MMSE-Based Linear Parallel Interference Cancellation in CDMA

Dongning Guo, Lars K. Rasmussen, Sumei Sun, Teng J. Lim & Christopher Cheah
Centre for Wireless Communications
20 Science Park Road, Singapore Science Park II
#02-34/37 Teletech Park, Singapore 117674
email: cwclkr@leonis.nus.edu.sg
FAX: +65 779 5441

Abstract— **In this paper we mathematically describe the linear parallel interference canceller (PIC) using matrix algebra. It is shown that the linear PIC, whether conventional or weighted, can be seen as a linear matrix filter applied directly to the received chip-matched filtered signal vector. It is then possible to get an analytical expression for the exact bit error rate and to derive necessary conditions on the eigenvalues of the code correlation matrix and the weighting factors to ensure convergence. The close relationship between the steepest descent method for minimising the mean squared error (MSE) and linear PIC is demonstrated and a modified PIC structure is suggested which converges to the MMSE detector rather than the decorrelator. Following the principles of the steepest descent method techniques are devised for optimising the choice of weighting factors with respect to the mean squared error. It is shown that only K (the number of users) PIC stages are required for the equivalent matrix filter to be identical to the MMSE filter. For fewer stages, $m < K$, one unique optimal choice of weighting factors exists which will lead to the minimum achievable MSE at the last stage.**

I. Introduction

Conventional single-user detection techniques are severely affected by MAI, making such systems interference limited. Traditional matched filter receivers for CDMA also require strict power control in order to alleviate the near-far problem where a high-powered user creates significant MAI for low-powered users. More advanced detection strategies can be adopted to improve performance. For practical implementation successive interference cancellation and parallel interference cancellation (PIC) schemes have been subjected to most attention. The first structure based on the principle of interference cancellation was the parallel multi-stage detector in [1]. A significant improvement to the PIC was suggested by Divsalar et al in [2] where they proposed a weighted cancellation scheme. Here the current decision statistic is a weighted sum of the previous decision statistic and the statistic resulting from interference cancellation based on current tentative decisions. They considered both linear and non-linear decisions functions based on joint ML considerations. An identical approach has been suggested by Suzuki and Takeuchi in [3] for a linear PIC.

In this paper we mathematically describe the linear PIC scheme using matrix algebra. Assuming a symbol synchronous system with short codes, we show that the linear PIC schemes correspond to linear matrix filtering that can be performed directly on the received chip-matched filtered signal vector. The approach applies to both conventional and weighted structures. It is then possible to get an analytical expression for the exact bit error rate (BER) and to derive necessary conditions on the eigenvalues of the code correlation matrix and the weighting factors to ensure convergence. The concept of weighted linear PIC resembles the concept of the steepest descent method (SDM) for updating adaptive filter weights to minimise the MSE. Here we demonstrate the close relationship between the two and present a new PIC structure which in fact is a modified version of the structures suggested in [2] and [3]. This new structure will ensure convergence to the performance of the MMSE detector rather than the decorrelating detector which other PIC structures generally converge to. Following the principles of the SDM, we derive the corresponding one-shot cancellation filters and devise techniques for optimising the choice of weighting factors (or equivalently step sizes for the SDM) with respect to the mean squared error (MSE). The case of having a fixed weighting factor was considered in [4]. It is here shown that only K PIC stages are required for the equivalent one-shot filter to be identical to the minimum MSE (MMSE) filter. For fewer stages, $m < K$, one unique optimal choice of step sizes exists which will lead to the minimum achievable MSE. Finally we demonstrate that for long codes, it is possible to find a set of weighting factors that will provide significant performance improvement for the linear PIC scheme as compared to the conventional structure.

The paper is organised as follows. In Section II, the uplink model is briefly described. The algebraic description of the conventional PIC scheme is presented in Section III together with performance analysis. The close connection to the SDM is explained in detail in Section IV and in Section V the optimisation of the step sizes is described and powerful techniques for obtaining these parameters devised. Numerical examples are presented in Section VI and the paper is completed by some concluding remarks.

II. System Model

In this section, the model for the uplink of the CDMA communication system considered throughout this paper is briefly described. The uplink model is based on a discrete-time symbol-synchronous CDMA system assuming single-path channels and the presence of complex stationary additive white Gaussian noise (AWGN) with zero mean and variance $\sigma^2 = N_0$.

A specific user in this K-user communication system transmits an M-ary PSK information symbol $d_k \in \{\exp(j2\pi(i-1)/M)\}$, $i = 1, 2, ..., M$, by multiplying the symbol with a q-ary spreading code $\mathbf{s}_k \in \mathbb{S}^N$, $\mathbb{S} = \{\varrho_1, ..., \varrho_q\}$, $\varrho_j \in \mathbb{C}$ of length N chips and then transmitting over an AWGN channel. The spreading codes

This work was supported in part by Oki Techno Centre (S'pore) Pte. Ltd., National University of Singapore and the National Science and Technology Board (NSTB), S'pore. The CWC is a National Research Centre funded by NSTB.

