Linear Parallel Interference Cancellation in Long-code CDMA Multiuser Detection

Dongning Guo, Lars K. Rasmussen and Teng Joon Lim

Abstract—**Parallel interference cancellation (PIC) is a promising detection technique for code-division multiple access (CDMA) systems. It has previously been shown that the weighted multistage PIC can be seen as an implementation of the steepest descent algorithm used to minimise the mean squared error (MSE). Following this interpretation, a unique set of weights, based on the eigenvalues of the correlation matrix, was found to lead to the minimum achievable MSE for a given number of stages in a short-code system. In this paper, we introduce a method for finding an appropriate set of time-invariant weights for systems using long codes. The weights are dependent on moments of the eigenvalues of the correlation matrix, exact expressions of which can be derived. This set of weights is optimal in the sense that it minimises the ensemble averaged MSE over all code-sets. The loss incurred by averaging rather than using the optimal, time-varying weights is practically negligible, since the eigenvalues of sample correlation matrices are tightly clustered in most cases of interest. The complexity required for computing the weights increases linearly with the number of users but is independent of the processing gain, hence online weight-updating is possible in a dynamic system. Simulation results show that a few stages is usually sufficient for near-MMSE performance.**

I. Introduction

In a code-division multiple access (CDMA) system, all frequency and time resources are allocated to all users simultaneously. To distinguish between users, each user is assigned a user-specific spreading code (signature) sequence for transmission. In short-code CDMA, the period of such a spreading code sequence spans a symbol interval, i.e., each user's spreading code remains the same for all symbols. In long-code CDMA, the spreading code has a period which is many times longer than a symbol interval. Consecutive segments of this long sequence, each spanning exactly one symbol interval, are then used for spreading consecutive symbols. The statistical properties of such segments of the long spreading code resemble those of randomly selected sequences. Long-code CDMA is therefore also referred to as random-code CDMA. The wideband CDMA proposals for third-generation cellular mobile communication, as well as IS-95, are all based on long-code CDMA [1], [2], [3].

D. Guo and T. J. Lim are with the Centre for Wireless Communications, 20 Science Park Road, #02-34/37 Teletech Park, Singapore Science Park II, Singapore 117674 (email: {cwcguodn, cwclimtj}@leonis.nus.edu.sg, FAX: +65 779 5441). L. K. Rasmussen was with the Centre for Wireless Communications, Singapore. He is currently with the Department of Computer Engineering, Chalmers University of Technology, SE-412 96 Göteborg, Sweden (email: larsr@ce.chalmers.se, FAX: +46 31 772 3663). This work was supported in part by Oki Techno Centre (S'pore) Pte Ltd and the National Science and Technology Board (NSTB), Singapore. The Centre for Wireless Communications is a National Research Centre funded by NSTB.

By selecting mutually orthogonal codes for all users, the conventional matched-filter detector achieves single-user performance for each user. It is however not possible to maintain orthogonality in a mobile environment, hence the multiple access interference (MAI) that results may degrade the performance of a CDMA system severely. Moreover, the conventional detector suffers from a near-far problem in which the signal component from a weak user may be buried in the MAI from a strong user [4]. It is known that advanced detection techniques may be used in the uplink receiver in order to alleviate the MAI and thereby increase overall capacity and loosen the requirements for strict and fast power control.

In [5], Verdù developed the optimal (0,1)-constrained maximum-likelihood (ML) detector. This ML problem corresponds to a combinatorial quadratic minimisation which is known to be NP-hard [6]. It can only be solved by an exhaustive search, leading to a detection complexity that grows exponentially with the number of users. To address this complexity problem, a variety of sub-optimal detectors have been proposed [7]. For example, the linear decorrelating detector in [8] applies the inverse of the correlation matrix in order to decouple the data. It is known to be near-far resistant, but also to cause noise enhancement [9]. Another type of linear detector, the MMSE detector minimises the mean squared error (MSE) between detector output and the transmitted symbol [10]. This detector takes the background noise as well as the correlation between users into account and therefore generally performs better than the decorrelator in terms of bit-error-rate (BER) [11]. Both the decorrelating detector and the MMSE detector face the task of matrix inversion which can be prohibitively complex for a large number of users. A number of strategies have been developed for approximating these detectors. Adaptive detectors based on algorithms such as the LMS algorithm [12], the RLS algorithm [13] and Kalman filtering [14] have been suggested while iterative techniques such as the steepest descent and the conjugate gradient iterations have been proposed in [15], [16].

For practical implementation, interference cancellation schemes have been subject to most attention. These techniques rely on simple processing elements constructed around the matched filter concept. Varanasi and Aazhang proposed a multi-stage parallel interference cancellation (PIC) structure in [17]. The linear version of this structure has been shown by Elders-Boll *et al.* to be equivalent to the Jacobi iteration for solving a set of linear equations [16]. A linear PIC therefore represents an efficient way of implementing linear detectors. A significant improvement to PIC was suggested in [18] and [19] independently where a weighted cancellation scheme was proposed for both linear and non-linear PIC. In [20] the MMSE detector is approximated by a polynomial of the correlation matrix. This approach is in fact equivalent to a bit-level implementation of a weighted linear PIC. The linear PIC approach has been further described and analysed in detail in [21] and [22] where it was demonstrated that weighted linear PIC can be seen as a realisation of the steepest descent optimisation method [23] for minimising the MSE. This is in turn also equivalent to the steepest decent iteration for solving a set of linear equations [24]. It was shown in [22] that for a short-code system with a given number of cancellation stages, a unique choice of weights exists which leads to the minimum achievable MSE. These weighting parameters are dependent on the eigenvalues of the channel correlation matrix.

In long-code CDMA, the spreading codes change for every symbol interval. Hence the optimal set of weights that leads to the minimum achievable MSE for each symbol interval must be re-computed and updated symbol by symbol. Unfortunately the eigenvalue decomposition involved is prohibitively complex for implementation. Instead we consider using a fixed set of weights designed to give the best *average* performance. This is feasible since the eigenvalues of randomly selected correlation matrices are seen to be clustered around certain values. As the system becomes larger, the clustering gets increasingly tight [25]. In fact the penalty of averaging rather than using the optimal, time-varying weights is negligible in most cases of interest.

