ISIT 2009, Seoul, Korea, June 28 - July 3, 2009

# A Systematic Space-Time Code Design and Its Maximum-Likelihood Decoding for Combined Channel Estimation and Error Correction

Chia-Lung Wu*, Mikael Skoglund†, Po-Ning Chen* and Yunghsiang S. Han‡
*Dept. of Comm. Eng., National Chiao-Tung Univ., Taiwan, ROC
Email: clwu@banyan.cm.nctu.edu.tw, qponing@mail.nctu.edu.tw
†School of Electrical Engineering, Royal Institute of Technology, Sweden
Email: skoglund@ee.kth.se
‡Graduate Institute of Comm. Eng., National Taipei Univ., Taiwan, ROC
Email: yshan@mail.ntpu.edu.tw

***Abstract*—Several previous works have confirmed that a joint design that combines channel estimation, channel coding and space-time transmission can improve the system performance over that of a separate design. These conclusions are however in general based on unstructured solutions obtained using computer search. The coding gain of these joint designs is therefore limited by both the computer-searchable "short" code length and the compromise between "suboptimal" performance and "high" complexity of their optimal decoding.**

**At this background, we propose a systematic space-time code construction for joint channel estimation and error correction for a two-transmit-antenna and half-rate system. Also proposed is its *maximum-likelihood* decoder that follows a priority-first search principle. Our systematic code construction, together with a fairly low-complexity optimal decoder, then allows one to work with longer codes with no sacrifice in performance. For codes of short block length, our simulations illustrate that the codes we propose have comparable performance to the best computer-searched codes. For codes of long block lengths that are almost beyond the searchable range of existing computer systems, our codes are still better than some reference designs based on separate channel estimation and error correction components.**

## I. Introduction

Coding and transmission schemes for noncoherent receivers used in multiple-input multiple-output (MIMO) flat-fading channels can be roughly classified into two categories.[1] Schemes in the first category devise the space-time *constellations* for a given noncoherent receiver structure using computer search [1], [3], [10], while schemes in the second category couple the well-known space-time block codes with blind detection [11], [12], [15]. A brief summary of these schemes is as follows.

In [1], Beko *et al.* propose a two-phase code design approach, where the first phase produces a rough space-time code constellation that is subsequently refined in the second phase through a search-based geodesic descent optimization algorithm (GDA). In [3], Borran *et al.* uses the Kullback-Leibler distance as a design criterion to partition the signal space into several subsets, resulting in a reduction of number of parameters to be computer-searched. The authors in [10] construct unitary space-time signals by random search upon a Fourier-based structure, which only requires optimizing $L-1$ parameters instead of $L(L-1)/2$ in the correlation matrix, where $L$ is the number of space-time signals.

On the other hand, [11], [12] and [15] incorporate blind detection to existing space-time block codes. Based on the semidefinite relaxation (SDR) approach, an efficient suboptimal blind detection scheme is also suggested by Ma *et al.* in [11]. Later in [12], Ma further addresses the necessary properties for the family of orthogonal space-time block codes that can well co-work with blind detection.

Two main problems of designing codes or signal constellations based on unconstrained computer-search are that the design complexity is in general high, especially for codes of long block length, and the codes often need to be redesigned when design assumptions change. Moreover, their decoding depends mostly on operationally intensive exhaustive search, which further prevents their practical use in the case of long block lengths. Obviously, these problems can be solved by realizing a systematic code construction and its respective low-complexity decoder. Such an approach designed under two-transmit-antenna and half-rate condition is presented in this paper.

Furthermore, one main difference between our work and the existing works on combining known space-time block codes with blind detection, is that we aim at achieving a coding gain in contrast to targeting only improved diversity gains at maximum rate.

The paper is organized in the following fashion. Section II introduces the system model. Section III presents our code

[1] There are some notable papers that deal with similar problems, but cannot be classified into the two categories. For example, both [5] and [6] consider the so-called *training codes* that incorporate training symbols into their codewords. As anticipated, the receiver estimates the channel coefficients via training symbols. Such designs are very different from ours, which combines channel estimation and error correction by adopting joint maximum-likelihood decoding at the receiver. In [4], a noncoherent code is constructed through a mapping from coherent code. The code structure however only allows for a suboptimal efficient decoder.

978-1-4244-4313-0/09/$25.00 ©2009 IEEE

design scheme that is devised based on the unitary and full-rank properties. Section IV derives the maximum-likelihood metric that can be used by priority-first search decoding. Simulations are summarized and discussed in Section V.

