Founded 1905

LINEAR PARALLEL INTERFERENCE CANCELLATION IN CDMA

BY

GUO DONGNING (B.S.)

DEPARTMENT OF ELECTRICAL ENGINEERING

A THESIS SUBMITTED

FOR THE DEGREE OF MASTER OF ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

1998

To Mum and Dad.

Acknowledgements

Should this thesis be proved to have any value, or I happen to have any contribution to the academic world in my future life, it should mainly be attributed to my supervisor, Dr. Lars Rasmussen, who had not only led me into the beautiful and exciting world of communications, but also re-aroused my long-forgotten childhood dream of becoming a scientist. I am the one that has benefited the most from the open research environment that Lars has been advocating with so much effort. But for his advice and encouragements I would not have accomplished this work, which I am sure to be proud of for many years.

I want to express my gratitude to Mum and Dad, who would be so pleased. All through these years I keep feeling the strong effect of the characters I have inherited from them. Their love and expectation are the most important propelling forces for the completion of this thesis.

I would like to thank Associate Professor Lye Kin Mun, Dr. Lim Teng Joon, Dr. Paul Alexander and Ms. Sun Sumei for their help as well as advice, and my colleagues, friends and fellow students for all the valuable discussions as well as happy times. Thanks to Luo Jian at M.I.T. for helping me solve a mathematical problem important to this thesis.

The support of the Centre for Wireless Communications and the Electrical Engineering Department of the National University of Singapore is highly appreciated.

Contents

List of Figures

Summary

Code-division multiple access (CDMA) has been decided the air interface technology for third-generation cellular mobile communication systems. For practical implementation interference cancellation has been subject to most attention among all CDMA multiuser detection techniques. In this thesis weighted linear parallel interference cancellation (PIC) is mathematically described and studied. It is shown to be equivalent to a linear matrix filter applied directly to the received chip-matched filtered signal vector. Expression for the exact bit error rate is obtained and conditions on the eigenvalues of the code correlation matrix and the weighting factors to ensure convergence are derived. Linear PIC is found to be a realisation of the steepest descent method for minimising the mean squared error (MSE) and a modified PIC structure is suggested which converges to the MMSE detector rather than the decorrelator. It is shown that for a K-user system only K PIC stages are required for the equivalent matrix filter to be identical to the MMSE filter. For fewer stages, techniques are devised to find a unique optimal choice of weighting factors which will lead to the minimum achievable MSE at the last stage. For long-code CDMA without power control, an algorithm is proposed for updating the set of weighting factors that will lead to the minimum achievable ensemble average of the MSE over random codes. The on-line computational complexity increases linearly with the number of users but is independent of the processing gain.

Chapter 1

Introduction

1.1 Background

The last decade has witnessed a dramatic expansion of wireless technology as a basic tool for fulfilling increasing demand for mobile communications. With the available frequency resources being saturated quickly, how to share the bandwidth efficiently among users becomes of most concern. Code-division multiple access (CDMA) is one of the methods serving such a demand well.

In CDMA, all frequency and time resources are allocated to all users simultaneously. The users are multiplexed by distinct codes rather than by orthogonal frequency bands, as in frequency-division multiple access (FDMA), or by orthogonal time slots, as in time-division multiple access (TDMA) [1]. CDMA is rooted on spread spectrum techniques [2], where each user transmits his signal using a bandwidth much larger than the data rate. This expansion in bandwidth results in frequency diversity, which is advantageous with respect to the frequency selectivity of the mobile radio channels. As another advantage of CDMA, the multiple access interference (MAI) experienced at each receiver input is generated by a larger number of signal sources than in systems not employing CDMA. This larger interferer diversity reduces the fluctuations of the total interfering power caused by both long- and short-term fading, and thus results in a much smoother channel. CDMA also has the advantages of voice activity, privacy, soft capacity limit, soft hand-off capability, and the most important — greater uplink capacity [3]. All these technological merits make CDMA the air-interface technology for next-generation cellular services [4], as well as personal communications services (PCS) and Future Public Land Mobile Telecommunications Services (FPLMTS).

Two major spreading schemes exist, namely Direct Sequence (DS) spreading and Frequency Hopping (FH). In DS-CDMA, each user is assigned a unique code sequence upon which the data sequence to be transmitted is modulated. In FH-CDMA, each user transmits his data on a narrow-band frequency slot, which changes according to a preassigned pattern, unique to each user. In this thesis, only DS-CDMA is studied since it is deemed to be more suitable for mobile communications. Hereafter CDMA is referred to as DS-CDMA unless otherwise stated.

CDMA can also be divided into short-code CDMA and long-code CDMA depending on the period of spreading codes. If the period equals a symbol interval, i.e., the spreading code remains the same symbol by symbol, it is called short-code CDMA, otherwise it is called long-code CDMA. In long-code CDMA, the spreading code is usually generated by

shift registers and has a period several orders longer than a symbol interval. Therefore the statistical properties of spreading codes resemble those of random generated chips, hence it is also called random-code CDMA.

