

# Multuser TH-precoding for TDD-CDMA over multipath channels

I. Berenguer, A. Høst-Madsen and X. Wang

**Abstract:** Nonlinear precoding schemes for downlink time-division duplex-CDMA systems over multipath fading channels, are considered. First, the capacity results of a downlink CDMA system with either multiuser detection or precoding, were obtained and compared. It is seen that the two schemes exhibit similar capacity regions for both sum rate and maximum equal rate, which motivates the development of efficient nonlinear transmitter precoding techniques to reduce the receiver complexity at the mobile units without degrading the system performance. We then develop both bit-wise and chip-wise Tomlinson-Harashima (TH) multiuser precoding methods for downlink CDMA with multipath, to remove multi-user interference, inter-chip interference and inter-symbol interference. Efficient algorithms for multiuser power loading and ordering are also developed. Implementation of the proposed TH-precoding schemes in time-varying channels based on channel prediction is addressed as well. Simulations results are provided to demonstrate the effectiveness of the proposed techniques in suppressing interference in downlink CDMA.

## 1 Introduction

Multuser detection (MUD) techniques are considered powerful for interference suppression in CDMA systems, especially in uplinks, where the base-station receiver has the knowledge of all users' spreading sequences and channel states, and can perform sophisticated signal processing [1]. In the downlinks, however, the mobile receiver typically has only the knowledge of its own spreading sequence and channel state. Although adaptive linear MUD (either training-based or blind) can be employed for such scenario, the performance can be limited because of the linear constraint on the detector. Moreover, the power constraint of the mobile calls for simple receiver processing. On the other hand, the precoding schemes for downlink CDMA effectively transfer the signal processing for interference suppression from the mobile receiver to the base-station transmitter. This approach is feasible if the base-station can estimate the downlink channels of all users [e.g. in systems employing time-division duplexing (TDD) where the uplink and downlink channels are reciprocal]. In [2], a precoding method was proposed which is essentially an implementation of the RAKE receiver at the transmitter. Hence this approach does not attempt to mitigate the multiple-access interference (MAI). Recently, different linear precoding techniques have been proposed to combat MAI and inter-chip interference but without considering inter-symbol interference (ISI) [3, 4]. If the ISI is present then the complexity of these techniques becomes

prohibitive since the dimension of the matrix filter is proportional to the data frame length multiplied by the number of users (i.e. block processing). More recently, bit-wise linear precoding methods have been proposed to reduce the precoding complexity in the presence ISI [5].

The downlink CDMA is a special case of broadcast channels. There has been significant recent interest in characterising the capacity of broadcast channels. In particular, it has been shown that when the interference is non-causally known to the transmitter and unknown to the receiver, the capacity is the same as if the interference were not present – a result known as 'dirty paper coding'. These results were originally proved for Gaussian channels [6], and have been generalised to other types of causal interference [7–9]. Several practical suboptimal implementations of dirty paper coding based on Tomlinson-Harashima (TH) precoding [10, 11] have been proposed, for example for digital subscriber line systems [12], for multi-antenna systems [13, 14] and CDMA systems [15].

In this paper, we first obtain the capacity regions of a downlink CDMA system employing either MUD or precoding. It is seen that these two approaches provide similar capacity regions, but not equivalent, suggesting that precoding can potentially achieve similar performance offered by MUD and reduce the complexity of the mobile user. This motivates the development of practical approaches to 'dirty-paper coding' for downlink TDD-CDMA systems that will be presented in the second part of the paper.

In the second part of the paper, we extend the nonlinear precoding method in [15] to systems with simpler receivers, that is without channel state information (CSI) and with ISI. Note that the work in [15] assumes that each user implements a RAKE receiver and hence assumes the knowledge of CSI at the receiver. In CDMA systems, we have more degrees of freedom: the combination of (a) the spreading at the transmitter, (b) despreading at the receiver, (c) operations at the receiver and (d) precoding operation for cancelling MUI, ICI and ISI can be implemented in many different ways given different results, as shown in this paper. In particular, in this paper, we propose a new

© The Institution of Engineering and Technology 2007

doi:10.1049/iet-com:20060510

Paper first received 31st August 2006 and in revised form 22nd January 2007

I. Berenguer is with the Department of Engineering, University of Cambridge, Cambridge, UK

A. Høst-Madsen is with the Department of Electrical Engineering, University of Hawaii, Honolulu, Hawaii

X. Wang is with the Department of Electrical Engineering, Columbia University, New York, NY, USA

E-mail: wangx@ee.columbia.edu

chip-wise precoding scheme that combines the spreading and TH-precoding operations. Our results show that with a similar complexity at the transmitter, our chip-wise solution with only a fixed matched filter (to the original spreading sequence) at the receiver can outperform the TH precoder with a RAKE receiver (i.e. with CSI at the receiver) proposed in [14, 15]. Furthermore, efficient algorithms for multiuser power loading and ordering are developed. Implementation of the proposed TH-precoding schemes in time-varying channels based on channel prediction is also addressed.

## 2 Downlink capacity regions of MUD and precoding

In TDD systems, the uplink channel and the downlink channel for each individual user is the same. Hence the base station can use the uplink channel information to perform preprocessing for the downlink and thereby transfer sophisticated signal processing from the receiver end to the transmitter end, that is to replace MUD (receiver processing) by precoding (transmitter processing). In this section, we present and compare the capacity results of precoding and MUD in the downlink of a CDMA system. These two approaches for downlink CDMA are illustrated schematically in Fig. 1. Note that this is a special case of the multiple-input multiple-output (MIMO) broadcast channel for which recent progressive developments [7, 16–19] have led to what is considered as the final solution to the capacity region for the general broadcast channel [20]. In what follows, we use results from [7, 20] to obtain and compare the capacity regions of precoding and MUD in the downlink of a CDMA system.

Consider a synchronous CDMA system with  $K$  users signalling over a real-valued AWGN channel. Let  $f_k$  and  $s_k$  be the channel gain and the spreading signature of the  $k$ th user, respectively. Denote  $\mathbf{S} = [s_1, \dots, s_K]$ . Then  $\mathbf{R} = \mathbf{S}^T \mathbf{S}$  is the  $K \times K$  cross-correlation matrix of the spreading waveforms of all users.

### 2.1 Multiuser detection

Assume that the users are ordered according to their path gains so that  $f_1 \geq f_2 \geq \dots \geq f_K$ . The received signal at the  $k$ th mobile receiver is given by  $\mathbf{r}_k = f_k \sum_{\ell=1}^K x_\ell s_\ell + \mathbf{n}_k$ , where  $\mathbf{n}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ . Note that in this case symbols from different users  $x_1, \dots, x_K$  are independently encoded. Denote  $\mathbf{x} = [x_1, \dots, x_K]^T$ . A sufficient

statistic for  $\mathbf{x}$  is the output of a bank of matched-filters [1]

$$\mathbf{y}_k \triangleq [s_1^T \mathbf{r}_k, s_2^T \mathbf{r}_k, \dots, s_K^T \mathbf{r}_k]^T = f_k \mathbf{R} \mathbf{x} + \mathbf{v}_k \quad (1)$$

with  $E\{\mathbf{v}_k \mathbf{v}_k^T\} = \mathbf{R}$ . The  $k$ th user then makes a decision on its own data  $x_k$  based on  $\mathbf{y}_k$ . Denote  $\rho_k$  as the  $k$ th column of  $\mathbf{R}$  and

$$\mathbf{Q}_k \triangleq \mathbf{R} + f_k^2 \sum_{\ell=1}^{k-1} P_\ell \rho_\ell \rho_\ell^T \quad (2)$$

where  $P_k \triangleq E\{x_k^2\}$ . Denote  $P_T \triangleq \sum_{k=1}^K P_k$  as the total transmit power. We have the following result regarding an outer bound on the rate region.

*Proposition 1:* Consider the channel model (1) and suppose that each user's data is encoded independently. Then the multiuser rate tuple  $(R_1, \dots, R_K)$  must satisfy

$$R_k \leq \frac{1}{2} \log(1 + P_k f_k^2 \rho_k^T \mathbf{Q}_k^{-1} \rho_k), \quad k = 1, \dots, K \quad (3)$$

for some  $P_k = 1, \dots, K$ , satisfying  $P_k \geq 0$  and  $\sum_{k=1}^K P_k = P_T$ .

*Proof:* Define  $\mathbf{y}'_k = \mathbf{y}_k / f_k$ . Then we can rewrite the following equivalent model to (1)

$$\mathbf{y}'_1 = \mathbf{R} \mathbf{x} + \mathbf{v}'_1 \quad \text{and} \quad \mathbf{y}'_k = \mathbf{y}'_{k-1} + \mathbf{v}'_k, \quad k = 2, \dots, K \quad (4)$$

where the  $\mathbf{v}'_1, \dots, \mathbf{v}'_K$  are independent, zero-mean Gaussian vectors, and  $E\{\mathbf{v}'_k \mathbf{v}'_k^T\} = f_k^{-2} \mathbf{R}$ . The model (4) is the same as the aligned degraded broadcast channel model in [20]. The difference is that here each  $x_k$  is encoded independently. This corresponds to the model in [20] with  $\mathbf{B}_i$  zero except for the  $i$ th diagonal element and  $\mathbf{S}$  a diagonal matrix. It can be checked that the proof still applies with these restrictions, and we therefore obtain a rate given in [20], which here becomes

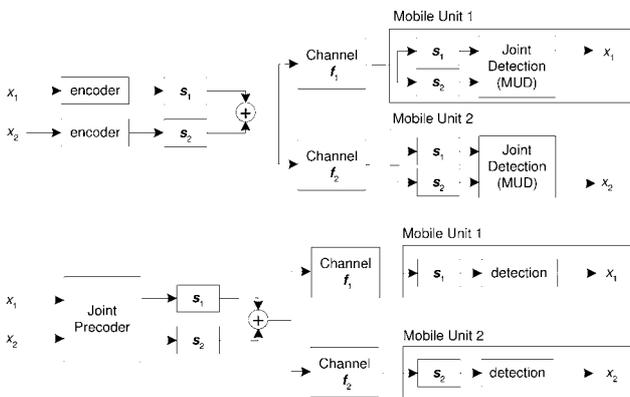
$$\begin{aligned} R_k &\leq \frac{1}{2} \log \left( \frac{\det \left( \sum_{\ell=1}^k P_\ell \rho_\ell \rho_\ell^T + \sum_{\ell=1}^k E\{\mathbf{v}'_\ell \mathbf{v}'_\ell^T\} \right)}{\det \left( \sum_{\ell=1}^{k-1} P_\ell \rho_\ell \rho_\ell^T + \sum_{\ell=1}^k E\{\mathbf{v}'_\ell \mathbf{v}'_\ell^T\} \right)} \right) \\ &= \frac{1}{2} \log \left( \frac{\det(\mathbf{Q}_k + f_k^2 P_k \rho_k \rho_k^T)}{\det \mathbf{Q}_k} \right) \end{aligned} \quad (5)$$

Let  $\mathbf{F}_k$  be a Cholesky factor of  $\mathbf{Q}_k$ , that is  $\mathbf{F}_k \mathbf{F}_k^T = \mathbf{Q}_k$ . Then

$$\begin{aligned} &\det(\mathbf{Q}_k + P_k f_k^2 \rho_k \rho_k^T) \\ &= \det \left[ \mathbf{F}_k \left( \mathbf{I} + \left( \sqrt{P_k} f_k \mathbf{F}_k^{-1} \rho_k \right) \left( \sqrt{P_k} f_k \mathbf{F}_k^{-1} \rho_k \right)^T \right) \mathbf{F}_k^T \right] \\ &= (1 + P_k f_k^2 \rho_k^T \mathbf{Q}_k^{-1} \rho_k) \det \mathbf{Q}_k \end{aligned} \quad (6)$$

where in (6) we used the following identity  $\det(\mathbf{A}\mathbf{B}) = \det(\mathbf{B}\mathbf{A}) = \det(\mathbf{A}) \det(\mathbf{B})$ , and  $\det(\mathbf{I} + \mathbf{a}\mathbf{a}^T) = \mathbf{a}^T \mathbf{a} + 1$ , where  $\mathbf{a}$  is a vector. Substituting (6) into (5) we obtain (3).  $\square$