In this work, superscripts "$H$" and "$T$" are specifically reserved for the matrix operations of Hermitian transpose and transpose, respectively.

## II. System Model

We consider an MIMO system with $A_{\rm T}$ transmit antennas and $A_{\rm R}$ receive antennas. The $N \times A_{\rm R}$ complex received matrix $\mathbb{Y} = [\boldsymbol{y}_1\ \boldsymbol{y}_2\ \ldots\ \boldsymbol{y}_{A_{\rm R}}]$ is then given by

$$\mathbb{Y} = \mathbb{B}\mathbb{H} + \mathbb{N},$$

where $\mathbb{B} = [\boldsymbol{b}_1\ \boldsymbol{b}_2\ \ldots\ \boldsymbol{b}_{A_{\rm T}}]$ is the $N \times A_{\rm T}$ transmitted code matrix, and $\mathbb{N} = [\boldsymbol{n}_1\ \boldsymbol{n}_2\ \ldots\ \boldsymbol{n}_{A_{\rm R}}]$ is an $N \times A_{\rm R}$ zero-mean complex Gaussian matrix with independent and identically distributed elements and covariance matrix

$$E[\boldsymbol{n}_i \boldsymbol{n}_i^H] = \sigma^2 \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}_{N\times N}.$$

Also, $\boldsymbol{b}_i = [b_{1,i}\ b_{2,i}\ \ldots\ b_{N,i}]^T$ is the bipolar codeword transmitted by antenna $i$ with each $b_{i,n} \in \{\pm 1/\sqrt{A_{\rm T}}\}$. Likewise, $\boldsymbol{y}_j = [y_{1,j}\ y_{2,j}\ \ldots\ y_{N,j}]^T$ is the received vector at the $j$th receive antenna.

Because $\mathbb{H}$ is assumed an unknown constant matrix, the Gaussian assumption on the additive noise matrix $\mathbb{N}$ immediately gives that the maximum-likelihood (ML) decision about the transmitted codeword should be made based on the generalized likelihood ratio test (GLRT) as

$$\begin{aligned} \hat{\mathbb{B}} &= \arg\min_{\mathbb{B}} \min_{\mathbb{H}} \|\mathbb{Y} - \mathbb{B}\mathbb{H}\|^2 \\ &= \arg\min_{\mathbb{B}} \|\mathbb{Y} - \mathbb{B}\hat{\mathbb{H}}\|^2 \\ &= \arg\min_{\mathbb{B}} \|(\mathbb{I}_N - \mathbb{P}_B)\mathbb{Y}\|^2, \end{aligned} \tag{1}$$

where $\hat{\mathbb{H}} \triangleq (\mathbb{B}^T\mathbb{B})^{-1}\mathbb{B}^T\mathbb{Y}$ is the least-square estimate of $\mathbb{H}$ with respect to codeword $\mathbb{B}$ and received matrix $\mathbb{Y}$, and

$$\mathbb{P}_B \triangleq \mathbb{B}(\mathbb{B}^T\mathbb{B})^{-1}\mathbb{B}^T$$

is a function of the codeword $\mathbb{B}$. Here, $\mathbb{I}_N$ denotes an $N \times N$ identity matrix.

## III. Code Design

### A. Criteria for Good Codes

Several criteria for good codes have been proposed in the literature [1], [7], [8], [16]. We will in particular center on two of them: unitary and pairwise full-rank.

Firstly, it has been derived in [16] that unitary codewords, i.e., $\mathbb{B}^T\mathbb{B} = (N/A_{\rm T}) \cdot \mathbb{I}_{A_{\rm T}}$, can maximize the average signal-to-noise ratio (SNR) regardless of the statistics on $\mathbb{H}$. It has also been shown that when $\mathbb{H}$ is zero-mean complex Gaussian distributed, a unitary signal maximizes the capacity [14] and minimizes the union bound of word error rate (WER) [2] at high SNR. These results suggest that a good code can perhaps be constructed by collecting unitary codewords.