0-7803-4281-X/97/$10.00 ©1998 IEEE

transmitted by each user in any given symbol interval are assumed to be symbol-synchronous. Note that we have assumed that $\mathbf{s}_k^{\mathrm{H}}\mathbf{s}_k = 1$. To make the notation less cumbersome, we assume perfect power control. The output of a chip-matched filter is then expressed as a linear combination of spreading codes, specifically, the chip matched filtered received vector, $\mathbf{r}$, is a column vector of length N, encompassing the transmissions for all users. The received vector $\mathbf{r}$ is hence described through matrix algebra as $\mathbf{r} = \mathbf{A}\mathbf{d} + \mathbf{n}$, where $\mathbf{A} = (\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_K)$ and $\mathbf{d} = (d_1, d_2, \cdots, d_K)^{\mathsf{T}}$. The sampled noise corrupting the output of the chip-matched filter is independent in each sample since the chip waveforms are assumed to be rectangular. We therefore obtain a noise vector $\mathbf{n}$ where each sample is circularly complex Gaussian distributed with zero mean and variance N_0. The matched filtered statistic is then obtained as $\mathbf{y} = \mathbf{A}^{\mathrm{H}}\mathbf{r} = \mathbf{A}^{\mathrm{H}}\mathbf{A}\mathbf{d}+\mathbf{A}^{\mathrm{H}}\mathbf{n} = \mathbf{R}\mathbf{d}+\mathbf{z}$, where $\mathrm{E}\{\mathbf{z}\mathbf{z}^{\mathrm{H}}\} = \sigma^2\mathbf{R}$.

III. Algebraic Description of Conventional PIC

In a conventional PIC structure, previous tentative decisions are used to estimate the interference for cancellation. The structure is described by the diagram in Fig. 1 with $\mu_{j,k} = 1$ and $\alpha = 0$. The decision statistic at stage $(m+1)$ for user k is then

$$\begin{aligned} y_{m+1,k} &= \mathbf{s}_k^{\mathrm{H}}\left(\mathbf{r} - \sum_{\substack{i=1 \\ i\neq k}}^{K} \mathbf{s}_i \hat{d}_{m,i}\right) \\ &= \mathbf{s}_k^{\mathrm{H}}\left(\mathbf{r} - \sum_{i=1}^{K} \mathbf{s}_i \hat{d}_{m,i}\right) + \hat{d}_{m,k}, \end{aligned}$$

where $y_{1,k} = \mathbf{s}_k^{\mathrm{H}}\mathbf{r} = y_k$ and $\hat{d}_{m,k}$ is a tentative decision for user k at stage m. In a linear PIC structure, $\hat{d}_{m,k} = y_{m,k}$, and defining $\mathbf{y}_m = (y_{m,1}, y_{m,2}, ..., y_{m,K})^{\mathsf{T}}$, the scheme is described by

$$\begin{aligned} \mathbf{y}_{m+1} &= \mathbf{A}^{\mathrm{H}}(\mathbf{r} - \mathbf{A}\mathbf{y}_m) + \mathbf{y}_m = \mathbf{y} - \mathbf{R}\mathbf{y}_m + \mathbf{y}_m \\ &= \mathbf{y} + (\mathbf{I} - \mathbf{R})\mathbf{y}_m, \end{aligned} \tag{1}$$

where $\mathbf{y}_1 = \mathbf{A}^{\mathrm{H}}\mathbf{r} = \mathbf{y}$. Using this recursion, we can express $\mathbf{y}_m$ as

$$\mathbf{y}_m = \sum_{i=1}^{m}(\mathbf{I} - \mathbf{R})^{(i-1)}\mathbf{A}^{\mathrm{H}}\mathbf{r} = \mathbf{G}_m^{\mathrm{H}}\mathbf{r},$$

where $\mathbf{G}_m = \mathbf{A}\sum_{i=1}^{m}(\mathbf{I} - \mathbf{R})^{(i-1)}$. This is the equivalent one-shot cancellation filter for conventional linear PIC. Since $\mathbf{G}_m$ is a linear filter, the noise in $\mathbf{y}_m$ is still Gaussian with correlation matrix $\mathrm{E}\{\mathbf{G}_m^{\mathrm{H}}\mathbf{n}\mathbf{n}^{\mathrm{H}}\mathbf{G}_m\} = \sigma^2\mathbf{G}_m^{\mathrm{H}}\mathbf{G}_m$. We can therefore analytically calculate the BER for user k at stage m using the same techniques as for the conventional matched filter detector. Specifically for BPSK systems,

$$P_b(m,k) = \frac{1}{2^{K-1}} \sum_{\substack{\mathbf{d}\in\{-1,1\}^K \\ d_k=1}} Q\left(\frac{\mathrm{Re}\{\mathbf{g}_{m,k}^{\mathrm{H}}\mathbf{A}\mathbf{d}\}}{\sigma\|\mathrm{Re}\{\mathbf{g}_{m,k}\}\|}\right), \tag{2}$$

where $\mathbf{g}_{m,k}$ is the k^{th} column of $\mathbf{G}_m$.

