MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

Non-Coherent Grassmann TCM Design for Physical-Layer Network Coding in Bidirectional MIMO Relaying Systems

Koike-Akino, T.

TR2011-032 June 2011

Abstract

AbstractWe investigate a bidirectional relaying system which uses physical layer network coding and noncoherent multiinput multioutput (MIMO) signal processing. With noncoherent codes on Grassmannian manifold, a receiver employing generalized likelihood ratio test (GLRT) algorithm offers the maximumlikelihood performance even without any channel state information (CSI). We propose a new family of noncoherent Grassmann codes which enjoy a significant coding gain by introducing a trelliscoded modulation (TCM) through an affinelattice convolution with exponential mapping. We develop a design method of individual TCM code-books for multiple users in noncoherent two-way relaying channels with network coding. Since the proposed scheme does not require CSIs, it effectively deals with timevarying fading channels.

IEEE International Conference on Communications (ICC)

Non–Coherent Grassmann TCM Design for Physical–Layer Network Coding in Bidirectional MIMO Relaying Systems

Toshiaki Koike-Akino
Mitsubishi Electric Research Laboratories (MERL), 201 Broadway, Cambridge, MA 02139, U.S.A.
Email: koike@merl.com

***Abstract*—We investigate a bidirectional relaying system which uses physical–layer network coding and non–coherent multi–input multi–output (MIMO) signal processing. With non–coherent codes on Grassmannian manifold, a receiver employing generalized likelihood ratio test (GLRT) algorithm offers the maximum–likelihood performance even without any channel state information (CSI). We propose a new family of non–coherent Grassmann codes which enjoy a significant coding gain by introducing a trellis–coded modulation (TCM) through an affine–lattice convolution with exponential mapping. We develop a design method of individual TCM codebooks for multiple users in non–coherent two-way relaying channels with network coding. Since the proposed scheme does not require CSIs, it effectively deals with time–varying fading channels.**

I. Introduction

In the last decade, multi–way relaying which exploits network coding [1] at the physical layer has received a significant amount of attention [2–14]. The author has optimized signalling constellations for network–coded bidirectional relaying in [10, 11], and extended it for convolutionally–coded systems in [12], for adaptive modulations in [13], and for multi–input multi–output (MIMO) systems in [14]. We have found that physical–layer network coding should be adaptively changed according to the channel state information (CSI), and that non-linear network coding can significantly improve data throughput. The most of literature, including our previous works [10–14], has considered coherent detections which require CSI at receivers (a.k.a. CSIR). In this paper, we propose a new approach of network trellis–coded modulation (NetTCM) for non–CSIR scenarios by using non–coherent MIMO signalling.

Without CSI, we require *non–coherent* communications. Some information–theoretical studies on non–coherent MIMO communications, e.g. [15, 16], have motivated the signal design of non–coherent codes, which include unitary space–time constellations [17–19], exponential mapping Grassmann codes [20, 21], non–parametric Grassmann codes [22, 23], and differential space–time modulations [24, 25]. It was shown in [15] that unitary space–time codes asymptotically achieve the non–coherent channel capacity. For such codes, the maximum–likelihood (ML) performance can be offered by a generalized likelihood ratio test (GLRT) decoding [26].

The contribution of this paper includes a codebook design suited for non–coherent bidirectional MIMO relaying systems employing physical–layer network coding. We propose a new non–coherent TCM which uses affine–lattice convolution and exponential mapping on the Grassmannian manifold. Using the gradient approach, the design method jointly optimizes network coding and non–coherent Grassmann TCM in an iterative fashion. Through computer simulations, we demonstrate that the proposed NetTCM scheme significantly improves performance in non–CSIR scenarios.

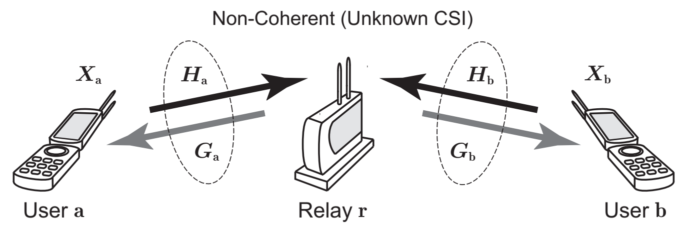

Fig. 1. Non–coherent bidirectional MIMO relaying systems.

Notations: Throughout the paper, we describe matrices and vectors by bold–face italic letters in capital cases and small cases, respectively. Let $\boldsymbol{X} \in \mathbb{C}^{m\times n}$ be a complex–valued $(m \times n)$–dimensional matrix, where $\mathbb{C}$ denotes the complex field. The notations $\boldsymbol{X}^*$, $\boldsymbol{X}^{\mathrm{T}}$, $\boldsymbol{X}^\dagger$, $\boldsymbol{X}^{-1}$, $\mathrm{tr}[\boldsymbol{X}]$, $\det[\boldsymbol{X}]$, and $\|\boldsymbol{X}\|$ represent the complex conjugate, the transpose, the Hermitian transpose, the inverse, the trace, the determinant, and the Frobenius norm of $\boldsymbol{X}$, respectively. The operator $\mathrm{vec}[\cdot]$ denotes the vector–operation which stacks all columns of a matrix into a single column vector in a left–to–right fashion, and the operator $\otimes$ stands for the Kronecker product of two matrices. The set of real numbers is denoted by $\mathbb{R}$, and $\boldsymbol{I}_m \in \mathbb{R}^{m\times m}$ denotes an m–dimensional identity matrix. The m–element integer ring is written as $\mathbb{Z}_m \triangleq \{0, 1, \ldots, m-1\}$.