Secondly, it is better to have full-rank codeword pairs, where a pair of codewords, $\mathbb{B}(i)$ and $\mathbb{B}(j)$, is said to be *pair-wisely full-rank* if

$$\mathrm{rank}\left(\begin{bmatrix}\mathbb{B}(i) & \mathbb{B}(j)\end{bmatrix}\right) = 2A_{\rm T},$$

subject to $N \geq 2A_{\rm T}$. This is because at fairly high SNR, the average error probability is well approximated by the sum of pair-wise word error rates, namely, the union bound [1]. Also at fairly high SNR, the pair-wise word error is in turn well approximated by

$$\begin{aligned} &\Pr\left(\hat{\mathbb{B}} = \mathbb{B}(j) \middle| \mathbb{B}(i) \text{ transmitted}\right) \\ &\qquad\approx \mathcal{Q}\left(\frac{1}{\sqrt{2}}\|\mathbb{H}\|\sqrt{\lambda_{\min}(\mathbb{L}_{ij})}\right) \end{aligned} \tag{2}$$

where

$$\mathbb{L}_{ij} \triangleq \mathbb{I}_{A_{\rm R}} \otimes \left(\mathbb{B}(i)^T\left(\mathbb{I}_N - \mathbb{P}_{B(j)}\right)\mathbb{B}(i)\right),$$

and $\lambda_{\min}(\mathbb{L}_{ij})$ is the smallest eigenvalue of $\mathbb{L}_{ij}$. Here, "$\otimes$" indicates the Kronecker product, and $\mathcal{Q}(x) \triangleq \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-t^2/2}dt$ is the area under the tail of a standard Gaussian probability density function. Hence, if $[\mathbb{B}(i)\ \mathbb{B}(j)]$ do not achieve full column rank, we can obtain by [8] that

$$\det\left|\mathbb{B}(i)^T\left(\mathbb{I}_N - \mathbb{P}_{B(j)}\right)\mathbb{B}(i)\right| = 0.$$

This subsequently implies that $\lambda_{\min}(\mathbb{L}_{ij}) = 0$, and (2) will be close to $1/2$ at fairly high SNR, which is a situation that a good code should avoid.

Therefore, a code that satisfies both the above criteria should guarantee a good pairwise-error-based union bound (which in turn hints to have a good performance). This viewpoint will be confirmed by the subsequent simulations.

### B. The Proposed Code Design

Denote the information sequence by $\boldsymbol{k} = [k_1\ k_2\ \ldots\ k_K]^T$, where $k_i \in \{\pm 1\}$. The corresponding codeword is then proposed to be

$$\mathbb{B} = \frac{1}{\sqrt{A_{\rm T}}}\begin{bmatrix} \boldsymbol{k} & \boldsymbol{k}\odot\boldsymbol{s} \\ -\boldsymbol{k}\odot\boldsymbol{s} & \boldsymbol{k} \end{bmatrix}$$

where "$\odot$" denotes the Hadamard product, and

$$\boldsymbol{s} = \begin{cases} \begin{bmatrix} \mathbf{1}_{K-\lceil K/2\rceil} \\ -\mathbf{1}_{\lceil K/2\rceil} \end{bmatrix}, & \text{if } k_1 = -1 \\ \begin{bmatrix} \mathbf{1}_{K-\lceil K/2\rceil} \\ -\mathbf{1}_{\lceil K/2\rceil} \end{bmatrix} \odot \boldsymbol{d}, & \text{otherwise.} \end{cases}$$

In the above equation, $\mathbf{1}_k$ represents a $k \times 1$ all-one vector, and $\boldsymbol{d} \triangleq [(-1)^0\ (-1)^1\ \ldots\ (-1)^{K-1}]^T$.

It can be easily examined that the unitary criterion is satisfied, i.e., $\mathbb{B}^T\mathbb{B} = (N/A_{\rm T}) \cdot \mathbb{I}_{A_{\rm T}}$. It remains to show that the code just introduced satisfies pair-wise full-rank criterion.

Let $\mathbb{A}_{i,j} \triangleq \mathbb{B}(i)^T\mathbb{B}(j)$. Then for the validity of the pair-wise full-rank criterion, it suffices to prove that

$$\det\left|\mathbb{I}_{A_{\mathrm{T}}} - \frac{1}{(N/A_{\mathrm{T}})^2}\mathbb{A}_{i,j}\mathbb{A}_{i,j}^T\right| \neq 0, \quad (3)$$

for $1 \leq i,j \leq 2^K$ with $i \neq j$. By denoting respectively the $k$th eigenvalue and $k$th eigenvector of $\mathbb{A}_{i,j}\mathbb{A}_{i,j}^T$ by $\lambda_k$ and $\boldsymbol{u}_k$, the validity of (3) can be verified by showing that $\lambda_k \neq (N/A_{\mathrm{T}})^2$ for every $k$ because

$$\begin{aligned}(3) &\Leftrightarrow \det\left|\mathbb{I}_{A_{\mathrm{T}}} - \frac{1}{(N/A_{\mathrm{T}})^2}\sum_{k=1}^{A_{\mathrm{T}}}\lambda_k\boldsymbol{u}_k\boldsymbol{u}_k^T\right| \neq 0\\ &\Leftrightarrow \det\left|\sum_{k=1}^{A_{\mathrm{T}}}\left(1-\frac{\lambda_k}{(N/A_{\mathrm{T}})^2}\right)\boldsymbol{u}_k\boldsymbol{u}_k^T\right| \neq 0.\end{aligned}$$