IV. Steepest Descent Method vs PIC

A linear detector $\mathbf{G}$ is a linear matrix filter that gives the following estimate of the transmitted data symbols, $\mathbf{x} = \mathbf{G}^{\mathrm{H}}\mathbf{r} = \mathbf{G}^{\mathrm{H}}(\mathbf{A}\mathbf{d} + \mathbf{n})$. The corresponding MSE is given by

$$\begin{aligned} J_{\text{MMSE}} &= \mathrm{E}\{\|\mathbf{x} - \mathbf{d}\|^2\} = \mathrm{E}\{\|\mathbf{G}^{\mathrm{H}}\mathbf{r} - \mathbf{d}\|^2\} \\ &= \mathrm{E}\{\mathbf{r}^{\mathrm{H}}\mathbf{G}\mathbf{G}^{\mathrm{H}}\mathbf{r} - \mathbf{d}^{\mathrm{H}}\mathbf{G}^{\mathrm{H}}\mathbf{r} - \mathbf{r}^{\mathrm{H}}\mathbf{G}\mathbf{d} + \mathbf{d}^{\mathrm{H}}\mathbf{d}\}, \end{aligned}$$

and the gradient with respect to $\mathbf{G}^{\mathrm{H}}$, assuming that $\mathrm{E}\{\mathbf{d}\mathbf{d}^{\mathrm{H}}\} = \mathbf{I}$, is [5]

$$\nabla J_{\text{MMSE}} = 2\frac{\partial J_{\text{MMSE}}}{\partial \mathbf{G}^{\mathsf{T}}} = 2\left[\mathbf{G}^{\mathrm{H}}(\mathbf{A}\mathbf{A}^{\mathrm{H}} + \sigma^2\mathbf{I}) - \mathbf{A}^{\mathrm{H}}\right].$$

The steepest descent method gives the following recursion for approaching the MMSE filter,

$$\mathbf{G}_{m+1}^{\mathrm{H}} = \mathbf{G}_m^{\mathrm{H}} - \frac{\mu_{m+1}}{2}\nabla J_{\text{MMSE}},$$

where μ_{m+1} is a variable step size dependent on the current stage. We then have that

$$\mathbf{G}_{m+1}^{\mathrm{H}} = \mathbf{G}_m^{\mathrm{H}} - \mu_{m+1}\left[\mathbf{G}_m^{\mathrm{H}}(\mathbf{A}\mathbf{A}^{\mathrm{H}} + \sigma^2\mathbf{I}) - \mathbf{A}^{\mathrm{H}}\right], \tag{3}$$

where $\mathbf{G}_0 = \mathbf{0}$. The non-recursive form is then

$$\mathbf{G}_m^{\mathrm{H}} = \sum_{i=1}^{m}\mu_i \prod_{j=i+1}^{m}(\mathbf{I} - \mu_j(\mathbf{R} + \sigma^2\mathbf{I}))\mathbf{A}^{\mathrm{H}}. \tag{4}$$

The BER performance can then be found for BPSK modulation formats using (2). In case we use a fixed step size,

$$\mathbf{G}_m^{\mathrm{H}} = \mu\sum_{i=0}^{m-1}(\mathbf{I} - \mu(\mathbf{R} + \sigma^2\mathbf{I}))^i\mathbf{A}^{\mathrm{H}}.$$

Observe that $(\mathbf{I} - \mu_j(\mathbf{R} + \sigma^2\mathbf{I}))\mathbf{A}^{\mathrm{H}}\mathbf{A}\mathbf{A}^{\mathrm{H}} = \mathbf{R}(\mathbf{I} - \mu_j(\mathbf{R} + \sigma^2\mathbf{I}))\mathbf{A}^{\mathrm{H}}$ and so, from (4) $\mathbf{G}_m^{\mathrm{H}}\mathbf{A}\mathbf{A}^{\mathrm{H}} = \mathbf{R}\mathbf{G}_m^{\mathrm{H}}$. Substituting this expression into (3) and post-multiplying with $\mathbf{r}$ gives

$$\mathbf{y}_{m+1} = \mu_{m+1}\mathbf{A}^{\mathrm{H}}\mathbf{r} + (\mathbf{I} - \mu_{m+1}(\mathbf{R} + \sigma^2\mathbf{I}))\mathbf{y}_m. \tag{5}$$

For $\mu_{m+1} = 1$ and $\sigma^2 = 0$, Eqns. (1) and (5) are identical[1]. The MMSE detector implemented using steepest descent updates may be seen as a modified linear parallel interference canceller as illustrated in Fig. 1. When $\alpha = \sigma^2$, this structure implements the algorithm of Eqn. (5) exactly. The reason for introducing an arbitrary real-valued weighting factor α becomes clear later on.

