On the Coding Scheme for Joint Channel Estimation and Error Correction over Block Fading Channels

Chia-Lung Wu*, Po-Ning Chen* and Yunghsiang S. Han†‡, Yan-Xiu Zheng§
*Dept. of Comm. Eng., National Chiao-Tung Univ., Taiwan, R.O.C.
Email: clwu@banyan.cm.nctu.edu.tw, qponing@mail.nctu.edu.tw
†Graduate Institute of Comm. Eng., National Taipei Univ., Taiwan, R.O.C.
Email: yshan@mail.ntpu.edu.tw
‡Dept. of Computer Science and Inform. Eng., National Chi Nan Univ., Taiwan, R.O.C.
§Inform. & Commun. Research Lab., Industrial Technology Research Institute of Taiwan, R.O.C.
Email: zhengyanxiu@itri.org.tw

***Abstract*—In this work, we propose a novel systematic code construction scheme for joint channel estimation and error correction for channels with independently varying fading subblocks. Unlike the existing noncoherent codes that are designed with the help of computer search, a code of desired code length and code rate can be directly generated with our coding scheme. We then compare our codes with the three-times-repetitive $(12, 6)$ code proposed by Xu *et al.* for use of channel quality indicator (CQI) in uplink control for IEEE 802.16m. Simulations show that our constructed $(36, 6)$ code has comparable performance to Xu's code when channel coefficients changes randomly in every 12 symbols. If the channel taps remain constant in the entire coding block of length 36, our code outperforms Xu's code by 0.7 dB. This indicates that the new constructed code adapts more robustly to the two simulated scenarios. For frequency selective channels of unit memory order, our simulation results suggest that our code that takes in consideration the varying characteristic of channels can achieve better performance at median-to-high signal-to-noise ratio over the computer-searched, union-bound-minimized code of length less than the varying subblock size. A side advantage of our code construction scheme is that its systematic structure makes it maximum-likelihoodly decodable by the priority-first search algorithm. The decoding complexity is therefore significantly decreased in contrast to that of exhaustive decoder for the structureless computer-searched codes.**

I. INTRODUCTION

The coding technique that combines channel estimation and error correction has received general attention recently. Several previous works [1]–[5], [9], [11], [12] have substantiated that such a joint noncoherent design can improve the system performance over existing separate designs. Theoretical evidence that the coherent channel capacity and noncoherent channel capacity almost coincide to each other at median SNR range such as 30 dB further suggests the potential of such technique [15].

The error correcting code design that jointly considers channel estimation is especially useful in situation when either the fading is rapid enough to preclude a good estimate of channel taps or the cost of implementing the channel estimators is high. One example is the reliable delivery of often short-in-length control signal such as channel quality indicator (CQI) in a highly mobile environment.

At this background, Xu *et al.* proposed a novel nonlinear coding scheme suitable for blind noncoherent detection of the transmitted control signal to the 802.16m standard body [13]. In the proposal, the uplink CQI information is encoded using a $(12, 6)$ code. The codeword will then be repeatedly transmitted three times (perhaps through different OFDM channels) in order to further benefit from diversity gain (which can be equivalently regarded as a $(36, 6)$ coding scheme).

Since most of the existing blind-detectable noncoherent codes are designed with the help of computer search, they exhibit no apparent structure for efficient decoding. The operation-intensive exhaustive search therefore becomes the only decoding option, of which the dramatically increasing decoding complexity prevents its practical use for codes of long codeword length or high code rate.

In this work, we take a different approach in such code design. Based on self-orthogonality framework, we propose a *systematic* (N, K) coding scheme that can deal with any given N and K for channels with possibly varying channel coefficients in a coding block. It is an extension of our previous work that targets the blind detection over channels with static (i.e., constant) channel coefficients during the transceiving of a codeword [14]. Simulations show that our constructed $(36, 6)$ code has almost the same performance as Xu's three-times-repetitive $(12, 6)$ code when the channel independently varies its coefficients three times in a coding block. In case the channel remains constant during the entire coding block, our constructed code has 0.7 dB performance improvement over Xu's code.