The rate in Proposition 1 is achievable with a multiuser detector that performs serial interference cancellation on weak users and linear MMSE interference suppression on strong users. In particular, user  $k$  can decode the data intended for users  $k+1, \dots, K$  as user  $k$  receives the same signal but with higher SNR. Suppose that user  $k$  has decoded users  $k+1, \dots, K$ . It then subtracts the signals



**Fig. 1** Schematic illustration of MUD and precoding in downlink CDMA systems

of these users from  $\mathbf{y}_k$  in (1) to obtain

$$\begin{aligned}\tilde{\mathbf{y}}_k &= \mathbf{y}_k - f_k \sum_{\ell=k+1}^K \boldsymbol{\rho}_\ell x_\ell = f_k \sum_{\ell=1}^k \boldsymbol{\rho}_\ell x_\ell + \mathbf{v}_k \\ &= f_k \boldsymbol{\rho}_k x_k + f_k \sum_{\ell=1}^{k-1} \boldsymbol{\rho}_\ell x_\ell + \mathbf{v}_k\end{aligned}\quad (7)$$

It now applies a linear MMSE filter on  $\tilde{\mathbf{y}}_k$ . Note that the covariance matrix of the noise and interference is given by  $\mathbf{Q}_k$  in (2). The linear MMSE filter output is given by

$$\hat{x}_k = \mathbf{m}_k^T \tilde{\mathbf{y}}_k, \quad \text{with} \quad \mathbf{m}_k = f_k \mathbf{Q}_k^{-1} \boldsymbol{\rho}_k \quad (8)$$

This gives a rate

$$\begin{aligned}R_k &= \frac{1}{2} \log \left( 1 + \frac{(\boldsymbol{\rho}_k^T \mathbf{Q}_k^{-1} \boldsymbol{\rho}_k)^2 P_k f_k^2}{\boldsymbol{\rho}_k^T \mathbf{Q}_k^{-1} \boldsymbol{\rho}_k} \right) \\ &= \frac{1}{2} \log (1 + P_k f_k^2 \boldsymbol{\rho}_k^T \mathbf{Q}_k^{-1} \boldsymbol{\rho}_k)\end{aligned}\quad (9)$$

when  $K = 2$ , denote  $\rho \triangleq \rho_{1,2} = \rho_{2,1}$ . The above rate region can be easily evaluated as

$$R_1 = \frac{1}{2} \log (1 + P_1 f_1^2) \quad (10)$$

$$\begin{aligned}R_2 &= \frac{1}{2} \log \left( 1 + \frac{1 + P_1 f_2^2 (1 - \rho^2)}{1 + P_1 f_2^2} P_2 f_2^2 \right) \\ &= \frac{1}{2} \log \left( 1 + \left( 1 - \rho^2 \frac{P_1 f_2^2}{1 + P_1 f_2^2} \right) P_2 f_2^2 \right)\end{aligned}\quad (11)$$

## 2.2 Precoding

In systems employing precoding, each downlink user simply applies a filter matched to its own spreading sequence. The output of this matched-filter is given by  $y_k = f_k \boldsymbol{\rho}_k^T \mathbf{x} + \mathbf{v}_k$ , where  $\mathbf{x}$  is the precoded vector. Stacking the output of the matched-filters of all users in a single vector, we then obtain the downlink precoding signal model:  $\mathbf{y} = \mathbf{A} \mathbf{R} \mathbf{x} + \mathbf{v}$ , where  $\mathbf{A} = \text{diag}(f_1, \dots, f_K)$  and  $E\{\mathbf{v} \mathbf{v}^T\} = \mathbf{I}$ . This is similar to a multiple-antenna broadcast channel [7]. However, note that here the power of the transmitted signal is  $E\{\mathbf{x}^T \mathbf{R} \mathbf{x}\}$ . Therefore the power constraint is  $E\{\mathbf{x}^T \mathbf{R} \mathbf{x}\} \leq P_T$ , which is different from [7]. This is easily fixed: let  $\mathbf{F}$  be the Cholesky factor of  $\mathbf{R}$ , that is  $\mathbf{F} \mathbf{F}^T = \mathbf{R}$ . Define  $\mathbf{u} = \mathbf{F}^T \mathbf{x}$ . Then  $\mathbf{u}$  should satisfy the power constraint  $E\{\mathbf{u}^T \mathbf{u}\} \leq P_T$ , and we can write the received signal as

$$\mathbf{y} = \mathbf{A} \mathbf{R} \mathbf{F}^{-T} \mathbf{u} + \mathbf{v} = \mathbf{A} \mathbf{F} \mathbf{u} + \mathbf{v} \quad (12)$$

This is the same as the model in [7] for a broadcast channel with  $K$  antennas at the base station and one antenna at each terminal, with  $\mathbf{H} \triangleq \mathbf{A} \mathbf{F}$ . In [7], Costa's 'dirty-paper coding' was suggested, and very recently in [20] it was shown that this scheme actually gives the capacity region. Hence the results in [7] apply to the current problem and the following rate is achievable [17]

$$R_k = \frac{1}{2} \log \frac{\det(\mathbf{I} + \mathbf{h}_k^T (\sum_{\ell=k}^K \boldsymbol{\Sigma}_\ell) \mathbf{h}_k)}{\det(\mathbf{I} + \mathbf{h}_k^T (\sum_{\ell=k+1}^K \boldsymbol{\Sigma}_\ell) \mathbf{h}_k)} \quad (13)$$

where  $\boldsymbol{\Sigma}_1, \dots, \boldsymbol{\Sigma}_K$  are positive semidefinite matrices satisfying  $\text{tr}(\sum_{k=1}^K \boldsymbol{\Sigma}_k) \leq P_T$  and  $\mathbf{h}_k^T$  is the  $k$ th row of  $\mathbf{H}$ . The capacity region, as proved in [20], is the convex union over all matrices  $\boldsymbol{\Sigma}_1, \dots, \boldsymbol{\Sigma}_K$  and all orderings of the users. Unfortunately, except for the two-user case solved

in [7], no closed-form solution for the capacity region has been found.

For the case of  $K = 2$ , we can obtain explicit expressions for the capacity region. In [7], it was proved that the capacity region is given by

$$R_1 \leq \log \left( 1 + \frac{\alpha c_1 P_T}{1 + q(1 - \alpha) c_1 P_T} \right) \quad (14)$$

$$R_2 \leq \log (1 + p(1 - \alpha) c_1 P_T), \quad 0 \leq q \leq 1, 0 \leq \alpha \leq 1 \quad (15)$$

where  $c_1 \triangleq \mathbf{h}_1^T \mathbf{h}_1$ ,  $c_2 \triangleq |\det(\mathbf{H})|^2 / c_1$ ,  $c_3 \triangleq |\mathbf{h}_1^T \mathbf{h}_2|^2 / c_1^2$  and

$$p = \left( \sqrt{c_3 q} + \sqrt{\frac{c_2}{c_1} (1 - q)} \right)^2 \quad (16)$$

We can obtain a more explicit expression as follows. First set equality in (14) and solve for  $q$ , to obtain

$$q = \frac{2^{R_1} - 1 - \alpha c_1 P_T}{(1 - \alpha) c_1 P_T (1 - 2^{R_1})} \quad (17)$$

Then substitute (17) into (16) and in (15) with equality, and solve for  $\alpha$  from  $dR_2/d\alpha = 0$ , to obtain the unique solution

$$\alpha = \frac{-c_1^2 P_T + 2^{R_1} c_1^2 c_3 P_T - 2^{2R_1} c_2 - c_1 c_3 + 2^{3R_1} c_2 + 2^{R_1} c_1 c_3}{(c_1 c_3 + 2^{R_1} c_2) 2^{R_1} c_1 P_T} \quad (18)$$

Substituting (18) into (14) and (15), we obtain

$$R_1 \leq \log (1 + c_1 P_T) \quad (19)$$

$$R_2 \leq \log \left( 1 + \frac{c_1^2 c_3 P_T + c_1 c_3 + 2^{R_1} \times (c_1 c_2 P_T + c_2 - c_1 c_3 - c_2 2^{R_1})}{c_1 2^{R_1}} \right) \quad (20)$$

Now substituting (19) with equality into (20), and using the definitions of  $c_i$  and  $\mathbf{H} \triangleq \mathbf{A} \mathbf{F}$ , after some straightforward but tedious simplifications, we obtain

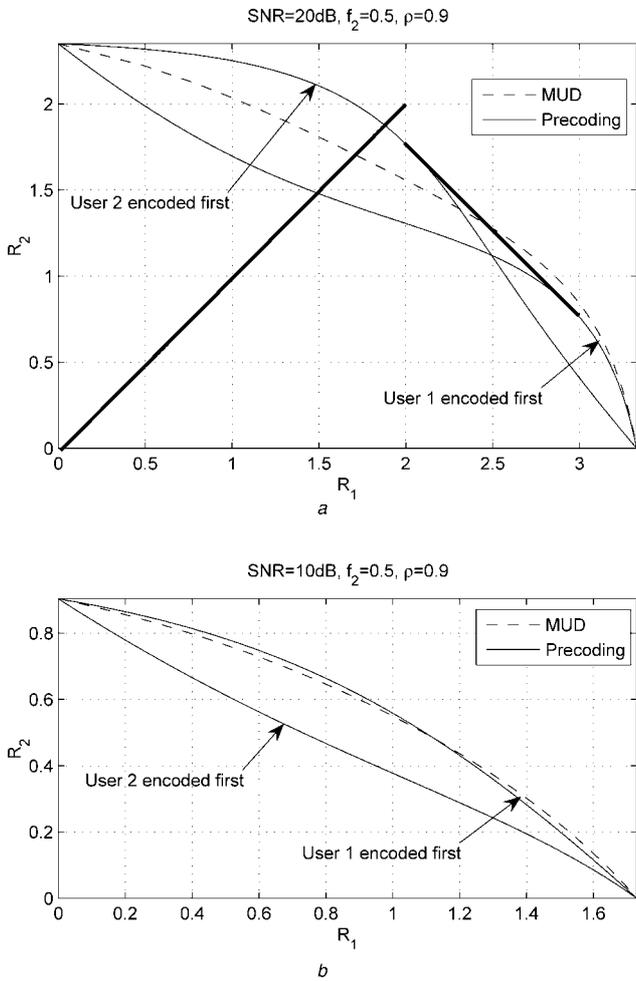
$$R_1 \leq \frac{1}{2} \log (1 + P_1 f_1^2) \quad (21)$$

$$R_2 \leq \frac{1}{2} \log \left( 1 + \left( 1 - \rho^2 \frac{P_1 f_1^2}{1 + P_1 f_1^2} \right) P_2 f_2^2 \right) \quad (22)$$

If we swap the order of the users, we obtain another region, and the total region is the convex closure of these two regions.