1.1.1 Single-User Detection

In conventional single-user detection, hard decisions are made simply according to the matched filter outputs. By selecting mutually orthogonal codes for all users, they each achieve interference free single-user performance. It is however not possible to maintain orthogonality at the receiver in a mobile environment and thus multiple access interference (MAI) arises. It is well known that the MAI can degrade the bit error rate (BER) performance of a CDMA system severely. The conventional detector also suffers from a ubiquitous near-far problem in practice, which means that when the received signal energies are very dissimilar, the signal component from a weak user may be buried in the MAI from a strong user, even if the signature waveforms have very low cross-correlations [5]. Rigid power control is required to ensure that all user signals arrive at about the same power at the receiver, as is the case of IS-95, a North American cellular standard [6]. However, the near-far problem is not an inherent weak point of CDMA systems. Rather, it is the inability of the conventional single-user receiver to exploit the structure of the MAI, i.e., treating MAI as Gaussian noise is a kind of loss of information, which may in turn result in severe performance loss.

1.1.2 Multi-User Detection

More advanced detection strategies can be adopted to improve performance [7, 8]. In [9] Verdú developed the optimal complexity-unconstrained maximum-likelihood (ML) detector for coordinated multiuser CDMA. This detector performs an exhaustive search over the constrained space of all possible hypotheses. The inherent complexity however increases exponentially with the number of users, rendering the optimal ML detector impractical.

The inherent trellis structure of the optimal ML detector has motivated the development of a series of sub-optimal trellis-searching detectors (TSD), such as metric-first, depth-first and breadth-first algorithms. The breadth-first algorithms, e.g., the M-algorithm and the T-algorithm, are especially promising. In [10] a general recursive, additive maximum-likelihood metric has been developed by Rasmussen *et al.* based on pre-filtering which results in a causal MAI pattern. This provides a benchmark performance for all complexity-constrained sub-optimal TSD. Assuming a whitening filter to be used in [10] and only one path is retained in trellis-searching, the TSD reduces to Duel-Hallen's decorrelating decision feedback detector (DDFD) in [11]. The advantage of a TSD is that its performance can always be improved by retaining more paths in the M-algorithm, or by using looser threshold in the T-algorithm, till eventually it approaches the ML detector. However, due to the complexity entailed, TSD is not likely to be implemented in the next few years.

The most fundamental group of multi-user detectors is linear multi-user detectors, which apply a linear mapping to the soft output of the conventional detector to reduce the MAI seen by each user. An important type of linear detector, the decorrelating or zero-forcing detector, applies the inverse of the correlation matrix in order to decouple the data. It obviates knowledge of the received signal energies and is shown to be near-far resistant [12]. A disadvantage of this detector is that it causes noise enhancement [13]. On the other hand, the MMSE detector, which minimises the mean squared error (MSE) between detector output

and transmitted sequences, takes into account the background noise and utilises knowledge of the received signal energies, and therefore generally performs better in terms of BER. Both the decorrelating detector and the MMSE detector face the task of implementing matrix inversion, which has a complexity cubic to the number of users. A variety of adaptive strategies have been developed for approximating these detectors [14], based on algorithms such as the LMS algorithm, the RLS algorithm, the steepest descent method (SDM) and the profound as well as powerful Kalman filtering [15, 16], to name a few.

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. In one of the earliest articles on this subject, Varanasi and Aazhang proposed a parallel multi-stage structure [17]. The detector selects in each stage the most likely transmitted symbol for each user in parallel assuming that the decisions made for all the other users in the previous stage are correct. Hence it is termed parallel interference cancellation (PIC) in the literature. The linear version of this structure has been shown to be an implementation of the Jacobi iteration for calculating matrix inversion by Elders-Boll *et al.* in [18]. Other interference cancellation techniques are successive interference cancellation (SIC) [19] and hybrid interference cancellation (HIC) [20, 21]. SIC differs from PIC in that it takes a serial rather than parallel approach. HIC belongs to the family of group detectors [22], in which users are divided into groups and detection is performed in parallel within each group but serial among groups. Analysis of linear interference canceller reveals that such detectors tend to converge to the decorrelating detector. It is worth mentioning that the iterative detection in [23] happen to implement interference cancellation from the adaptive filtering point of view.

A significant improvement to the PIC scheme was suggested by Divsalar and Simon in [24] where they proposed a weighted cancellation scheme. Here the current set of decision statistics is a weighted sum of the previous set of decision statistics and the statistics resulting from interference cancellation based on current tentative decisions. They considered both linear and non-linear decision functions based on joint ML considerations. An identical approach has been suggested in [25] for a linear PIC. A different type of linear multistage detector has been proposed by Moshavi *et al.* in [26], where the outputs from all stages are weighted and combined to result in the final decision statistics. This approach, although happened to be equivalent to the weighted linear PIC, does not give much insight into the detector's behaviour in a stage-by-stage manner, and more importantly, is not comparable to the weighted PIC when non-linear decisions are made in the intermediate stages.

Simulations have shown that a linear weighted PIC may outperform the decorrelating detector and even be comparable to the MMSE detector. It is still unclear in the literature why weighted cancellation dramatically improves performance. The weighting factors are so far selected empirically, and there is no known algorithm for finding suitable weighting factors efficiently. For long-code systems, very little has been reported for multiuser detection. This thesis attempts to solve these problems and the results are first summarised in the next section before detailed discussions in the following chapters.

1.2 Contributions

This thesis studies in detail weighted linear parallel interference cancellation based on a symbol synchronous CDMA channel corrupted by additive white Gaussian noise (AWGN).

The contributions can be summarised as follows.

1. It is shown that linear PIC schemes, either conventional or weighted, correspond to linear matrix filtering that can be performed directly on the received chip-matched filtered signal vector. An expression for the exact bit error rate is obtained. Requirements on the eigenvalues of the code correlation matrix and the weighting factors to ensure convergence are derived. This was first presented in [27].