II. Non–Coherent MIMO Two–Way Relaying

A. Two–Way Relaying Systems

We consider a non–coherent bidirectional relaying system in which there are three users, Alice a, Bob b, and Richard r, as depicted in Fig. 1. The radio transceivers for those users are equipped with M_{a}, M_{b}, and M_{r} antennas, respectively, for MIMO communications. Alice and Bob wish to exchange data by help of Richard, supposed that the direct wireless link between Alice and Bob is seriously attenuated. We focus on two–stage relaying protocol with physical–layer network coding.

At the first stage termed *multiple–access (MA) stage*, Alice and Bob simultaneously transmit data to Richard. Richard then broadcasts network–coded data to Alice and Bob in the following stage termed *broadcast (BC) stage*. In this paper, we investigate non–coherent communication systems wherein any users do not have CSI either for transmitting or receiving. For such a scenario, we use non–coherent Grassmann codes which do not need any pilot or training sequence.

B. Multiple–Access (MA) Stage

The source data $s_k(n)$ to be transmitted from the user $k \in \{\mathsf{a}, \mathsf{b}\}$ at the n–th symbol is drawn from an integer set $\mathbb{Z}_{Q_k}$ where Q_k denotes the alphabet size. The D–symbol data sequence $\boldsymbol{s}_k \triangleq [s_k(0), s_k(1), \ldots, s_k(D-1)]$ is sequentially buffered into an N_k memory as

$$\boldsymbol{s}_k(n) = \begin{bmatrix} s_k(n) & s_k(n-1) & \cdots & s_k(n-N_k+1) \end{bmatrix}^{\mathrm{T}}, \quad (1)$$

where N_k denotes the constraint length in symbol for trellis coding (thus, the number of trellis states becomes $Q_k^{N_k-1}$). The terminal user $k \in \{\mathsf{a}, \mathsf{b}\}$ transmits the multi–dimensional (or, space–time) TCM signals at the n-th symbol ($n \in \mathbb{Z}_D$):

$$\boldsymbol{X}_k(n) = \boldsymbol{\mathcal{C}}_k\big(\boldsymbol{s}_k(n)\big) \in \mathbb{C}^{M_k \times L}, \quad (2)$$

where the constellation mapper $\boldsymbol{\mathcal{C}}_k(\cdot)$ selects a codeword from the non–coherent Grassmann codebook $\mathbb{X}_k$:

$$\boldsymbol{X}_k(n) \in \mathbb{X}_k \triangleq \big\{\boldsymbol{\mathcal{X}}_k[0], \ldots, \boldsymbol{\mathcal{X}}_k[Q_k^{N_k}-1]\big\}. \quad (3)$$

The Grassmann TCM consists of L sub–symbols per codeword. We assume $L \geq 2(M_\mathsf{a} + M_\mathsf{b})$. The q-th codeword is denoted by $\boldsymbol{\mathcal{X}}_k[q]$. We will later describe a design method of the TCM codebook for network–coded relaying systems.

During the MA stage, Richard receives

$$\begin{aligned} \boldsymbol{Y}_\mathsf{r}(n) &= \boldsymbol{H}_\mathsf{a}(n)\boldsymbol{X}_\mathsf{a}(n) + \boldsymbol{H}_\mathsf{b}(n)\boldsymbol{X}_\mathsf{b}(n) + \boldsymbol{Z}_\mathsf{r}(n) \\ &= \underbrace{\begin{bmatrix} \boldsymbol{H}_\mathsf{a}(n) & \boldsymbol{H}_\mathsf{b}(n) \end{bmatrix}}_{\boldsymbol{H}_\mathsf{ab}(n)} \underbrace{\begin{bmatrix} \boldsymbol{X}_\mathsf{a}(n) \\ \boldsymbol{X}_\mathsf{b}(n) \end{bmatrix}}_{\boldsymbol{X}_\mathsf{ab}(n)} + \boldsymbol{Z}_\mathsf{r}(n), \end{aligned} \quad (4)$$