It is possible to write Eqn. (4) (with σ^2 replaced by α) in terms of a "steady-state" solution corrupted by some disturbance,

$$\mathbf{G}_m^{\mathrm{H}} = \left(\mathbf{I} - \prod_{i=1}^{m}(\mathbf{I} - \mu_i(\mathbf{R} + \alpha\mathbf{I}))\right)(\mathbf{R} + \alpha\mathbf{I})^{-1}\mathbf{A}^{\mathrm{H}}, \tag{6}$$

where $(\mathbf{R} + \alpha\mathbf{I})^{-1}\mathbf{A}^{\mathrm{H}}$ is the "steady-state" filter. We therefore have $\mathbf{y}_m = \mathbf{y}_\infty - \mathbf{e}_m$ where $\mathbf{y}_\infty$ is the "steady-state" filter output and $\mathbf{e}_m$ is some excess transient error related to the m^{th} stage,

$$\mathbf{e}_m = \left(\prod_{i=1}^{m}(\mathbf{I} - \mu_i(\mathbf{R} + \alpha\mathbf{I}))\right)(\mathbf{R} + \alpha\mathbf{I})^{-1}\mathbf{A}^{\mathrm{H}}\mathbf{r}.$$

[1] For varying μ_m and $\sigma^2 = 0$, (5) also describes the structure in [2]. For a fixed $\mu \neq 1$ and $\sigma^2 = 0$, (5) describes the structure in [3].

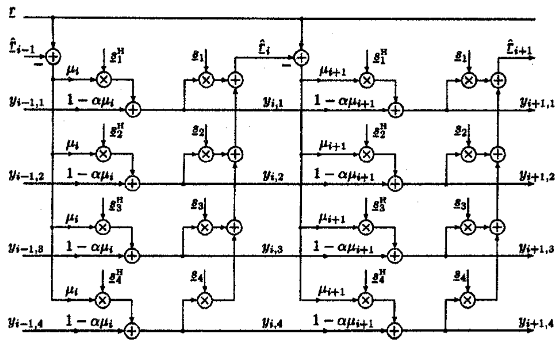

Fig. 1. The modified weighted PIC structure for a 4-user case. Stages i and $(i+1)$ are shown. When $\alpha = 0$ and $\mu_j = 1$, the structure is identical to conventional PIC.

Based on (6) it is clear that a sufficient (but not necessary) condition[2] for convergence is,

$$-1 < 1 - \mu_i(\lambda_k + \alpha) < 1 \Rightarrow 0 < \mu_i < \frac{2}{\lambda_K + \alpha},$$

where λ_K is the largest eigenvalue. The "steady-state" filter in this case would be: $\mathbf{G}_{\infty}^{\mathrm{H}} = (\mathbf{R} + \alpha\mathbf{I})^{-1}\mathbf{A}^{\mathrm{H}}$, which in case $\alpha = 0$ is the decorrelating filter and in case $\alpha = \sigma^2$ is the MMSE filter. For the conventional PIC, $\alpha = 0$ and $\mu = 1$. In this case we have convergence only if $\lambda_K < 2$ which is not true in general. This accounts for the unstable behaviour of the conventional PIC [4].

To follow the link to the SDM and the MMSE detector, we consider the MSE for the PIC structure. Using the properties of the trace operator, we can write the MSE in the m^{th} stage as

$$\begin{aligned} J(\boldsymbol{\mu}, \alpha, m) &= \sum_{k=1}^{K} \frac{\sigma^2}{\lambda_k + \sigma^2} + \sum_{k=1}^{K} \frac{\lambda_k(\lambda_k + \sigma^2)}{(\lambda_k + \alpha)^2} \\ &\quad \times \left| \frac{\sigma^2 - \alpha}{\lambda_k + \sigma^2} - \prod_{i=1}^{m}(1 - \mu_i(\lambda_k + \alpha)) \right|^2 \quad (7) \\ &= J_{\mathrm{MMSE}} + J_{\mathrm{ex}}(\boldsymbol{\mu}, \alpha, m), \end{aligned}$$

where $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, ..., \lambda_K)$ are the eigenvalues of $\mathbf{R}$ sorted in increasing order and $\boldsymbol{\mu} = (\mu_1, \mu_2, ..., \mu_m)^{\mathsf{T}}$. In (7) the first term on the RHS is the MMSE, while the second term represents the degradation with respect to the MMSE filter. For a fixed step size, we have J_{ex} expressed as

$$\begin{aligned} J_{\mathrm{ex}}(\mu, \alpha, m) &= \sum_{k=1}^{K} \frac{\lambda_k(\lambda_k + \sigma^2)}{(\lambda_k + \alpha)^2} \\ &\quad \times \left| \frac{\sigma^2 - \alpha}{\lambda_k + \sigma^2} - (1 - \mu(\lambda_k + \alpha))^m \right|^2. \end{aligned}$$