2. It is demonstrated that the weighted linear PIC essentially realises the steepest descent method (SDM) for updating adaptive filters to minimise the mean squared error. A new PIC structure that converges to the MMSE detector rather than the decorrelating detector which other PIC structures generally converge to is suggested. These results are presented in [28] as well as submitted paper [29].

3. A solution to the unique optimal set of weighting factors that leads to the minimum achievable MSE given a limited number of PIC stages is derived for short-code CDMA. This is also included in [28, 29].

4. It is demonstrated that for systems employing long codes, a set of fixed weighting factors will lead to the minimum achievable ensemble average of the MSE over random codes. An algorithm with the complexity linear to the number of users but independent of the processing gain is proposed for computing and on-line updating the optimal set of parameters with no power control assumption [30].

1.3 Notation

Throughout this thesis scalars are lower-case, random variables are upper-case, vectors are bold font lower-case, and matrices are bold font upper-case, unless otherwise stated. Subscripting is dropped where no ambiguities arise. The symbols $(\cdot)^*$, $(\cdot)^\top$, $(\cdot)^{\mathrm{H}}$, $(\cdot)^{-1}$, $(\cdot)^+$, $|\cdot|$, $\|\cdot\|$, $\mathrm{tr}\{\cdot\}$, $\mathrm{Re}\{\cdot\}$, $\mathrm{Im}\{\cdot\}$, $\mathrm{E}\{\cdot\}$ and $\det\{\cdot\}$ are the complex conjugate, transposition, complex conjugate transposition, inversion, pseudo inversion, absolute value, Euclidean vector-norm, trace, real part, imaginary part, expectation and determinant operators respectively, and the delimiter $\{\cdot\}^y$ defines a space of dimension y. All vectors are defined as column vectors with row vectors represented by transposition. $\mathbb{Z}$ denotes the set of integers, $\mathbb{R}$ the set of real numbers, $\mathbb{C}$ the set of complex numbers, and the following notation is used for the product of square matrices,

$$\prod_{i=n_1}^{n_2} \mathbf{X}_i = \begin{cases} \mathbf{X}_{n_2}\mathbf{X}_{n_2-1}\cdots\mathbf{X}_{n_1+1}\mathbf{X}_{n_1} & \text{if } n_1 \leq n_2, \\ \mathbf{I} & \text{if } n_1 > n_2. \end{cases} \tag{1.1}$$

Subscripting is done according to the following conventions. Variables independent of the detector stage are, when needed, subscripted with a user index, e.g., x_k. The first subscript on variables dependent on the detector stage, e.g., $x_{i,k}$, denotes the current stage, the second subscript the user index.

1.4 Outline

The thesis is organised as follows. Chapter 2 serves as an introduction to CDMA as well as a mathematical framework for the rest of the thesis. The algebraic description of the

conventional PIC scheme is presented in Chapter 3 together with discussion on convergence issues. Weighted linear PIC is studied in Chapter 4, where the close connection of PIC to the SDM is explained, the optimisation of the step sizes is described and powerful techniques for obtaining these parameters devised. The application of the weighted PIC to long codes as well as an efficient algorithm for computing the optimal weighting factors is addressed in Chapter 5. The thesis is then completed by some concluding remarks in Chapter 6.

Chapter 2

Introduction to CDMA

In this chapter we will discuss the background of CDMA schemes. The signal and channel models for a multiuser CDMA communication system are described in section 2.1. The discrete time model based on a symbol synchronous system studied throughout this thesis is then introduced in section 2.2. Discussions on eigenvalue decomposition of the correlation matrix follows in section 2.3. Finally we briefly describe a few fundamental but important receivers in section 2.4.

2.1 CDMA Signal and Channel Models

A low-pass equivalent model for a K-user CDMA system is depicted in Fig. 2.1. Stationary single-path channels and the presence of stationary additive white Gaussian noise are assumed. Each user transmits every information symbol by modulating on his unique signature waveform. Consider a block of $(M+1)$ symbols, $d_k(0), d_k(1), \cdots, d_k(M)$, transmitted by user k. The transmitted signal is described by

$$s_k(t) = \sum_{i=0}^{M} d_k(i) g_k^{(i)}(t - iT) \tag{2.1}$$

where $g_k^{(i)}(t)$ is user k's signature waveform of duration T at the i^{th} symbol interval. $g_k^{(i)}(t)$ may be expressed as

$$g_k^{(i)}(t) = \sum_{n=1}^{N} s_k^{(i)}(n) p\left(t - (n-1)T_c\right), \quad 0 \le t \le T \tag{2.2}$$

where $\left\{ s_k^{(i)}(n), n = 1, 2, \cdots, N \right\}$ is a pseudo-noise (PN) code sequence consisting of N chips and $p(t)$ is the chip waveform of duration $T_c = T/N$, the chip interval. Assume that $p(t)$ has unit energy, and is 0 outside $[0, T_c]$, hence

$$\int_{-\infty}^{\infty} |p(t)|^2 dt = \int_{0}^{T_c} |p(t)|^2 dt = 1. \tag{2.3}$$

In a short-code system, $g_k^{(i)}(t) = g_k(t)$ for $i = 0, 1, \cdots, M$, i.e., the spreading waveform remains the same for every symbol interval, otherwise it is a long-code system.