where $\boldsymbol{Y}_\mathsf{r}(n) \in \mathbb{C}^{M_\mathsf{r} \times L}$, $\boldsymbol{H}_\mathsf{a}(n) \in \mathbb{C}^{M_\mathsf{r} \times M_\mathsf{a}}$, $\boldsymbol{H}_\mathsf{b}(n) \in \mathbb{C}^{M_\mathsf{r} \times M_\mathsf{b}}$, and $\boldsymbol{Z}_\mathsf{r}(n) \in \mathbb{C}^{M_\mathsf{r} \times L}$ denote the received signal, the MIMO channel matrix from Alice to Richard, the MIMO channel matrix from Bob to Richard, and the additive noise at the n-th time instance. The compound channel matrix $\boldsymbol{H}_\mathsf{ab}(n) \in \mathbb{C}^{M_\mathsf{r} \times (M_\mathsf{a}+M_\mathsf{b})}$ may rapidly change along a time instance n and, hence, it is hard for receivers to accurately estimate the channels. The joint codeword $\boldsymbol{X}_\mathsf{ab}(n) \in \mathbb{X}_\mathsf{a} \times \mathbb{X}_\mathsf{b}$ shall be well designed so that Richard can reliably decode it in non–coherent multiple–access channels.

The GLRT decoding [26] (namely, blind ML decoding for non–coherent receivers) is performed to obtain the most–likely estimates, $\hat{\boldsymbol{s}}_\mathsf{a}$ and $\hat{\boldsymbol{s}}_\mathsf{b}$, as follows:

$$\begin{aligned} \big\{\hat{\boldsymbol{s}}_\mathsf{a}, \hat{\boldsymbol{s}}_\mathsf{b}\big\} &= \mathop{\arg\min}_{\hat{\boldsymbol{s}}_\mathsf{a}, \hat{\boldsymbol{s}}_\mathsf{b}} \inf_{\hat{\boldsymbol{H}}_\mathsf{ab}(n)} \sum_{n \in \mathbb{Z}_D} \Big\| \boldsymbol{Y}_\mathsf{r}(n) - \hat{\boldsymbol{H}}_\mathsf{ab}(n)\hat{\boldsymbol{X}}_\mathsf{ab}(n) \Big\|^2 \\ &= \mathop{\arg\min}_{\hat{\boldsymbol{s}}_\mathsf{a}, \hat{\boldsymbol{s}}_\mathsf{b}} \sum_{n \in \mathbb{Z}_D} \Big\| \boldsymbol{Y}_\mathsf{r}(n) \hat{\boldsymbol{X}}_\mathsf{ab}^{\perp}(n) \Big\|^2, \end{aligned} \quad (5)$$

where $\hat{\boldsymbol{X}}_\mathsf{ab}^{\perp}(n)$ is an orthogonal projection matrix of an estimate $\hat{\boldsymbol{X}}_\mathsf{ab}(n)$ such that $\hat{\boldsymbol{X}}_\mathsf{ab}(n)\hat{\boldsymbol{X}}_\mathsf{ab}^{\perp}(n) = \mathbf{0}$. An orthogonal projection is written as

$$\hat{\boldsymbol{X}}_\mathsf{ab}^{\perp}(n) = \boldsymbol{I}_L - \hat{\boldsymbol{X}}_\mathsf{ab}^{\dagger}(n)\Big(\hat{\boldsymbol{X}}_\mathsf{ab}(n)\hat{\boldsymbol{X}}_\mathsf{ab}^{\dagger}(n)\Big)^{-1}\hat{\boldsymbol{X}}_\mathsf{ab}(n), \quad (6)$$

which is an idempotent matrix whose eigenvalue is either one or zero (its rank is at most $L - M_\mathsf{a} - M_\mathsf{b}$). The joint projection set is denoted as $\mathbb{X}_\mathsf{ab}^{\perp} \triangleq \big\{\boldsymbol{\mathcal{X}}_\mathsf{ab}^{\perp}[0], \ldots, \boldsymbol{\mathcal{X}}_\mathsf{ab}^{\perp}[Q_\mathsf{a}^{N_\mathsf{a}} Q_\mathsf{b}^{N_\mathsf{b}} - 1]\big\}$. Note that the orthogonal projection matrix is predetermined from the original codebooks $\mathbb{X}_\mathsf{a}$ and $\mathbb{X}_\mathsf{b}$. This GLRT sequence estimation along the trellis–state diagram enables Richard to decode TCM signals even without CSI if the matrix $\hat{\boldsymbol{X}}_\mathsf{ab}(n)\hat{\boldsymbol{X}}_\mathsf{ab}^{\dagger}(n)$ is non–singular. We propose a design method of such a codebook to maximize the codeword distance.

C. Broadcast (BC) Stage with Network Coding

With the ML estimates $\hat{s}_\mathsf{a}(n)$ and $\hat{s}_\mathsf{b}(n)$, Richard first generates a combined data $s_\mathsf{r}(n)$ using a network coding function $f(\cdot)$. We may use a network coding based on a modulo addition over $\mathbb{Z}_{Q_\mathsf{r}}$:

$$s_\mathsf{r}(n) = f\big(\hat{s}_\mathsf{a}(n), \hat{s}_\mathsf{b}(n)\big) = \hat{s}_\mathsf{a}(n) + \hat{s}_\mathsf{b}(n) \bmod Q_\mathsf{r}, \quad (7)$$