## 2.3 Numerical comparisons

We next provide some numerical results comparing the downlink CDMA capacity regions for MUD and precoding with  $K = 2$ . First, we notice that the two capacity expressions (10) and (11), and (21) and (22) are very similar. Fig. 2 shows typical rate regions, one for high SNR and one for low SNR. There are two curves for the precoding case because of the dependency on user ordering (the capacity region is the convex union) and one curve for the MUD case. It is seen that the regions are quite similar. The maximum sum rate is slightly larger for MUD, whereas the maximum equal rate (i.e.  $R_1 = R_2$ ) is slightly larger for precoding. This turns out to be general, as the following numerical results show. Fig. 3 shows the sum rate as a function of  $f_2/f_1$  and  $\rho \in [0, 1]$  with  $f_1$  fixed. It is seen that the sum rate for MUD is consistently



**Fig. 2** Comparisons of rate regions for MUD and precoders with  $K = 2$

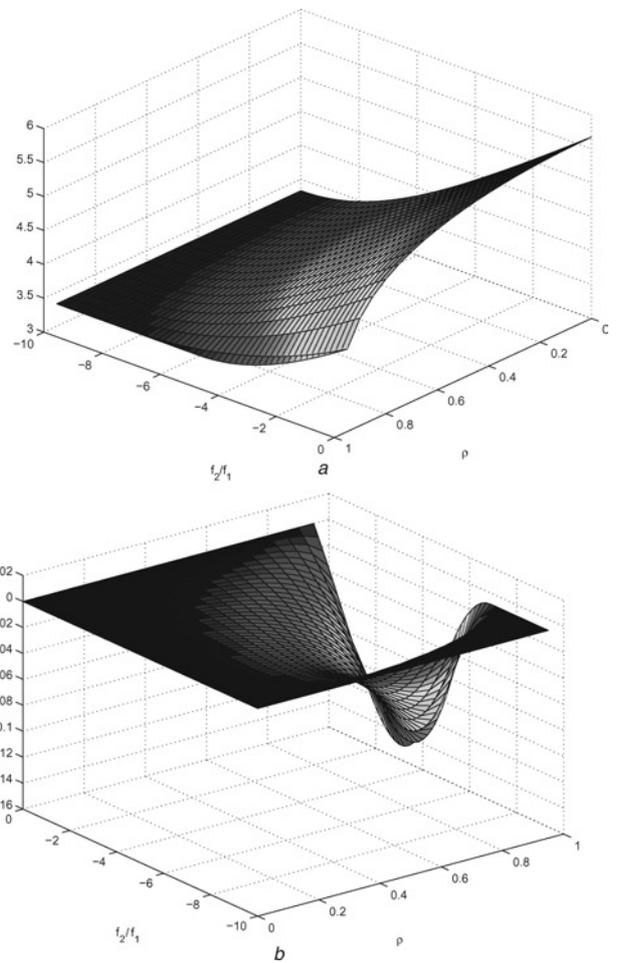
a High SNR  
b Low SNR

better, but only slightly. Similar observation can be made for maximum equal rate in Fig. 4. In summary, it is seen that precoding can potentially provide similar capacity as MUD, which motivates the development of practical transmitter precoding techniques as an alternative to MUD to reduce the complexity of the mobile receiver. In the following sections, we consider more realistic scenarios (i.e. frequency selective channels and larger number of users) for which we propose suboptimal approaches to ‘dirty-paper coding’ based on the TH precoding technique.

### 3 TH precoding in downlink CDMA

#### 3.1 Downlink CDMA system model

We consider a  $K$ -user discrete-time downlink CDMA system signalling over multipath channels. Denote  $b_k[i] \in \mathcal{A}$  as the information symbol of the  $k$ th user transmitted during the  $i$ th symbol interval, where  $\mathcal{A}$  is a finite constellation set; and  $\mathbf{b}[i] = [b_1[i], \dots, b_K[i]]^T$ . Denote the symbol by symbol precoding operation as  $\mathbf{x}[i] = \Psi(\mathbf{b}[i], \dots, \mathbf{b}[i - \nu + 1])$ , where  $\mathbf{x}[i]$  is the  $K \times 1$  precoded symbol vector based on  $\nu$  information symbol vectors. Denote  $N$  as the spreading factor and  $\mathbf{s}_k = [s_{k,1}, \dots, s_{k,N}]^T$  as the spreading waveform of the  $k$ th user. Then the signal transmitted from the base station during the  $i$ th symbol interval can be written as  $\mathbf{p}[i] =$



**Fig. 3** Sum rate for precoding and MUD as a function of  $\rho$  and  $f_2/f_1$  (SNR is 20 dB for user 1)

a Sum rate for precoding and MUD (almost equal, so the two surfaces are nearly indistinguishable)  
b Difference between sum rates for precoding and MUD ( $R_{\text{precoding}} - R_{\text{MUD}}$ )

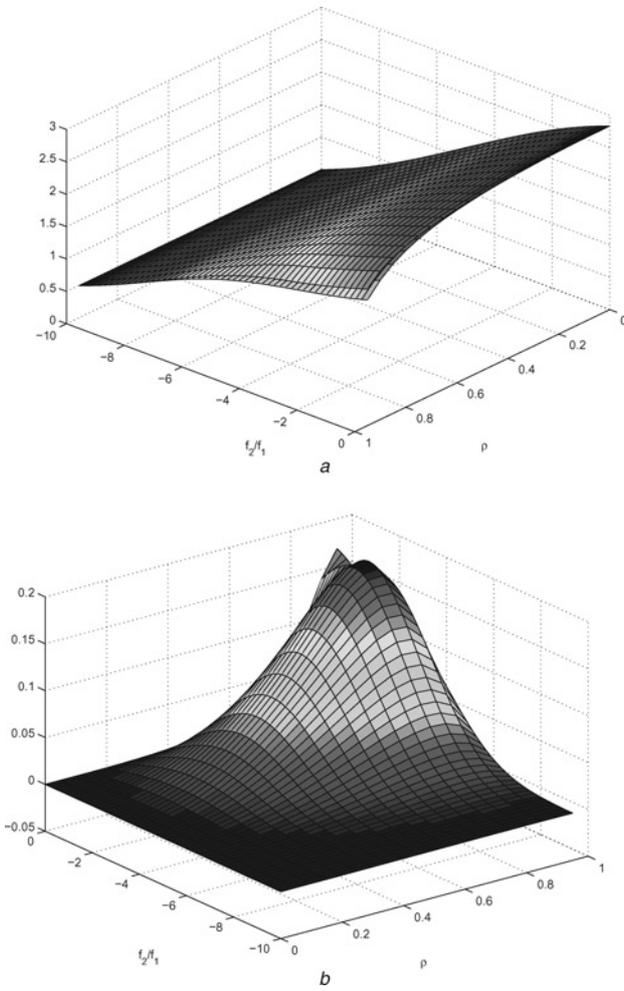
$\mathbf{S}\mathbf{x}[i]$ , where  $\mathbf{S} = [s_1, s_2, \dots, s_K]$ . The vector  $\mathbf{p}[i]$  is passed through a parallel-to-serial converter and transmitted over the wireless channel. The path delays are assumed to be integral multiples of the chip interval. Denote the multipath channel seen by the  $k$ th user as  $\mathbf{f}_k = [f_{k,1}, f_{k,2}, \dots, f_{k,L}]^T$ , where  $L$  is the number of resolvable paths and  $f_{k,l}$  is the complex fading gain corresponding to the  $l$ th path of the  $k$ th user. We assume that  $L \leq N$ , so that the delay spread is at most one symbol interval.

Denote  $\mathbf{r}_k[i]$  as the  $N \times 1$  received signal vector by the  $k$ th user during the  $i$ th symbol interval (i.e.  $N$  consecutive chip intervals). Then

$$\mathbf{r}_k[i] = \mathbf{D}_k \mathbf{S}\mathbf{x}[i] + \bar{\mathbf{D}}_k \mathbf{S}\mathbf{x}[i - 1] + \mathbf{n}_k[i] \quad (23)$$

where  $\mathbf{n}_k[i] \sim \mathcal{N}_c(\mathbf{0}, \sigma_n^2 \mathbf{I}_N)$  is the complex white Gaussian noise vector at the  $k$ th receiver,

$$\mathbf{D}_k = \begin{bmatrix} f_{k,1} & 0 & \dots & 0 \\ f_{k,2} & f_{k,1} & 0 & \ddots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & f_{k,L} & \dots & f_{k,1} \end{bmatrix}_{N \times N} \quad \text{and}$$



**Fig. 4** Equal rate for precoding and MUD as a function of  $\rho$  and  $f_2/f_1$  (SNR is 20 dB for user 1)

a Equal rate for precoding and MUD (almost equal, so the two surfaces are nearly indistinguishable)

b Difference between equal rates for precoding and MUD ( $R_{\text{precoding}} - R_{\text{MUD}}$ )

$$\bar{\mathbf{D}}_k = \begin{bmatrix} 0 & \dots & f_{k,L} & \dots & f_{k,2} \\ 0 & \dots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & f_{k,L} \\ 0 & \dots & \dots & 0 & 0 \end{bmatrix}_{N \times N}$$

At the  $k$ th mobile receiver, a matched-filter is applied to the received signal  $\mathbf{r}_k[i]$  with this user's signature waveform, that is  $y_k[i] = \mathbf{s}_k^H \mathbf{r}_k[i]$ . By stacking the matched-filter output from all users into a single vector we have

$$\underbrace{\begin{bmatrix} \mathbf{s}_1^H \mathbf{r}_1[i] \\ \mathbf{s}_2^H \mathbf{r}_2[i] \\ \vdots \\ \mathbf{s}_K^H \mathbf{r}_K[i] \end{bmatrix}}_{\mathbf{y}[i]} = \underbrace{\begin{bmatrix} \mathbf{s}_1^H \mathbf{D}_1 \mathbf{S} \\ \mathbf{s}_2^H \mathbf{D}_2 \mathbf{S} \\ \vdots \\ \mathbf{s}_K^H \mathbf{D}_K \mathbf{S} \end{bmatrix}}_{\mathbf{H}} \mathbf{x}[i] + \underbrace{\begin{bmatrix} \mathbf{s}_1^H \bar{\mathbf{D}}_1 \mathbf{S} \\ \mathbf{s}_2^H \bar{\mathbf{D}}_2 \mathbf{S} \\ \vdots \\ \mathbf{s}_K^H \bar{\mathbf{D}}_K \mathbf{S} \end{bmatrix}}_{\bar{\mathbf{H}}} \mathbf{x}[i-1] + \underbrace{\begin{bmatrix} \mathbf{s}_1^H \mathbf{n}_1[i] \\ \mathbf{s}_2^H \mathbf{n}_2[i] \\ \vdots \\ \mathbf{s}_K^H \mathbf{n}_K[i] \end{bmatrix}}_{\mathbf{v}[i]} \quad (25)$$