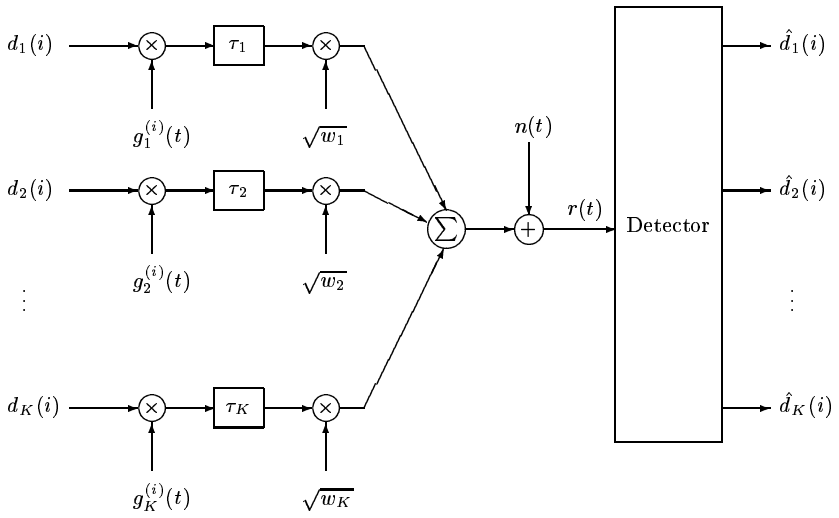

Figure 2.1: Low-pass equivalent model for a CDMA system.

For simplicity, both the symbols and the chips are assumed to adopt Phase Shift Keying (PSK) modulation, i.e.,

$$d_k(i) \in \mathbb{D} = \{\exp(j\pi(2m-1)/p) \mid m = 1, 2, \cdots, p\}, \tag{2.4}$$

and

$$s_k^{(i)}(n) \in \mathbb{S} = \left\{ \frac{1}{\sqrt{N}} \exp(j\pi(2m-1)/q) \mid m = 1, 2, \cdots, q \right\}. \tag{2.5}$$

Note that above definition in (2.5) implies that all signature waveforms also have unit energy, i.e.,

$$\int_0^T \left| g_k^{(i)}(t) \right|^2 dt = \sum_{n=1}^{N} \left| s_k^{(i)}(n) \right|^2 \int_0^T |p(t-(n-1)T_c)|^2 dt = 1. \tag{2.6}$$

The composite transmitted signal for the K users is expressed as

$$s(t) = \sum_{k=1}^{K} \sqrt{w_k} s_k(t-\tau_k) = \sum_{k=1}^{K} \sqrt{w_k} \sum_{i=0}^{M} d_k(i) g_k^{(i)}(t - iT - \tau_k) \tag{2.7}$$

where $\{w_k\}$ are the received signal energies per symbol and $\{\tau_k \in [0, T)\}$ are the transmission delays.

In a multiuser CDMA system, three levels of synchronism exist. A system where the spreading waveforms from all users arrive at the receiver synchronously is termed symbol synchronous. In this case, $\tau_k = 0$ for $k = 1, 2, \cdots, K$. Conversely, if there is no timing control, the system is said to be chip asynchronous. The level in between them is chip synchronous where although the chip boundaries of all users are aligned, the symbol boundaries are not. In a chip synchronous system τ_k is a multiple of T_c.

The transmitted signal is assumed to be corrupted by AWGN. Hence the received signal may be expressed as

$$r(t) = s(t) + n(t) \tag{2.8}$$

where $n(t)$ is the noise with a power spectral density of $\sigma^2 = N_0$.

2.2 Discrete Time Model for a Symbol Synchronous System

For simplicity, a symbol synchronous system is considered throughout this thesis. Treating τ_k as 0, the received signal is simplified as

$$\begin{aligned} r(t) &= \sum_{k=1}^{K} \sqrt{w_k} \sum_{j=0}^{M} d_k(j) g_k^{(j)}(t - jT) + n(t) \\ &= \sum_{k=1}^{K} \sqrt{w_k} \sum_{j=0}^{M} d_k(j) \sum_{l=1}^{N} s_k^{(j)}(l) p(t - jT - (l-1)T_c) + n(t). \end{aligned} \tag{2.9}$$

It is clear that the signal space is spanned by a set of orthonormal signal basis functions

$$\left\{ f_l^{(j)}(t) = p(t - jT - (l-1)T_c) \;\middle|\; j = 0, 1, \cdots, M,\ l = 1, 2, \cdots, N \right\}, \tag{2.10}$$

which consists of time-shifted versions of the chip waveform. The received signal is seen to be a linear combination of the signal basis functions. Its equivalent vector representation is then obtained through matched filtering [31, Sec. 4-2],

$$\begin{aligned} r_n^{(i)} &= \int_0^T r(t) p^*(t - iT - (n-1)T_c) dt \\ &= \sum_{k=1}^{K} \sqrt{w_k} \sum_{j=0}^{M} d_k(j) \sum_{l=1}^{N} s_k^{(j)}(l) \int_0^T p(t - jT - (l-1)T_c) p^*(t - iT - (n-1)T_c) dt \\ &\quad + \int_0^T n(t) p^*(t - iT - (n-1)T_c) dt \\ &= \sum_{k=1}^{K} \sqrt{w_k} \sum_{j=0}^{M} d_k(j) \sum_{l=1}^{N} s_k^{(j)}(l) \delta(i-j) \delta(n-l) + z_n^{(i)} \\ &= \sum_{k=1}^{K} \sqrt{w_k} d_k(i) s_k^{(i)}(n) + z_n^{(i)} \end{aligned} \tag{2.11}$$