Xu's code is specifically designed for a frequency-nonselective OFDM system, while our systematic code construction scheme can also be applied in a frequency selective environment. Our simulation results indicate that with a proper design, a blind-detectable noncohrent code can be made robust for channels whose taps may vary more often than a coding block.

A side advantage of our code construction scheme is that its systematic structure makes it maximum-likelihoodly decodable by the priority-first search algorithm. Thus, when being

978-1-4244-5213-4/09/ $26.00 ©2009 Crown

compared with the operation-intensive exhaustive decoder, the decoding complexity is greatly reduced especially when codes of longer code length is adopted.

Throughout this work, superscripts "H" and "T" are specifically reserved for the matrix operations of Hermitian transpose and transpose, respectively [8].

II. System Model

Suppose that a codeword $\boldsymbol{b} = [b_1 \ \cdots \ b_N]^T$ is transmitted over a block fading channel of memory order P, of which channel coefficients may vary in every Q symbols, where $b_i \in \{\pm 1\}$ and $Q > P$. By letting $L \triangleq N + P$ and $M \triangleq \lceil L/Q \rceil$, the system can be modelled by:

$$\boldsymbol{y} = \mathbb{B}\boldsymbol{h} + \boldsymbol{n},$$

where $\boldsymbol{n}$ is zero-mean white Gaussian distributed,

$$\boldsymbol{h} \triangleq \begin{bmatrix} \boldsymbol{h}_1^H & \boldsymbol{h}_2^H & \cdots & \boldsymbol{h}_M^H \end{bmatrix}^H$$

with $\boldsymbol{h}_k \triangleq [h_{0,k} \ \ h_{1,k} \ \cdots \ h_{P,k}]^T$, and

$$\mathbb{B} \triangleq \mathbb{B}_1 \oplus \mathbb{B}_2 \oplus \cdots \oplus \mathbb{B}_M$$

with $\mathbb{B}_k \triangleq [\mathbf{0}_{Q\times P} \ \ \mathbb{I}_Q][\boldsymbol{b}_k \ \ \tilde{\mathbb{E}}\boldsymbol{b}_k \ \cdots \ \tilde{\mathbb{E}}^P\boldsymbol{b}_k]$. Here, $\mathbf{0}_{Q\times P}$ represents a $Q \times P$ all-zero matrix, $\mathbb{I}_Q$ is a $Q \times Q$ identity matrix, $\boldsymbol{b}_k \triangleq [b_{(k-1)Q-P+1} \ \cdots \ b_{(k-1)Q+1} \ \cdots \ b_{kQ}]^T$ is a portion of the transmitted codeword $\boldsymbol{b}$,

$$\tilde{\mathbb{E}} \triangleq \begin{bmatrix} 0 & 0\cdots & 0 & 0 \\ 1 & 0 \ddots & 0 & 0 \\ 0 & 1 \ddots & 0 & 0 \\ 0 & 0\cdots & 1 & 0 \end{bmatrix}_{(Q+P)\times(Q+P)}$$

equates the logical left-shift operator, and "$\oplus$" is the direct sum operator for two matrices.[1] Also, for notational convenience, we take $n_j = 0$ for $j > L$, and $b_j = 0$ for $j \le 0$ and $j > N$. Under such system setting, $\boldsymbol{y}$ is an $MQ \times 1$ received vector with $y_j = 0$ for $j > L$.

It can be derived that the joint maximum-likelihood decoder [3], [12] upon the reception of $\boldsymbol{y}$ is given by:

$$\hat{\boldsymbol{b}} \ = \ \arg\max_{\boldsymbol{b}\in\mathcal{C}} \sum_{k=1}^{M} \left\| \boldsymbol{y}_k \boldsymbol{y}_k^H - \mathbb{P}_{B_k} \right\|^2, \qquad (1)$$

where $\boldsymbol{y}_k \triangleq [y_{(k-1)Q+1} \ \ y_{(k-1)Q+2} \ \cdots \ y_{kQ}]$ is the output portion affected by $\boldsymbol{b}_k$, and $\mathbb{P}_{B_k} \triangleq \mathbb{B}_k(\mathbb{B}_k^T\mathbb{B}_k)^{-1}\mathbb{B}_k^T$. In the above derivation, we assume that the receiver, although it knows nothing about $\boldsymbol{h}$, has perfect knowledge about the values of P and Q.