Different optimality criteria in averaging over all code-sets can be adopted. In this paper we minimise the ensemble averaged MSE over all possible channel matrices. This strategy is in general very close to minimising the bit-error-rate [11]. Moreover, the MSE is a quadratic function of the filter tap weights and therefore has a unique global minimum. Following the approach in [26], we demonstrate that given the number of stages, a unique optimal set of weights exists, which leads to the minimum achievable ensemble averaged MSE at the last stage of a weighted PIC.

For an m-stage PIC, the weights depend on the first $2m$ moments of the eigenvalues of channel correlation matrix. Previously, asymptotic analysis of the eigenvalue distribution as the size of the multiuser system goes to infinity has been presented in [25], [27]. Here, we demonstrate a method for deriving the exact expressions for the moments of the eigenvalues, which is also known as the moments of the correlation matrix [28]. The moments are found to be polynomials of the processing gain, the number of active users and the received signal energies. The computational complexity of calculating the weights increases only linearly with the number of users. Hence, it can be implemented on-line given a moderate number of PIC stages. It should be noted that weight updates are only required if the number of active users or the received signal energies change.

The paper is organised as follows. The following section briefly introduces the CDMA uplink model. Section III de-

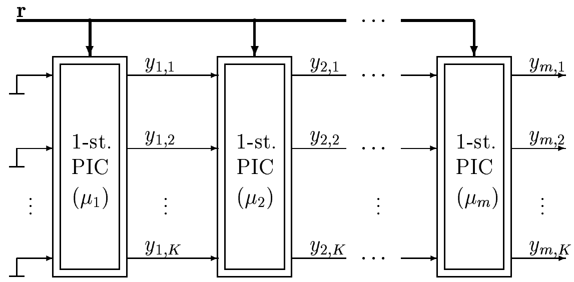

Fig. 1. General structure for a K-user, m-stage PIC.

scribes the PIC structure and presents the proposed method for obtaining the optimal weights. The moments of the correlation matrix are considered in Section IV where exact expressions are derived. Simulation results are shown in Section V and Section VI concludes the paper.

II. Uplink Model

A specific user in a K-user communication system transmits an M-ary PSK information symbol $d_k \in \{\exp(j(2p-1)\pi/M)\}$, $p = 1, 2, ..., M$, by multiplying the symbol with a q-ary PSK spreading code $\mathbf{s}_k$ of length N chips and then transmitting over an AWGN channel, i.e., $\mathbf{s}_k = [s_{1k}, s_{2k}, \cdots, s_{Nk}]^\top/\sqrt{N}$, where $s_{nk} \in \{\exp(j(2p-1)\pi/q)\}$, $p = 1, 2, \ldots, q$. The spreading codes transmitted by each user in any given symbol interval are assumed to be symbol-synchronous. Note that we have assumed that $\mathbf{s}_k^{\mathrm{H}}\mathbf{s}_k = 1$. Also denote the received signal energy of user k by w_k. The output of a chip-matched filter is then expressed as a weighted linear combination of spreading codes, $\mathbf{r} = \mathbf{A}\mathbf{d} + \mathbf{n} \in \mathbf{C}^N$, where $\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_K) = \left(\sqrt{w_1}\mathbf{s}_1, \sqrt{w_2}\mathbf{s}_2, \cdots, \sqrt{w_K}\mathbf{s}_K\right)$, $\mathbf{d} = (d_1, d_2, \cdots, d_K)^\top$, and $\mathbf{n}$ is a noise vector where each sample is independently, circularly complex Gaussian distributed with zero mean and variance $\sigma^2 = \frac{N_0}{2}$. The received signal-to-noise ratio (SNR) of user k can then be defined as $\beta_k = w_k/N_0$.

III. Linear Weighted Parallel Interference Cancellation

The general structure for an m-stage PIC is illustrated in Fig. 1. The detailed structure of one PIC stage with weighting parameter μ_i is depicted in Fig. 2, where α is a non-negative parameter to be discussed later on. Note that all thick lines represent vectors of length N.

It was pointed out in [22] that the weighted PIC structure is essentially a realisation of the steepest descent optimisation method (SDOM) for iteratively approaching the MMSE estimate. The weights in the structure then correspond to the variable step sizes in the SDOM. Consider one particular symbol interval, where the spreading codes used are randomly chosen from all possible code-sets. It has been shown that the set of output decision statistics for all K users at stage i is determined recursively by [22]:

$$\begin{aligned}\mathbf{y}_i &= (\mathbf{I} - \mu_i(\mathbf{R} + \alpha\mathbf{I}))\,\mathbf{y}_{i-1} + \mu_i\mathbf{A}^{\mathrm{H}}\mathbf{r}, &(1)\\ &= (1 - \alpha\mu_i)\mathbf{y}_{i-1} + \mu_i\mathbf{A}^{\mathrm{H}}(\mathbf{r} - \mathbf{A}\mathbf{y}_{i-1}). &(2)\end{aligned}$$

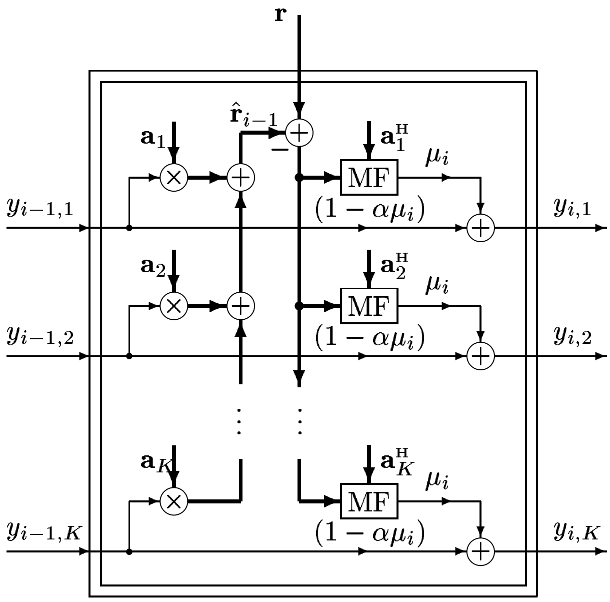

Fig. 2. The structure for stage i, employing weight μ_i. MF — matched filtering.

where $\mathbf{R} = \mathbf{A}^{\mathrm{H}}\mathbf{A}$ is the correlation matrix of the received spreading waveform. Note that (2) essentially corresponds to the weighted interference cancellation structure depicted by Fig. 1, in which the interference is reconstructed and cancelled in each stage to improve the decision statistics.

