Joint FEC Encoder and Linear Precoder Design for MIMO Systems with Antenna Correlation

Chongbin Xu[1], Peng Wang[2], Zhonghao Zhang[1], and Li Ping[1]
[1]Department of Electronic Engineering, City University of Hong Kong, Hong Kong
[2]School of Electrical and Information Engineering, the University of Sydney, Sydney, Australia
xchongbin2@cityu.edu.hk, peng.wang@sydney.edu.au, zhozhang@student.cityu.edu.hk, eeliping@cityu.edu.hk

***Abstract*—We study transmissions in multiple-input multiple-output (MIMO) systems with antenna correlation. We focus on joint forward error correction (FEC) encoder and linear precoder design based on channel covariance information at the transmitter (CCIT). We aim to optimize the system performance using the extrinsic information transfer (EXIT) chart type curve matching principle. By adopting a Hadamard precoding technique, we show that, the EXIT chart type curve of the precoded system is asymptotically determined by the channel correlation matrix at the transmitter as the number of receive antennas tends to infinity. The encoder-precoder curve matching can be made asymptotically accurate even in the lack of full channel state information at the transmitter (CSIT). Excellent performance based on this strategy is demonstrated by simulation results in systems with only a moderately large number of receive antennas.**

Keywords—MIMO; antenna correlation; channel covariance information; extrinsic information transfer chart

I. INTRODUCTION

For multiple-input multiple-output (MIMO) systems, channel state information (CSI) at the transmitter (CSIT) is important for performance improvement [1]-[2]. However, CSIT acquisition may be costly, especially for MIMO systems with a larger number of antennas [3]-[4]. For MIMO systems with antenna correlation, low-cost options have been studied to use channel covariance information at the transmitter (CCIT) [1]-[2]. A suboptimal but simple CCIT scheme, referred to as statistical water-filling (SWF), has been developed in [5]-[6]. However, so far discussions on SWF are mainly focused on mutual information analysis. There is less investigation on the implementation of SWF in practically coded systems.

A direct approach to SWF is adaptive modulation, in which different constellations are used to realize different rates on different eigen-directions. The demodulation complexity of this approach increases exponentially with the number of bits carried by each symbol, which can be an obstacle in practice. The problem becomes especially stringent in a large MIMO system where high rate transmission is highly desirable.

In this paper, we discuss a low-cost alternative involving uniform constellation, in which the FEC encoder and linear precoder are jointly designed assuming iterative receiver. It is well known that the extrinsic information transfer (EXIT) chart type curve matching principle can be used to optimize the performance of systems with iterative receiver [7]-[8]. Here the question is how to perform curve matching based on CCIT only with a relatively low complexity.

Using a Hadamard precoding technique [9]-[12], we show that, the EXIT chart type curve of the precoded system is asymptotically determined by the channel correlation matrix at the transmitter as the number of receive antennas tends to infinity. In this case, the encoder-precoder curve matching can be made asymptotically accurate even in the lack of full CSIT. Excellent performance based on this strategy is demonstrated by simulation results even for a moderately large number of receive antennas.

The proposed scheme in this paper is applicable to MIMO systems with more receive antennas than the transmit ones and relatively strong correlation at the transmitter side. This is especially suitable for uplink transmissions.

II. PRELIMINIARIES

A. System Model

Consider an MIMO system with N_{R} receive antennas and N_{T} transmit antennas. The received $N_{\mathrm{R}}\times 1$ signal vector, $\boldsymbol{r}$, is given by

$$\boldsymbol{r} = \boldsymbol{H}\boldsymbol{y} + \boldsymbol{\eta} \tag{1}$$

where $\boldsymbol{H}$ is an $N_{\mathrm{R}}\times N_{\mathrm{T}}$ channel transfer matrix, $\boldsymbol{y}$ an $N_{\mathrm{T}}\times 1$ transmitted signal vector, and $\boldsymbol{\eta}$ an $N_{\mathrm{R}}\times 1$ vector of independent and identically distributed (i.i.d.) complex additive white Gaussian noise (AWGN) samples with mean 0 and variance $\sigma^2 = 1$.We will assume that $\boldsymbol{y}$ has a zero mean $E[\boldsymbol{y}] = \boldsymbol{0}$ and a power constraint P_T, i.e.,

$$\mathrm{tr}\{\boldsymbol{Q}\} \le P_T \tag{2}$$

where $\boldsymbol{Q} = E[\boldsymbol{y}\boldsymbol{y}^{\mathrm{H}}]$ is the transmission covariance matrix and $\mathrm{tr}\{\cdot\}$ denotes the trace operation. Throughout this paper, we will always assume $E[\boldsymbol{H}] = \boldsymbol{0}$.