[1]For two matrices $\mathbb{A}$ and $\mathbb{B}$, the direct sum of $\mathbb{A}$ and $\mathbb{B}$ is defined as

$$\mathbb{A} \oplus \mathbb{B} = \begin{bmatrix} \mathbb{A} & \mathbf{0} \\ \mathbf{0} & \mathbb{B} \end{bmatrix}.$$

III. Code Design

We summarize the proposed code construction scheme [14] for $P = 0$ (frequency nonselective) and $P = 1$ (frequency selective) in the following algorithm.

Step 1. Fix $b_1 = -1$,[2] and choose a certain integer Δ defined later. Find 2^K codewords of the (N, K) code by repeating Steps 2–4 for $0 \le i \le 2^K - 1$.

Step 2. Let $\rho_{\min} = 0$ and $\rho = i \cdot \Delta$.

Step 3. For $\ell = 2$ to N, assign the ℓ-th bit of the i-th codeword, b_ℓ, according to that if $\rho < \rho_{\min} + \gamma_\ell$, then $b_\ell = -1$; else, $b_\ell = 1$ and $\rho_{\min} = \rho_{\min} + \gamma_\ell$, where

$$\gamma_\ell = |\mathcal{A}_P(b_1, \ldots, b_{\ell-1}, b_\ell = -1)|,$$

which will be defined shortly.

Step 4. Store the ith codeword $\boldsymbol{b}$, and goto Step 2 for the next codeword until all 2^K codewords are selected.

Now, as far as the code design for frequency nonselective channels is concerned, $\mathcal{A}_0(b_1, \ldots, b_\ell)$ is simply the set of all binary ± 1-sequences of length N, whose first ℓ bits are assigned as the arguments indicate, and which at the same time satisfy that

$$\begin{cases} \mathbb{B}_k^T\mathbb{B}_k = Q \text{ for } 1 \le k < M \\ \mathbb{B}_M^T\mathbb{B}_M = N - (M-1)Q. \end{cases} \qquad (2)$$

For channels of memory order $P = 1$, the conditions to define $\mathcal{A}_1(b_1, \ldots, b_\ell)$ are the same as those to define $\mathcal{A}_0(b_1, \ldots, b_\ell)$ except that condition (2) is replaced with

$$\begin{cases} \mathbb{B}_1^T\mathbb{B}_1 = \begin{bmatrix} Q & c_1 \\ c_1 & Q-1 \end{bmatrix} \\ \mathbb{B}_k^T\mathbb{B}_k = \begin{bmatrix} Q & c_k \\ c_k & Q \end{bmatrix} \text{ for } 2 \le k \le M-1 \\ \mathbb{B}_M^T\mathbb{B}_M = \begin{bmatrix} N-(M-1)Q & c_M \\ c_M & N-[(M-1)Q-1]^+ \end{bmatrix} \end{cases} \qquad (3)$$

where $c_1 = -((Q-1) \bmod 2)$, $c_2 = -(Q \bmod 2)$, $c_M = -((N-[(M-1)Q-1]^+ - 1) \bmod 2)$, and $[x]^+ \triangleq \max\{x, 0\}$. Note that conditions (2) and (3) are devised based on the self-orthogonal codeword property that guarantees to maximize the system signal-to-noise ratio regardless of the statistics of $\boldsymbol{h}$ [14].