Assuming $\mathbf{y}_0$ to be $\mathbf{0}$, this relationship can also be described by a one-shot linear matrix filter as $\mathbf{y}_m = \mathbf{G}_m^{\mathrm{H}}\mathbf{r}$, where

$$\mathbf{G}_m^{\mathrm{H}} = \left[\mathbf{I} - \prod_{i=1}^{m}(\mathbf{I} - \mu_i(\mathbf{R}+\alpha\mathbf{I}))\right](\mathbf{R}+\alpha\mathbf{I})^{-1}\mathbf{A}^{\mathrm{H}}. \quad (3)$$

Note that if the product term can be reduced to (close to) $\mathbf{0}$, the detector would be (approximately) $\mathbf{G}^{\mathrm{H}} = (\mathbf{R}+\alpha\mathbf{I})^{-1}\mathbf{A}^{\mathrm{H}}$. We can thus realise the decorrelator by making $\alpha = 0$, or the MMSE detector with $\alpha = \sigma^2$.

Assume that $\lambda_1, \lambda_2, ..., \lambda_K$ are the eigenvalues[1] of $\mathbf{R}$ and $\boldsymbol{\mu} = (\mu_1, \mu_2, ..., \mu_m)^{\top}$ is a vector formed by the m weights for an m-stage PIC. It is then possible to get a closed-form expression for the MSE as a function of the weighting parameters of the PIC structure:

$$J^{(m)}(\boldsymbol{\mu},\alpha) = \mathsf{E}\left\{\|\mathbf{G}_m^{\mathrm{H}}\mathbf{r} - \mathbf{d}\|^2\right\} = J_{\mathrm{MMSE}} + J_{\mathrm{ex}}^{(m)}(\boldsymbol{\mu},\alpha),$$

where the expectation is taken over the probability density function (pdf) of the noise and that of the data vector $\mathbf{d}$. Here

$$J_{\mathrm{MMSE}} = \sum_{k=1}^{K}\frac{\sigma^2}{\lambda_k+\sigma^2} \quad (4)$$

is the minimum mean squared error which can be achieved only by the MMSE filter and does not depend on the weights.

[1] They are obviously non-negative real since $\mathbf{R}$ is positive semi-definite.

The second term, called the excess MSE, represents the degradation with respect to the MMSE detector. It can be expressed as

$$J_{\mathrm{ex}}^{(m)}(\boldsymbol{\mu},\alpha) = \sum_{k=1}^{K}\frac{\lambda_k(\lambda_k+\sigma^2)}{(\lambda_k+\alpha)^2}\left|\frac{\sigma^2-\alpha}{\lambda_k+\sigma^2} - \prod_{i=1}^{m}(1-\mu_i(\lambda_k+\alpha))\right|^2 \quad (5)$$

In [22], the value of $\boldsymbol{\mu}$ that minimises $J_{\mathrm{ex}}^{(m)}(\boldsymbol{\mu},\alpha)$ was found to be unique and a function of α as well as the eigenvalues of $\mathbf{R}$. For long-code systems, however, it is not feasible to compute the optimal set of weights for every symbol interval. Instead we consider using a fixed set of parameters that will minimise the ensemble average of the excess MSE over random codes. This excess MSE is described by

$$\mathcal{J}_{\mathrm{ex}}^{(m)}(\boldsymbol{\mu},\alpha) = \mathsf{E}\left\{J_{\mathrm{ex}}^{(m)}(\boldsymbol{\mu},\alpha)\right\}, \quad (6)$$

where the expectation is now taken over the pdf of the spreading code matrix $\mathbf{A}$.

Given a number of stages, m, which is smaller than K due to limitations on the overall receiver complexity, our objective is to find the global minimum of $\mathcal{J}_{\mathrm{ex}}^{(m)}(\boldsymbol{\mu},\alpha)$ with respect to $\boldsymbol{\mu}$ and α. We first assume that α is given and minimise (6) with respect to $\boldsymbol{\mu}$. For notational simplicity, we define $\phi_k = \lambda_k + \sigma^2$, $\gamma_k = \lambda_k + \alpha$, and $\boldsymbol{\gamma}_k = \left(1, \gamma_k, ..., \gamma_k^{m-1}\right)^{\top}$ for $k = 1, 2, \ldots, K$. We also introduce the following mapping,

$$T: \ \mathbf{C}^m \to \mathbf{C}^m \ \text{ given by } \mathbf{x} = T(\boldsymbol{\mu})$$

where $\mathbf{x} = (x_1, x_2, ..., x_m)^{\top}$ and

$$\begin{aligned} x_1 &\triangleq (-1)(\mu_1+\mu_2+\cdots+\mu_m) \\ x_2 &\triangleq (-1)^2(\mu_1\mu_2+\mu_1\mu_3+\cdots+\mu_{m-1}\mu_m) \\ \vdots & \qquad \cdots \qquad \vdots \\ x_m &\triangleq (-1)^m\mu_1\mu_2\cdots\mu_m. \end{aligned} \quad (7)$$