where $z_n^{(i)}$ is a zero-mean independent Gaussian random variable with variance of σ^2. Note that the chip-matched filter output in the i^{th} symbol interval, $r_n^{(i)}$, depends only on the data and spreading codes of this particular symbol interval. In other words, the channel is memoryless in symbol level. Therefore it is sufficient to consider an arbitrary symbol interval so that we can drop index i in our notation, i.e.,

$$r_n = \sum_{k=1}^{K} \sqrt{w_k} d_k s_k(n) + z_n. \tag{2.12}$$

It is possible to rewrite (2.12) in vector representation as

$$\mathbf{r} = \sum_{k=1}^{K} \sqrt{w_k} d_k \mathbf{s}_k + \mathbf{n} \tag{2.13}$$

where $\mathbf{r} = (r_1, r_2, \cdots, r_N)^\top$ is the chip-matched filtered received signal vector corresponding to a full symbol interval, $\mathbf{s}_k = (s_k(1),\ s_k(2),\ \cdots, s_k(N))^\top$ is user k's spreading code vector which satisfies $\mathbf{s}_k^{\mathrm{H}}\mathbf{s}_k = 1$, and $\mathbf{n} = (z_1, z_2,\ \cdots,\ z_N)^\top$ is a noise vector where each sample is independently, circularly complex Gaussian distributed with zero mean and variance σ^2. The output of a chip-matched filter is then seen to be a weighted linear combination of spreading codes.

The received vector $\mathbf{r}$ can also be described through matrix algebra as

$$\mathbf{r} = \mathbf{A}\mathbf{d} + \mathbf{n} \tag{2.14}$$

where $\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_K) = \left(\sqrt{w_1}\mathbf{s}_1, \sqrt{w_2}\mathbf{s}_2, \cdots, \sqrt{w_K}\mathbf{s}_K\right)$ is an $N \times K$ matrix and $\mathbf{d} = (d_1, d_2, \cdots, d_K)^\top$ is the data vector. The matrix $\mathbf{A}$ contains all information about the multiple access channel and may be expressed as

$$\mathbf{A} = \mathbf{S}\mathbf{W} \tag{2.15}$$

where $\mathbf{S} = (\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_K)$ is a $N \times K$ matrix formed by all users' spreading code vectors and $\mathbf{W} = \mathrm{diag}(\sqrt{w_1}, \sqrt{w_2}, \cdots, \sqrt{w_K})$ is a $K \times K$ diagonal matrix formed by the received signal amplitudes.

2.3 Correlation Matrix and Its Eigenvalue Decomposition

In this section we discuss some useful properties of the CDMA channel correlation matrix. We concentrate on the eigenvalue analysis which will be utilised intensively to derive many results throughout this thesis.

In the literature the cross-correlation of 2 users' spreading codes is usually defined as

$$\rho_{ij} = \mathbf{s}_i^{\mathrm{H}}\mathbf{s}_j = \sum_{n=1}^{N} s_i^*(n) s_j(n). \tag{2.16}$$

Here we define the cross-correlation of the received spreading waveforms as

$$r_{ij} = \mathbf{a}_i^{\mathrm{H}}\mathbf{a}_j = \sqrt{w_i w_j}\mathbf{s}_i^{\mathrm{H}}\mathbf{s}_j = \sqrt{w_i w_j}\rho_{ij}. \tag{2.17}$$

Clearly received signal energies are considered in r_{ij}. These cross-correlations then form the *correlation matrix* as

$$\mathbf{R} = \mathbf{A}^{\mathrm{H}}\mathbf{A} = \mathbf{W}\mathbf{S}^{\mathrm{H}}\mathbf{S}\mathbf{W}, \tag{2.18}$$

which has the element in its i^{th} row and j^{th} column as r_{ij}. Here we call $\mathbf{S}^{\mathrm{H}}\mathbf{S}$ the *code correlation matrix*, to make a difference[1].

[1]It is more often in the literature to define $\mathbf{S}^{\mathrm{H}}\mathbf{S}$ as the *correlation matrix*.

Since $\mathbf{R}$ is a Hermitian matrix, it must have an eigenvalue decomposition [32] as

$$\mathbf{R} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathrm{H}} \tag{2.19}$$

where $\mathbf{U}$ is a unitary matrix satisfying $\mathbf{U}\mathbf{U}^{\mathrm{H}} = \mathbf{I}$ that contains the eigenvectors of $\mathbf{R}$ while $\mathbf{\Lambda}$ is a diagonal matrix formed by the K eigenvalues of $\mathbf{R}$,

$$\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_K). \tag{2.20}$$

Since a Hermitian matrix is positive semi-definite, all the eigenvalues of $\mathbf{R}$ are non-negative. If $\mathbf{A}$ is full-rank, which corresponds to that the spreading codes from all K users are linearly independent, $\mathbf{R}$ has only positive eigenvalues and is said to be positive definite.