V. Optimisation of Weighting Factors

In this section we consider the choice of optimal step sizes with respect to minimising the MSE given that we only have a few stages ($m \ll \infty$). To simplify the derivations, we assume that $\alpha = \sigma^2$. It is later shown that if $\alpha = \sigma^2$, the optimal weighting factors are real as opposed to the case when $\alpha \neq \sigma^2$, where the optimal weighting factors can be complex.

[2] For a fixed $\mu_i = \mu$, it is a necessary condition.

Consider Eqn. (7) with $\alpha = \sigma^2$. We then have $J_{\mathrm{ex}}(\boldsymbol{\mu}, \alpha = \sigma^2, m) = J_{\mathrm{ex}}(\boldsymbol{\mu}, m)$ expressed as

$$J_{\mathrm{ex}}(\boldsymbol{\mu}, m) = \sum_{k=1}^{K} \frac{\lambda_k}{(\lambda_k + \sigma^2)} \prod_{i=1}^{m} |1 - \mu_i(\lambda_k + \sigma^2)|^2. \quad (8)$$

Assuming that $m \geq K$ and letting $\phi_k = \lambda_k + \sigma^2$, we can write out (8) as

$$\begin{aligned} &J_{\mathrm{ex}}(\boldsymbol{\mu}, m) = \\ &\frac{\lambda_1}{\phi_1}(|\underline{1 - \mu_1\phi_1}||1 - \mu_2\phi_1| \cdot\cdot |1 - \mu_K\phi_1| \cdot\cdot |1 - \mu_m\phi_1|)^2 \\ &+\frac{\lambda_2}{\phi_2}(|1 - \mu_1\phi_2||\underline{1 - \mu_2\phi_2}| \cdot\cdot |1 - \mu_K\phi_2| \cdot\cdot |1 - \mu_m\phi_2|)^2 \\ &\qquad \vdots \qquad\qquad \cdots \qquad\qquad \vdots \\ &+\frac{\lambda_K}{\phi_K}(|1 - \mu_1\phi_K||1 - \mu_2\phi_K| \cdot\cdot |\underline{1 - \mu_K\phi_K}| \cdot\cdot |1 - \mu_m\phi_K|)^2. \end{aligned}$$

Obviously we can now make $J_{\mathrm{ex}}(\boldsymbol{\mu}, m)$ zero by selecting the step sizes in such a way that the underlined factors above are zero. It is therefore clear that we can reach the MMSE solution if

$$\mu_i = \frac{1}{\lambda_i + \sigma^2}, \quad i = 1, 2, ...K \Rightarrow J_{\mathrm{ex}}(\boldsymbol{\mu}, m) = 0. \quad (9)$$

So the linear PIC needs exactly K stages to implement the MMSE detector.

Usually the number of stages is significantly less than K so the objective is to minimise $J_{\mathrm{ex}}(\boldsymbol{\mu}, m)$ given this constraint, i.e., we are seeking the global minimum of (8) with respect to $\boldsymbol{\mu}$.

Since (8) is a $2m$-order polynomial in $\boldsymbol{\mu}$, it has at most $(2m - 1)$ stationary points. However, individual summands can be re-written in the following way:

$$\begin{aligned} &\frac{\lambda_k}{\phi_k} \prod_{i=1}^{m} |1 - \mu_i\phi_k|^2 = \\ &\qquad \frac{\lambda_k}{\phi_k} \left|1 + \phi_k x_1 + \phi_k^2 x_2 + \cdots + \phi_k^m x_m\right|^2, \quad (10) \end{aligned}$$

where

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

This corresponds to the following mapping where $\mathbf{x} = (x_1, x_2, ..., x_m)^{\mathsf{T}}$,