Now, let $(\boldsymbol{k}_i, \boldsymbol{s}_i)$ and $(\boldsymbol{k}_j, \boldsymbol{s}_j)$ be respectively the vector pairs that define codewords $\mathbb{B}(i)$ and $\mathbb{B}(j)$. Denote for convenience $\boldsymbol{c}_{j,i} = \boldsymbol{k}_j \odot \boldsymbol{s}_i$. We then prove that $\lambda_1 \neq (N/A_{\mathrm{T}})^2$ and $\lambda_2 \neq (N/A_{\mathrm{T}})^2$ by differentiating the following two cases.

*Case 1:* $\boldsymbol{s}_i = \boldsymbol{s}_j = \boldsymbol{s}$.

In this case,

$$\begin{aligned}\mathbb{B}(i)^T\mathbb{B}(j) &= \frac{1}{A_{\mathrm{T}}}\begin{bmatrix}\boldsymbol{k}_i^T\boldsymbol{k}_j + \boldsymbol{c}_{i,i}^T\boldsymbol{c}_{j,i} & \boldsymbol{k}_i^T\boldsymbol{c}_{j,i} - \boldsymbol{c}_{i,i}^T\boldsymbol{k}_j\\ \boldsymbol{c}_{i,i}^T\boldsymbol{k}_j - \boldsymbol{k}_i^T\boldsymbol{c}_{j,i} & \boldsymbol{k}_i^T\boldsymbol{k}_j + \boldsymbol{c}_{i,i}\boldsymbol{c}_{j,i}\end{bmatrix}\\ &= \begin{bmatrix}\boldsymbol{k}_i^T\boldsymbol{k}_j & 0\\ 0 & \boldsymbol{k}_i^T\boldsymbol{k}_j\end{bmatrix}\end{aligned}$$

Then,

$$\mathbb{A}_{i,j}\mathbb{A}_{i,j}^T = \begin{bmatrix}(\boldsymbol{k}_i^T\boldsymbol{k}_j)^2 & 0\\ 0 & (\boldsymbol{k}_i^T\boldsymbol{k}_j)^2\end{bmatrix}$$

So, $\lambda_1 = \lambda_2 = (\boldsymbol{k}_i^T\boldsymbol{k}_j)^2 < (N/A_{\mathrm{T}})^2$.

*Case 2:* $\boldsymbol{s}_i \neq \boldsymbol{s}_j$.

In this case,

$$\begin{aligned}\mathbb{B}(i)^T\mathbb{B}(j) &= \frac{1}{A_{\mathrm{T}}}\begin{bmatrix}\boldsymbol{k}_i^T\boldsymbol{k}_j + \boldsymbol{c}_{i,i}^T\boldsymbol{c}_{j,j} & \boldsymbol{k}_i^T\boldsymbol{c}_{j,j} - \boldsymbol{c}_{i,i}^T\boldsymbol{k}_j\\ \boldsymbol{c}_{i,i}^T\boldsymbol{k}_j - \boldsymbol{k}_i^T\boldsymbol{c}_{j,j} & \boldsymbol{k}_i^T\boldsymbol{k}_j + \boldsymbol{c}_{i,i}^T\boldsymbol{c}_{j,j}\end{bmatrix}\\ &= \frac{1}{A_{\mathrm{T}}}\begin{bmatrix}\boldsymbol{k}_i^T(\boldsymbol{k}_j + \boldsymbol{k}_j\odot\boldsymbol{d}) & \boldsymbol{k}_i^T(\boldsymbol{c}_{j,j} - \boldsymbol{c}_{j,i})\\ -\boldsymbol{k}_i^T(\boldsymbol{c}_{j,j} - \boldsymbol{c}_{j,i}) & \boldsymbol{k}_i^T(\boldsymbol{k}_j + \boldsymbol{k}_j\odot\boldsymbol{d})\end{bmatrix},\end{aligned}$$

which gives

$$\mathbb{A}_{i,j}\mathbb{A}_{i,j}^T = \begin{bmatrix}c & 0\\ 0 & c\end{bmatrix},$$

where $c \triangleq (1/A_{\mathrm{T}})^2(|\boldsymbol{k}_i^T(\boldsymbol{k}_j + \boldsymbol{k}_j\odot\boldsymbol{d})|^2 + |\boldsymbol{k}_i^T(\boldsymbol{c}_{j,j} - \boldsymbol{c}_{j,i})|^2)$. Accordingly, $\lambda_1 = \lambda_2 = c < (N/A_{\mathrm{T}})^2$.