This mapping is known as the elementary symmetric polynomial transform, or equivalently as the disjunctive normal form [29, pg. 371]. Clearly it is not one-to-one since we have in general $m!$ different $\boldsymbol{\mu}$'s that lead to the same $\mathbf{x}$. Note that the product component in (5) can now be written as

$$\prod_{i=1}^{m}(1-\mu_i(\lambda_k+\alpha)) = 1 + \sum_{i=1}^{m}x_i(\lambda_k+\alpha)^i = 1 + \gamma_k\cdot\boldsymbol{\gamma}_k^{\top}\mathbf{x}. \quad (8)$$

Based on (8), we can then express the excess MSE as a quadratic function of $\mathbf{x}$ as

$$\mathcal{J}_{\mathrm{ex}}^{(m)}(\mathbf{x},\alpha) = \mathsf{E}\left\{\sum_{k=1}^{K}\lambda_k\phi_k\left|\frac{1}{\phi_k}+\boldsymbol{\gamma}_k^{\top}\mathbf{x}\right|^2\right\}. \quad (9)$$

Differentiating with respect to $\mathbf{x}^*$, we have the gradient of the excess MSE as

$$\frac{\partial \mathcal{J}_{\mathrm{ex}}^{(m)}(\mathbf{x},\alpha)}{\partial \mathbf{x}^*} = \mathsf{E}\left\{\sum_{k=1}^{K} \lambda_k \phi_k \boldsymbol{\gamma}_k \left(\frac{1}{\phi_k} + \boldsymbol{\gamma}_k^\top \mathbf{x}\right)\right\}. \quad (10)$$

Equating the above to zero gives the minimum of $\mathcal{J}_{\mathrm{ex}}^{(m)}(\mathbf{x},\alpha)$ as the solution to

$$\mathbf{C}\mathbf{x} = -\mathbf{p} \quad (11)$$

where

$$\mathbf{C} = \mathsf{E}\left\{\sum_{k=1}^{K} \lambda_k \phi_k \boldsymbol{\gamma}_k \boldsymbol{\gamma}_k^\top\right\} \in \mathbf{R}^{(m\times m)} \quad (12)$$

and

$$\mathbf{p} = \mathsf{E}\left\{\sum_{k=1}^{K} \lambda_k \boldsymbol{\gamma}_k\right\} \in \mathbf{R}^{m}. \quad (13)$$

Here $\mathbf{C}$ is an expectation taken over a set of positive semi-definite Hermitian matrices. It is clear that $\mathbf{C}$ is positive definite since for any non-zero vector $\mathbf{z}$,

$$\mathbf{z}^\top \mathbf{C}\mathbf{z} = \mathsf{E}\left\{\sum_{k=1}^{K} \lambda_k \phi_k (\boldsymbol{\gamma}_k^\top \mathbf{z})^2\right\} > 0. \quad (14)$$

Hence the unique real minimum in $\mathbf{x}$ is obtained as

$$\hat{\mathbf{x}} = -\mathbf{C}^{-1}\mathbf{p}. \quad (15)$$

We can then find the corresponding equivalent minimum in $\boldsymbol{\mu}$ by considering the following polynomial,

$$\begin{aligned} p(\mu) &= \mu^m + \hat{x}_1 \mu^{m-1} + \hat{x}_2 \mu^{m-2} + \cdots + \hat{x}_m && (16)\\ &= (\mu - \hat{\mu}_1)(\mu - \hat{\mu}_2)\cdots(\mu - \hat{\mu}_m) && (17) \end{aligned}$$

which has exactly m roots, $(\hat{\mu}_1, \hat{\mu}_2, \cdots, \hat{\mu}_m)$, that are related to $\hat{\mathbf{x}}$ through the elementary symmetric polynomial transform. It can be easily shown that the set of all permutations of these roots gives the complete set of solutions to the problem $T(\hat{\boldsymbol{\mu}}) = \hat{\mathbf{x}}$. In other words, any vector $\hat{\boldsymbol{\mu}} = (\hat{\mu}_1, \hat{\mu}_2, \cdots, \hat{\mu}_m)^\top$ consisting of a permutation of the m roots of (16) leads to the global minimum of $\mathcal{J}_{\mathrm{ex}}^{(m)}(\boldsymbol{\mu},\alpha)$, and no other minimum exists.

It can be shown that the parameter α has no influence on the minimum achievable MSE, i.e., for any $\alpha \geq 0$, there exists a unique set of weights (dependent on α) that will give the same minimum of $\mathcal{J}_{\mathrm{ex}}^{(m)}(\boldsymbol{\mu},\alpha)$. However, if $\alpha = \sigma^2$, the optimal weights are always real; otherwise they can be complex numbers [30]. This affects the implementation complexity.

The problem that remains is the computation of $\mathbf{C}$ and $\mathbf{p}$. Good numerical estimates can be found for $\mathbf{C}$ and $\mathbf{p}$ based on Monte Carlo averaging over random codes. Such an approach however, is very complex and can not be carried out on-line for a dynamic system. Instead an analytical approach based on statistical moments of the eigenvalues can be used.

IV. Moments of the Correlation Matrix

It is helpful here to define the r^{th} order moment of the correlation matrix as

$$M_r = \mathsf{E}\left\{\frac{1}{K}\sum_{k=1}^{K} (\lambda_k)^r\right\}. \quad (18)$$

Expanding (11) and dividing both sides by K, we get

$$\begin{bmatrix} c_2 & c_3 & \cdots & c_{m+1} \\ c_3 & c_4 & \cdots & c_{m+2} \\ \vdots & \vdots & & \vdots \\ c_{m+1} & c_{m+2} & \cdots & c_{2m} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_m \end{bmatrix}, \quad (19)$$

where

$$\begin{aligned} c_i &= \mathsf{E}\left\{\frac{1}{K}\sum_{k=1}^{K} \lambda_k \phi_k \gamma_k^{i-2}\right\} \\ &= \sum_{r=2}^{i} \binom{i-2}{r-2} \alpha^{i-r} \left(M_r + \sigma^2 M_{r-1}\right), \\ &\qquad i = 2, 3, \ldots, 2m, \end{aligned} \quad (20)$$

and

$$p_i = \mathsf{E}\left\{\frac{1}{K}\sum_{k=1}^{K} \lambda_k \gamma_k^{i-1}\right\} = \sum_{r=1}^{i} \binom{i-1}{r-1} \alpha^{i-r} M_r, \quad i = 1, 2, \ldots, m. \quad (21)$$

Obviously the elements of $\mathbf{C}$ and $\mathbf{p}$ are determined by the first $2m$ moments of the correlation matrix. It is difficult, if not impossible, to get a closed-form expression for M_r, where r is an integer variable. In fact there is no known general expression. However, given a particular integer value of r, M_r can be derived as follows.