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

This mapping is neither onto nor one-to-one since we have $m!$ different $\boldsymbol{\mu}$ for which $\mathbf{x} = T(\boldsymbol{\mu})$. Based on (10), we can re-write (8) as

$$\begin{aligned} J_{\mathrm{ex}}(\mathbf{x}, m) &= \sum_{k=1}^{K} \frac{\lambda_k}{\phi_k} \left|1 + \sum_{i=1}^{m} \phi_k^i x_i\right|^2 \\ &= \sum_{k=1}^{K} \frac{\lambda_k}{\phi_k} \left|1 + \boldsymbol{\varphi}_k^{\mathsf{T}}\mathbf{x}\right|^2, \end{aligned}$$

where $\varphi_k = (\phi_k, \phi_k^2, ..., \phi_k^m)^\top$. Differentiating with respect to $\mathbf{x}$ and equating to zero gives

$$\begin{aligned}\frac{dJ_{\text{ex}}(\mathbf{x},m)}{d\mathbf{x}} &= \sum_{k=1}^{K}\frac{\lambda_k}{\phi_k}\varphi_k\varphi_k^\top\mathbf{x} + \sum_{k=1}^{K}\frac{\lambda_k}{\phi_k}\varphi_k = 0\\ \Rightarrow & \sum_{k=1}^{K}\frac{\lambda_k}{\phi_k}\varphi_k\varphi_k^\top\mathbf{x} = -\sum_{k=1}^{K}\frac{\lambda_k}{\phi_k}\varphi_k\\ \Rightarrow & \ \mathbf{C}\mathbf{x} = -\mathbf{c}. \qquad (11)\end{aligned}$$

Since the eigenvalues of a Hermitian matrix are real, it is clear that if $\mathbf{C}$ is non-singular, we have a unique real minimum in $\mathbf{x}$ as $\hat{\mathbf{x}} = -\mathbf{C}^{-1}\mathbf{c}$. Clearly, $\mathbf{C}$ is positive semi-definite since $\mathbf{x}^\top\mathbf{C}\mathbf{x} = \sum_{k=1}^{K}\frac{\lambda_k}{\phi_k}(\varphi_k^\top\mathbf{x})^2 \geq 0$. Furthermore, it can be shown that $\mathbf{C}$ is positive definite if and only if m or more of the eigenvalues in Λ are distinct ($\lambda_i \neq \lambda_j$) [6]. Otherwise, $\mathbf{C}$ is singular and there are multiple solutions, any of which leads to the corresponding MMSE solution. So when we have a valid solution $\hat{\mathbf{x}}$, we find the corresponding $m!$ equivalent minima in $\boldsymbol{\mu}$, by considering the following polynomial,

$$\begin{aligned}p(\mu) &= (\mu-\mu_1)(\mu-\mu_2)\cdots(\mu-\mu_m) \qquad (12)\\ &= \mu^m + x_1\mu^{m-1} + x_2\mu^{m-2} + \cdots + x_m.\end{aligned}$$

which has exactly m roots. Substituting the solution to (11) into (12) gives us a polynomial where the m roots are the optimal step sizes, $\hat{\boldsymbol{\mu}}$. Since we have m step sizes, we then have $m!$ different choices of $\hat{\boldsymbol{\mu}}$ that all lead to the same $J_{\text{ex}}(\hat{\boldsymbol{\mu}}, K)$. The order in which the m step sizes are applied however, has a significant influence on the MSE performance at intermediate stages. Depending on the desired behaviour for intermediate stages, different criteria for the step size ordering can be adopted. In this paper we have chosen to order the optimal step sizes according to a recursive minimisation of $J_{\text{ex}}(\hat{\boldsymbol{\mu}}_i, i)$ 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 ordering is obtained by selecting at stage i the step size $\mu_i \in M_i$ which is closest to $\tilde{\mu}_i$, where M_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 $\tilde{\mu}_i = \arg\min_{\mu_i} J_{\text{ex}}(\hat{\boldsymbol{\mu}}_i, i)$. The closest step size to $\tilde{\mu}_i$ is the best choice since $J_{\text{ex}}(\hat{\boldsymbol{\mu}}_i, i)$, given that all previous $\hat{\mu}_j$ is already chosen, is a quadratic function in μ_i. It can be shown that if $\alpha = \sigma^2$, the m roots must be real and hence $\boldsymbol{\mu}$ has only real elements[6]. For $\alpha \neq \sigma^2$, both negative and complex elements can occur in $\boldsymbol{\mu}$.

So far in this section, we have assumed that $\alpha = \sigma^2$ and based the derivations on (8). The derivations presented here can also be done based on (7) for a general α. In this case, the global minimum $\hat{\mathbf{x}}(\alpha)$, which satisfies $\partial J_{\text{ex}}(\mathbf{x}, \alpha, m)/\partial x_i = 0$ for $1 \leq i \leq m$, must be a function of α. We can therefore express the minimum achievable MSE, $J_{\text{ex,min}}$, as a function of α as $J_{\text{ex,min}}(\alpha) = J_{\text{ex}}(\hat{\mathbf{x}}(\alpha), \alpha, m)$. After some manipulations we can express the derivative of $J_{\text{ex}}(\mathbf{x}, \alpha, m)$ with respect to α as