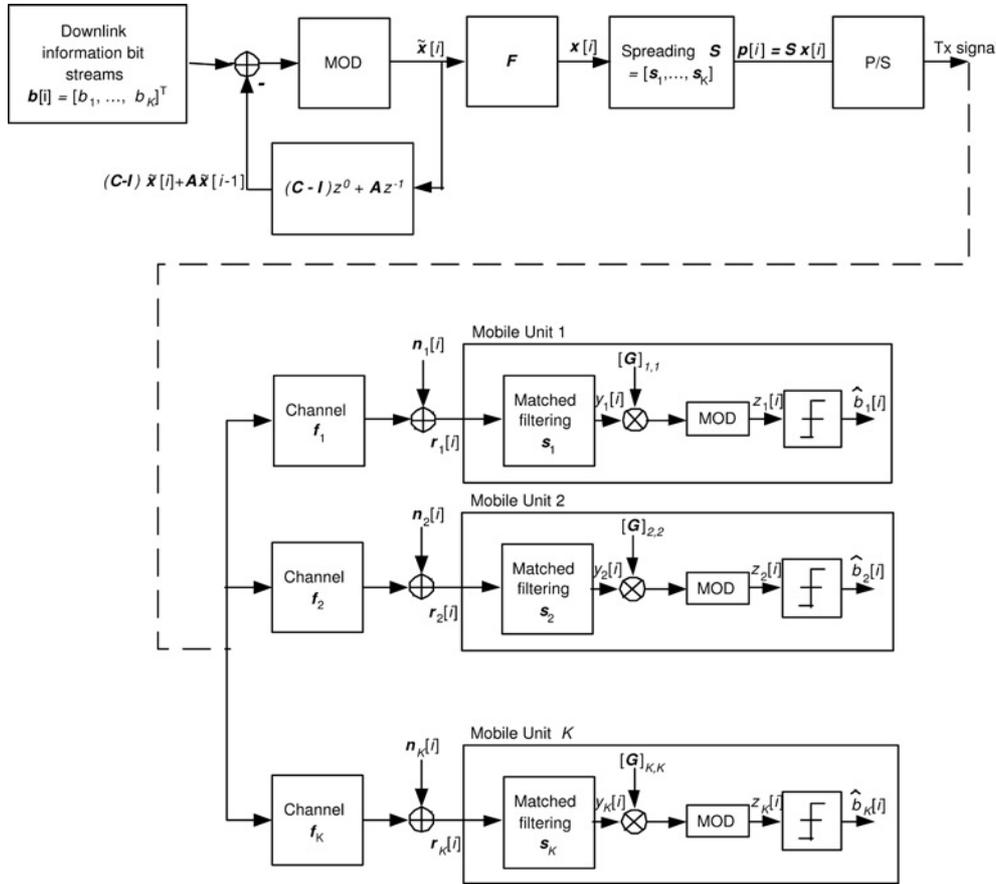
A different situation is when, instead of applying a filter matched to the original spreading waveform (i.e. channel independent), the receiver implements a RAKE receiver as proposed in [15]. The main difficulty is that each receiver must also estimate the channel to apply the RAKE receiver, thus, increasing the number of pilot symbols and the complexity of the receiver. We discuss this method only for comparison, since we seek precoding solutions with simple receivers without CSI at the receiver. The RAKE receiver can be implemented with a matched filter using the normalised composite spreading waveform (i.e. the  $k$ th composite spreading waveform is given by the convolution of the channel and the original spreading waveform  $\bar{\mathbf{s}}_k = \mathbf{f}_k * \mathbf{s}_k$ ) instead of the original spreading waveform. With our notation, the normalised  $k$ th composite spreading waveform is given by  $\bar{\mathbf{s}}_k = (1/\|\mathbf{D}_k \mathbf{s}_k\|) \mathbf{D}_k \mathbf{s}_k$ , where we have limited the convolution to  $N$  chip samples. Let us define TH-bit-original and TH-bit-RAKE to differentiate between the TH solutions that will be obtained when the receiver applies a filter matched to the original spreading waveform or the composite waveform, respectively.

The problem of precoder design is how to best choose an appropriate precoding function  $\Psi(\cdot)$  so that the output vector  $\mathbf{y}[i]$  is as close as possible to the transmitted data vector  $\mathbf{b}[i]$ .

### 3.2 Bit-wise multiuser THP (TH-bit-original and TH-bit-RAKE)

If the ISI term  $\bar{\mathbf{H}}\mathbf{x}[i-1]$  were not present in (25) (this is the case when a guard interval is inserted between consecutive symbols), then the TH precoding scheme in [14] for MIMO systems can be directly applied here. In such a case, the TH precoder consists of a feedforward (FF) filter matrix  $\mathbf{F}$  and a feedback (FB) filter matrix  $(\mathbf{C} - \mathbf{I})$ , which are obtained in the following way. Denote the LQ factorisation of the matrix  $\mathbf{H}$  as  $\mathbf{H} = \mathbf{W}\mathbf{F}^H$ , where  $\mathbf{F}$  is unitary and  $\mathbf{W}$  is lower triangular. The purpose of the FF matrix  $\mathbf{F}$  is to convert the interference into a causal form without increasing the transmit power. This permits the cancellation of the causal interference using the FB filter matrix  $(\mathbf{C} - \mathbf{I})$ . For the interference cancellation to be possible,  $\mathbf{C}$  needs to be monic lower triangular. To obtain  $\mathbf{C}$ , decompose  $\mathbf{W} = \mathbf{G}^{-1}\mathbf{C}$ , where  $\mathbf{G}$  is the diagonal matrix that makes  $\mathbf{C}$  monic, that is  $\mathbf{G} = \text{diag}(w_{1,1}^{-1}, \dots, w_{K,K}^{-1})$  where  $w_{i,i}$  denotes the  $i$ th diagonal element in  $\mathbf{W}$ . Denote  $\tilde{\mathbf{x}}[i]$  as the output of the FB filter. Then we have  $\tilde{\mathbf{x}}[i] = \mathbf{b}[i] - (\mathbf{C} - \mathbf{I})\tilde{\mathbf{x}}[i]$ , and consequently, the equivalent FB operation is  $\tilde{\mathbf{x}}[i] = \mathbf{C}^{-1}\mathbf{b}[i]$ . Thus, the input data symbols  $\mathbf{b}[i]$  are first passed through the FB filter  $\mathbf{C}^{-1}$  and then through the FF filter  $\mathbf{F}$ , that is  $\mathbf{x}[i] = \mathbf{F}\mathbf{C}^{-1}\mathbf{b}[i]$ , followed by spreading (cf. Fig. 5).

**Feedback and modulo operations:** Owing to the lower triangular structure of the matrix  $\mathbf{C}$ , the output of the FB filter  $\tilde{x}_k$ ,  $k = 1, \dots, K$ , is successively generated from the input data symbols  $b_k[i] \in \mathcal{A}$ , and the previous output of the FB filter,  $\tilde{x}_\ell[i], \ell = 1, \dots, k-1$ , and  $\tilde{x}_k[i] = b_k[i] - \sum_{\ell=1}^{k-1} c_{k,\ell} \tilde{x}_\ell[i]$ ,  $k = 1, \dots, K$ . To prevent an increase in transmit power, a modulo operation with respect to  $\mathcal{A}$  is applied. For example, for  $M$ -QAM constellations, the modulo operation corresponds to adding integer multiples of  $2\sqrt{M}$  to the real and the imaginary parts of  $b_k[i]$ , so that the resulting output signal falls in the range of  $\mathcal{A}$ .



**Fig. 5** Bit-wise TH-precoded downlink CDMA system over multipath channels

Then the output of the FB filter becomes

$$\tilde{x}_k[i] = b_k[i] + d_k[i] - \sum_{\ell=1}^{k-1} c_{k,\ell} \tilde{x}_\ell[i], \quad k = 1, \dots, K \quad (26)$$

where  $d_k[i] \in \{2\sqrt{M}(d_I + jd_Q)|d_I, d_Q \in \mathbb{Z}\}$ . That is, instead of feeding back  $b_k[i]$ , the symbols  $v_k[i] = b_k[i] + d_k[i]$  are passed through  $C^{-1}$ . If the receiver applies the same modulo operation, then the effect is cancelled.

*Cascade of operations:* At the  $k$ th user's receiver, a matched-filter  $s_k$ , a scalar operation  $g_k = \mathbf{G}[k, k] = w_{k,k}^{-1}$  and the same modulo operation as applied at the transmitter are applied to the received signal  $r_k[i]$ . Therefore without considering the modulo operation, the end-to-end operation for all  $K$  users is given by

$$z[i] = \mathbf{G}(\mathbf{H}\mathbf{F}\mathbf{C}^{-1}\mathbf{b}[i] + \mathbf{v}[i]) = \mathbf{b}[i] + \mathbf{v}[i] \quad (27)$$

and the  $k$ th user makes a decision on  $b_k[i]$  based on the decision statistic  $z_k[i]$ . Note that the scalar gains  $g_k$ ,  $k = 1, \dots, K$ , can be either estimated at the mobile receiver (automatic gain control) or broadcast by the base station.

*TH-precoding with ISI:* Consider now (25) without dropping the ISI term. In addition to the FF and FB filters discussed earlier based on the decomposition  $\mathbf{H} = \mathbf{G}^{-1}\mathbf{C}\mathbf{F}^H$ , another the FB filter is employed to cancel the ISI term  $\bar{\mathbf{H}}\mathbf{x}[i-1]$ . Suppose that the previously precoded symbol  $\tilde{\mathbf{x}}[i-1] = \mathbf{F}^H\mathbf{x}[i-1]$ , is first filtered by a filter  $\mathbf{A}$  and then subtracted from the current data symbol  $\mathbf{b}[i]$ , as shown in Fig. 5. To find the matrix  $\mathbf{A}$  that minimises the mean-square error (MSE), consider the error signal at the

decision device

$$e[i] = \underbrace{\mathbf{G}\mathbf{H}\mathbf{F}\mathbf{C}^{-1}(\mathbf{b}[i] - \mathbf{A}\mathbf{F}^H\mathbf{x}[i-1]) + \mathbf{G}\mathbf{v}[i] + \mathbf{G}\bar{\mathbf{H}}\mathbf{x}[i-1] - \mathbf{b}[i]}_{z[i]} \quad (28)$$

By the orthogonality principle,  $E\{e\mathbf{z}^H\} = \mathbf{0}$ , which leads to  $(\mathbf{G}\bar{\mathbf{H}} - \mathbf{G}\mathbf{H}\mathbf{F}\mathbf{C}^{-1}\mathbf{A}\mathbf{F}^H) = \mathbf{0}$ , that is  $\mathbf{A} = \mathbf{G}\bar{\mathbf{H}}\mathbf{F}$ . Hence the end-to-end cascade of operations is given by

$$\begin{aligned} z[i] &= \underbrace{\mathbf{G}}_{\text{rx}} \left( \underbrace{\mathbf{G}^{-1}\mathbf{C}\mathbf{F}^H}_{\text{channel}} \left( \underbrace{\mathbf{F}}_{\text{FF}} \underbrace{\mathbf{C}^{-1}}_{\text{cancel,FB}} \right) \right) \\ &\quad \times \left( \underbrace{\mathbf{b}[i] - \mathbf{G}\bar{\mathbf{H}}\mathbf{x}[i-1]}_{\text{cancel ISI}} \right) + \underbrace{\bar{\mathbf{H}}\mathbf{x}[i-1]}_{\text{ISI channel}} + \mathbf{v}[i] \\ &= \mathbf{b}[i] + \mathbf{G}\mathbf{v}[i] \end{aligned} \quad (29)$$

where the modulo operation is not included for clarity. The transmitter and receiver diagram for the bit-wise TH-precoded downlink CDMA system is shown in Fig. 5.

**3.2.1 Bit-wise multiuser THP with composite waveform (TH-bit-preRAKE):** In the previous scheme, the spreading at the transmitter was implemented with the spreading matrix  $\mathbf{S}$  and the receiver implemented a matched filter using either the original spreading waveform (i.e. no CSI at the receiver) or the composite spreading waveform (i.e. RAKE receiver with CSI at the receiver). Here, in order to simplify the implementation of the mobile units, we only consider a filter matched to the original spreading waveform (i.e. no CSI at the receiver) and we allow spreading at the transmitter using the composite spreading waveform. The composite spreading waveform

is the convolution of the channel response and the spreading code. That is, at the transmitter, we consider the spreading matrix  $\hat{\mathbf{S}} = [\hat{s}_1, \dots, \hat{s}_K]$ , where  $\hat{s}_i = (1/\|\mathbf{D}_i \mathbf{s}_i\|)\mathbf{D}_i \mathbf{s}_i$ ,  $i = 1, \dots, K$  and the implementation in Fig. 5 is valid by substituting the transmit spreading operation by  $\hat{\mathbf{S}}$ . This operation is equivalent to a pre-RAKE and we denote this scheme as TH-bit-preRAKE. Note that the transmit power does not vary if the individual spreading waveforms are normalised: Assuming an i.i.d. data sequence  $b_k[i]$ ,  $k = 1, \dots, K$  from a  $M$ -QAM constellation, each output of the modulo operator  $\tilde{x}_k[i]$  can be well approximated by an i.i.d. uniform distribution within the shaping region  $[M, M)$  [21, 22]. Then, the transmit signal can be written as  $\mathbf{p}[i] = \hat{\mathbf{S}}\mathbf{F}\tilde{\mathbf{x}}[i]$ , with power  $E\{\|\hat{\mathbf{S}}\mathbf{F}\tilde{\mathbf{x}}[i]\|^2\} = M/(M-1)/\text{tr}\{\hat{\mathbf{S}}\hat{\mathbf{S}}^H\}$ , where we have used that  $\mathbf{F}$  is unitary.