We end this section by commenting that our design can be viewed as a high-dimensional variation of Alamouti codes. Hence, the unitary property is satisfied simply by the Alamouti code structure. By properly introducing the additional Hadamard product, our code can further fulfill the pairwise full-rank property.

## IV. PRIORITY-FIRST SEARCH DECODING

In this section, we will derive the recursive decoding metric that can be used by the priority-first search algorithm [9]. Since the metric proposed is nondecreasing along every path in the code tree, the optimality of the decoding result is certified [16].

Continuing the derivation in (1) by noting that $\|\mathbb{P}_B\|^2 = A_{\mathrm{T}}$, we obtain

$$\begin{aligned}\hat{\mathbb{B}} &= \arg\min_{\mathbb{B}} -\sum_{j=1}^{A_{\mathrm{R}}}\boldsymbol{y}_j^H\mathbb{P}_B\boldsymbol{y}_j\\ &= \arg\min_{\mathbb{B}} -\mathrm{tr}\left(\mathbb{P}_B\sum_{j=1}^{A_{\mathrm{R}}}\boldsymbol{y}_j\boldsymbol{y}_j^H\right)\\ &= \arg\min_{\mathbb{B}} -\mathrm{tr}\left(\frac{A_{\mathrm{T}}}{N}\mathbb{B}\mathbb{B}^T\sum_{j=1}^{A_{\mathrm{R}}}\boldsymbol{y}_j\boldsymbol{y}_j^H\right)\\ &= \arg\min_{\mathbb{B}} -\mathrm{tr}\left(\mathbb{W}\sum_{i=1}^{A_{\mathrm{T}}}\boldsymbol{b}_i\boldsymbol{b}_i^T\right),\end{aligned}$$

where $\mathbb{W} \triangleq \mathbf{Re}\{(A_{\mathrm{T}}/N)\sum_{j=1}^{A_{\mathrm{R}}}\boldsymbol{y}_j\boldsymbol{y}_j^H\}$, and $\mathrm{tr}(\cdot)$ is the trace matrix operation. By letting

$$\mathbb{M}_1 \triangleq \boldsymbol{k}\boldsymbol{k}^T + (\boldsymbol{k}\odot\boldsymbol{s})(\boldsymbol{k}\odot\boldsymbol{s})^T,$$

and

$$\mathbb{M}_2 \triangleq \boldsymbol{k}(\boldsymbol{k}\odot\boldsymbol{s})^T - (\boldsymbol{k}\odot\boldsymbol{s})\boldsymbol{k}^T,$$

we have

$$\sum_{i=1}^{A_{\mathrm{T}}}\boldsymbol{b}_i\boldsymbol{b}_i^T = \frac{1}{A_{\mathrm{T}}}\begin{bmatrix}\mathbb{M}_1 & \mathbb{M}_2^T\\ \mathbb{M}_2 & \mathbb{M}_1\end{bmatrix}.$$

This reduces the decoding criterion to

$$\hat{\boldsymbol{k}} = \arg\min_{\boldsymbol{k}}\{-\mathrm{tr}(\mathbb{M}_1\mathbb{D}) - \mathrm{tr}(\mathbb{M}_2\mathbb{E})\},$$

where $\mathbb{D} \triangleq \mathbb{W}_{1,1} + \mathbb{W}_{2,2}$, $\mathbb{E} \triangleq \mathbb{W}_{1,2} - \mathbb{W}_{1,2}^H$, and $\mathbb{W}_{1,1}$, $\mathbb{W}_{1,2}$ and $\mathbb{W}_{2,2}$ are the corresponding submatrices of

$$\mathbb{W} = \begin{bmatrix}\mathbb{W}_{1,1} & \mathbb{W}_{1,2}\\ \mathbb{W}_{1,2}^H & \mathbb{W}_{2,2}\end{bmatrix}.$$