B. Kronecker Model

The discussions hereafter are based on the Kronecker model, which has been widely discussed and validated under certain practical channel conditions [1]-[2], [13]-[14]. This model is given by [13]-[14]

$$\boldsymbol{H} = \boldsymbol{C}_{\mathrm{R}}^{1/2}\boldsymbol{H}_{\mathrm{w}}\boldsymbol{C}_{\mathrm{T}}^{1/2} \tag{3}$$

where $\boldsymbol{H}_{\mathrm{w}}$ is an $N_{\mathrm{R}}\times N_{\mathrm{T}}$ matrix whose elements are i.i.d. complex Gaussian random variables with distribution $CN(0,1)$, and $\boldsymbol{C}_{\mathrm{R}}$ and $\boldsymbol{C}_{\mathrm{T}}$ are, respectively, $N_{\mathrm{R}}\times N_{\mathrm{R}}$ and $N_{\mathrm{T}}\times N_{\mathrm{T}}$ Hermitian matrices modeling the antenna correlation at the receiver and transmitter. We adopt the following normalizations

$$\mathrm{tr}\{\boldsymbol{C}_{\mathrm{R}}\} = N_{\mathrm{R}} \text{ and } \mathrm{tr}\{\boldsymbol{C}_{\mathrm{T}}\} = N_{\mathrm{T}}. \tag{4}$$

C. Statistical Water-Filling (SWF)

In the following, we will briefly discuss some transmission schemes based on mutual information analysis. This will be useful in the discussions on the joint FEC encoder and linear precoder design in the next section.

Denote by

$$R(\boldsymbol{H},\boldsymbol{Q}) = \log_2 \det(\boldsymbol{I}_{N_R} + \boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^H) \tag{5}$$

the achievable rate of the system in (1) with channel realization $\boldsymbol{H}$ and transmission covariance matrix $\boldsymbol{Q}$. Under CCIT assumption, instantaneous CSI for each channel realization is unavailable at the transmitter and so $\boldsymbol{Q}$ can only be designed based on the known $\boldsymbol{C}_T$ and $\boldsymbol{C}_R$. In this case, the ergodic capacity is defined by

$$C_{\text{CCIT}} = \max_{\text{tr}\{\boldsymbol{Q}\}\le P_T} E_{\boldsymbol{H}}\left[R(\boldsymbol{H},\boldsymbol{Q})\right] \tag{6}$$

where the expectation is over the distribution of $\boldsymbol{H}$ conditioned on given $\boldsymbol{C}_T$ and $\boldsymbol{C}_R$. The key here is to find $\boldsymbol{Q}$ that maximizes the average mutual information $E_{\boldsymbol{H}}[R(\boldsymbol{H},\boldsymbol{Q})]$. In general, this is a highly complicated problem.

A sub-optimal option has been developed in [5]-[6] by maximizing the Jensen bound of $E_{\boldsymbol{H}}[R(\boldsymbol{H},\boldsymbol{Q})]$, as illustrated below. Note that

$$\begin{aligned} E_{\boldsymbol{H}}\left[R(\boldsymbol{H},\boldsymbol{Q})\right] &= E\left[\log_2 \det(\boldsymbol{I}_{N_T} + \boldsymbol{Q}\boldsymbol{H}^H\boldsymbol{H})\right] \\ &\le \log_2 \det\left(\boldsymbol{I}_{N_T} + \boldsymbol{Q}E[\boldsymbol{H}^H\boldsymbol{H}]\right) \\ &= \log_2 \det\left(\boldsymbol{I}_{N_T} + N_R\boldsymbol{Q}\boldsymbol{C}_T\right) \end{aligned} \tag{7}$$

Denote the eigenvalue decompositions of $\boldsymbol{Q}$ and $\boldsymbol{C}_T$ by

$$\boldsymbol{Q} = \boldsymbol{U}_Q\boldsymbol{D}_Q\boldsymbol{U}_Q^H, \tag{8}$$

$$\boldsymbol{C}_T = \boldsymbol{U}_T\boldsymbol{D}_T\boldsymbol{U}_T^H. \tag{9}$$

It can be shown that the optimal $\boldsymbol{Q}$ to maximize the Jensen bound in (7) is given by

$$\boldsymbol{U}_Q = \boldsymbol{U}_T. \tag{10}$$

and $\boldsymbol{D}_Q$ generated by solving the following problem:

$$\max_{\boldsymbol{D}_Q} \text{imize:}\quad \log_2 \det\left(\boldsymbol{I}_{N_T} + N_R\boldsymbol{D}_Q\boldsymbol{D}_T\right) \tag{11a}$$

$$\text{subject to:}\quad \text{tr}\{\boldsymbol{D}_Q\} \le P_T. \tag{11b}$$

Solving (11) is straightforward; it is equivalent to water-filling over a parallel channel with channel gains given by the diagonal elements of $N_R\boldsymbol{D}_T$. This method is referred to as statistical water-filling (SWF) in [5], which can obtain a very good mutual information performance.