### 3.3 Chip-wise multiuser THP (TH-chip-original)

We next propose a chip-wise TH precoding scheme that effectively combines precoding and spreading and at the receiver it only implements a matched filter to the original spreading waveform  $s_k$  (i.e. CSI at the mobile units is not required). We denote this scheme as TH-chip-original. The diagram for this scheme is shown in Fig. 6. It is seen that the precoder takes as input the  $K \times 1$  symbol vector  $\mathbf{b}[i]$  and produces as output the  $N \times 1$  chip vector  $\mathbf{p}[i]$  that is transmitted through the channel. At the  $k$ th user's receiver, the  $N \times 1$  received signal vector corresponding to  $\mathbf{p}[i]$  is given by

$$\mathbf{r}_k[i] = \mathbf{D}_k \mathbf{p}[i] + \bar{\mathbf{D}}_k \mathbf{p}[i-1] + \mathbf{n}_k[i] \quad (30)$$

At each receiver  $k$ , the matched-filter  $s_k$  is applied to  $\mathbf{r}_k[i]$ . By stacking the outputs of all  $K$  matched-filters we obtain

$$\begin{bmatrix} s_1^H \mathbf{r}_1[i] \\ s_2^H \mathbf{r}_2[i] \\ \vdots \\ s_K^H \mathbf{r}_K[i] \end{bmatrix} = \underbrace{\begin{bmatrix} s_1^H \mathbf{D}_1 \\ s_2^H \mathbf{D}_2 \\ \vdots \\ s_K^H \mathbf{D}_K \end{bmatrix}}_{\mathbf{H}} \mathbf{p}[i] + \underbrace{\begin{bmatrix} s_1^H \bar{\mathbf{D}}_1 \\ s_2^H \bar{\mathbf{D}}_2 \\ \vdots \\ s_K^H \bar{\mathbf{D}}_K \end{bmatrix}}_{\bar{\mathbf{H}}} \mathbf{p}[i-1] + \underbrace{\begin{bmatrix} s_1^H \mathbf{n}_1[i] \\ s_2^H \mathbf{n}_2[i] \\ \vdots \\ s_K^H \mathbf{n}_K[i] \end{bmatrix}}_{\mathbf{v}[i]} \quad (31)$$

Note that different from the previous section, here  $\mathbf{H}$  is not a square matrix but has dimension  $K \times N$  with  $N \geq K$ . Similarly as before, to apply TH-precoding, we perform the LQ decomposition on  $\mathbf{H} = \mathbf{W}\mathbf{F}^H = \mathbf{G}^{-1}\mathbf{C}\mathbf{F}^H$ . The decomposition is easily obtained applying the Gram-Schmidt

orthogonalisation procedure on the rows of  $\mathbf{H}$ , where the resulting orthonormal vectors form the columns of  $\mathbf{F}$  of dimension  $N \times K$  with  $\mathbf{F}^H \mathbf{F} = \mathbf{I}_K$ . The Gram-Schmidt coefficients define the  $K \times K$  lower triangular matrix  $\mathbf{W}$ . The diagonal matrix  $\mathbf{G} = \text{diag}(w_{1,1}^{-1}, \dots, w_{K,K}^{-1})$  converts  $\mathbf{W}$  into the monic lower triangular matrix  $\mathbf{C}$ . In this way,  $\mathbf{F}$  and  $\mathbf{C} - \mathbf{I}$  are the FF and FB filter matrices, respectively, and the FB matrix  $\mathbf{A} = \mathbf{G}\bar{\mathbf{H}}\mathbf{F}^H$  cancels the inter-symbol interference, as shown in Fig. 6. The  $k$ th diagonal element in  $\mathbf{G}$  corresponds to the scalar gain applied at the  $k$ th user's receiver.

### 3.4 Power loading and ordering

**3.4.1 Power loading:** It is seen from (29) that the noise at each user's receiver is amplified by the corresponding diagonal element of  $\mathbf{G} = \text{diag}(w_{1,1}^{-1}, \dots, w_{K,K}^{-1})$  resulting in different SNR (hence BER) performance among users. Power loading can be employed to enforce the same performance across users (or the required SNR per user). That is, the symbol vector  $\mathbf{b}[i]$  is first multiplied by a diagonal matrix  $\mathbf{\Gamma} = \text{diag}(\gamma_1, \dots, \gamma_K)$  with  $\gamma_k^2$  denoting the power assigned to user  $k$ . The modulo operation for each user then needs to take the loading value into account since the distance between the constellation points is scaled by it. Given the total transmit power  $P_T$ , we then need to solve for  $\gamma_1, \dots, \gamma_K$  such that  $\sum_{k=1}^K \gamma_k^2 = P_T$ , and  $\gamma_k^2 w_{k,k}^2 = \eta$ ,  $\forall k$ . The solution is

$$\begin{aligned} \gamma_k^2 &= \frac{w_{k,k}^{-2}}{\sum_{k=1}^K w_{k,k}^{-2}} P_T, \quad k = 1, \dots, K, \quad \text{and} \\ \eta &= \frac{P_T}{\sum_{k=1}^K w_{k,k}^{-2}} \end{aligned} \quad (32)$$

The base station can broadcast the common constant value  $\eta$  to all mobile receivers and then the receivers can adjust their respective  $w_{k,k}$  to obtain the required  $\gamma_k$  value in the modulo operator. Therefore the loading operation only requires the transmission of a constant value  $\eta$  common to all mobile users.

Next we show the effect of loading and spreading on the average transmit power. Note that as in traditional TH-precoding with  $M$ -QAM constellations, the modulo operation of TH-precoding in the presence of MUI and ISI generates almost i.i.d. outputs and enhances the transmit power by a factor of  $\beta = M/(M-1)$  [21, 22]. When the power loading solution  $\mathbf{\Gamma}$  is applied to any of the previous bit-wise solutions, the average transmit power becomes  $E\{\|\hat{\mathbf{S}}\mathbf{F}\mathbf{\Gamma}\tilde{\mathbf{x}}\|^2\} = M/(M-1)\text{tr}(\hat{\mathbf{S}}\mathbf{F}\mathbf{\Gamma}\mathbf{\Gamma}^H\mathbf{F}^H\hat{\mathbf{S}}^H) = M/(M-1)\text{tr}(\hat{\mathbf{S}}\mathbf{\Gamma}\mathbf{\Gamma}^H\hat{\mathbf{S}}^H)$ . Therefore after applying loading with  $\mathbf{\Gamma}$  a diagonal matrix, and if the individual spreading waveforms are normalised, then the average transmit power becomes  $M/(M-1)/\text{tr}(\hat{\mathbf{S}}\mathbf{\Gamma}\mathbf{\Gamma}^H\hat{\mathbf{S}}^H) = M/(M-1)\sum_{k=1}^K \gamma_k^2 = M/(M-1)$

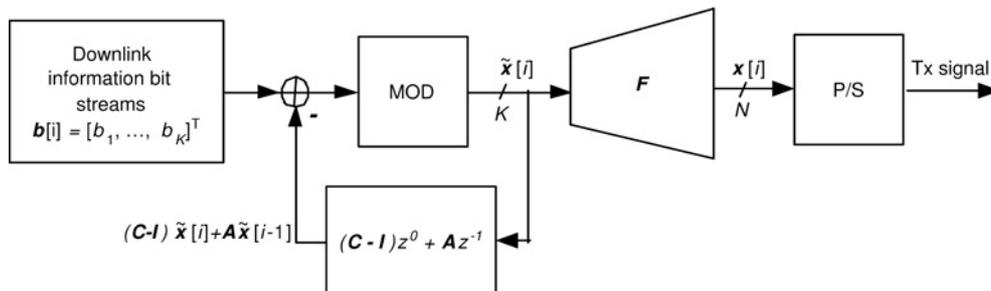


Fig. 6 Chip-wise TH-precoded downlink CDMA system over multipath channels

$P_T$ . On the other hand, assuming that  $E\{|b_k [i]|^2\} = 1$ , then the bit error probability of each user can be well approximated by  $P_e = \alpha Q(\sqrt{(\eta/\sigma_n^2)})$ , where  $\alpha$  accounts for the increase in the number of nearest neighbours due to the modulo operation (e.g. in QPSK  $\alpha = 2$ ) [10, 11].

We next show that when orthogonal spreading sequences are employed, that is when  $\mathbf{S}^T \mathbf{S} = \mathbf{I}_K$ , then we have  $w_{k,k}^{(b)} \leq w_{k,k}^{(c)}$  for  $k = 1, \dots, K$ , and therefore  $\eta^{(b)} \leq \eta^{(c)}$ , where the superscripts  $b$  and  $c$  denote bit-wise and chip-wise precoders, respectively. First, comparing (25) and (31), we have  $\mathbf{H}^{(b)} = \mathbf{H}^{(c)} \mathbf{S}$ . Let  $\mathbf{u}_{K+1}, \dots, \mathbf{u}_N$  be  $(N-K)$  orthonormal vectors in  $\mathcal{V}^\perp \triangleq \mathbb{R}^N \setminus \text{span}(\mathbf{S})$ . Define the unitary matrix  $\mathbf{S}' = [\mathbf{s}_1, \dots, \mathbf{s}_K, \mathbf{u}_{K+1}, \dots, \mathbf{u}_N] = [\mathbf{S}, \mathbf{U}]$  and let

$$\mathbf{X} = [\mathbf{H}^{(b)}, \mathbf{H}^{(c)} \mathbf{U}] = \mathbf{H}^{(c)} \mathbf{S}' \quad (33)$$

Since  $\mathbf{S}'$  is a unitary transformation, the rows in  $\mathbf{X}$  and  $\mathbf{H}^{(c)}$  maintain the norm and the angles. Therefore if the  $K \times (N-K)$  block matrix  $\mathbf{H}_c \mathbf{U}$  has any non-zero row (i.e. the projection of the rows of  $\mathbf{H}_c$  onto  $\text{span}(\mathbf{U})$  is non-zero), the norm of the corresponding row in  $\mathbf{H}^{(b)}$  will be smaller than in  $\mathbf{H}^{(c)}$ . Now consider the LQ factorisation  $\mathbf{H}^{(c)} = \mathbf{W}^{(c)} \mathbf{F}^{(c)H}$ , obtained using Gram-Schmidt on the rows of  $\mathbf{H}^{(c)}$ , that is  $\{\mathbf{h}_k^{(c)T}\}_{k=1}^K$ . Each value  $w_{k,k}^{(c)}$  can be obtained as follows. Assume that at the  $k$ th step of the Gram-Schmidt algorithm the orthonormal vectors  $\mathbf{f}_1^{(c)}, \dots, \mathbf{f}_{k-1}^{(c)}$  (i.e. first  $k$  columns in  $\mathbf{F}^{(c)}$ ) have been obtained from  $\mathbf{h}_1^{(c)}, \dots, \mathbf{h}_{k-1}^{(c)}$ , and denote  $\mathcal{U}_{k-1} = \text{span}\{\mathbf{f}_1^{(c)}, \dots, \mathbf{f}_{k-1}^{(c)}\}$ . Then, by simple inspection from the structure of the LQ factorisation  $w_{k,k}^{(c)}$  is the norm of  $\tilde{\mathbf{f}}_k^{(c)} = \text{proj}_{\mathcal{U}_{k-1}^\perp} \{\mathbf{h}_k^{(c)}\}$ , where  $\mathcal{U}_{k-1}^\perp = \mathbb{R}^N \setminus \mathcal{U}_{k-1}$  and  $\tilde{\mathbf{f}}_k^{(c)} = \mathbf{f}_k^{(c)} / w_{k,k}^{(c)}$ . That is

$$w_{k,k}^{(c)} = \left\| \mathbf{h}_k^{(c)} - \text{proj}_{\mathcal{U}_{k-1}} \{\mathbf{h}_k^{(c)}\} \right\| = \left\| \tilde{\mathbf{f}}_k^{(c)} \right\| \quad (34)$$

On the other hand, the diagonal elements of  $\mathbf{W}^{(b)}$  are similarly obtained from  $[\mathbf{H}^{(b)}, \mathbf{0}_{K, N-K}]$ . Then, using (33) and (34) we obtain

$$w_{k,k}^{(b)} = w_{k,k}^{(c)} - \left\| \text{proj}_{\mathcal{V}^\perp} \{\tilde{\mathbf{f}}_k^{(c)}\} \right\| \quad (35)$$

and hence  $w_{k,k}^{(b)} \leq w_{k,k}^{(c)}$ . Note that when  $N = K$  and orthogonal spreading sequences are employed,  $\mathbf{S}$  is unitary and  $w_{k,k}^{(b)} = w_{k,k}^{(c)}$  for all  $k$ , and hence  $\eta^{(b)} = \eta^{(c)}$  [cf. Fig. 9].

