the transfer matrix. Further works include how to characterize such dependence and how to design coding schemes based on the superposition structure.

Acknowledgment

This work was supported in part by the National Basic Research Program of China Grant 2011CBA00300, 2011CBA00301, the National Natural Science Foundation of China Grant 61033001, 61061130540, 61073174, and a grant from the University Grants Committee of the Hong Kong Special Administrative Region, China (Project No. AoE/E-02/08).

References

[1] S.-Y. R. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Trans. Inform. Theory*, vol. 49, no. 2, pp. 371–381, Feb. 2003.

[2] T. Ho, B. Leong, M. Medard, R. Koetter, Y. Chang, and M. Effros, "The benefits of coding over routing in a randomized setting," in *Proc. IEEE ISIT '03*, Jun. 2003.

[3] S. Yang, J. Meng, and E.-h. Yang, "Coding for linear operator channels over finite fields," in *Proc. IEEE Inte. Symp. on Information Theory ISIT '10*, Austin, USA, Jun. 2010.

[4] S. Yang and R. W. Yeung, "Coding for a network coded fountain," in *Proc. IEEE ISIT '11*, Saint Petersburg, Russia, 2011.

[5] ——, "Batched sparse codes," 2012. [Online]. Available: http://arxiv.org/abs/1206.5365

[6] R. Koetter and F. R. Kschischang, "Coding for errors and erasures in random network coding," *IEEE Trans. Inform. Theory*, vol. 54, no. 8, pp. 3579–3591, Aug. 2008.

[7] D. Silva, F. Kschischang, and R. Koetter, "A rank-metric approach to error control in random network coding," *IEEE Trans. Inform. Theory*, vol. 54, no. 9, pp. 3951–3967, Sept. 2008.

[8] D. Silva and F. Kschischang, "On metrics for error correction in network coding," *Information Theory, IEEE Transactions on*, vol. 55, no. 12, pp. 5479 –5490, dec. 2009.

[9] M. Gadouleau and Z. Yan, "Packing and covering properties of subspace codes for error control in random linear network coding," *Information Theory, IEEE Transactions on*, vol. 56, no. 5, pp. 2097 –2108, may 2010.

[10] R. W. Nóbrega and B. F. Uchôa-Filho, "Multishot codes for network coding: bounds and a multilevel construction," in *Proc. IEEE ISIT'09*, Jul. 2009.

[11] ——, "Multishot codes for network coding using rank-metric codes," in *Wireless Network Coding Conference (WiNC), 2010 IEEE*, june 2010.

[12] T. M. Cover, "Broadcast channels," *Information Theory, IEEE Transactions on*, vol. IT-18, pp. 2–14, 1972.

[13] G. E. Andrews, *The theory of partitions*, ser. vol. 2, Encyclopedia of mathematics and its applications. Addison-Wesley Pub. Co., 1976.

[14] M. Siavoshani, S. Mohajer, C. Fragouli, and S. Diggavi, "On the capacity of noncoherent network coding," *Information Theory, IEEE Transactions on*, vol. 57, no. 2, pp. 1046 –1066, feb. 2011.

[15] R. W. Nóbrega, B. F. Uchôa-Filho, and D. Silva, "On the capacity of multiplicative finite-field matrix channels," in *Proc. IEEE ISIT'11*, Saint Petersburg, Russia, August 2011.

[16] R. W. Nóbrega, D. Silva, and B. F. Uchôa-Filho, "On the capacity of multiplicative finite-field matrix channels," 2011. [Online]. Available: http://arxiv.org/abs/1105.6115

[17] S. Yang, S.-W. Ho, J. Meng, and E.-h. Yang, "Linear operator channels over finite fields," 2010. [Online]. Available: http://arxiv.org/abs/1002.2293v1

[18] ——, "Symmetry properties and subspace degradations of linear operator channels over finite fields," 2012. [Online]. Available: http://arxiv.org/abs/1108.4257

[19] T. M. Cover and J. A. Thomas, *Elements of information theory*, 2nd ed. Wiley, 2006.

[20] I. Csiszár and J. Körner, *Information theory*. Cambridge University Press, 2011.

[21] S. Yang, S.-W. Ho, J. Meng, and E.-h. Yang, "Optimality of subspace coding for linear operator channels over finite fields," in *Proc. IEEE ITW '10*, Cario, Egypt, Jan. 2010.