In the sequel, we focus on how to design an efficient transceiver scheme based on SWF in practically coded systems with a relatively low complexity.

III. Joint FEC Encoder and Linear Precoder Design

As mentioned in the introduction, the direct approach to SWF involves modulation constellations of different sizes. The complexity of this approach can be a serious concern. In this section, we discuss a low-cost alternative involving a uniform constellation with jointly designed FEC encoder and linear precoder. We will show that the SWF performance can be approximately realized by this joint design.

A. Hadamard Precoding

Assume that $\boldsymbol{x}$ is a sequence of FEC coded elements with independent and identical distribution and normalized power 1. Denote by $\boldsymbol{P}$ the linear precoder. The transmitted signal $\boldsymbol{y}$ in (1) is given by

$$\boldsymbol{y} = \boldsymbol{P}\boldsymbol{x} \tag{12a}$$

Let $\boldsymbol{V}_{\text{Had}}$ be a Hadamard matrix with a proper size. (See Fig. 1.) For convenience, we will call

$$\boldsymbol{P} = \boldsymbol{U}_Q\boldsymbol{D}_Q^{1/2}\boldsymbol{V}_{\text{Had}} \tag{12b}$$

a Hadamard precoder. Similar precoder structures have been discussed for channels without CSIT in [9] and for channels with full CSIT in [10]-[11]. In this paper, our focus is on its application in MIMO systems with CCIT. As we will see below, the use of $\boldsymbol{V}_{\text{Had}}$ in (12) facilitates the EXIT chart type curve matching for SWF.

The matrices $\boldsymbol{U}_Q$ and $\boldsymbol{D}_Q$ in (12b) can be simply obtained from the SWF scheme, i.e., they are given by (10) and (11). However, it can be further jointly optimized with the underlying FEC code, as will be detailed in Section III-E.

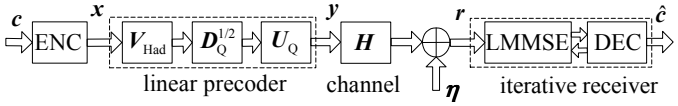


Fig. 1. Illustration of the whole system. ENC stands for encoder, and DEC for decoder.

B. Iterative LMMSE Detection

Substituting (12) into (1), the received signal is given by

$$\boldsymbol{r} = \boldsymbol{H}\boldsymbol{U}_Q\boldsymbol{D}_Q^{1/2}\boldsymbol{V}_{\text{Had}}\boldsymbol{x} + \boldsymbol{\eta}. \tag{13}$$

Applying the standard linear minimum-mean-square-error (LMMSE) detection to $\boldsymbol{r}$ in (13), we have

$$\hat{\boldsymbol{x}} = E[\boldsymbol{x}] + v\left(\tilde{\boldsymbol{H}}\boldsymbol{D}_Q^{1/2}\boldsymbol{V}_{\text{Had}}\right)^H \boldsymbol{R}^{-1}\left(\boldsymbol{r} - \tilde{\boldsymbol{H}}\boldsymbol{D}_Q^{1/2}\boldsymbol{V}_{\text{Had}}E[\boldsymbol{x}]\right) \tag{14a}$$

where $\tilde{\boldsymbol{H}} = \boldsymbol{H}\boldsymbol{U}_Q$, $E[\boldsymbol{x}]$ is the *a priori* mean of $\boldsymbol{x}$, $v\boldsymbol{I}$ is the *a priori* covariance for $\boldsymbol{x}$ and

$$\boldsymbol{R} \triangleq \text{cov}(\boldsymbol{r},\boldsymbol{r}^H) = v(\tilde{\boldsymbol{H}}\boldsymbol{D}_Q^{1/2}\boldsymbol{V}_{\text{Had}})(\tilde{\boldsymbol{H}}\boldsymbol{D}_Q^{1/2}\boldsymbol{V}_{\text{Had}})^H + \sigma^2\boldsymbol{I}. \tag{14b}$$

Initially, if we do not have any information on $\boldsymbol{x}$, we set $E[\boldsymbol{x}] = \boldsymbol{0}$ and $v\boldsymbol{I} = \boldsymbol{I}$ (recalling that the elements in $\boldsymbol{x}$ have average power of 1).