A. Deriving the Moments

Consider chip n of spreading waveform for user k as a random variable, denoted by S_{nk}, to distinguish it from a realisation s_{nk}. For a long-code system all the chips S_{nk}, $n = 1, 2, \cdots, N$, $k = 1, 2, \cdots, K$, are mutually independent random variables, each uniformly distributed over the q-ary constellation. The corresponding chip sample observed at the receiver may be expressed as $A_{nk} = \sqrt{w_k} S_{nk}$. Then the following properties obviously hold,

$$\begin{aligned} \mathsf{E}\left\{A_{n_1 k_1}^* A_{n_2 k_2}\right\} &= \frac{\sqrt{w_{k_1} w_{k_2}}}{N} \delta(n_1 - n_2)\cdot\delta(k_1 - k_2) \\ &= \begin{cases} w_{k_1}/N & \text{if } n_2 = n_1 \text{ and } k_2 = k_1, \\ 0 & \text{otherwise,} \end{cases} \end{aligned} \quad (22)$$

and

$$\mathsf{E}\left\{(A_{nk})^r\right\} = \left(\frac{w_k}{N}\right)^{\frac{r}{2}} \sum_{m=0}^{q-1} e^{2\pi m r/q} = \begin{cases} \left(\frac{w_k}{N}\right)^{\frac{r}{2}} & \text{if } r/q \in \mathbf{Z}, \\ 0 & \text{otherwise.} \end{cases} \quad (23)$$

The correlation between user i and user j's spreading codes is also a random variable

$$R_{ij} = \mathbf{a}_i^{\mathrm{H}} \mathbf{a}_j = \sum_{n=1}^{N} A_{ni}^* A_{nj}, \tag{24}$$

which is the element of row i, column j of $\mathbf{R}$.

Considering the definition of the r^{th} order moment, we have

$$M_r = \mathsf{E}\left\{ \frac{1}{K} \sum_{k=1}^{K} (\lambda_k)^r \right\} = \frac{1}{K} \mathsf{E}\left\{ \mathrm{tr}\left\{ \Lambda^r \right\} \right\} = \frac{1}{K} \mathsf{E}\left\{ \mathrm{tr}\left\{ \mathbf{R}^r \right\} \right\}, \tag{25}$$

where the trace of $\mathbf{R}^r$ can be expressed as

$$\mathrm{tr}\left\{ \mathbf{R}^r \right\} = \sum_{k_1=1}^{K} \sum_{k_2=1}^{K} \cdots \sum_{k_r=1}^{K} R_{k_1 k_2} R_{k_2 k_3} \cdots R_{k_{r-1} k_r} R_{k_r k_1}. \tag{26}$$

It then follows that

$$M_r = \frac{1}{K} \sum_{k_1=1}^{K} \sum_{k_2=1}^{K} \cdots \sum_{k_r=1}^{K} \mathsf{E}\left\{ R_{k_1 k_2} R_{k_2 k_3} \cdots R_{k_{r-1} k_r} R_{k_r k_1} \right\} \tag{27}$$

$$= \frac{1}{K} \sum_{k_1=1}^{K} \sum_{k_2=1}^{K} \cdots \sum_{k_r=1}^{K} \cdot \sum_{n_1=1}^{N} \sum_{n_2=1}^{N} \cdots \sum_{n_r=1}^{N} \mathsf{E}\left\{ A_{n_1 k_1} A_{n_1 k_2}^* A_{n_2 k_2} A_{n_2 k_3}^* \cdots A_{n_r k_r} A_{n_r k_1}^* \right\}. \tag{28}$$

Here, A_{nk} are independent random variables selected from a scaled q-ary PSK[2] constellation with equal probability. Based on the statistical properties of the code-set given in (22) and (23), only terms containing all complex conjugate pairs and/or q-powers of the variables A_{nk} are relevant [28]. It is therefore possible to obtain M_r through evaluation of the expectation over all combinations of indices. This involves a grouping of the indices into equivalence classes. Details of this grouping and evaluation can be found in [30]. As the expectation is taken over all code-sets, M_r only depends on N, K, and the received signal energies and not on specific spreading codes. In fact M_r is shown to be a polynomial in N and K as well as the first r moments of the received signal energies. Here the r^{th} order moment of the energies is defined as

$$\mathcal{E}_r = \sum_{k=1}^{K} (w_k)^r. \tag{29}$$

In the Appendix we list the exact expressions for the first 6 moments obtained by computer-aided symbolic manipulations assuming BPSK spreading.

[2]This approach is not confined to PSK spreading only. It is applicable for arbitrary spreading schemes, provided that the statistical property of the spreading codes are known. Furthermore, influences of asynchronism and multi-path fading can also be incorporated here.