Since the decision criterion is intact by adding a constant independent of the codewords,

$$\begin{aligned}\hat{\boldsymbol{k}} = \arg\min_{\boldsymbol{k}}\Bigg(&\sum_{m=1}^{K}C_m - \frac{1}{4}\sum_{m=1}^{K}\sum_{n=1}^{K}k_mk_n(1+s_ms_n)d_{m,n}\\ &-\frac{1}{4}\sum_{m=1}^{K}\sum_{n=1}^{K}k_mk_n(s_m-s_n)e_{m,n}\Bigg), \quad (4)\end{aligned}$$

where

$$C_m \triangleq \sum_{n=1}^{m-1}(|d_{m,n}| + |e_{m,n}|) + \frac{1}{2}|d_{m,m}|,$$

and $d_{m,n}$ and $e_{m,n}$ are respectively the elements in matrices $\mathbb{D}$ and $\mathbb{E}$ and can be expressed as

$$d_{m,n} = \begin{cases} \dfrac{A_{\mathrm{T}}}{N} \displaystyle\sum_{j=1}^{A_{\mathrm{R}}} \mathbf{Re}\{y_{m,j} y_{n,j}^* + y_{m+K,j} y_{n+K,j}^*\}, \\ \qquad\qquad\qquad \text{for } 1 \le m, n \le N \\ 0, \qquad\qquad\qquad \text{otherwise} \end{cases}$$

$$e_{m,n} = \begin{cases} \dfrac{A_{\mathrm{T}}}{N} \displaystyle\sum_{j=1}^{A_{\mathrm{R}}} \mathbf{Re}\{y_{m,j} y_{n+K,j}^* - y_{m+K,j} y_{n,j}^*\}, \\ \qquad\qquad\qquad \text{for } 1 \le m, n \le N \\ 0, \qquad\qquad\qquad \text{otherwise} \end{cases}$$

Finally, the decoding metric $f$ inside the parenthesis of (4) can be computed recursively as

$$f(\boldsymbol{k}_{(\ell)}) = g(\boldsymbol{k}_{(\ell)}) - \gamma(\boldsymbol{k}_{(\ell)}),$$

where $\boldsymbol{k}_{(\ell)} = [k_1\ k_2\ \ldots\ k_\ell]^T$,

$$g(\boldsymbol{k}_{(\ell+1)}) = g(\boldsymbol{k}_{(\ell)}) - \beta(\boldsymbol{k}_{(\ell+1)}),$$

$$\beta(\boldsymbol{k}_{(\ell+1)}) = k_{\ell+1} \frac{\sqrt{A_{\mathrm{T}}}}{N} \sum_{r=1}^{A_{\mathrm{R}}} \mathbf{Re}\left\{ \sum_{t=0}^{1} y_{\ell+1+tK,r} \sum_{i=0}^{1}\sum_{j=0}^{1} (-1)^p s_{\ell+1}^q \cdot u_{i,j}^{(r)}(\boldsymbol{k}_{(\ell)}) \right\},$$

$$\gamma(\boldsymbol{k}_{(\ell)}) = -\sum_{m=\ell+1}^{K} \sum_{n=\ell+1}^{m} \left( |d_{m,n}|\, \mathbb{1}\{s_m = s_n\} + |e_{m,n}|\, \mathbb{1}\{s_m \neq s_n\} \right) + \frac{\sqrt{A_{\mathrm{T}}}}{N} \sum_{m=\ell+1}^{K} \left| \sum_{r=1}^{A_{\mathrm{R}}} \mathbf{Re}\left\{ \sum_{t=0}^{1} y_{m+tK,r} \times \sum_{i=0}^{1}\sum_{j=0}^{1} (-1)^p s_m^q \cdot u_{i,j}^{(r)}(\boldsymbol{k}_{(\ell)}) \right\} \right|,$$

$p = \lfloor t + (-1)^t (i+j)/2 \rfloor$, $q = t + |i-j|(-1)^t$, and

$$u_{i,j}^{(r)}(\boldsymbol{k}_{(\ell+1)}) = u_{i,j}^{(r)}(\boldsymbol{k}_{(\ell)}) + \frac{1}{\sqrt{A_{\mathrm{T}}}} k_{\ell+1} s_{\ell+1}^i y_{\ell+1+jK,r}^*.$$

In the above equation, $\mathbb{1}\{\cdot\}$ denotes the set indicator function.