It remains to determine the integer Δ. In order to have adequate number of codewords selected, Δ must satisfy

$$\Delta \le \frac{|\mathcal{A}_P(b_1 = -1)|}{2^K - 1}. \qquad (4)$$

We however found that letting Δ be the largest integer satisfying (4) as we did in [14] may not generate the alphabetically uniform-pick code with the best error performance. In certain cases, the second largest integer satisfying (4) is indeed a better choice. Further investigation that follows along this direction suggests that a better choice of Δ will yield a

[2]Codeword $\boldsymbol{b}$ and $\{\mathbb{P}_{B_k}\}_{k=1}^M$ in (1) is not one-to-one corresponding unless the first element of $\boldsymbol{b}$, namely b_1, is fixed. We thus fix $b_1 = -1$ in our code design.

code with larger minimum pairwise distance in the sense of (1), i.e., $\sum_{k=1}^{M} \|\mathbb{P}_{\bar{B}_k} - \mathbb{P}_{B_k}\|^2$, where $\{\bar{B}_k\}_{k=1}^{M}$ and $\{B_k\}_{k=1}^{M}$ respectively correspond to codewords $\bar{\boldsymbol{b}}$ and $\boldsymbol{b}$.

It may not be practical to examine the minimum pairwise distance for all 2^K codewords for the determination of the best Δ. Instead, we choose K codewords as representatives. These representative codewords correspond to $\rho = 2^j\Delta$ for $0 \le j \le K-1$. Subject to (4), we then adopt the Δ that minimizes the pairwise distance among these K codewords.

When $N > K+4$ and $P = 0$, the proposed process of determining Δ is indeed equivalent to that the Δ-th codeword must be of the form

$$[\underbrace{-1 \cdots -1}_{K+1}\ 1\ 1\ \boldsymbol{u}\ 1],$$

where $\boldsymbol{u}$ is a maximum-length shift-register sequence. In other words, the first $K+3$ bits are fixed as $[-1\ \cdots\ -1\ \ 1\ \ 1]$, and the last bit is always equal to 1. This is because under $P = 0$, all binary ± 1-sequences satisfy (2), which results in that $(2^{j+1}\Delta)$-th codeword is exactly the logical left-shift of $(2^j\Delta)$-th codeword.

We close this section by pointing out that the size of set $\mathcal{A}_P(b_1, \ldots, b_\ell)$ has explicit formula for $P = 0$ and $P = 1$. It is given below for which the derivation is omitted.

$$\mathcal{A}_P(b_1, \ldots, b_\ell) = \begin{cases} 2^{N-\ell}, & \text{for } P = 0; \\ \begin{pmatrix} Q - (\ell \bmod Q) \\ \frac{Q-(\ell \bmod Q)+c_\tau - m_\ell}{2} \end{pmatrix} \mathbf{1}\{|c_\tau - m_\ell| \le Q - (\ell \bmod Q)\} \\ \times \left[\prod_{k=\tau+1}^{M-1} \begin{pmatrix} Q \\ \frac{Q+c_{k+1}}{2} \end{pmatrix}\right] \begin{pmatrix} N - [(M-1)Q-1]^+ - 1 \\ \frac{N-[(M-1)Q-1]^+-1+c_M}{2} \end{pmatrix}, & \text{for } 1 \le \tau < M \text{ and } P = 1; \\ \begin{pmatrix} N - [(M-1)Q-1]^+ - 1 \\ \frac{N-[(M-1)Q-1]^+-1+c_M-m_\ell}{2} \end{pmatrix} \\ \times \mathbf{1}\{|c_M - m_\ell| \le N - [(M-1)Q-1]^+ - 1\}, & \text{for } \tau = M \text{ and } P = 1, \end{cases}$$

where $\tau = \lfloor \ell/Q \rfloor + 1$, and

$$m_\ell = \begin{cases} 0, \\ \text{for } \ell = 1 \text{ or } (\ell = [(\tau-1)Q-1]^+ + 1 \text{ and } 2 \le \tau \le M); \\ b_1 b_2 + \cdots + b_{\ell-1} b_\ell, & \text{for } 1 < \ell < Q; \\ b_{[(\tau-1)Q-1]^+ + 1} b_{[(\tau-1)Q-1]^+ + 2} + \cdots + b_{\ell-1} b_\ell, \\ \qquad \text{for } [(\tau-1)Q-1]^+ + 1 < \ell < \tau Q \text{ and } 2 \le \tau \le M. \end{cases}$$

IV. Optimal Priority-First Search Decoding

In this section, we will derive two decoding metrics that can be used by the priority first search algorithm [6], [7]. Both metrics will lead to the optimal maximum-likelihood decoding. The difference is that the first metric f_1 can be computed on-the-fly, and will therefore cause much less delay in the decoding. For the evaluation of the second metric f_2, however, one needs to know *all* received symbols, but its computational complexity is much less than that of f_1.