As a Hermitian matrix, $\mathbf{R}$ has the following important properties:

1. If $\mathbf{R}$ is positive definite, $\mathbf{R}^{-1}$ exists and shares the same eigenvectors as $\mathbf{R}$, i.e.,

$$\mathbf{R}^{-1} = \mathbf{U}\mathbf{\Lambda}^{-1}\mathbf{U}^{\mathrm{H}}; \tag{2.21}$$

2. For any polynomial $f(\cdot)$, $f(\mathbf{R})$ shares the same eigenvectors as $\mathbf{R}$, and its eigenvalues are $f(\lambda_1), f(\lambda_2), \cdots, f(\lambda_K)$, i.e.,

$$f(\mathbf{R}) = \mathbf{U}\mathrm{diag}(f(\lambda_1), f(\lambda_2), \cdots, f(\lambda_K))\mathbf{U}^{\mathrm{H}}; \tag{2.22}$$

3. The trace of $\mathbf{R}$ is the sum of its eigenvalues, i.e.,

$$\mathrm{tr}\{\mathbf{R}\} = \sum_{k=1}^{K} \lambda_k. \tag{2.23}$$

Clearly for any polynomial $f(\cdot)$,

$$\mathrm{tr}\{f(\mathbf{R})\} = \sum_{k=1}^{K} f(\lambda_k). \tag{2.24}$$

The above properties will be used frequently in deriving the main results in this thesis.

2.4 Fundamental Detectors

In this section we briefly discuss several fundamental detectors frequently referred to throughout this thesis.

2.4.1 Conventional Single-user Detector

The conventional single-user detector simply performs spreading waveform matched filtering on the received signal vector. This kind of linear filtering is efficiently expressed by an inner product in vector representation. The matched filtered statistic for user k is obtained as

$$\begin{aligned} y_k &= \mathbf{a}_k^{\mathrm{H}}\mathbf{r} = \mathbf{a}_k^{\mathrm{H}}\mathbf{A}\mathbf{d} + \mathbf{a}_k^{\mathrm{H}}\mathbf{n} \\ &= w_k d_k + \sum_{\substack{i=1 \\ i\neq k}}^{K} \sqrt{w_i w_k}\rho_{ik} d_i + z_k \\ &= w_k d_k + \mathrm{MAI}_k + z_k \end{aligned} \tag{2.25}$$

where MAI_k is the multiple access interference seen by user k and z_k represents white Gaussian noise with variance $\mathrm{E}\{|z_k|^2\} = w_k\sigma^2$. The soft output of all K users may be expressed as a vector $\mathbf{y} = \mathbf{A}^{\mathrm{H}}\mathbf{r}$. It is more often in the literature to define the conventional detector as code-matched filtering, i.e., $\mathbf{y}^{(c)} = \mathbf{S}^{\mathrm{H}}\mathbf{r}$. Clearly, $\mathbf{y} = \mathbf{W}\mathbf{y}^{(c)}$. Since $\mathbf{W}$ is assumed to be known, both definitions lead to the same decision and hence are equivalent. In this thesis we adopt the former definition to facilitate later discussions.

The decision for user k is made according to the position of y_k in the data constellation as if the MAI were Gaussian distributed. It is not difficult to obtain an expression for the exact bit error rate. Specifically for a BPSK symbol modulation,

$$P_k(\sigma) = \frac{1}{2^{K-1}} \sum_{\substack{d_i \in \{-1,1\} \\ i \neq k}} Q\left(\frac{\sqrt{w_k} + \sum_{\substack{i=1 \\ i \neq k}}^{K} \sqrt{w_i}\rho_{ik} d_i}{\sigma} \right). \tag{2.26}$$

Clearly, if the signature sequences are orthogonal ($\rho_{ik} = 0$, if $i \neq k$), the MAI vanishes so that each user enjoys a single-user channel and the conventional detector is optimal. In a cellular environment, however, this is not the case. One reason is that most channels contain some degree of asynchronism and it is theoretically not possible to design codes that are orthogonal over all possible delays. Another reason is that signature waveforms are subject to random distortion because of multi-path and frequency selective fading and can not always be orthogonal at the receiver.

On the other hand, since orthogonality is not perfectly maintained, the MAI can become excessive if the power levels of the signals from interfering users are sufficiently larger than that of the desired user. As mentioned in Chapter 1, this situation is generally called the near-far problem in multiuser communications.

In [33, Sec. 3.5] Verdú defined the asymptotic multiuser efficiency as

$$\begin{aligned} \eta_k &= \sup\left\{ 0 \leq r \leq 1 \ : \ \lim_{\sigma \to 0} P_k(\sigma) / Q\left(\frac{\sqrt{r w_k}}{\sigma} \right) \right\} \\ &= \frac{2}{w_k} \lim_{\sigma \to 0} \sigma^2 \log \frac{1}{P_k(\sigma)}. \end{aligned} \tag{2.27}$$

It measures the slope with which $P_k(\sigma)$ vanishes (in logarithmic scale) in the high signal-to-noise ratio region. The robustness against the near-far problem achieved by multiuser detectors can then be measured by the multiuser efficiency minimised over the received energies of all the other users:

$$\overline{\eta}_k = \inf_{\substack{w_j > 0 \\ j \neq k}} \eta_k. \tag{2.28}$$

It is commonly termed near-far resistance. A detector is near-far resistant if $\overline{\eta}_k > 0$. It is well known that the conventional detector is not near-far resistant. Hence it is usually necessary to have some type of power control for conventional single-user detection.