B. Determining the Order of the Weights

Since we have m weights, we have $m!$ different permutations that all lead to the same MSE at the last stage. The order in which the m weights are applied however, has a significant influence on the MSE performance at intermediate stages. Following the approach in [26], we have chosen to order the optimal weights according to a recursive minimisation of $\mathcal{J}_{\mathrm{ex}}^{(i)}(\hat{\boldsymbol{\mu}}_i, \alpha)$ for $i = 1, 2, ..., m$ where $\hat{\boldsymbol{\mu}}_i = (\hat{\mu}_1, \hat{\mu}_2, ..., \hat{\mu}_{i-1}, \mu_i)^{\top}$. Such an order is obtained by selecting at stage i the weight $\mu_i \in \mathcal{V}_i$ which is closest to $\tilde{\mu}_i$. Here, $\mathcal{V}_i$ denotes the set of $(m - i + 1)$ elements of $\hat{\boldsymbol{\mu}}$ which have not been used in the first $i - 1$ stages, and

$$\tilde{\mu}_i = \arg\min_{\mu_i} \mathcal{J}_{\mathrm{ex}}^{(i)}(\hat{\boldsymbol{\mu}}_i, \alpha). \tag{30}$$

The closest weight to $\tilde{\mu}_i$ is the best choice since $\mathcal{J}_{\mathrm{ex}}^{(i)}(\hat{\boldsymbol{\mu}}_i, \alpha)$, given that all previous $\hat{\mu}_j$ have already been chosen, is a quadratic function in μ_i. It has been found that $\tilde{\mu}_i$ can also be determined based on the first $2i$ moments of the correlation matrix [30].

C. Computational Complexity

Assuming that the received signal energies, the noise variance as well as N and K are known, the procedure of computing the weights includes evaluation of the engergy moments, $\mathcal{E}_1, \mathcal{E}_2, \cdots, \mathcal{E}_{2m}$, the eigenvalue moments, $M_1, M_2, \cdots, M_{2m}$, and then $\mathbf{C}$ and $\mathbf{p}$. $\hat{\mathbf{x}}$ is then solved from (15) and the weights are obtained as the roots of (16). Finally an appropriate order of the weights is determined. The total computational complexity of weight updating is of the order of $(4mK + 6m^3 + 19m^2)$ floating point operations. It is obviously independent of the processing gain and linear in the number of active users. This complexity is in fact negligible in comparison to $O(mKN)$, which is the complexity of performing code-matched filtering for all K users in an m-stage interference canceller. It follows that weight-updating can be done on-line and does not noticeably increase the overall system complexity.

V. Numerical Results

The numerical examples considered in this section are based on a symbol-synchronous system with $K = 15$ users. BPSK modulation and spreading formats are assumed and randomly generated long codes with a processing gain of $N = 31$ are considered. The parameter α is set to 0 in all examples.

Fig. 3 shows the stage-by-stage BER performance of a PIC and the loss incurred by ensemble averaging over random codes, as compared to exact selection of parameters for each symbol interval. An SNR of 7 dB is assumed for all users, i.e., $\beta_k = 7$ dB for $k = 1, \cdots, K$. We observe that the conventional PIC scheme diverges since the largest eigenvalue of $\mathbf{R}$ for a $K = 15$, $N = 31$ system, is almost always greater than 2 [30]. The eigenvalue criterion for convergence ($\lambda_{\max} < 2$) has been shown in [21]. Divergence can be overcome by a proper choice of weights, optimised for an SNR of 7 dB and ordered as described

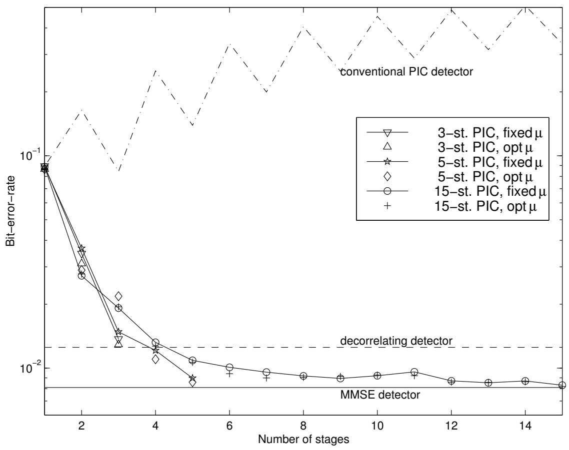

Fig. 3. BER performance as a function of the number of stages. Both the cases of a fixed set of weights as well as optimal weights for each symbol interval are included. Here, SNR=7 dB, $N = 31$ and $K = 15$.

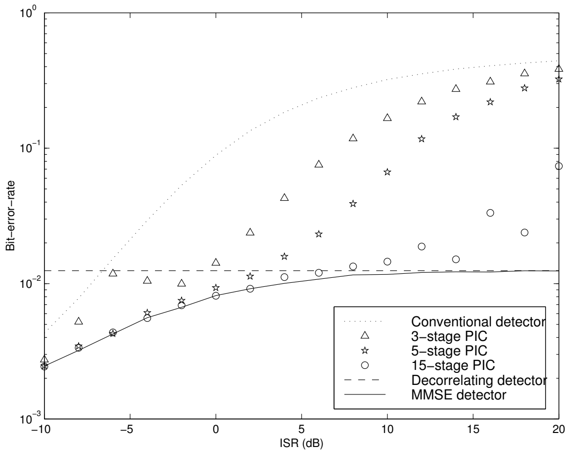

Fig. 4. Near-far ability of the weighted linear PIC for a long-code system. $N = 31$, $K = 15$ and the SNR for user 1 is $\beta_1 = 7$ dB while for all other users $\beta_k = 7+$ISR dB. The performance of user 1 is shown.

in section IV. The BER performance of a PIC using this fixed set of weights are represented by the solid lines in Fig. 3. Significant improvement over the conventional detector can be achieved using merely 3 stages. A 5-stage detector performs much better than the decorrelator and gives close to average MMSE performance while 15-stage PIC gives virtually MMSE performance. The performance of a PIC that makes use of the optimal set of weights corresponding to the instantaneous spreading codes in each symbol interval is also shown in the figure for comparison. It is clear that the penalty of making a compromise over all code-sets is negligible.