Continuing the derivation from (1) based on $\mathbb{B}_k^T \mathbb{B}_k = \mathbb{G}_k$ for $1 \le k \le M$, we establish that:

$$\hat{\boldsymbol{b}} = \arg\min_{\boldsymbol{b} \in \mathcal{C}} \frac{1}{2} \sum_{k=1}^{M} \sum_{m=1}^{Q+P} \sum_{n=1}^{Q+P} \left(-w_{m,n,k} b_{(k-1)Q-P+m} b_{(k-1)Q-P+n}\right)$$

where for $1 \le m, n \le Q+P$,

$$w_{m,n,k} = \sum_{i=0}^{P} \sum_{j=0}^{P} \delta_{i,j,k} \mathrm{Re}\{\tilde{y}_{m+i,k} \tilde{y}_{n+j,k}^*\},$$

$$\tilde{\boldsymbol{y}}_k \triangleq [\mathbf{0}_{1\times P}\ \boldsymbol{y}_k^H\ \mathbf{0}_{1\times P}]^H = [\tilde{y}_{1,k}\ \cdots\ \tilde{y}_{Q+2P,k}]^T,$$

and $\delta_{i,j,k}$ is the (i,j)-th entry[3] of matrix $\mathbb{D}_k \triangleq \mathbb{G}_k^{-1}$. By adding a constant $\frac{1}{2}\sum_{k=1}^{M}\sum_{m=1}^{Q+P}\sum_{n=1}^{Q+P} |w_{m,n,k}|$ to the decoding criterion, the on-the-fly metric f_1 that suits for the recursive computation of the priority-first search is given by:

$$f_1(b_1, \ldots, b_\ell) = f_1(b_1, \ldots, b_{\ell-1}) + \begin{cases} \alpha_{s,k} - b_\ell \sum_{i=0}^{P}\sum_{j=0}^{P} \delta_{i,j,k} \mathrm{Re}\{\tilde{y}_{s+i,k} \cdot u_{j,k}(b_1, \ldots, b_\ell)\}, \\ \qquad \text{for } P < s \le Q; \\ \alpha_{r,k} - b_\ell \sum_{i=0}^{P}\sum_{j=0}^{P} \delta_{i,j,k} \mathrm{Re}\{\tilde{y}_{r+i,k} \cdot u_{j,k}(b_1, \ldots, b_\ell)\} \\ +\alpha_{s,k+1} \\ -b_\ell \sum_{i=0}^{P}\sum_{j=0}^{P} \delta_{i,j,k+1} \mathrm{Re}\{\tilde{y}_{s+i,k+1} \cdot u_{j,k+1}(b_1, \ldots, b_\ell)\}, \\ \qquad \text{otherwise,} \end{cases}$$

where $s \triangleq [(\ell + P - 1) \bmod Q] + 1$, $r \triangleq s + Q$, $k \triangleq \max\{\lceil \ell/Q \rceil, 1\}$,

$$\alpha_{s,k} \triangleq \sum_{n=1}^{s-1} |w_{s,n,k}| + \frac{1}{2} |w_{s,s,k}|,$$

and

$$u_{j,k}(b_1, \ldots, b_{\ell+1}) = u_{j,k}(b_1, \ldots, b_\ell) + \frac{1}{2}\left(b_\ell \tilde{y}_{s+j,k}^* + b_{\ell+1} \tilde{y}_{s+j+1,k}^*\right)$$

with initial values $f_1(b_1, \ldots, b_\ell) = 0$ for $\ell = 0$, and $u_{j,k}(b_1, \ldots, b_{(k-1)Q-P+1}) = 0$ for $0 \le j \le P$ and $1 \le k \le M$. The low-complexity decoding metric f_2 is given by

$$f_2(b_1, \ldots, b_\ell) = f_1(b_1, \ldots, b_\ell) + h(b_1, \ldots, b_\ell),$$