2.4.2 Maximum-Likelihood Detector

The maximum-likelihood (ML) detector selects the most probable transmitted data sequence given the received signal observed over the entire time interval of transmission [9, 31]. Consider a symbol synchronous system for simplicity, the detector performs an exhaustive search over

the constrained space $\mathbb{D}^K$ of all possible hypotheses and chooses $\mathbf{d}$ which maximises the probability that $\mathbf{d}$ was transmitted given that $\mathbf{r}$ is received, i.e.,

$$\hat{\mathbf{d}} = \arg \max_{\mathbf{d} \in \mathbb{D}^K} p(\mathbf{d}|\mathbf{r}). \tag{2.29}$$

It can be easily shown that the soft output of the conventional detector, $\mathbf{y} = \mathbf{A}^{\mathrm{H}}\mathbf{r}$, is a sufficient decision statistics since the ML metric can be computed based on $\mathbf{y}$ as well as on $\mathbf{r}$. The ML detector is near-far resistant. Unfortunately it has an inherent complexity that increases exponentially with the number of users which renders it impractical. Various suboptimal multiuser detectors have been proposed to ease the implementation. Among them linear detectors are subject to most attention because of their simplicity.

2.4.3 Decorrelating Detector

The decorrelating detector (or decorrelator) applies the inverse of the correlation matrix onto the soft output of the conventional detector in order to decouple the data,

$$\mathbf{y} = \mathbf{G}_{\mathrm{dec}}^{\mathrm{H}}\mathbf{r} = \mathbf{R}^{-1}\mathbf{A}^{\mathrm{H}}\mathbf{r} = \mathbf{R}^{-1}\mathbf{A}^{\mathrm{H}}(\mathbf{A}\mathbf{d} + \mathbf{n}) = \mathbf{d} + \mathbf{z}, \tag{2.30}$$

where $\mathrm{E}\{\mathbf{z}\mathbf{z}^{\mathrm{H}}\} = \sigma^2\mathbf{R}^{-1}$. Since the decorrelator completely eliminates the MAI, it has a near-far resistance of 1. A disadvantage of this detector is that it causes noise enhancement, i.e., $(\mathbf{R}^{-1})_{kk}$ is always greater than or equal to 1 [13].

In case that linearly dependent spreading codes are used, the inverse of $\mathbf{R}$ should be replaced by any of its pseudo inverse [12, 13]. A user can be decorrelated if and only if his spreading code is independent of all the other users'.

2.4.4 MMSE Detector

The mean squared error of a linear detector $\mathbf{G}^{\mathrm{H}}$ is defined as

$$J = \mathrm{E}\left\{\|\mathbf{G}^{\mathrm{H}}\mathbf{r} - \mathbf{d}\|^2\right\}, \tag{2.31}$$

which indicates the averaged squared Euclidean distance of the soft output of the detector to the original data symbols transmitted. The MMSE detector minimises the MSE and is found to be

$$\mathbf{G}_{\mathrm{MMSE}}^{\mathrm{H}} = \left(\mathbf{R} + \sigma^2\mathbf{I}\right)^{-1}\mathbf{A}^{\mathrm{H}}. \tag{2.32}$$

Since the MMSE detector takes into account the background noise and utilises knowledge of the received signal powers, it generally performs better than the decorrelator in terms of BER.

Note that the MSE is not an equivalent index to the BER. Therefore the MMSE detector is not necessarily the optimal linear detector in the BER sense. However, in [34] Poor and Verdú showed that the MAI-plus-noise at the MMSE detector output contending with the demodulation of a desired user is approximately Gaussian in most cases of interest. As a result, the BER performance of the MMSE detector has no noticeable difference to that of the best achievable by a linear detector.

Moreover, the signal-to-interference ratio of the soft output for user k of any linear detector can be expressed as

$$\mathrm{SIR}_k = \mathrm{MSE}_k{}^{-1} - 1. \tag{2.33}$$

Since the multiuser interference in the soft output is approximately Gaussian (at least when the detector is close to the MMSE detector), we can expect that minimising the MSE is very close to minimising the BER. Hence the MSE is always treated as a benchmark for optimising linear detectors.

It is not practical to implement the decorrelating detector and the MMSE detector directly because of the heavy computation load of matrix inversion. In the following chapters we are going to study linear parallel interference cancellation techniques which approximate the above detectors in the MSE sense through multi-stage processing.

2.5 Summary

In this chapter we discuss the background of CDMA schemes. The signal and channel models of the CDMA communication system are first described. A mathematical framework for the derivation of proposed detectors in the following chapters has been introduced.

Chapter 3

Conventional Parallel Interference Cancellation

In this chapter, the conventional parallel interference canceller is analysed based on matrix algebra. An equivalent one-shot detector is obtained in section 3.1. It is then shown in section 3.2 that the linear conventional PIC converges if and only if all eigenvalues of the code correlation matrix are less than 2, which is not always true. Section 3.3 contains numerical results and section 3.4 summarises this chapter.

3.1 Algebraic Description of Conventional PIC

One of the first structure based on the principle of interference cancellation was the parallel multi-stage detector in [17], termed conventional PIC throughout this thesis. The general structure of a multi-stage conventional PIC detector is described by the diagram in Fig. 3.1, where each stage is depicted by Fig. 3.2.

The chip-matched filtered received signal vector $\mathbf{r}$ is fed into the detector, where

$$\mathbf{r} = \mathbf{SWd} + \mathbf{n}. \tag{3.1}$$

Interference cancellation is carried out stage-by-stage based on tentative decision outputs of the previous stage and decisions are made on the resulting decision statistics. At the first stage, previous tentative decisions are assumed to be 0. Note that thick lines in both figures represent a vector of length N, the processing gain.