with an alphabet size of $Q_\mathsf{r} = \max(Q_\mathsf{a}, Q_\mathsf{b})$. The network coding function will be optimized according to the TCM codebooks. The network–coded data $s_\mathsf{r}(n)$ is buffered to construct an N_r–symbol constraint vector $\boldsymbol{s}_\mathsf{r}(n) \triangleq [s_\mathsf{r}(n), \ldots, s_\mathsf{r}(n - N_\mathsf{r} + 1)]^{\mathrm{T}}$, and it is mapped on the Grassmann TCM:

$$\boldsymbol{X}_\mathsf{r}(n) = \boldsymbol{\mathcal{C}}_\mathsf{r}\big(\boldsymbol{s}_\mathsf{r}(n)\big) \in \mathbb{C}^{M_\mathsf{r} \times L'}, \quad (8)$$

where the codeword is drawn from the codebook $\mathbb{X}_\mathsf{r} \triangleq \{\boldsymbol{\mathcal{X}}_\mathsf{r}[0], \ldots, \boldsymbol{\mathcal{X}}_\mathsf{r}[Q_\mathsf{r}^{N_\mathsf{r}} - 1]\}$. The sub–symbol length L' can be optimized. The network TCM (NetTCM) signals are broadcasted from Richard to Alice and Bob, who in turn receive

$$\boldsymbol{Y}_\mathsf{a}(n) = \boldsymbol{G}_\mathsf{a}(n)\boldsymbol{X}_\mathsf{r}(n) + \boldsymbol{Z}_\mathsf{a}(n), \quad (9)$$

$$\boldsymbol{Y}_\mathsf{b}(n) = \boldsymbol{G}_\mathsf{b}(n)\boldsymbol{X}_\mathsf{r}(n) + \boldsymbol{Z}_\mathsf{b}(n), \quad (10)$$

respectively. The BC channel matrices $\boldsymbol{G}_\mathsf{a}(n) \in \mathbb{C}^{M_\mathsf{a} \times M_\mathsf{r}}$ and $\boldsymbol{G}_\mathsf{b}(n) \in \mathbb{C}^{M_\mathsf{b} \times M_\mathsf{r}}$ change in time and, thus, can be independent of the MA channels, $\boldsymbol{H}_\mathsf{a}(n)$ and $\boldsymbol{H}_\mathsf{b}(n)$.

With the side information $\boldsymbol{s}_\mathsf{a}$, the GLRT sequence estimation for Alice is employed as follows

$$\hat{\boldsymbol{s}}_\mathsf{b}' = \mathop{\arg\min}_{\hat{\boldsymbol{s}}_\mathsf{b}: \hat{\boldsymbol{s}}_\mathsf{r} = f(\boldsymbol{s}_\mathsf{a}, \hat{\boldsymbol{s}}_\mathsf{b})} \sum_{n \in \mathbb{Z}_D} \Big\| \boldsymbol{Y}_\mathsf{a}(n) \hat{\boldsymbol{X}}_\mathsf{r}^{\perp}(n) \Big\|^2. \quad (11)$$

In an analogous way, Bob obtains the ML estimate $\hat{\boldsymbol{s}}_\mathsf{a}'$ using the own information $\boldsymbol{s}_\mathsf{b}$.

III. Non–Coherent Grassmann TCM Design

A number of non–coherent codes have been reported, e.g., unitary space–time codes [17–19], Grassmann codes with exponential mapping [20, 21], Grassmann packing codes with numerical optimization [22, 23], and differential modulations [24, 25]. In this paper, we incorporate non–coherent space–time block codes with space–time trellis codes by using the exponential mapping and affine–lattice convolution techniques.

A. Design Criteria: Pairwise Error Probability

During the MA stage, a pairwise error probability is expressed as

$$\begin{aligned}&\Pr\big(\{\boldsymbol{X}_{\mathsf{ab}}(n)\} \to \{\boldsymbol{X}'_{\mathsf{ab}}(n)\} \mid \{\boldsymbol{H}_{\mathsf{ab}}(n)\}\big)\\ &\simeq \frac{1}{2}\mathrm{erfc}\sqrt{\frac{\sum_n \big\|\boldsymbol{H}_{\mathsf{ab}}(n)\boldsymbol{X}_{\mathsf{ab}}(n)\boldsymbol{X}'^{\perp}_{\mathsf{ab}}(n)\big\|^2}{4\sigma^2}},\end{aligned} \tag{12}$$

given instantaneous channels $\{\boldsymbol{H}_{\mathsf{ab}}(n)\}$. Here, σ^2 denotes the noise variance. The error probability averaged over *i.i.d.* Rayleigh fading channels of unity power is given as