On the other hand, when the spreading sequences  $\mathbf{S}$  are non-orthogonal, it is not true that  $w_{k,k}^{(b)} \leq w_{k,k}^{(c)}$ . However, we conjecture that  $\eta^{(b)} \leq \eta^{(c)}$  still holds.

**3.4.2 User ordering:** We can optimise the system BER performance by optimising the diagonal elements of the matrix  $\mathbf{W}$  such that the common SNR of all users  $\eta$  is maximised. Notice that  $\mathbf{W}$  is obtained from the LQ decomposition of  $\mathbf{H}$ . The LQ decomposition is essentially the Gram-Schmidt orthogonalisation of the rows of  $\mathbf{H}$ . The  $k$ th diagonal element of  $\mathbf{W}$  is the length of the projection of the  $k$ th row vector of  $\mathbf{H}$  onto the orthogonal complement of the space spanned by the first  $(k-1)$  row vectors already orthogonalised. Different ordering in the orthogonalisation process results in different diagonal values of  $\mathbf{W}$ , and hence different values of  $\eta$ . Let  $\mathcal{P}$  be the set of the  $K!$  possible  $K \times K$  row permutation matrices. Then for any  $\mathbf{P} \in \mathcal{P}$ ,  $\mathbf{P}\mathbf{H}$  is a row-permuted version of  $\mathbf{H}$ , which corresponds to a particular ordering of the  $K$  users in TH-precoding. Denote  $w_{k,k}(\mathbf{P})$  as the  $k$ th diagonal element of  $\mathbf{W}$  resulting from the LQ decomposition of  $\mathbf{P}\mathbf{H}$ . Then the optimal row

permutation matrix is given by

$$\mathbf{P}_{\text{opt}} = \arg \max_{\mathbf{P} \in \mathcal{P}} \frac{P_T}{\sum_{k=1}^K w_{k,k}^{-2}(\mathbf{P})} = \arg \min_{\mathbf{P} \in \mathcal{P}} \sum_{k=1}^K w_{k,k}^{-2}(\mathbf{P}) \quad (36)$$

With the optimal permutation  $\mathbf{P}_{\text{opt}}$ , the following modifications are needed at the transmitter and receiver: (1) Perform the LQ decomposition as  $\mathbf{P}\mathbf{H} = \mathbf{W}\mathbf{F}^H$ , or  $\mathbf{H} = \mathbf{P}^T \mathbf{G}^{-1} \mathbf{C}\mathbf{F}^H$ ; (2) Apply  $\mathbf{G}\mathbf{P}$  at the receiver (i.e. apply the scalar gains according to the optimal order); (3) The FB matrix for removing the ISI becomes  $\mathbf{A} = \mathbf{G}\mathbf{P}\mathbf{H}$ . With these modifications, the cascade of operations becomes

$$\begin{aligned} z[i] &= \underbrace{\mathbf{G}\mathbf{P}}_{\text{rx}} \left( \underbrace{\mathbf{P}^T \mathbf{G}^{-1} \mathbf{C}\mathbf{F}^H}_{\text{channel}} \left( \underbrace{\mathbf{F}}_{\text{FF}} \underbrace{\mathbf{C}^{-1}}_{\text{cancel, FB}} \right) \right) \\ &\quad \times \left( \underbrace{\mathbf{I}\mathbf{b}[i] - \mathbf{G}\mathbf{P}\bar{\mathbf{H}}\mathbf{x}[i-1]}_{\text{cancel ISI}} \right) + \underbrace{\bar{\mathbf{H}}\mathbf{x}[i-1]}_{\text{ISI channel}} + v[i] \\ &= \mathbf{I}\mathbf{b}[i] + \mathbf{G}\mathbf{P}v[i] \end{aligned} \quad (37)$$

Note that the matrices  $\mathbf{G}$ ,  $\mathbf{F}$  and  $\mathbf{C}$  are obtained from  $\mathbf{P}\mathbf{H}$ .

Clearly an exhaustive search solution to (36) is computationally prohibitive. Note that a different ordering in the orthogonalisation results in a different order in the cancellation of the interference. Other systems employing non-linear interference cancellation with ordering (for which low complexity suboptimal ordering solutions have been proposed) include: (a) successive multiuser interference cancellation receivers [23]; or (b) multiple antenna BLAST-like receivers [13, 24, 25]. We next propose a suboptimal algorithm for an approximate solution to (36) that performs especially well in the chip-wise precoder when  $N > K$ . The proposed algorithm resembles the BLAST-like ordering algorithm in [25]. First note that  $\prod_{k=1}^K w_{k,k}^2$  is invariant to the permutation matrix  $\mathbf{P}$ . This result is easily proved recalling that  $\mathbf{P}\mathbf{H} = \mathbf{W}\mathbf{F}^H$ , with orthonormal columns in  $\mathbf{F}$ , then

$$\begin{aligned} \det(\mathbf{H}\mathbf{H}^H) &= \det(\mathbf{P}^T) \det(\mathbf{W}) \det(\mathbf{W}^H) \det(\mathbf{P}) \\ &= \prod_{k=1}^K w_{k,k}^2 \end{aligned} \quad (38)$$

We first consider the simplest case with  $K = 2$  users, then  $\mathbf{H}$  contains two rows denoted by  $\mathbf{h}_1^T$  and  $\mathbf{h}_2^T$ . Without loss of generality, assume that  $\|\mathbf{h}_2\| < \|\mathbf{h}_1\|$ . Next we show that to maximise the objective function in (36), we should start with  $\mathbf{h}_2^T$ , that is start by orthogonalising the row with minimum  $w_{k,k}$ . Recall that  $w_{k,k}$  is the length of the projection of the  $k$ th row of  $\mathbf{H}$  onto the orthogonal complement of the subspace spanned by the previous  $(k-1)$  rows already orthogonalised. Then we need to show that

$$\begin{aligned} &\frac{1}{\|\mathbf{h}_2\|^2} + \frac{1}{\|\mathbf{h}_1 - (\mathbf{h}_2^H \mathbf{h}_1 / \|\mathbf{h}_2\|^2) \mathbf{h}_2\|^2} \\ &< \frac{1}{\|\mathbf{h}_1\|^2} + \frac{1}{\|\mathbf{h}_2 - (\mathbf{h}_1^H \mathbf{h}_2 / \|\mathbf{h}_1\|^2) \mathbf{h}_1\|^2} \end{aligned} \quad (39)$$

From (38), the products of the denominators on both sides

in (39) are equal. Therefore (39) is equivalent to

$$\left\| \mathbf{h}_1 - \frac{\mathbf{h}_2^H \mathbf{h}_1}{\|\mathbf{h}_2\|^2} \mathbf{h}_2 \right\|^2 + \|\mathbf{h}_2\|^2 < \left\| \mathbf{h}_2 - \frac{\mathbf{h}_1^H \mathbf{h}_2}{\|\mathbf{h}_1\|^2} \mathbf{h}_1 \right\|^2 + \|\mathbf{h}_1\|^2$$

$$\Leftrightarrow \frac{|\mathbf{h}_2^H \mathbf{h}_1|^2}{\|\mathbf{h}_2\|^2} > \frac{|\mathbf{h}_1^H \mathbf{h}_2|^2}{\|\mathbf{h}_1\|^2} \quad (40)$$

which is true by the assumption that  $\|\mathbf{h}_2\| < \|\mathbf{h}_1\|$ .

When  $K < 2$ , we adopt the greedy solution given in Algorithm 1 that at the  $k$ th iteration, orthogonalises the row with minimum  $w_{k,k}$ . In other words, the algorithm selects the row that is the closest to the subspace spanned by the rows already chosen. In the algorithm,  $\mu_{p,j} = \hat{\mathbf{h}}_p^H \mathbf{h}_j$  and  $\Theta_i$  represents the subset of rows already orthogonalised up to the  $i$ th step. Note that besides finding the ordering  $\mathbf{P}$ , the algorithm also provides the LQ decomposition  $\mathbf{PH} = \mathbf{WF}^H$ , since  $\mathbf{W}$  is given by the GS coefficients  $\mu_{ij}$  and the  $i$ th row of  $\mathbf{F}$  is given by  $\hat{\mathbf{h}}_i$ . Clearly the complexity of the above search algorithm is  $\mathcal{O}(K^2)$ , which is significantly lower than the  $\mathcal{O}(K!)$  complexity of the exhaustive search method.

*Algorithm 1:* Greedy ordering and LQ decomposition

```

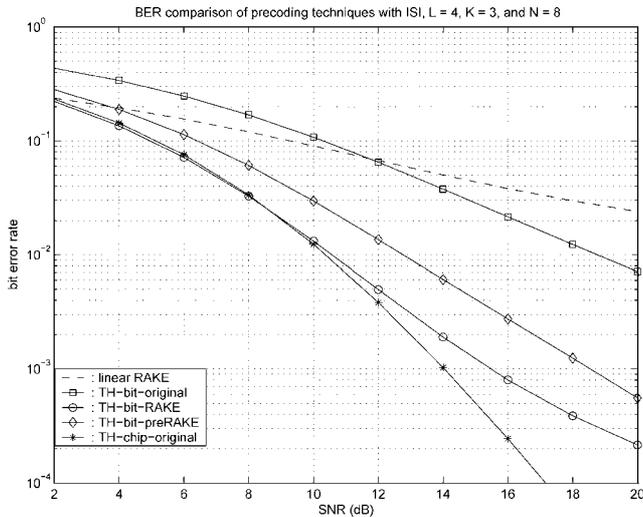
INPUT: row vectors  $\mathbf{h}_1^T, \dots, \mathbf{h}_K^T$  in  $\mathbf{H}$ 
 $\mathbf{P} = \mathbf{0}_{K \times K}$ 
 $k_1 = \arg \min_{i \leq K} \{\|\mathbf{h}_i\|\}$ ;
 $\hat{\mathbf{h}}_1 = \mathbf{h}_{k_1} / \|\mathbf{h}_{k_1}\|$ ;  $\mathbf{P}(1, k_1) = 1$ ;  $\Theta_1 = \{k_1\}$ ;
FOR  $i = 2: K$ 
  FOR EVERY  $j \in \{\{1, \dots, K\} \setminus \Theta_{i-1}\}$ 
     $\mathbf{u}_j = \mathbf{h}_j - \sum_{p=1}^{i-1} \mu_{p,j} \hat{\mathbf{h}}_p$ ;
  END FOR
   $k_i = \arg \min_j \{\|\mathbf{u}_j\|\}$ ;
   $\hat{\mathbf{h}}_i = \mathbf{u}_{k_i} / \|\mathbf{u}_{k_i}\|$ ;  $\Theta_i = \{\Theta_{i-1}\} \cup \{k_i\}$ ;
   $\mathbf{P}(i, k_i) = 1$ ;
END FOR
OUTPUT: matrix  $\mathbf{P}$  and LQ decomposition of  $\mathbf{PH}$ .