Fig. 4 shows the BER performance of a weighted PIC detector in a near-far environment, as compared with that of the conventional detector, the decorrelator and the MMSE detector. The SNR of the first user is $\beta_1 = 7$ dB while the remaining 14 users have the same SNR of $\beta_k = \beta_1+$ISR, for $k = 2, \cdots, K$ where ISR denotes the interference-to-signal

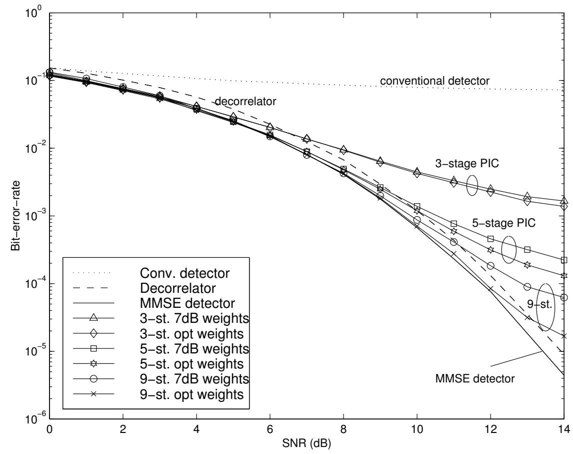

Fig. 5. BER performance and sensitivity versus SNR using long codes. Weights optimised for 7 dB are used for 0-14 dB, in comparison to when the weights are optimised for the actual SNR. Here, $N = 31$, $K = 15$ and $\beta_k = \beta$ (perfect power control) is assumed.

ratio in dB. The curves show the BER of user 1 only. It is observed that the PIC performs better than the conventional detector but worse than the MMSE detector. As the number of stages increases, the ability to combat the near-far environment improves.

The set of weights depends on the received signal energies (w_k) and the noise variance (σ^2), or equivalently, the SNRs (β_k). It is therefore of interest to investigate the sensitivity of the BER-performance to correct estimates of the SNR. This sensitivity is illustrated in Fig. 5. Perfect power control is assumed, i.e., $\beta_k = \beta$ for all k's. The weights determined for an assumed SNR of 7 dB are used for actual SNR's from 0 to 14 dB. It is compared to the case where the weights are optimised for the encountered SNR. As the number of stages increases, the sensitivity also increases slightly. For 3-stage, 5-stage and 9-stage PIC structures, a set of weighting factors determined for 7 dB works well for a wide range of SNR.

VI. Concluding Remarks

In this paper, we have proposed a weighted linear parallel interference cancellation structure for multiuser detection in long-code CDMA. Using a set of fixed weights found by averaging over the ensemble of (random) long codes, the detector achieves close to the MMSE performance in a few stages. The penalty of averaging rather than using the optimal, time-varying weights is virtually negligible. The weights depend on the moments of the eigenvalues of the code correlation matrix. Exact expressions for the moments are found to be polynomials in N, K and all users' received signal energies. In a dynamic system, the weights can be updated on-line as either the number of users or the received energies change. The involved complexity increases only linearly with the number of active users and is independent of the processing gain. Significant performance improvements are observed and the near-far problem is substantially alleviated as compared to the conventional PIC. For as few as 3 stages, it is possible to get close to the MMSE performance.

Acknowledgement

The authors would like to acknowledge the assistance provided by Dr. Paul D. Alexander and Ms. Sumei Sun.

Appendix
The First Six Moments

It is possible to derive moments of any order by evaluating the expectation in (28) over all combination of indices. Assuming BPSK spreading, the expressions for the first 6 moments are listed below. These moments are sufficient for computing the time-invariant weights for a 3-stage PIC. Detailed derivation as well as 4 higher order moments can be found in [30].

$$M_1 = \frac{1}{K}\mathcal{E}_1 \tag{31}$$

$$M_2 = \frac{1}{KN}[\mathcal{E}_1^2 + \mathcal{E}_2(N-1)] \tag{32}$$

$$M_3 = \frac{1}{KN^2}[\mathcal{E}_1^3 + \mathcal{E}_1\mathcal{E}_2(3N-3) + \mathcal{E}_3(N^2-3N+2)] \tag{33}$$

$$\begin{aligned} M_4 &= \frac{1}{KN^3}[\mathcal{E}_1^4 + \mathcal{E}_1^2\mathcal{E}_2(6N-6) + \mathcal{E}_1\mathcal{E}_3(5N^2-13N+8) \\ &+ \mathcal{E}_2^2(N^2-2N+1) + \mathcal{E}_4(N^3-6N^2+9N-4)] \end{aligned} \tag{34}$$

$$\begin{aligned} M_5 &= \frac{1}{KN^4}[\mathcal{E}_1^5 + \mathcal{E}_1^3\mathcal{E}_2(10N-10) + \mathcal{E}_1^2\mathcal{E}_3(14N^2-34N+20) \\ &+ \mathcal{E}_1\mathcal{E}_2^2(6N^2-11N+5) + \mathcal{E}_1\mathcal{E}_4(8N^3-36N^2+48N-20) \\ &+ \mathcal{E}_2\mathcal{E}_3(2N^3-9N^2+7N) \\ &+ \mathcal{E}_5(N^4-10N^3+25N^2-20N+4)] \end{aligned} \tag{35}$$

$$\begin{aligned} M_6 &= \frac{1}{KN^5}[\mathcal{E}_1^6 + \mathcal{E}_1^4\mathcal{E}_2(15N-15) \\ &+ \mathcal{E}_1^3\mathcal{E}_3(30N^2-70N+40) + \mathcal{E}_1^2\mathcal{E}_2^2(20N^2-35N+15) \\ &+ \mathcal{E}_1^2\mathcal{E}_4(30N^3-121N^2+151N-60) \\ &+ \mathcal{E}_1\mathcal{E}_2\mathcal{E}_3(19N^3-61N^2+40N+2) \\ &+ \mathcal{E}_1\mathcal{E}_5(12N^4-79N^3+163N^2-118N+22) \\ &+ \mathcal{E}_2^3(N^3-6N^2+8N-3) \\ &+ \mathcal{E}_2\mathcal{E}_4(2N^4-22N^3+34N^2-6N-8) \\ &+ \mathcal{E}_3^2(N^4-4N^3+2N^2+7N-6) \\ &+ \mathcal{E}_6(N^5-15N^4+55N^3-61N^2+8N+12)]. \end{aligned} \tag{36}$$

References

[1] F. Adachi, M. Sawahashi, and H. Suda, "Wideband DS-CDMA for next-generation mobile communications systems," *IEEE Communication Magazine*, vol. 36, pp. 56–69, Sept. 1998.