$$\begin{aligned}&\mathbb{E}_{\boldsymbol{H}_{\mathsf{ab}}(n)}\Big[\Pr\big(\{\boldsymbol{X}_{\mathsf{ab}}(n)\} \to \{\boldsymbol{X}'_{\mathsf{ab}}(n)\} \mid \{\boldsymbol{H}_{\mathsf{ab}}(n)\}\big)\Big]\\ &\simeq \begin{cases} \displaystyle\frac{1}{2}\prod_n\prod_m\left(1+\frac{\lambda_m\big[\boldsymbol{\Omega}(n)\big]}{4\sigma^2}\right)^{-M_{\mathsf{r}}}, & \text{(fast fading)},\\ \displaystyle\frac{1}{2}\prod_m\left(1+\frac{\lambda_m\big[\sum_n\boldsymbol{\Omega}(n)\big]}{4\sigma^2}\right)^{-M_{\mathsf{r}}}, & \text{(slow fading)},\end{cases}\end{aligned} \tag{13}$$

where $\lambda_m[\cdot]$ is the m-th eigenvalue of a matrix and

$$\boldsymbol{\Omega}(n) = \boldsymbol{X}_{\mathsf{ab}}(n)\boldsymbol{X}'^{\perp}_{\mathsf{ab}}(n)\boldsymbol{X}^{\dagger}_{\mathsf{ab}}(n). \tag{14}$$

For high and low SNRs, we should respectively maximize the determinant and the trace in fast fading as follows

$$d^2_{\det} = \min\prod_n \det\big[\boldsymbol{\Omega}(n)\big], \tag{15}$$

$$d^2_{\mathrm{tr}} = \min\sum_n \mathrm{tr}\big[\boldsymbol{\Omega}(n)\big]. \tag{16}$$

In slow fading channels, while the trace criterion does not differ from the one above, the determinant criterion becomes

$$d'^2_{\det} = \min\det\big[\sum_n \boldsymbol{\Omega}(n)\big]. \tag{17}$$

Note that the codewords pair such that $f(\boldsymbol{s}_{\mathsf{a}}, \boldsymbol{s}_{\mathsf{b}}) = f(\boldsymbol{s}'_{\mathsf{a}}, \boldsymbol{s}'_{\mathsf{b}})$ does not incur decoding error at the destination with the network coding function $f(\cdot)$. In this paper, we design the TCM codebook for network coding to minimize the pairwise error probability based on the above–mentioned criteria.

B. Gradient Optimization of NetTCM

Now, we numerically optimize non–coherent NetTCM codebook in a similar way of [10, 23, 27] with a TCM extension. For numerical Grassmann packing, we adopt the gradient method to optimize the determinant and the trace criteria.

Given the minimum determinant metric $d^2_{\det}$ for an erroneous path in the trellis–state diagram, the gradient is expressed over log–domain as follows

$$\frac{\partial \log d^2_{\det}}{\partial \boldsymbol{X}^{*}_{\mathsf{ab}}(n)} = \boldsymbol{\Omega}^{-1}(n)\boldsymbol{X}_{\mathsf{ab}}(n)\boldsymbol{X}'^{\perp}_{\mathsf{ab}}(n), \tag{18}$$

$$\frac{\partial \log d^2_{\det}}{\partial \boldsymbol{X}'^{*}_{\mathsf{ab}}(n)} = -\boldsymbol{X}'^{+}_{\mathsf{ab}}(n)\boldsymbol{X}^{\dagger}_{\mathsf{ab}}(n)\boldsymbol{\Omega}^{-1}(n)\boldsymbol{X}_{\mathsf{ab}}(n)\boldsymbol{X}'^{\perp}_{\mathsf{ab}}(n). \tag{19}$$

Here, $[\cdot]^{+}$ denotes the pseudo inverse. In slow fading channels, the gradient of the determinant metric is modified by replacing $\boldsymbol{\Omega}(n)$ with $\sum_n \boldsymbol{\Omega}(n)$. For the minimum trace metric d^2_{tr}, we have the gradient

$$\frac{\partial d^2_{\mathrm{tr}}}{\partial \boldsymbol{X}^{*}_{\mathsf{ab}}(n)} = \boldsymbol{X}_{\mathsf{ab}}(n)\boldsymbol{X}'^{\perp}_{\mathsf{ab}}(n), \tag{20}$$

$$\frac{\partial d^2_{\mathrm{tr}}}{\partial \boldsymbol{X}'^{*}_{\mathsf{ab}}(n)} = -\boldsymbol{X}'^{+}_{\mathsf{ab}}(n)\boldsymbol{X}^{\dagger}_{\mathsf{ab}}(n)\boldsymbol{X}_{\mathsf{ab}}(n)\boldsymbol{X}'^{\perp}_{\mathsf{ab}}(n). \tag{21}$$

Our joint design method of TCM codebook and network coding function based on the gradient optimization is described below:

1: Generate random codewords $\{\boldsymbol{\mathcal{X}}_{\mathsf{a}}[q] : q \in \mathbb{Z}_{Q_{\mathsf{a}}^{N_{\mathsf{a}}}}\}$ and $\{\boldsymbol{\mathcal{X}}_{\mathsf{b}}[q] : q \in \mathbb{Z}_{Q_{\mathsf{b}}^{N_{\mathsf{b}}}}\}$ such that $\|\boldsymbol{\mathcal{X}}_k[q]\|^2 = M_k$
2: Set initial network coding function $f(s_{\mathsf{a}}, s_{\mathsf{b}}) = s_{\mathsf{a}} + s_{\mathsf{b}} \pmod{Q_{\mathsf{r}}}$ for $s_{\mathsf{a}} \in \mathbb{Z}_{Q_{\mathsf{a}}}$ and $s_{\mathsf{b}} \in \mathbb{Z}_{Q_{\mathsf{b}}}$
3: Calculate $\boldsymbol{\Omega}(n)$ for any erroneous pairs in the trellis-state diagram
4: Search for the worst pair which has the minimum metric in determinant or trace, and yields different network codewords, $f(\boldsymbol{s}_{\mathsf{a}}, \boldsymbol{s}_{\mathsf{b}}) \neq f(\boldsymbol{s}'_{\mathsf{a}}, \boldsymbol{s}'_{\mathsf{b}})$
5: Calculate the gradient $\boldsymbol{\nabla}$ for the pair
6: Update codewords as $\boldsymbol{\mathcal{X}}_k[q] \leftarrow \boldsymbol{\mathcal{X}}_k[q] + \beta\boldsymbol{\nabla}$, where $\beta \in \mathbb{R}$ is a stepsize factor which is optimized by line searching to maximize the determinant or trace metric
7: Normalize the energy such that $\|\boldsymbol{\mathcal{X}}_k[q]\|^2 = M_k$
8: Repeat from 3 until convergence of TCM
9: Optimize network coding function $f(\cdot)$ based on updated metrics by closest–neighbor clustering method [10, 11]
10: Repeat from 3 until convergence of network coding

Using multiple initial codewords or small perturbations of optimized codebook, the gradient method yields well–designed codebook. Note that a pairwise error probability during the BC stage can be also minimized by designing the TCM codebook $\mathbb{X}_{\mathsf{r}}$ in a similar way.

C. Non–Coherent TCM with Exponential Mapping

In order to boost the convergence of the gradient optimization, we propose a new family of Grassmann TCM design by introducing an affine–lattice convolution with exponential mapping. As in [20, 21], each codeword is mapped on the Grassmannian manifold such that $\boldsymbol{\mathcal{X}}_k[q]\boldsymbol{\mathcal{X}}^{\dagger}_k[q] = \boldsymbol{I}_{M_k}$ by exponential mapping

$$\boldsymbol{X}_k(n) = \begin{bmatrix}\boldsymbol{I}_{M_k} & \boldsymbol{0}_{M_k\times(L-M_k)}\end{bmatrix}\exp\left(\begin{bmatrix}\boldsymbol{0}_{M_k} & \boldsymbol{B}_k(n)\\ -\boldsymbol{B}^{\dagger}_k(n) & \boldsymbol{0}_{L-M_k}\end{bmatrix}\right), \tag{22}$$

where $\boldsymbol{B}_k(n) \in \mathbb{C}^{M_k\times(L-M_k)}$ is a multi–dimensional affine–lattice convolution:

$$\mathrm{vec}\big[\boldsymbol{B}_k(n)\big] = \boldsymbol{\Theta}_k\boldsymbol{s}_k(n) + \boldsymbol{\theta}_k. \tag{23}$$

The lattice generator matrix $\boldsymbol{\Theta}_k \in \mathbb{C}^{M_k(L-M_k)\times N_k}$ and the affine shift vector $\boldsymbol{\theta}_k \in \mathbb{C}^{M_k(L-M_k)\times 1}$ are optimized later. The chief difference from the one proposed in [20, 21] lies in the extension to trellis coding by the affine–lattice convolution.

D. Optimization for Exponential Mapping Grassmann Codes

We use the aforementioned gradient method to design the exponential mapping Grassmann TCM by optimizing the lattice generating matrix $\boldsymbol{\Theta}_k$ and the affine shift vector $\boldsymbol{\theta}_k$. The gradient of the determinant or trace metrics in terms of $\boldsymbol{\Theta}_k$ and $\boldsymbol{\theta}_k$ is obtained as follows

$$\frac{\partial \eta}{\partial \gamma^*} = \sum_m \mathrm{tr}\Big[\frac{\partial \boldsymbol{X}_{\mathsf{ab}}^{\dagger}(m)}{\partial \gamma^*}\frac{\partial \eta}{\partial \boldsymbol{X}_{\mathsf{ab}}^{*}(m)}\Big], \qquad (24)$$

where $\eta \in \{d_{\det}^2, d_{\mathrm{tr}}^2, d_{\det}'^2\}$ is an optimizing metric, and $\gamma \in \{\boldsymbol{\Theta}_k, \boldsymbol{\theta}_k\}$ is a parameter to be optimized. The affine convolution with exponential mapping can significantly reduce the total number of parameters to be optimized from $Q_{\mathsf{a}}^{N_{\mathsf{a}}} M_{\mathsf{a}} L + Q_{\mathsf{b}}^{N_{\mathsf{b}}} M_{\mathsf{b}} L$ to $M_{\mathsf{a}}(L - M_{\mathsf{a}})(N_{\mathsf{a}} + 1) + M_{\mathsf{b}}(L - M_{\mathsf{b}})(N_{\mathsf{b}} + 1)$; e.g., from 512 to 36 for $Q_{\mathsf{a}} = Q_{\mathsf{b}} = 4$, $L = 8$, $M_{\mathsf{a}} = M_{\mathsf{b}} = 2$ and $N_{\mathsf{a}} = N_{\mathsf{b}} = 2$. It results in a faster convergence in the gradient optimization.