We rewrite (14a) into a symbol-wise form as

$$\hat{x}(i) = v\Omega(i,i)x(i) + \xi(i) \tag{15a}$$

where $\Omega(i,i)$ is the ith diagonal element of the following matrix

$$\boldsymbol{\Omega} \triangleq \left((\tilde{\boldsymbol{H}}\boldsymbol{D}_Q^{1/2}\boldsymbol{V}_{\text{Had}})^H\boldsymbol{R}^{-1}(\tilde{\boldsymbol{H}}\boldsymbol{D}_Q^{1/2}\boldsymbol{V}_{\text{Had}})\right)_{\text{diag}}, \tag{15b}$$

and $\xi(i)$ is an interference-plus-noise term. Following the treatments in [15], we approximate $\xi(i)$ by an AWGN sample with variance [16, Eq. (18)]

$$\text{Var}(\xi(i)) = v\Omega(i,i)(1 - v\Omega(i,i))v. \tag{15c}$$

We can then estimate $x(i)$ by the symbol-wise model in (15a).

The estimate of $\boldsymbol{x}$ is then processed by a FEC decoder using *a posteriori* probability (APP) decoding. The decoding output can be used to update the values of $E[\boldsymbol{x}]$ and v for the

sequence $\boldsymbol{x}$. Then LMMSE detection can be performed again. This process continues iteratively.

In the above detection scheme, the interference among different symbols in $\boldsymbol{x}$ is simply treated as additive noise (included in $\xi(i)$) in the symbol-wise demodulation in (15a). This interference can be suppressed gradually during the iterative process. However, the residual interference may still affect performance noticeably if not treated properly. We will outline an optimization procedure later in Sections III-D and III-E to minimize this interference effect, which is equivalent to maximize the signal to interference-plus-noise ratio (SINR), during the iterative process.

C. Detection Complexity

We now briefly discuss the complexity issue. Assume that a constellation of size 2^M is applied to the symbols of $\boldsymbol{x}$ in (12). The complexity related to the detection operations in Section III-B mainly lies in the calculation of $\boldsymbol{R}^{-1}$ in (14) with complexity $O(N_R N_T \cdot \min(N_R, N_T))$ and the symbol-wise demodulation in (15) with complexity $O(2^M)$. Although here modulation is uniform across all symbols, the scheme is still adaptive to channel conditions, which is achieved by controlling the power allocation matrix $\boldsymbol{D}_Q$ in (12).

As a comparison, consider an alternative based on adaptive modulation, in which constellations of different sizes are involved to realize different transmission rates on different eigen-directions. Recall that the demodulation complexity increases exponentially with M (i.e., the number of bits carried by a symbol) since the constellation size is 2^M. If M is variable over different symbols, 2^M can be very large for some symbols when high rate transmissions are considered. The complexity then becomes a serious problem.

D. Transfer Function for LMMSE Detection

With the use of the Hadamard matrix $\boldsymbol{V}_{\text{Had}}$ in (12), we can make the following approximation [11]-[12]

$$\boldsymbol{\Omega} \approx \omega \boldsymbol{I} \tag{16}$$

where $\omega = N_T^{-1}\text{tr}\{\boldsymbol{\Omega}\}$. Then, using (16) and (15c), we can compute the SINR in $\hat{x}(i)$ in (15a) as

$$\rho(i) = \frac{|v\Omega(i,i)|^2}{\text{Var}(\xi(i))} = \omega/(1-v\omega), \quad \forall i. \tag{17a}$$

Then the LMMSE module can be characterized by a function

$$\rho = \phi(v) \equiv \omega/(1-v\omega). \tag{17b}$$

The transfer function in (17) can be simplified in the asymptotic case, which will be useful for the optimization procedure discussed later. To show this, we assume that $\boldsymbol{C}_R$ does not contain dominant eigenvalues when $N_R \to \infty$. i.e.,

$$\lim_{N_R \to \infty} \lambda_n(\boldsymbol{C}_R)/\text{tr}\{\boldsymbol{C}_R\} = 0, \forall n \tag{18}$$

where $\lambda_n(\boldsymbol{C}_R)$ is the nth eigenvalue of matrix $\boldsymbol{C}_R$. (The assumption holds provided that there is sufficient spacing among receive antennas. This is possible for uplink transmissions where the physical size of a base station is usually much larger than that of a mobile terminal.)