[3]When $N - (M-1)Q = 0$, the designated $\mathbb{B}_M^T \mathbb{B}_M$ in (3) has no inverse. In such case, we redefine

$$\mathbb{D}_M \triangleq \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{N-[(M-1)Q-1]^+-1} \end{bmatrix}.$$

where

$$h(b_1,\ldots,b_\ell) \triangleq \begin{cases} \sum_{m=s+1}^{Q+P} \alpha_{m,k} - \sum_{m=s+1}^{Q+P} |v_{m,k}(b_1,\ldots,b_\ell)| - \beta_{s,k} & \text{for } P < s \le Q; \\ \sum_{m=s+1}^{Q+P} \alpha_{m,k+1} - \sum_{m=s+1}^{Q+P} |v_{m,k+1}(b_1,\ldots,b_\ell)| - \beta_{s,k+1} \\ \quad + \sum_{m=r+1}^{Q+P} \alpha_{m,k} - \sum_{m=r+1}^{Q+P} |v_{m,k}(b_1,\ldots,b_\ell)| - \beta_{r,k} & \text{otherwise,} \end{cases}$$

where s, r and k are defined the same as for $f_1(\cdot)$,

$$v_{m,k}(b_1,\ldots,b_\ell) = v_{m,k}(b_1,\ldots,b_{\ell-1}) + w_{s,m,k} b_\ell,$$

and

$$\beta_{s,k} \quad = \quad \beta_{s-1,k} - \sum_{n=s+1}^{Q+P} |w_{s,n,k}| - \frac{1}{2}|w_{s,s,k}|$$

with initial values $v_{m,k}(b_1,\ldots,b_{(k-1)Q-P+1}) = 0$ and $\beta_{0,k} = \sum_{m=1}^{Q+P} \alpha_{m,k}$.

V. Simulation Results

In our simulations, the channel parameters follow those in [12], where $\boldsymbol{h}$ is zero-mean complex-Gaussian distributed with $E[\boldsymbol{h}\boldsymbol{h}^H] = (1/(P+1))\mathbb{I}_{P+1}$.

We first compare our constructed $(36,6)$ code with Xu's three-times-repetitive $(12,6)$ code over frequency nonselective channels. As shown in Figure 1, the two codes has comparable performance when channel coefficients vary independently in every 12 symbols. In case the channel coefficients remain constant over the entire coding block, the proposed $(36,6)$ code performs 0.7 dB better than Xu's code as shown in Figure 2. It should be emphasized that when $P=0$, $\mathcal{A}_P(b_1,\ldots,b_\ell)$ is irrelevant to the design parameter Q; hence, the $(36,6)$ code in Figure 1 is identical to the one used in Figure 2. This indicates that the proposed $(36,6)$ code can adapt more robustly to the two simulated scenarios than Xu's code.

Figures 3 simulates three half-rate codes over frequency selective channels of memory order 1, in which the channel coefficients vary independently in every 15 symbols. The three codes are identified by $(28,14)(Q=29)$, $(28,14)(Q=15)$ and CS$(14,7)$, which respectively denote the constructed $(28,14)$ code with design parameter $Q=29$ (i.e., assuming at the design stage, the channel coefficients remain constant during the entire decoding block $L = N+P = 28+1 = 29$), the constructed $(28,14)$ code with design parameter $Q = 15$ (i.e., assuming the channel coefficients vary in every 15 symbols at the design stage), and the computer-searched (hence, structureless) $(14,7)$ code that minimizes the union bound derived based on the assumption that the channel taps remains constant during the decoding block (i.e., $Q = L = N+P = 14+1 = 15$, which is exactly the simulated channel).