## V. Simulation Results

In this section, we compare the performance of the code constructed in Section III with the codes obtained by computer search. The criterion used in the simulated annealing code search algorithm follows that in [2] (also, [7] and [8]). We take $A_{\mathrm{R}} = 1$ in our simulations, and assume that $\mathbb{H}$ is zero-mean complex Gaussian with $E[\mathbb{H}\mathbb{H}^H] = (1/A_{\mathrm{T}})\mathbb{I}_{A_{\mathrm{T}}}$. The average SNR is then give by

$$\frac{E[\mathbb{H}^H\mathbb{H}]}{(K/N)\sigma^2} = \frac{2}{\sigma^2}.$$

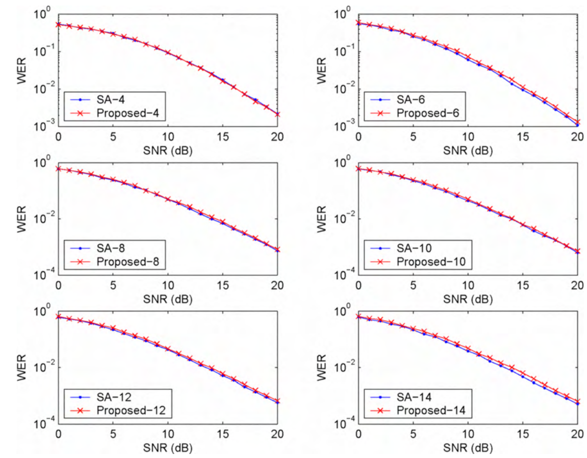


Fig. 1. Comparison of word error rates (WERs) between the codes constructed in Section III-B (Proposed-N) and the codes obtained from simulated annealing search (SA-N). The codeword lengths are taken to be equal to $N = 4, 6, 8, 10, 12$ and $14$.

Figure 1 shows that the best (structureless) computer-searched codes only have about 0.4 dB advantage over the constructed codes for $N = 4, 6, \ldots, 12$.

We also compare our code with a multiple-antenna system that uses the $(17, 12, 3)$ nonlinear channel code[2] in combination with the Alamouti code and a 7-bit training sequence. In particular, the code bits are mapped to the two transmit antennas using the Alamouti code before its transmission, and the receiver will estimate $\mathbb{H}$ in terms of a least square estimator based on the 7 training bits. The result in Figure 2 illustrates that this communication system performs 0.7 dB worse than the constructed code. In a technically infeasible situation that assumes the receiver can achieve a perfect estimate of $\mathbb{H}$ with merely 7 training bits, the typical communication system outperforms the constructed code by only 0.5 dB.

We would like to emphasize that to search the best code by computers for codeword length greater than 14 is very operational intensive even if there are only two transmit antennas. For example, it took about three weeks to cool down the simulated-annealing search when $N = 14$ and $A_{\mathrm{T}} = 2$. It can be anticipated that the search time will grow exponentially with the code word length. Thus, the systematic code construction that we propose may be a good alternative as far as long code is concerned.

Figure 3 shows the decoding complexity of the priority-first search decoder for constructed code of length 24. The complexity is defined as the average number of node expansions per information bit. Since the number of node expansions is half of the number of tree branch metrics computed (i.e., two recursions of $f$-function values), the equivalent complexity of exhaustive decoding is correspondingly $(2^{K+1} - 1) \cdot A_{\mathrm{T}}/K$. In the case of $(24, 12)$ code with two transmit antennas, this number is equal to 1365.17. It is then clear from the figure

[2]The $(17, 12, 3)$ code we adopt here is formed by taking out some code words from the $(17, 12 \cdot \log_2(20/16), 3)$ code in [13].

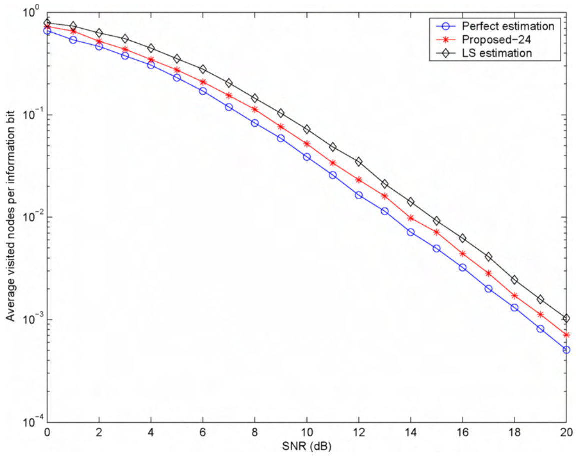


Fig. 2. Comparison of WERs among the codes constructed in Section III-B (Proposed-24) and the system using a (17,12) nonlinear code in combination with the Alamouti code and a 7-bit training sequence. The codeword length is equal to $N = 24$.