Proposition 1: Assume the Hadamard precoder (12) with $\boldsymbol{U}_Q = \boldsymbol{U}_T$. When (18) holds, for $N_R \to \infty$, we have

$$\phi(v) \approx \omega_\infty(v)/(1-v\omega_\infty(v)) \tag{19a}$$

where

$$\omega_\infty(v) \equiv \frac{1}{N_T}\left[\sum_n \frac{D_T(n,n)D_Q(n,n)}{vD_T(n,n)D_Q(n,n)+\sigma^2/N_R}\right]. \tag{19b}$$

Proof: See Appendix A.

Proposition 1 shows that $\phi(v)$ is asymptotically determined by $\boldsymbol{C}_T$. This leads to a key conclusion of this paper that for an MIMO system with a sufficiently large N_R, an efficient transmitter can be designed without the explicit knowledge of the complete channel coefficients. Knowing $\boldsymbol{C}_T$ is sufficient for this purpose. This fact greatly relieves the burden of channel state information acquiring at the transmitter.

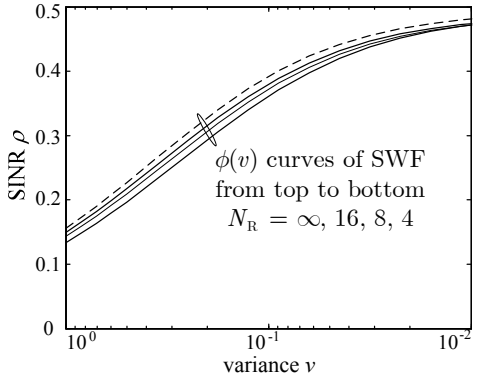


Fig. 2. The $\phi(v)$curves of SWF for different N_R values. $N_T = 8$, $\rho_T = 0.8$, $\rho_R = 0.5$, and $E_b/N_0 = -10\log_{10}(N_R)$.

To illustrate the convergence speed for LMMSE detection related to Proposition 1, we consider numerical examples. For simplicity, we adopt the following exponential model [17]-[18] to characterize the antenna correlation, i.e.,

$$\boldsymbol{C}_R(m,n) = \rho_R^{|m-n|}e^{j(m-n)\theta_R} \text{ and } \boldsymbol{C}_T(m,n) = \rho_T^{|m-n|}e^{j(m-n)\theta_T}$$

where $j \equiv \sqrt{-1}$, ρ_{BS} and $\rho_{MT} \in [0, 1]$ are respectively receive and transmit correlation factors, and θ_{BS} and θ_{MT} are uniformly distributed over $[0, 2\pi)$.

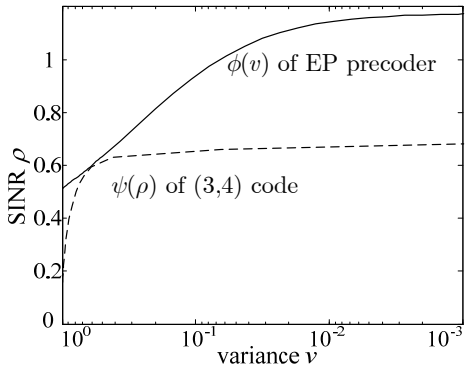
Figure 2 provides a numerical example. The $\phi(v)$ curves for the actual channel $\boldsymbol{H}$ and the asymptotical result in (19) are provided. Noting the term σ^2/N_R in the denominator in (19b), we set $E_b/N_0 = -10\log_{10}(N_R)$ in Fig. 2 for different N_R values for a fair comparison. We can see that the curves are quite close for all the N_R values considered. This property is crucial for our discussions in the next subsection.

E. Transmitter Optimization

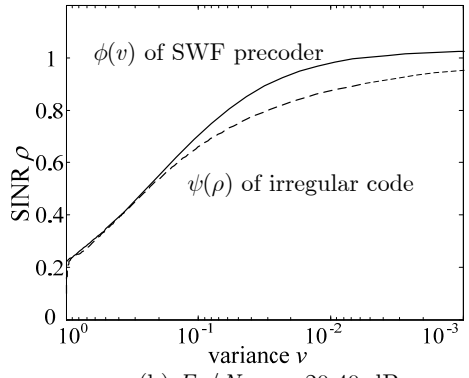
The performance of the FEC decoder is determined by its input SINR (denoted by ρ) and the quality of its output can be measured by the variance (denoted by v). This can be characterized by a function $v = \psi(\rho)$ that can be produced by pre-simulation [19]. This function is the counterpart of $\rho = \phi(v)$ discussed earlier.

Similar to [7]-[8], we can show that the performance of the

iterative receiver in Fig. 1 is characterized by a fixed point of the two transfer functions $\rho = \phi(v)$ and $v = \psi(\rho)$. Given $\rho = \phi(v)$, we can carefully design an irregular LDPC code such that $\psi(\rho)$ matches $\phi(v)$. This topic has been discussed in [20]-[21] and details will be omitted.