$$\frac{\partial J_{\text{ex}}(\mathbf{x},\alpha,m)}{\partial\alpha} = \sum_{j=2}^{m}(j-1)x_j\frac{\partial J_{\text{ex}}(\mathbf{x},\alpha,m)}{\partial x_{j-1}},$$

which is obviously zero at $\mathbf{x} = \hat{\mathbf{x}}(\alpha)$. So the specific α-value has no influence on the minimum achievable MSE. For any α we can find a corresponding $\boldsymbol{\mu}$ that will give us the achievable MMSE for any m-stage PIC. It can be shown that for $\alpha = \sigma^2$ it is always true that the corresponding weighting factors are real [6]. For $\alpha \neq \sigma^2$ it is however not always true. As briefly mentioned earlier, the optimal weighting factors can now be both negative and complex. This is not a serious problem since the structure in Fig. 1 can easily accommodate negative or complex weighting factors.

Regardless of α, the optimal step sizes, whether real or complex, are still dependent on σ^2. The sensitivity of the performance of the PIC to a mismatch in σ^2 is investigated through simulations in the following section.

VI. Numerical Results

The numerical examples considered in this section are based on a symbol-synchronous system with 15 equal-power users. BPSK modulation and spreading formats are assumed and only short codes with processing gain 31 are considered. The performance of the detector is illustrated as a function of the number of PIC stages at a noise level of 7 dB.

Fig. 2 shows the performance of PIC with optimised weighting factors. A randomly selected set of short codes

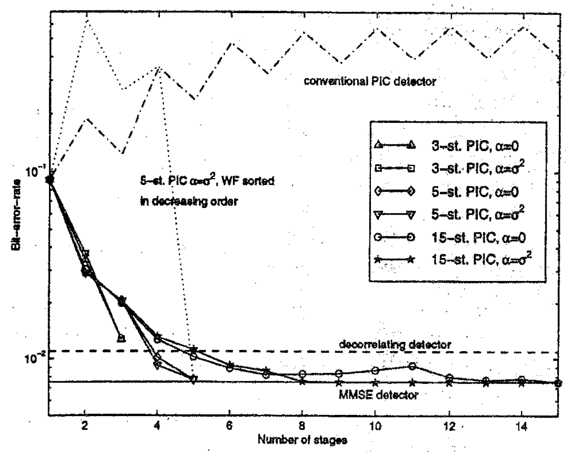

Fig. 2. Stage-by-stage performance of PIC detector with optimal weighting factors (WF) using short codes.

is used where the corresponding correlation matrix has the eigenvalues (0.15236, 0.22025, 0.26652, 0.35013, 0.57906, 0.63268, 0.77314, 0.88426, 0.93251, 1.1702, 1.4478, 1.5469, 1.6528, 1.8221, 2.5693). Since the largest eigenvalue is greater than 2, the condition for convergence [4] for the conventional PIC is violated and in Fig. 2 we observe divergence as well as the well-known ping-pong effect. For the right choice of weighting factors however, it is clear that a 15-stage PIC can achieve exactly the MMSE performance no matter what α is chosen in the structure. 5-stage PIC gives close to MMSE performance and even 3-stage PIC shows considerable improvement over the conventional detector[3]. The weighting factors are ordered according to the criterion described in Section V, forcing the MSE to decrease the most, stage-by-stage. It is observed that the BER decreases monotonically when $\alpha = \sigma^2$, but not necessarily so for $\alpha \neq \sigma^2$. When the weighting factors are sorted in decreasing order, the BER performance fluctuates around 0.50 for all but the first

[3]The performance of the conventional detector is identical to the PIC performance in the first stage.

and last stage. After the last stage, the BER is of course identical for all orderings. This is illustrated for a 5-stage PIC with $\alpha = \sigma^2$ in Fig. 2. Obviously the ordering of the weighting factors is vital for performance in the intermediate stages.

The weighting factors are determined based on a specific working SNR. The sensitivity of the detector performance at all SNRs (0-14 dB), to the choice of the working SNR is illustrated in Fig. 3 and 4. The same set of short

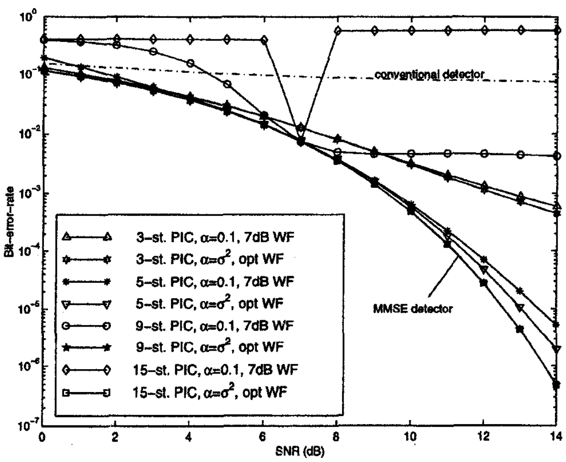

Fig. 3. BER performance and sensitivity vs. SNR using short codes. Weighting factors (WF) optimised for 7 dB and $\alpha = 0.099763$ (simplified to $\alpha = 0.1$ in the figure) are used for 0-14 dB, in comparison to when weighting factors are optimised for each working SNR and corresponding α.