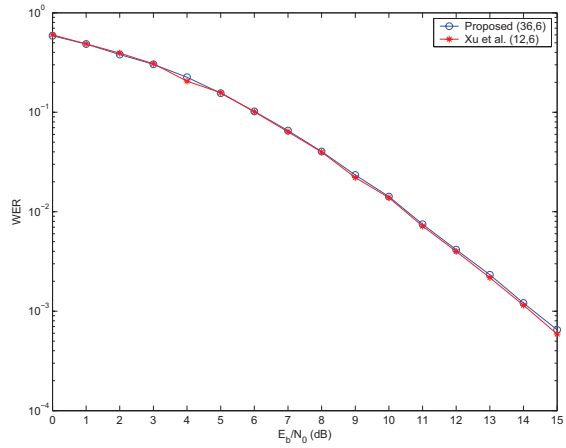

Fig. 1. Word error rates (WERs) for the constructed $(36,6)$ code and Xu's three-times-repetitive $(12,6)$ code over flat fading channel with coefficients varying independently in every 12 symbols.

As anticipated, $(28,14)(Q=29)$ seriously degrades since its assumption at the design stage does not match the characteristic of the true simulated channel. This suggests that the assumption that the channel coefficients remain constant in a coding block is very critical in the code design, and should be made with caution.

A striking result from Figure 3 is that the constructed $(28,14)(Q=15)$ code performs markedly better than the CS$(14,7)$ code at medium-to-high signal-to-noise ratios, despite that the CS$(14,7)$ code is the computer-optimized code specifically for the simulated channel. This suggests that when the channel memory order and varying characteristic are *prior* known (i.e., P and Q), performance gain can be obtained by enhancing the inter-Q-block correlation, and the system favors a longer code design.

In Table I, we summarize the decoding complexity for the $(28,14)(Q=15)$ code simulated in Figure 3, measured by the average number of node expansions per information bit. It shows, as previously mentioned, that the decoding metric f_2 requires less decoding efforts than the on-the-fly decoding metric f_1.

The performance of our constructed code can be further (slightly) improved if the codewords are selected uniformly from all feasible $(c_1, c_2, \cdots, c_M) \in \{-1,0,1\}^M$. For example, select only half (i.e., 2^{13}) of the codewords according to $c_1 = 0$ and $c_2 = -1$ for the $(28,14)(Q=15)$ code, and pick the remaining half of the codewords from those binary sequences satisfying (3) with $c_1 = 0$ and $c_2 = 1$. This however will slightly increase the decoding complexity. The trade-off between selecting codewords from fixed $(c_1,\ldots,c_M)$ or multiple $(c_1,\ldots,c_M)$'s is thus evident.

TABLE I
AVERAGE NUMBERS OF NODE EXPANSIONS PER INFORMATION BIT FOR THE $(28, 14)(Q = 15)$ CODE SIMULATED IN FIGURE 3.

SNR	3dB	4dB	5dB	6dB	7dB	8dB	9dB	10dB	11dB	12dB	13dB	14dB	15dB
f_1	1658	1367	1074	899	701	593	488	448	356	309	277	244	232
f_2	766	625	482	392	321	254	219	177	149	133	121	104	92
ratio of f_1/f_2	2.2	2.2	2.2	2.3	2.2	2.3	2.2	2.5	2.4	2.3	2.3	2.3	2.5

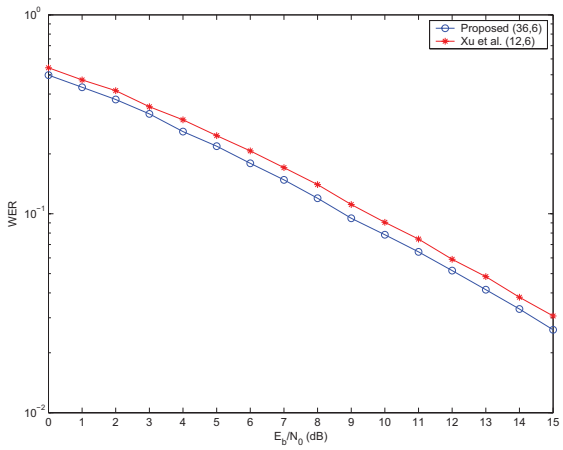

Fig. 2. Word error rates (WERs) for the constructed $(36, 6)$ code and Xu's three-times-repetitive $(12, 6)$ code over flat fading channel with coefficients unchanged during the transmission of a codeword.