(a) $E_b / N_0 = -14.60$ dB



(b) $E_b / N_0 = -20.40$ dB

Fig. 3. SINR-variance transfer curves. $N_T = 8$, $N_R = 64$, $\rho_T = 0.8$, and $\rho_R = 0.5$.

We now demonstrate the effectiveness of such an optimization using numerical results in Fig. 3. For reference, we consider a regular LDPC code with equal power (EP) allocation (in which $\boldsymbol{Q} = P_T/N_T \cdot \boldsymbol{I}$) in Fig. 3(a). Specifically, a rate-0.25 regular (3, 4) LDPC code with length 32768 is adopted, followed by QPSK modulation. The number of transmit antennas is $N_T = 8$ and the total transmission rate is 4 bits/symbol. In Fig. 3(a), we can observe an early crossing point between $\phi(v)$ for EP and $\psi(\rho)$ for the regular (3,4) LDPC code. This leads to considerable performance deterioration, as will be shown in Fig. 4.

In Fig. 3(b), we first generate $\phi(v)$ using the SWF technique described in Section II. We then generate a rate-0.25 irregular LDPC code with a $\psi(\rho)$ function that best matches $\phi(v)$. Note that we can further match the two functions by alternating between optimizing $\phi(v)$ for a given $\psi(\rho)$ and optimizing $\psi(\rho)$ for a given $\phi(v)$. However, we observed only limited gain in this way.

Figure 4 shows the simulated performance for the design examples in Fig. 3. We can see that though the LDPC code is designed based on the asymptotic analysis, it works well in systems even for $N_R = 8$. The SWF scheme obtains a significant improvement from the EP performance. The threshold given by the EXIT chart type analysis is less 0.5 dB away from the full CSIT capacity limit. Note that these results are obtained under the assumption of $N_R \geq N_T$ (recalling the condition $N_R \to \infty$ in Proposition 1). If this assumption does not hold, the performance of the SWF scheme can degrade and then full CSIT may be necessary for transmitter optimization.

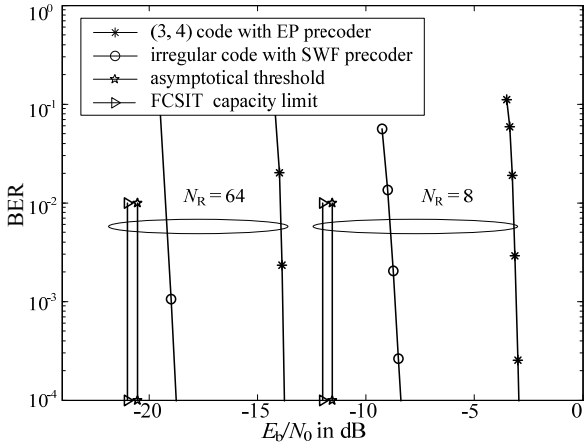


Fig. 4. Simulation performance of the proposed scheme. $N_T = 8$, $\rho_T = 0.8$, $\rho_R = 0.5$, and rate = 4 bits/symbol.

In the above discussion, the LDPC code is optimized for every realization of $\boldsymbol{C}_T$. In practice, when $\boldsymbol{C}_T$ changes, a different LDPC code is optimized. The transmitter needs to inform the receiver about the structure of this code. This incurs considerable complexity overhead.

An alternative is to fix coding for all realizations of $\boldsymbol{C}_T$ and employ adaptive precoding (by optimizing $\boldsymbol{D}_Q$ in (12)) for different $\boldsymbol{C}_T$. Note that, unlike adaptive coding, the transmitter does not need to inform the receiver about the precoder structure. This is because the receiver can estimate $\boldsymbol{HP}$ as an equivalent channel, which is sufficient for the purpose of data detection. Fig. 5 shows that the above approach is effective within a certain statistical range of $\boldsymbol{C}_T$, for which the code is optimized at $\rho_T = 0.8$ and the precoder is adaptive for different $\rho_T \in [0.7, 0.9]$. We can again see significant performance gain in this case.