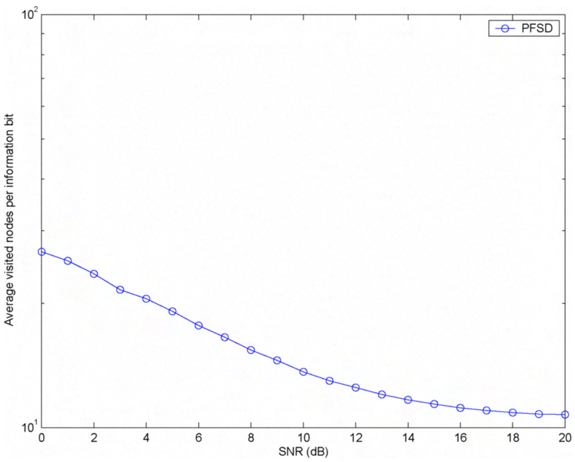


Fig. 3. The decoding complexity of the priority-first search decoder (PFSD) for the constructed code of length $N = 24$.

that the priority-first search decoder significantly improve the decoding complexity when it is compared with the exhaustive decoder.

## REFERENCES

[1] M. Beko, J. Xavier and V. A. N. Barroso, "Noncoherent communication in multiple-antenna system: receiver design and codebook construction," *IEEE Trans. Signal Process.*, vol. 55, no. 12, pp. 5703-5715, Dec. 2007.

[2] M. Brehler and M. K. Varanasi, "Asymptotic error probability analysis of quadratic receivers in Rayleigh fading channels with applications to a unified analysis of coherent and noncoherent space-time receivers," *IEEE Trans. Inform. Theory*, vol. 47, no. 6, pp. 2383-2399, Sept. 2001.

[3] M. J. Borran, A. Sabharwal and B. Aazhang, "On design criteria and construction of noncoherent space-time constellations," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2332-2351, Oct. 2003.

[4] A.M. Cipriano, J.-C. Belfiore, "noncoherent Codes over the Grassmannian", *IEEE Trans Wireless Commun.*, vol. 6, no. 10, pp 3657-3667, Oct. 2007.

[5] P. Dayal, M. Brehler and M. K. Varanasi, "Leveraging coherent space-time codes for noncoherent communication via training", *IEEE Trans. Inform. Theory*, vol. 50, no. 9, pp. 2058-2080, Sept. 2004.

[6] H. El Gamal, D. Aktas, M. O. Damen, "Noncoherent space-time coding: An algebraic perspective", *IEEE Trans. Inform. Theory*, vol. 51, no. 7, pp. 2380-2390, July 2005.

[7] J. Giese and M. Skoglund, "Space-time code design for unknown frequency-selective channels," *Proc. IEEE. Int. Conf. Acoust., Speech, Signal Process.*, Orlando, FL, USA, May 2002.

[8] J. Giese and M. Skoglund, "Single- and multiple-antenna constellations for communication over unknown frequency-selective fading channels," *IEEE Trans. Inform. Theory*, vol. 53, no. 4, pp. 1584-1594, April 2007.

[9] Y. S. Han and P.-N. Chen, "Sequential decoding of convolutional codes," *The Wiley Encyclopedia of Telecommunications*, edited J. Proakis, John Wiley and Sons, Inc., 2002.

[10] B. Hochwald, T. L. Marzetta, T. J. Richardson, W. Sweldens, and R. Urbanke, "Systematic design of unitary space-time constellations," *IEEE Trans. Inform. Theory*, vol. 46, no. 6, pp. 1962-1973, Sept. 2000.

[11] W.-K. Ma, B.-N. Vo, T. N. Davidson and P.-C. Ching, "Blind ML detection of orthogonal space-time block codes: efficient high-performance implementations," *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 738-751, Feb. 2006.

[12] W.-K. Ma, "Blind ML detection of orthogonal space-time block codes: Identifiability and code construction," *IEEE Trans. Signal. Process.*, vol. 55, no. 7, pp. 3312-3324, July 2007.

[13] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes Amsterdam*, The Netherlands: North-Holland, 1977.

[14] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat-fading," *IEEE Trans. Inform. Theory*, vol. 45, no. 1, pp. 139-157, Jan. 1999.

[15] V. Tarokh and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1456-1467, July 2007.

[16] C.-L. Wu, P.-N. Chen, Yunghsiang S. Han and M.-H. Kuo, "Maximum-likelihood priority-first search decodable codes for combined channel estimation and error correction," In review, *IEEE Trans. Inform. Theory*.