IV. Performance Evaluation

Now, we show the performance advantage of our optimized NetTCM over the conventional schemes through computer simulations. We use a sub–symbol length of $L = L' = 8$ for non–coherent Grassmann codes with a cardinality of $Q_{\mathsf{a}} = Q_{\mathsf{b}} = 4$. The constraint length is set to be $N_{\mathsf{a}} = N_{\mathsf{b}} = 2$. Every node uses two antennas, i.e., $M_{\mathsf{a}} = M_{\mathsf{b}} = M_{\mathsf{r}} = 2$. The transmission block sequence consists of $D = 64$ symbols. We assume that the channel is frequency–flat Rayleigh fading with the maximum Doppler frequency $f_{\mathrm{D}} T_{\mathrm{s}} = 1/200$ where T_{s} denotes the symbol duration. The CSI is not available at any users, and the GLRT decoding algorithm is employed. We assume the average SNR is identical for the link between Alice and Richard and that between Bob and Richard, for simplicity. The symbol timing is assumed to be synchronized for Alice and Bob.

Fig. 2 shows end–to–end bit error rate (BER) performance as a function of average SNR. The performance curve of original Grassmann codes in [20, 21] is presented as a reference. Since both Alice and Bob use the same Grassmann codes without TCM, the performance is severely degraded due to the multiple–access interference (MAI) during the MA stage. The MAI effect still degrades the performance of non–coherent TCMs which achieves coding gains if we use the identical codebooks. One can observe that our design method significantly improves performance. It is because the codebooks are well optimized to be near–orthogonal so that the MAI is effectively reduced at the relaying node. Our proposed Grassmann NetTCM offers additional 2.5 dB gains at a BER of 10^{-4}. It is expected that the performance is further improved by using high–order super–block GLRT decoding proposed in [27], for fast fading MIMO channels of non–CSIR scenarios.

V. Conclusion

In this paper, we proposed a blind network trellis–coded modulation (NetTCM) which jointly optimizes non–coherent Grassmann constellations, trellis coding, and network coding

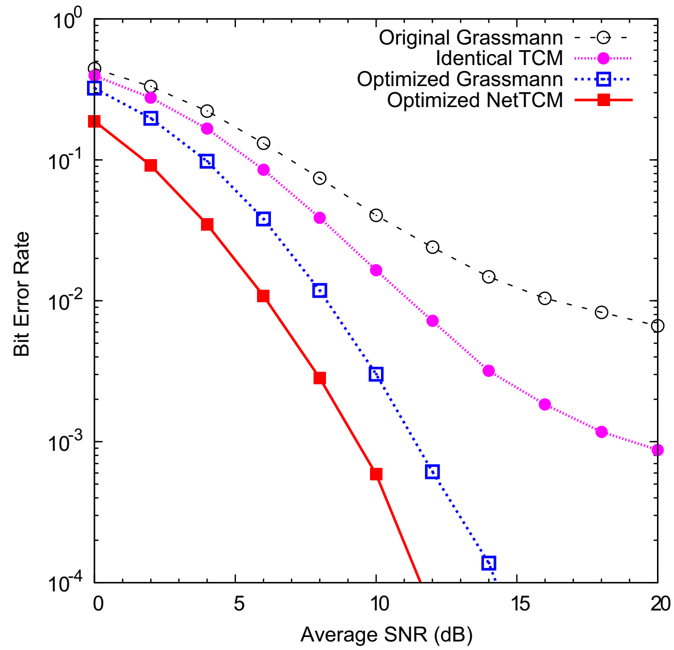

Fig. 2. End–to–end BER versus average SNR for bidirectional MIMO relaying systems in frequency–flat Rayleigh fading ($f_{\mathrm{D}} T_{\mathrm{s}} = 1/200$).

for bidirectional MIMO relaying systems in which CSI is not available at either transmitter or receiver. We optimized the blind NetTCM to minimize the pairwise error probability through the use of the gradient method as a practical sphere packing over the Grassmannian manifold. In addition, we introduced an affine–lattice convolution with exponential mapping to improve the convergence speed of the gradient method. It was demonstrated that the designed NetTCM offers a significant performance improvement over the conventional schemes for non–CSIR scenarios.

References

[1] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, “Network information flow,” *IEEE Trans. IT*, vol. IT-46, pp. 1204–1216, June 2000.

[2] Y. Wu, P. A. Chou, and S.-Y. Kung, “Information exchange in wireless networks with network coding and physical-layer broadcast,” Microsoft, Tech. Rep. MSR-TR-2004-78, Aug. 2004.