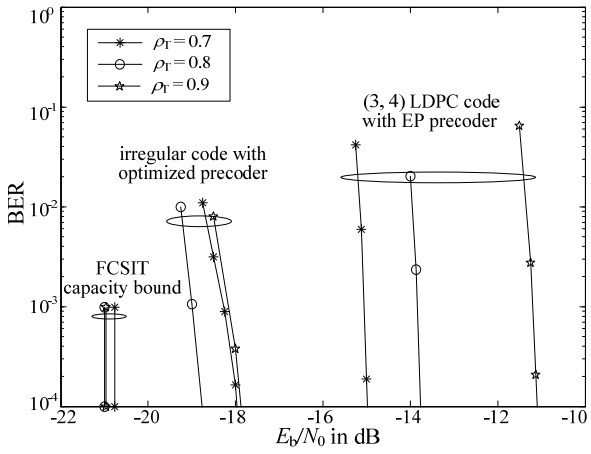


Fig. 5. Simulation performance of the proposed scheme with varying ρ_T, $N_T = 8$, $\rho_R = 0.5$, and rate = 4 bits/symbol.

IV. CONCLUSIONS

We have studied the joint FEC encoder and linear precoder design in MIMO systems with antenna correlation. We show that asymptotically accurate encoder-precoder matching can be made, only requiring the correlation matrix at the transmitter, in systems with $N_R \to \infty$. Excellent performance based on this strategy is demonstrated by simulation results.

ACKNOWLEDGMENT

This work was supported by a grant from the Research Grant Council of the Hong Kong SAR, China [Project No. CityU 118013].

APPENDIX A PROOF OF PROPOSITION 1

To prove (19), we first rewrite (16) as follows

$$\begin{aligned}\omega &= N_T^{-1}\mathrm{tr}\left\{(\tilde{\boldsymbol{H}}\boldsymbol{D}_Q^{1/2}\boldsymbol{V}_{\mathrm{Had}})^H\left(v\tilde{\boldsymbol{H}}\boldsymbol{D}_Q\tilde{\boldsymbol{H}}^H+\sigma^2\boldsymbol{I}\right)^{-1}(\tilde{\boldsymbol{H}}\boldsymbol{D}_Q^{1/2}\boldsymbol{V}_{\mathrm{Had}})\right\}\\ &= N_T^{-1}\mathrm{tr}\left\{(v\tilde{\boldsymbol{H}}\boldsymbol{D}_Q\tilde{\boldsymbol{H}}^H+\sigma^2\boldsymbol{I})^{-1}\tilde{\boldsymbol{H}}\boldsymbol{D}_Q\tilde{\boldsymbol{H}}^H\right\}\\ &= N_T^{-1}\sum_{n=1}^{N_R}\frac{\lambda_n(\tilde{\boldsymbol{H}}\boldsymbol{D}_Q\tilde{\boldsymbol{H}}^H)}{v\lambda_n(\tilde{\boldsymbol{H}}\boldsymbol{D}_Q\tilde{\boldsymbol{H}}^H)+\sigma^2}\\ &= N_T^{-1}\sum_{n=1}^{N_T}\frac{\lambda_n(\boldsymbol{D}_Q^{1/2}\tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{H}}\boldsymbol{D}_Q^{1/2})}{v\lambda_n(\boldsymbol{D}_Q^{1/2}\tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{H}}\boldsymbol{D}_Q^{1/2})+\sigma^2}\end{aligned} \tag{A1}$$

where $\lambda_n(\boldsymbol{A})$ represents the nth eigenvalue of matrix $\boldsymbol{A}$.

Recall that the elements of $\boldsymbol{H}_w$ are i.i.d. complex Gaussian random variables $CN(0,1)$. From the law of large numbers and the assumption in (18), we can show that

$$\lim_{N_R\to\infty}\left(\boldsymbol{H}_w^H\boldsymbol{C}_R\boldsymbol{H}_w\right)_{i,j}\to\begin{cases}\sum_{n=1}^{N_R}C_R(n,n), & i=j;\\ 0, & i\neq j.\end{cases}$$

where $(\boldsymbol{A})_{i,j}$ denotes the (i,j) element of matrix $\boldsymbol{A}$, $(\boldsymbol{A})_{i,:}$ the ith row of matrix $\boldsymbol{A}$, and $(\boldsymbol{A})_{:,j}$ the jth column of matrix $\boldsymbol{A}$. Hence

$$\boldsymbol{H}_w^H\boldsymbol{C}_R\boldsymbol{H}_w\to\mathrm{tr}\{\boldsymbol{C}_R\}\cdot\boldsymbol{I}_{N_T}=N_R\cdot\boldsymbol{I}_{N_T},\ N_R\to\infty. \tag{A2}$$