codes as for Fig. 2 is used. In Fig. 3 the weighting factors optimised for SNR of 7 dB and $\alpha = 0.099763$ (which corresponds to a σ^2 at 7 dB) are used for various SNRs from 0 to 14 dB. It is compared to the case when the weighting factors are optimised and α chosen for the SNR under which the system is supposed to be working. The BER performance of the MMSE detector is also shown. Similar tests are done in Fig. 4, where α is assumed to be 0. The system is observed to be practically insensitive

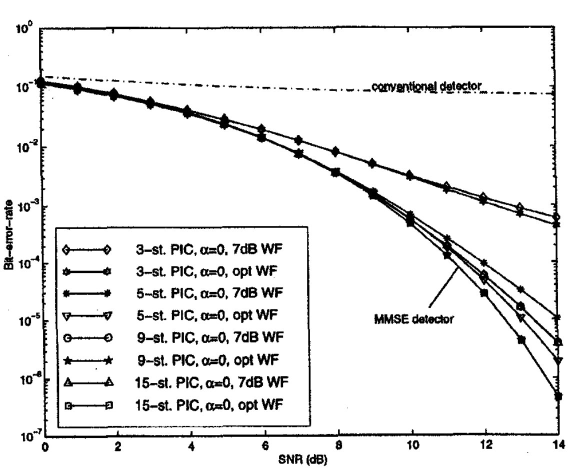

Fig. 4. BER performance and sensitivity vs. SNR using short codes. Weighting factors (WF) for 7 dB are used for 0-14 dB, in comparison to when weighting factors are optimised for each working SNR. $\alpha = 0$ is assumed in all cases.

to SNR variation when α is chosen to be 0. The set of weighting factors optimised for an SNR of 7 dB can virtually be used for any working SNR. When $\alpha = \sigma^2$ is used, the sensitivity increases substantially when the number of stages increases beyond 5. A PIC detector with more than 9 stages and 7 dB weighting factors will perform poorly for any working SNR other than 7 dB. This result implies that $\alpha = 0$ instead of $\alpha = \sigma^2$ should be used in a short-code PIC detector with a large number of stages. Using $\alpha = 0$ obviates estimation of the noise level but on the other hand increases detector complexity since the weighting factors would in general be complex.

VII. Concluding Remarks

In this paper, we have developed a matrix algebraic approach to linear PIC. It is shown that linear PIC is equivalent to a one-shot linear matrix filtering. A modified weighted PIC structure is suggested which has a performance that converges to the performance of the MMSE detector.

For an optimal choice of weighting factors, it is shown that for short codes, only K (the number of users) stages are necessary to exactly achieve the MMSE performance. For any number of stages, an analytical approach is derived for finding the optimal weighting factors that will obtain the achievable MMSE. An ordering of the weighting factors which will ensure the largest decrease in the MSE, stage by stage, is suggested and shown to provide a monotonically decreasing BER for the weighting factor $\alpha = \sigma^2$. The optimal weighting factors are dependent on the working SNR. It is demonstrated however, that for $\alpha = 0$, the detector performance is practically insensitive to a design mismatch. For $\alpha = \sigma^2$, the detector is however, quite sensitive when a large number of stages are used.

Acknowledgment

The authors would like to acknowledge the help and assistance provided by Mr. Ma Yao, Dr. P. D. Alexander, Mr. H. Sugimoto, Dr. T. Oyama, and Mr. Y. Matsumoto.

References

[1] 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.
[2] D. Divsalar, M. Simon and D. Raphaeli, "Improved parallel interference cancellation for CDMA," *IEEE Trans. Commun.*, vol. 46, pp. 258–268, Feb. 1998.
[3] T. Suzuki and Y. Takeuchi, "Real-time decorrelation scheme with multi-stage architechture for asynchronous DS/CDMA," *Technical Report of the IEICE*, SST96-10, SAT96-24, RCS96-34 (in Japanese), June 1996.
[4] L. K. Rasmussen, D. Guo, Y. Mao and T. J. Lim, "Aspects on Linear Parallel Interference Cancellation in CDMA," to be presented at the *IEEE International Symposium on Information Theory '98*, Aug. 1998.
[5] S. Haykin, *Adaptive Filter Theory*, 3rd edition, Prentice Hall, 1996.
[6] D. Guo, *Linear Parallel Interference Cancellation in CDMA*, M.Eng. Thesis, National University of Singapore, 1998.