```

## 4 Simulation results

### 4.1 TH-precoding with perfect channel knowledge

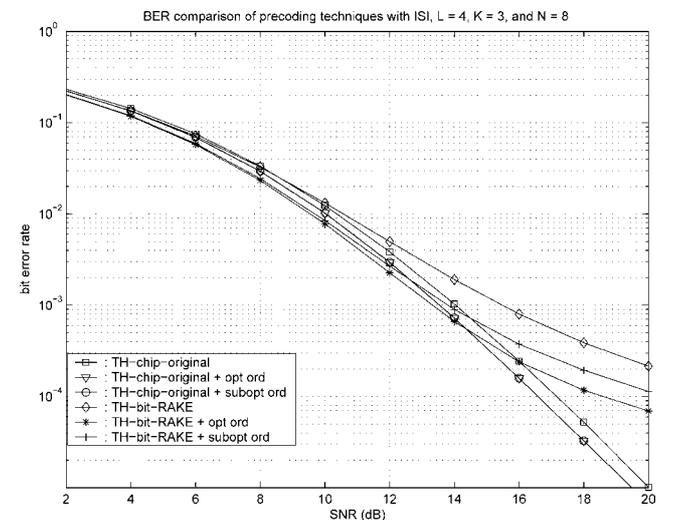
We first provide simulation results to compare the BER performance of the different precoding schemes:



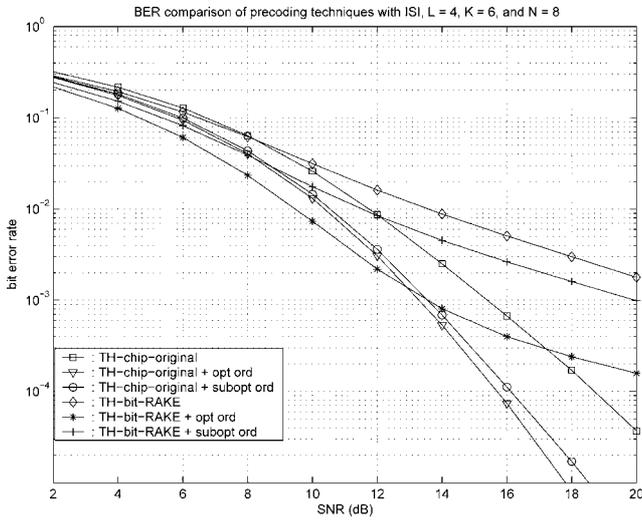
**Fig. 7** BER performance of different precoding schemes without ordering, with  $K = 3$  users, spreading gain  $N = 8$ , number of paths  $L = 4$

TH-bit-original, TH-bit-RAKE, TH-bit-preRAKE and TH-chip-original. Each user employs a normalised Hadamard sequences of length  $N = 8$  as its original spreading signature. The number of users is  $K = 3$ . All users employ QPSK modulation. We assume that each mobile user experiences an independent multipath channel  $\mathbf{f}_k = [f_{k,1}, \dots, f_{k,L}]^T$  with  $L = 4$  resolvable paths and the transmitter has perfect CSI of all users. The path gains are generated according to  $f_{k,i} \sim \mathcal{N}_c(0, 1/L)$ . For each data block, independent channel realisations are simulated for each user and the results are averaged over 5000 blocks. Fig. 7 shows the BER performance of the precoding schemes without ordering. Loading is employed in all the schemes. For comparison purposes, we also show the performance of the linear block-wise precoding method given in [4]. In the figure, the approximate BER formula  $P_e = \alpha Q(\sqrt{(\eta/\sigma_n^2)})$  matches very well with the simulated results. Nonlinear TH-precoders significantly outperform the linear precoder. Moreover, the complexity of the TH-precoders is substantially lower than that of the linear precoders since in multipath channels, linear precoders [4] operate blockwise with matrices of size  $KT \times KT$  (where  $T$  is the frame length) with  $\mathcal{O}((KT)^3)$  complexity of the inverse computation. In our proposed non-linear solution, the matrices have dimensions  $K \times K$ . Comparing bit-wise and chip-wise precoding schemes, the chip-wise precoder offers the best performance. Therefore it not only reduces the mobile unit complexity of the TH-bit-RAKE precoder without increasing the base station complexity, but also shows superior performance. Similar performance was obtained for different number of users and paths (the results are not provided here for brevity).

Fig. 8 compares the BER performance of the TH-chip-original and TH-bit-RAKE precoders with  $K = 3$  users. For both methods, we consider the cases of no-ordering, optimal ordering (i.e. exhaustive search) and the suboptimal ordering method given in Algorithm 1. Ordering has a significant effect on the TH-bit-RAKE precoder, whereas it makes a smaller difference to the chip-wise TH-precoder. Furthermore, in the chip-wise precoder there is no observable difference between the suboptimal greedy ordering algorithm and the optimal exhaustive search method. Therefore the chip-wise TH-precoder not only offers superior performance and



**Fig. 8** BER performance of different precoding schemes with  $K = 3$  users, spreading gain  $N = 8$ , number of paths  $L = 4$

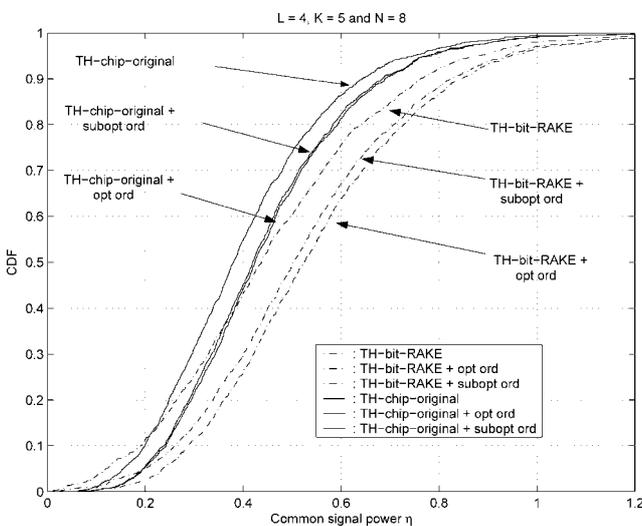


**Fig. 9** BER performance of different precoding schemes with  $K = 6$  users, spreading gain  $N = 8$ , number of paths  $L = 4$

lower mobile unit complexity, but also computationally less complex since suboptimal ordering is sufficient.

We repeat the simulations with the same parameters except that the number of users is increased to  $K = 6$ . Fig. 9 shows that both TH-precoding schemes perform very well even in highly loaded systems. When the number of users is high, ordering brings a significant improvement, specially to the TH-bit-RAKE precoder, although the complexity of the exhaustive search method becomes prohibitive (i.e. it involves computing  $K! = 720$  LQ decompositions of  $6 \times 6$  matrices). The greedy ordering algorithm performs well in the chip-wise precoder and it requires less than 6 LQ decompositions.

In Fig. 10, we illustrate the CDF of the common SNR  $\eta$  in (32) obtained with TH-bit-RAKE and TH-chip-original when  $L = 4$ ,  $N = 8$  and  $P_T / K = 1$ . For both methods, we show the  $\eta$  value for cases of no-ordering, optimal ordering and suboptimal ordering. Note that the performance of communication systems is usually dominated by outage events. Given an outage probability  $p_{\text{out}}$ , let us define the corresponding outage  $\eta_{\text{out}}$  as  $p_{\text{out}} = \Pr(\eta \leq \eta_{\text{out}})$ . In Fig. 10, it is seen that although on average, the TH-bit-RAKE precoder obtains better performance than the TH-chip-original precoder, it is



**Fig. 10** CDF of the common SNR value  $\eta$  Spreading gain  $N = 8$ ,  $K = 5$

more prone to outage (e.g. consider in the plot  $p_{\text{out}} = 3\%$ ), which explains the error rate behaviour in Figs. 7–9.

#### 4.2 TH-precoding with channel prediction

A crucial assumption in the development of the precoding techniques in the previous section is that the transmitter has perfect knowledge about the multipath channel states of all mobile users. In TDD wireless systems, the downlink CSI is available at the transmitter (which is estimated from the uplink transmission) as long as the coherence time of the channel is larger than the time difference between the uplink and downlink slots. On the other hand, in fast fading channels, the channel state that has been estimated during an uplink slot may have changed and the estimate may no longer be accurate for precoding in the next downlink slot. In this case, channel prediction techniques can be used to estimate the future downlink channel state from the current and previous uplink channel estimates, by exploiting the second-order statistics of the fading channel [26] (refer to [27] for a comparison of different channel prediction techniques). Assume that the complex Gaussian fading process of each channel path  $f_{k,i}(t)$  follows the Jakes' model [28] with the maximum Doppler spread  $f_d$ , that is, we have  $E\{f_{k,i}(t_1)f_{k,i}^*(t_2)\} = v_{k,i}^2 J_0(2\pi f_d |t_1 - t_2|)$ ,  $k = 1, \dots, K$ ;  $i = 1, \dots, L$ , where  $J_0(\cdot)$  is the zeroth-order Bessel function of the first kind.

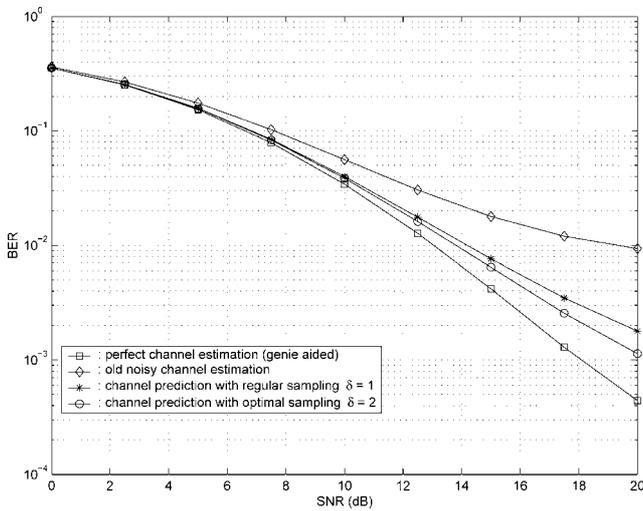
Assume that in the TDD system, the uplink and downlink slots are separated by  $T$  seconds; and the base station estimates the multipath channel of each user every uplink slot. We set the time of the latest channel estimation as the reference  $t = 0$ . Then the base station will estimate the channel state at times  $t \in \{0, -2T, -4T, \dots\}$ . We consider channel estimation based on pilot symbols and the channel estimate has the form  $\hat{f}_{k,i}(t) = f_{k,i}(t) + \xi_{k,i}(t)$ , where  $\xi_{k,i}(t) \sim \mathcal{N}_c(0, \gamma_{k,i}^2)$ . We assume that the base station estimates the channel once per slot and these estimates will be used to predict the channel for data precoding in the next downlink slot.