Then we have

$$\begin{aligned}&\boldsymbol{D}_Q^{1/2}\tilde{\boldsymbol{H}}^H\tilde{\boldsymbol{H}}\boldsymbol{D}_Q^{1/2}\\ &=\boldsymbol{D}_Q^{1/2}\left(\boldsymbol{C}_R^{1/2}\boldsymbol{H}_w\boldsymbol{U}_T\boldsymbol{D}_T^{1/2}\right)^H\left(\boldsymbol{C}_R^{1/2}\boldsymbol{H}_w\boldsymbol{U}_T\boldsymbol{D}_T^{1/2}\right)\boldsymbol{D}_Q^{1/2}\\ &=\boldsymbol{D}_Q^{1/2}\boldsymbol{D}_T^{1/2}\cdot\boldsymbol{U}_T^H\boldsymbol{H}_w^H\boldsymbol{C}_R\boldsymbol{H}_w\boldsymbol{U}_T\cdot\boldsymbol{D}_T^{1/2}\boldsymbol{D}_Q^{1/2}\\ &\to\boldsymbol{D}_Q\boldsymbol{D}_T\cdot N_R\cdot\boldsymbol{I}_{N_T}\end{aligned} \tag{A3}$$

Substituting (A3) in (A1), we complete the proof □

REFERENCES

[1] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Sel. Areas Commun*., vol. 21, no. 5, pp. 684-702, Jun. 2003.

[2] M. Vu, and A. Paulraj, "MIMO wireless linear precoding," *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 86-105, Sep. 2007.

[3] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590-3600, Nov. 2010.

[4] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, pp. 40-60, Jan. 2013.

[5] E. Yoon, J. Hansen, and A. Paulraj, "Space-frequency precoding with space-tap correlation information at the transmitter," *IEEE Trans. Commun.*, vol. 55, no. 9, pp. 1702-1711, Sep. 2007.

[6] X. Li, S. Jin, X. Gao, and K. K. Wong, "Near-optimal power allocation for MIMO channels with mean or covariance feedback," *IEEE Trans.Commun.*, vol. 58, no. 1, pp. 289-300, Jan. 2010.

[7] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1727-1737, Oct. 2001.

[8] S. ten Brink, G. Kramer, and A. Ashikhmin, "Design of low density parity-check codes for modulation and detection," *IEEE Trans. Commun.*, vol. 52, no. 4, pp. 670-678, Apr. 2004.

[9] P. Lusina, E. M. Gabidulin, and M. Bossert, "Efficient decoding of space-time Hadamard codes using the Hadamard transform," in *Proc. ISIT2001*, Washington DC, Jun. 24-29, 2001.

[10] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: a unified framework for convex optimization," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2381-2401, Sep. 2003.

[11] X. Yuan, C. Xu, Li Ping, and X. Lin, "Precoder design for multiuser MIMO ISI channels based on iterative LMMSE detection," *IEEE J. Sel. Topics Signal Process.*, vol. 3, no. 6, pp. 1118-1128, Dec. 2009.

[12] C. Xu, J. Zhang, F. Zheng, and Li Ping, "Linear precoding in distributed MIMO systems with partial CSIT," *EURASIP J. Adv. Signal Process.*, 2013:22, Feb. 2013.

[13] C. N. Chuah, D. N. C. Tse, and J. M. Kahn, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Inf. Theory*, vol. 48, pp. 637-650, Mar. 2002.

[14] J. P. Kermoal, L. Schumacher, K. I. Pedersen, P. E. Mogensen, and F. Fredriksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 6, pp. 1211-1226, Aug. 2002.

[15] X. Wang, and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Commun.*, vol. 47, no. 7, pp. 1046-1061, Jul. 1999.

[16] Li Ping, J. Tong, X. Yuan, and Q. Guo, "Superposition coded modulation and iterative linear MMSE detection," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 6, pp. 995-1004, Aug. 2009.

[17] J. Lee, H. Cho, H. Park, and Y. Lee, "Sum-rate capacity of correlated multi-user MIMO channels," *Inf. Theory and Applications Workshop*, San Diego, Jan. 2010.

[18] C. Martin, and B. Ottersten, "Asymptotic eigenvalue distributions and capacity for MIMO channels under correlated fading," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1350-1359, Jul. 2004.

[19] X. Yuan, Q. Guo, X. Wang, and Li Ping, "Evolution analysis of low-cost iterative equalization in coded linear systems with cyclic prefixes," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 2, pp. 301-310, Feb. 2008.

[20] K. R. Narayanan, D. N. Doan, and R. V. Tamma, "Design and analysis of LDPC codes for turbo equalization with optimal and suboptimal soft output equalizers," in *Proc. Allerton Conf. Commun., Control, and Computing, Monticello*, IL, pp. 737-746, Oct. 2002.

[21] X. Yuan, and Li Ping, "Achievable rates of coded linear systems with iterative MMSE detection," in *Proc. IEEE Globecom'09*, Honolulu, HI, USA, Nov.30 -Dec.4 2009.