[2] E. H. Dinan and B. Jabbari, "Spreading codes for direct sequence CDMA and wideband CDMA cellular networks," *IEEE Communication Magazine*, vol. 36, pp. 48–54, Sept. 1998.

[3] "Mobile station/base station compatibility standard for dual-mode wideband spread spectrum cellular system." TIA/EIA Interim Standard IS-95A, March 1995.

[4] R. Lupas and S. Verdú, "Near-far resistance of multiuser detectors in asynchronous channels," *IEEE Trans. Commun.*, vol. 38, pp. 496–508, April 1990.

[5] S. Verdú, "Minimum probability of error for asynchronous Gaussian multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 32, pp. 85–96, Jan. 1986.

[6] S. Verdú, "Computational complexity of optimum multiuser detection," *Algorithmica*, vol. 4, pp. 303–312, 1989.

[7] S. Verdú, *Multiuser Detection.* Cambridge University Press, 1998.

[8] R. Lupas and S. Verdú, "Linear multiuser detectors for synchronous code-division multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 35, pp. 123–136, Jan. 1989.

[9] T. J. Lim, D. Guo, and L. K. Rasmussen, "Noise enhancement in the family of decorrelating detectors for multiuser CDMA," in *Proceedings IEEE Asia-Pac. Conf. Comms./Int'l Conf. Comm. Systems (APCC/ICCS)*, pp. 401–405, Singapore, Nov. 1998.

[10] Z. Xie, R. T. Short, and C. K. Rushforth, "A family of suboptimum detector for coherent multiuser communications," *IEEE J. Selected Areas Commun.*, vol. 8, pp. 683–690, May 1990.

[11] H. V. Poor and S. Verdú, "Probability of error in MMSE multiuser detection," *IEEE Trans. Inform. Theory*, vol. 43, pp. 858–871, May 1997.

[12] U. Madhow and M. Honig, "MMSE interference suppression for direct sequence spread spectrum CDMA," *IEEE Trans. Commun.*, vol. 42, pp. 3178–3188, Dec. 1994.

[13] T. J. Lim and S. Roy, "Adaptive filters in multiuser (MU) CDMA detection," *Wireless Networks*, no. 4, pp. 307–318, 1998.

[14] T. J. Lim, L. K. Rasmussen, and H. Sugimoto, "An adaptive asynchronous multiuser CDMA detector based on the Kalman filter," *IEEE J. Selected Areas Commun.*, vol. 16, pp. 1711–1722, Dec. 1998.

[15] M. J. Juntti, B. Aazhang, and J. O. Lilleberg, "Iterative implementation of linear multiuser detection for dynamic asynchronous CDMA systems," *IEEE Trans. Commun.*, vol. 46, pp. 503–508, April 1998.

[16] H. Elders-Boll, H. D. Schotten, and A. Busboom, "Efficient implementation of linear multiuser detectors for asynchronous CDMA systems by linear interference cancellation," *European Trans. on Telecommun.*, vol. 9, pp. 427–438, Sept./Oct. 1998.

[17] M. K. Varanasi and B. Aazhang, "Multistage detection in asynchronous code-division multiple-access communications," *IEEE Trans. Commun.*, vol. 38, pp. 509–519, April 1990.

[18] D. Divsalar, M. Simon, and D. Raphaeli, "Improved parallel interference cancellation for CDMA," *IEEE Trans. Commun.*, vol. 46, pp. 258–268, Feb. 1998.

[19] T. Suzuki and Y. Takeuchi, "Real-time decorrelation scheme with multi-stage architecture for asynchronous DS/CDMA," *Technical Report of the IEICE*, vol. SST96-10, SAT96-24, RCS96-34, 1996. (in Japanese).

[20] S. Moshavi, E. G. Kanterakis, and D. L. Schilling, "Multistage linear receivers for DS-CDMA systems," *International Journal of Wireless Information Networks*, vol. 3, no. 1, pp. 1–17, 1996.

[21] L. K. Rasmussen, D. Guo, T. J. Lim, and Y. Ma, "Aspects on linear parallel interference cancellation in CDMA," in *Proceedings 1998 IEEE International Symposium on Information Theory*, p. 37, MIT, Cambridge, MA USA, Aug. 1998.

[22] D. Guo, L. K. Rasmussen, S. Sun, T. J. Lim, and C. Cheah, "MMSE–based linear parallel interference cancellation in CDMA," in *Proceedings IEEE Fifth International Symposium on Spread Spectrum Techniques and Applications*, vol. 3, pp. 917–921, Sun City, South Africa, Sept. 1998.

[23] R. Fletcher, *Practical Methods of Optimization.* John Wiley, 1987.

[24] O. Axelsson, *Iterative Solution Methods.* Cambridge University Press, 1994.

[25] A. J. Grant and P. D. Alexander, "Random sequence multisets for synchronous code-division multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2832–2836, Nov. 1998.

[26] D. Guo, L. K. Rasmussen, S. Sun, and T. J. Lim, "A matrix-algebraic approach to linear parallel interference cancellation in CDMA," *IEEE Trans. Commun.*, vol. 48, pp. 152–161, Jan. 2000.

[27] D. N. C. Tse and S. V. Hanly, "Linear multiuser receivers: effective interference, effective bandwidth and user capacity," *IEEE Trans. Inform. Theory*, vol. 45, pp. 622–640, March 1999.

[28] D. Jonsson, "Some limit theorems for the eigenvalues of a sample covariance matrix," *Journal of Multivariate Analysis*, vol. 12, pp. 1–38, Dec. 1982.

[29] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes.* North-Holland Publ. Company, 1978.

[30] D. Guo, "Linear parallel interference cancellation in CDMA," M.Eng. thesis, National University of Singapore, 1998.