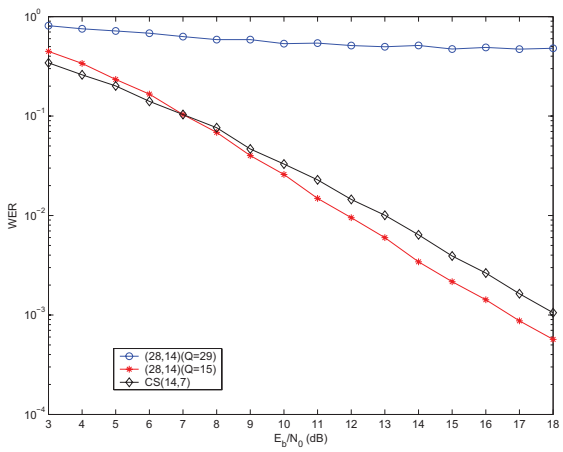

Fig. 3. Word error rates (WERs) for the $(28, 14)(Q = 29)$ code, the $(28, 14)(Q = 15)$ code and the CS$(14, 7)$ code over channels of memory order 1, whose coefficients varying independently in every 15 symbols.

VI. CONCLUSION

An extension of the code design for combined channel estimation and error correction to channels with independently varying fading subblocks is established in this work. This design can directly construct a code of any desired code length and code rate, of which the performance is shown to be comparable to the best computer-searched code for the channels simulated. Although we only derive the coding scheme and its decoding metric for a fixed Q, further extension to the situation that the channel coefficients $\boldsymbol{h}$ vary nonstationarily as the periods $Q_1, Q_2, \ldots, Q_M$ are not equal is straightforward. Such design may be suitable for, e.g., the frequency-hopping scheme of Global System for Mobile communications (GSM) and Universal Mobile Telecommunications System (UMTS), or the time-hopping scheme in IS-54, in which cases the channel coefficients change (or hop) at protocol-aware scheduled time [10].

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. 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.

[3] O. Coskun and K. M. Chugg, "Combined coding and training for unknown ISI channels," *IEEE Trans. Commun.*, vol. 53, no. 8, pp. 1310-1322, Aug. 2005.

[4] 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.

[5] 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, Apr. 2007.

[6] 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.

[7] Y. S. Han, C. R. P. Hartmann and C.-C. Chen, "Efficient priority-first search maximum-likelihood soft-decision decoding of linear block codes," *IEEE Trans. Inform. Theory*, vol. 39, no. 5, pp. 1514–1523, Sep. 1993.

[8] D. Harville, *Matrix Algebra From a Statistician's Perspective*, 1st edition, Springer, 2000.

[9] 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, Sep. 2000.

[10] R. Knopp and P. A. Humblet "On coding for block fading channels," *IEEE Trans. Inform. Theory* , vol. 46, no. 1, pp. 189-205, Jan. 2000.

[11] 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.

[12] M. Skoglund, J. Giese and S. Parkvall, "Code design for combined channel estimation and error protection," *IEEE Trans. Inform. Theory*, vol. 48, no. 5, pp. 1162-1171, May 2002.

[13] C. Xu, H. Sun, J.-K. Fwu, Q. Li, Y. Zhu, J. Wang, S. Zhang, Y. Gao, H. Yin, R. Vannithamby and S. Ahmadi, *IEEE 802.16m UL PHY Control: CQI Feedback Channel Design*, C80216m-UL_PHY_Ctrl-08_065r1, Intel Corporation, IEEE 802.16 UL PHY Control RG, 11th, Nov. 2008.

[14] C.-L. Wu, P.-N. Chen, Y. S. Han and S.-W. Wang, "A self-orthogonal code and its maximum-likelihood decoder for combined channel estimation and error protection," *Proc. IEEE International Symposium on Information Theory and its Applications*, Auckland, New Zealand, Dec. 2008.

[15] S.-H. Wu, "Effects of estimation errors on the noncoherent capacity over time-varying channels," *Proc. IEEE Vehicular Technology Conference*, Montreal, Canada, Sep. 2006.