[3] P. Popovski and H. Yomo, “Bi-directional amplification of throughput in a wireless multi-hop network,” *IEEE VTC2006*, Melbourne, May 2006.

[4] S. Katti, D. Katabi, W. Hu, H. Rahul, and M. Médard, “The importance of being opportunistic: Practical network coding for wireless environments,” *Allerton Conf. Commun., Control, and Comput.*, Sept. 2005.

[5] S. Zhang, S. C. Liew, and P. P. Lam, “Hot topic: Physical-layer network coding,” *MobiCom*, pp. 358–365, Sept. 2006,

[6] S. J. Kim, P. Mitran, and V. Tarokh, “Performance bounds for bidirectional coded cooperation protocols,” *IEEE Trans. IT*, vol. 54, no. 11, pp. 5235–5241, Nov. 2008.

[7] P. Popovski and H. Yomo, “Physical network coding in two-way wireless relay channels,” *IEEE ICC*, Glasgow, Scotland, June 2007.

[8] K. Narayanan, M. P. Wilson, and A. Sprintson, “Joint physical layer coding and network coding for bi-directional relaying,” Allerton Conference on Communication, Control and Computing, Monticello, 2007.

[9] B. Nazer and M. Gastpar, “Computation over multiple access channels,” *IEEE Trans. IT*, vol. 53, no. 10, pp. 3498–3516, Oct. 2007.

[10] T. Koike-Akino, P. Popovski, and V. Tarokh, “Denoising maps and constellations for wireless network coding in two–way relaying systems,” *IEEE GLOBECOM*, New Orleans, U.S.A., Nov.–Dec. 2008.

[11] T. Koike-Akino, P. Popovski and V. Tarokh, “Optimized constellations for two–way wireless relaying with physical network coding,” *IEEE JSAC*, vol. 27, no. 5, pp. 773–787, June 2009.

[12] T. Koike-Akino, P. Popovski, and V. Tarokh, “Denoising strategy for convolutionally–coded bidirectional relaying,” *IEEE ICC*, Dresden, Germany, June 2009.

[13] T. Koike-Akino, P. Popovski, and V. Tarokh, “Adaptive modulation and network coding with optimized precoding in two–way relaying,” *IEEE GLOBECOM*, Honolulu, Hawaii, Dec. 2009.

[14] T. Koike-Akino, “Adaptive network coding in two–way relaying MIMO systems,” *IEEE GLOBECOM*, Miami, Dec. 2010.

[15] T. Marzetta and B. Hochwald, “Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading,” *IEEE Trans. IT*, vol. 45, pp. 139–157, 1999.

[16] L. Zheng and D. N. C. Tse, “Communication on the Grassmann manifold: a geometric approach to the noncoherent multiple-antenna channel,” *IEEE Trans. IT*, vol. 48, no. 2, pp. 359–384, Feb. 2002.

[17] B. M. Hochwald, T. L. Marzetta, T. J. Richardson, W. Sweldens, and R. Urbanke, “Systematic design of unitary space–time constellations,” *IEEE Trans. IT*, vol. 46, no. 6, Sept. 2000.

[18] Y. Jing and B. Hassibi, “Unitary space-time modulation via Cayley transform,” *IEEE Trans. Signal Proc.*, vol. 51, no. 11, pp. 2891–2904, Nov. 2003.

[19] B. Hochwald and W. Sweldens, “Differential unitary space-time modulation,” *IEEE Trans. Commun.*, vol. 48, pp. 2041–2052, Dec. 2000.

[20] I. Kammoun, A. M. Cipriano, and J.-C. Belfiore, “Non-coherent codes over the Grassmannian,” *IEEE Trans. Wireless Commun.*, vol. 6, no. 10, pp. 3657–3667, Oct. 2007.

[21] I. Kammoun and J. Belfiore, “A new family of Grassmann space-time codes for non-coherent MIMO systems,” *IEEE Commun. Lett.*, vol. 7, no. 11, pp. 528–530, Nov. 2003.

[22] M. J. Borran, A. Sabharwal, and B. Aazhang, “On design criteria and construction of non-coherent space-time constellations,” *IEEE Trans. IT*, vol. 49, no. 10, pp. 2332–2351, Oct. 2003.

[23] M. Beko, J. Xavier, and V. Barroso, “Codebook design for non-coherent communication in multiple-antenna systems,” *IEEE ICASSP*, Toulouse, France, 2006.

[24] B. L. Hughes, “Differential space–time modulation,” *IEEE Trans. IT*, vol. 46, no. 7, pp. 2567–2578, Nov. 2000.

[25] V. Tarokh and M. Kim, “Existence and construction of noncoherent unitary space-time codes,” *IEEE Trans. IT*, vol. 25, no. 8, pp. 3112–3120, Dec. 2002.

[26] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*, New York: Addison-Wesley Publishing Co., 1990.

[27] T. Koike-Akino and P. Orlik, “High–order super–block GLRT for non–coherent Grassmann codes in MIMO–OFDM systems,” *IEEE GLOBECOM*, Miami, Dec. 2010.