Assume that after the current channel estimate at time  $t = 0$ , the base station predicts each channel path at time  $\tau$  which is called the prediction depth (e.g.  $\tau = T$  where  $T$  is the slot duration). The prediction is implemented using a  $P$ th order finite impulse response filter

$$\tilde{f}_{k,i}(\tau) = \sum_{p=0}^P \varphi_{k,i}(p) \hat{f}_{k,i}(-p2T) = \boldsymbol{\varphi}_{k,i}^H \hat{\mathbf{f}}_{k,i} \quad (41)$$

where  $\boldsymbol{\varphi}_{k,i} \triangleq [\varphi_{k,i}(0), \varphi_{k,i}(1), \dots, \varphi_{k,i}(P)]^T$ ,  $\hat{\mathbf{f}}_{k,i} \triangleq [\hat{f}_{k,i}(0), \hat{f}_{k,i}(-2T), \dots, \hat{f}_{k,i}(-P2T)]^T$ . The optimal filter that minimises the MSE  $\zeta_{\text{pred}} \triangleq E\{|f_{k,i}(\tau) - \tilde{f}_{k,i}(\tau)|^2\}$  is given by  $\boldsymbol{\varphi}_{k,i} = \mathbf{R}_{k,i}^{-1} \mathbf{r}_{k,i}$ , where the entries of  $\mathbf{R}_{k,i}$  and  $\mathbf{r}_{k,i}$  are given, respectively, by  $[\mathbf{R}_{k,i}]_{p,q} = v_{k,i}^2 J_0(2\pi f_d |p - q|2T) + \gamma_{k,i}^2 \delta_{p,q}$ , and  $[\mathbf{r}_{k,i}]_p = v_{k,i}^2 J_0(2\pi f_d (\tau + p2T))$ ,  $p, q = 0, 1, \dots, P$ .

In the prediction filter described earlier, we use estimates of the channel that have been sampled every  $2T$  seconds. This sampling rate is in general much higher than the required minimum Nyquist sampling rate equal to twice the Doppler frequency  $2f_d$ . It has been shown in [26] that such oversampling could be unfavourable when the order of the filter  $P$  is fixed. Assume that the base station is able to estimate the channel every  $2T$  seconds. Define the optimal sampling period as  $\delta 2T$ , where  $\delta$  is a positive integer. Then for fixed values of the prediction depth, noise variance, Doppler frequency and filter order, we can compute the MSE of the prediction filter  $\zeta_{\text{pred}}$  for different



**Fig. 11** BER performance of chip-wise TH-precoding in time-varying channels

integer values of  $\delta$  and select the one that minimises  $\zeta_{\text{pred}}$ . On the other hand, it has been observed that when the system parameters are fixed,  $\zeta_{\text{pred}}$  decreases with the order of the prediction filter  $P$ . However, after a certain filter order,  $\zeta_{\text{pred}}$  saturates since the noise in the previous channel estimates dominates in the MSE of the prediction error. Therefore it is convenient to evaluate the MSE expression for different values of  $P$  and choose the shortest one that brings  $\zeta_{\text{pred}}$  close to the saturation level.

As in the WCDMA TDD mode, we assume that the uplink and downlink are time multiplexed into a carrier centred at  $f_c = 2$  GHz. The frame length is 10 ms, which is subdivided into 15 slots that can be allocated for either uplink or downlink. Therefore the uplink and downlink transmission can be interleaved in bursts of  $T = 666.7 \mu\text{s}$ . As in Section 4.1, we consider  $N = 8$ ,  $L = 5$  and  $v_{k,i}^2 = 1/L$ . The fading process of each channel path is formed by samples of a stationary zero-mean complex Gaussian process with autocorrelation function  $J_0(2\pi f_d t)$  [28] and is generated according to the method described in [29]. We consider the performance of the TH-chip-original precoding technique with loading and ordering. We assume that all the mobile users are moving at  $v = 36$  km/h. The previous channel estimates  $\{\hat{f}_{k,i}(t), t = 0, -2T, \dots\}$  are given by the true channel values corrupted by complex Gaussian noise with variance  $\gamma_{k,i}^2 = 0.001$ . Evaluating the MSE expression  $\zeta_{\text{pred}}$  for different orders of the prediction filter, we find that a very short prediction filter with  $P = 2$  obtains good results. Evaluating  $\zeta_{\text{pred},\delta}$ , we find that slightly better results can be obtained if the channel is sampled with  $\delta = 2$ . We evaluate the results over ten different initial channel realisations. For each channel realisation, we consider 200 slots of length  $T = 666 \mu\text{s}$  (i.e. 200 channel variations) and in each slot, we send 1000 QPSK symbols per mobile user. In the results, we consider perfect channel estimation (genie aided), old channel estimation  $\tau$  seconds before and channel prediction with the optimal sampling ( $\delta = 2$ ) and with the regular sampling ( $\delta = 1$ ). Fig. 11 shows that the prediction algorithm gives very good results even considering that all the users are moving at  $v = 36$  km/h and the prediction is based on noisy channel estimates. Notice that without channel prediction and only using old channel estimates, the performance would decrease considerably in these scenarios representing very high mobility.

## 5 Conclusions

In this paper, we have first obtained the capacity results for downlink CDMA systems employing either MUD or transmitter precoding. It is seen from numerical examples that these two techniques offer comparable capacity regions. However, MUD at the downlink mobile receiver may not be practical due to the requirement that each mobile receiver should have the knowledge of all users' spreading sequences and channel states, as well as the limited signal processing capability of the mobile receiver. On the other hand, transmitter precoding is an attractive solution for systems employing time-division multiplexing, where uplink and downlink channels are reciprocal. We have developed nonlinear multiuser precoding algorithms based on the TH precoding technique. Our precoding algorithms effectively remove multiuser interference, inter-chip interference and inter-symbol interference in the downlink of CDMA systems. The main property of the proposed algorithms is that they can be implemented at either bit level or chip level, and they are considerably less complex compared with the linear block-wise linear precoders in the literature. Moreover, the proposed chip-wise precoder combines the precoding operation and the spreading operation, and offers performance improvement compared with the previously proposed bit-wise precoders. We have also proposed a suboptimal user ordering algorithm for power loading which further optimises the system performance. Channel prediction for precoding is also discussed. Simulation results have shown that the proposed precoding techniques offer excellent performance even in heavily loaded systems or time-varying scenarios.

## 6 References

- Verdu, S.: 'Multiuser detection' (Cambridge University Press, 1998)
- Esmailzadeh, R., Sourour, E., and Nakagawa, M.: 'Pre-rake diversity combining in time-division duplex CDMA mobile communications', *IEEE Trans. Veh. Tech.*, 1999, **48**, pp. 795–801
- Meurer, M., Baier, P.W., Weber, T., Lu, Y., and Papathanasiou, A.: 'Joint transmission: advantageous downlink concept for CDMA mobile radio systems using time division duplexing', *Electron. Lett.*, 2000, **36**, pp. 900–901
- Vojcic, B.R., and Jang, W.M.: 'Transmitter precoding in synchronous multiuser communications', *IEEE Trans. Commun.*, 1998, **46**, (10), pp. 1346–1355
- Georgoulis, S.L., and Cruickshank, D.G.M.: 'Transmitter-based inverse filters for reducing MAI and ISI mitigation in a CDMA-TDD downlink', *IEEE Trans. Wirel. Commun.*, 2004, **3**, (2), pp. 353–358
- Costa, M.: 'Writing on dirty paper', *IEEE Trans. Inform. Theory*, 1983, **29**, pp. 439–441
- Caire, G., and Shamai, S.: 'On the achievable throughput of a multi-antenna Gaussian broadcast channel', *IEEE Trans. Inform. Theory*, 2003, **49**, (7), pp. 1691–1706
- Cohen, A., and Lapidot, A.: 'The Gaussian watermarking game', *IEEE Trans. Inform. Theory*, 2002, **48**, pp. 1639–1667
- Erez, U., Shamai, S., and Zamir, R.: 'Capacity and lattice-strategies for cancelling known interference'. Proc. ISITA, Honolulu, HI, November 2000
- Harashima, H., and Miyakawa, H.: 'Matched transmission technique for channels with inter-symbol interference', *IEEE Trans. Commun.*, 1972, **20**, pp. 774–780
- Tomlinson, M.: 'New automatic equaliser employing modulo arithmetic', *Electron. Lett.*, 1971, **7**, pp. 138–139
- Ginis, G., and Cioffi, J.M.: 'Vectored transmission for digital subscriber line systems', *IEEE J. Sel. Areas Commun.*, 2002, **20**, (5), pp. 1085–1104
- Joham, M., Brehmer, J., and Utschick, W.: 'MMSE approaches for multiuser spatio-temporal Tomlinson-Harashima precoding'. Proc. 5th Int. ITG Conf. on Source and Channel Coding (SCC), Erlangen, Germany, January 2004.
- Windpassinger, C., Fischer, R.F.H., Vencel, T., and Huber, J.B.: 'Precoding in multi-antenna and multi-user communications', *IEEE Trans. Wirel. Commun.*, 2004, **3**, (4), pp. 1305–1316

- 15 Fischer, R.F.H., Windpassinger, C., Lampe, A., and Huber, J.B.: 'MIMO precoding for decentralized receivers'. Proc. Int. Symp. Information Theory (ISIT), Lausanne, Switzerland, June 2002
- 16 Boche, H., and Schubert, M.: 'A general duality theory for uplink and downlink beamforming'. Proc. IEEE Vehicular Technology Conf. (VTC), Vancouver, Canada, September 2002
- 17 Vishwanath, S., Jindal, N., and Goldsmith, A.: 'Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels', *IEEE Trans. Inform. Theory*, 2003, **49**, (10), pp. 2658–2668
- 18 Viswanath, P., and Tse, D.: 'Sum capacity of the vector Gaussian channel and uplink–downlink duality', *IEEE Trans. Inform. Theory*, 2003, **49**, (8), pp. 1912–1921
- 19 Yu, W., and Cioffi, J.: 'Sum capacity of the Gaussian vector broadcast channels', *IEEE Trans. Inform. Theory*, 2004, **50**, (9), pp. 1875–1892
- 20 Weingarten, H., Steinberg, Y., and Shamai, S. (Shitz): 'The capacity region of the gaussian multiple-input multiple-output broadcast channel', *IEEE Trans. Inform. Theory*, 2006, **52**, (9), pp. 3936–3964
- 21 Forney, G.D., and Eyuboglu, M.V.: 'Combined equalization and coding using precoding', *IEEE Commun. Mag.*, 1991, **29**, pp. 25–34
- 22 Trump, T., and Forssen, U.: 'On the statistical properties of Tomlinson filters' (Telecommunications Theory, Royal Institute of Technology, Stockholm, Sweden, March 1992)
- 23 Benvenuto, N., Carnevale, G., and Tomasin, S.: 'Energy optimization of CDMA transceivers using successive interference cancellation'. Proc. Globecom Dallas, IX, December 2004
- 24 Windpassinger, C., Vencel, T., and Fischer, R.: 'Precoding and loading for BLAST-like systems'. Proc. Int. Conf. Wireless and Optical Communication (WOC), Banff, Canada, July 2003
- 25 Wuebben, D., Boehnke, R., Rinas, J., Kuehn, V., and Kammeyer, K.D.: 'Efficient algorithm for decoding layered space-time codes', *IEE Electron. Lett.*, 2001, **37**, (22), pp. 1348–1350
- 26 Duel-Hallen, A., Hu, S., and Hallen, H.: 'Long-range prediction of fading signals: enabling adapting transmission for mobile radio channels', *IEEE Signal Proc. Mag.*, 2000, **17**, (3), pp. 62–75
- 27 Semmelrodt, S., and Kattenbach, R.: 'Investigation of different fading forecast schemes for flat fading radio channels'. Proc. Vehic. Technol. Conf. (VTC), Orlando, FL, October 2003
- 28 Jakes, W.C.: 'Microwave mobile communications' (Wiley, 1974)
- 29 Dent, P., Bottomley, G.E., and Croft, T.: 'Jakes fading model revisited', *Electron. Lett.*, 1993, **29**, (13), pp. 1162–1963