Consider the i^{th}-stage PIC shown in Fig. 3.2. Assuming decisions (either soft or hard) for all users are correct for stage $(i-1)$, the detector first reconstructs the interference to user k from all the interfering users based on their decisions and then cancels it out from the received signal. The residue signal for user k may be expressed as

$$\mathbf{r}_{i-1}^{(k)} = \mathbf{r} - \sum_{\substack{j=1 \\ j\neq k}}^{K} \mathbf{s}_j \hat{d}_{i-1,j} = \mathbf{r} - \sum_{j=1}^{K} \mathbf{s}_j \hat{d}_{i-1,j} + \mathbf{s}_k \hat{d}_{i-1,k}, \tag{3.2}$$

where $\hat{d}_{i,k}$ is a tentative decision made for user k at stage i based on the soft decision statistic, i.e.,

$$\hat{d}_{i,k} = \text{TDD}(y_{i,k}) \tag{3.3}$$

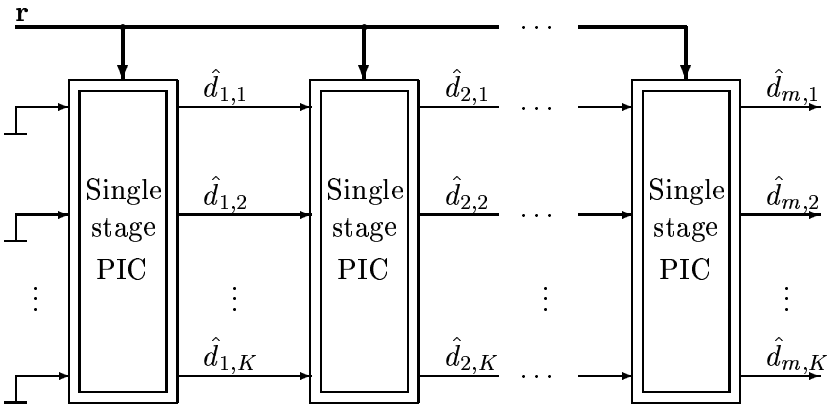

Figure 3.1: A general structure for a K-user, m-stage conventional PIC.

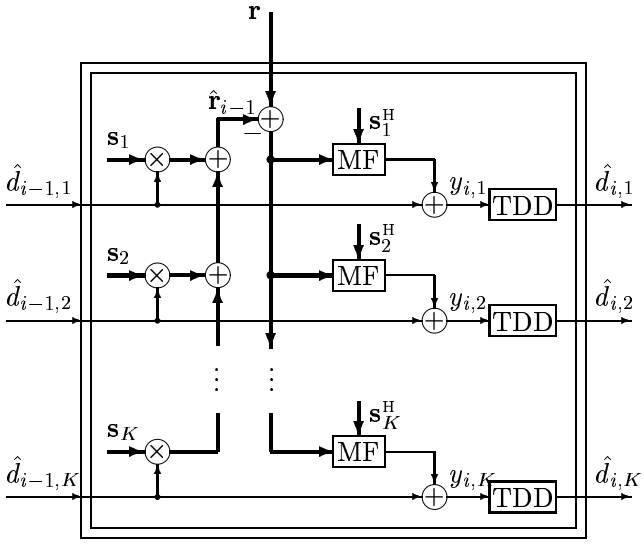

Figure 3.2: Single stage conventional PIC. MF — matched filtering. TDD — Tentative Decision Device.

where TDD denotes a tentative decision device. The residue signal is assumed to be interference free so that code-matched filtering can be performed to yield a current stage statistic for user k, which is expressed as

$$y_{i,k} = \mathbf{s}_k^{\mathrm{H}} \cdot \mathbf{r}_{i-1}^{(k)} = \mathbf{s}_k^{\mathrm{H}} \left(\mathbf{r} - \sum_{j=1}^{K} \mathbf{s}_j \hat{d}_{i-1,j} + \mathbf{s}_k \hat{d}_{i-1,k} \right) = \mathbf{s}_k^{\mathrm{H}} \left(\mathbf{r} - \sum_{j=1}^{K} \mathbf{s}_j \hat{d}_{i-1,j} \right) + \hat{d}_{i-1,k}. \quad (3.4)$$

Many types of decision functions can be used in the TDDs in the structure, such as hard decision function used in [17], the clip function, and hyperbolic tangent function, to name a few. The choice of TDD function has remarkable influence on the performance of a PIC detector. Since the hyperbolic tangent decision function is rooted on the maximum likelihood considerations [24, 35], it in general performs better than all the others. Throughout this thesis, however, we assume that a linear decision function is used in all intermediate stages, since a PIC making use of non-linear decision functions is very hard to analyse mathematically. Nevertheless, the study may give insight into the principles of interference cancellation techniques.

Assuming that in the above PIC structure linear soft decisions are made, i.e., $\hat{d}_{i,k} = y_{i,k}$, we have

$$y_{i,k} = \mathbf{s}_k^{\mathrm{H}} \left(\mathbf{r} - \sum_{j=1}^{K} \mathbf{s}_j y_{i-1,j} \right) + y_{i-1,k}. \quad (3.5)$$

Also define the vector of decision statistics as $\mathbf{y}_i = (y_{i,1}, y_{i,2}, ..., y_{i,K})^{\top}$, the set of decision statistics at stage $(i+1)$ is then described by

$$\begin{aligned} \mathbf{y}_i &= \mathbf{S}^{\mathrm{H}}(\mathbf{r} - \mathbf{S}\mathbf{y}_{i-1}) + \mathbf{y}_{i-1} \\ &= \mathbf{S}^{\mathrm{H}}\mathbf{r} - \mathbf{S}^{\mathrm{H}}\mathbf{S}\mathbf{y}_{i-1} + \mathbf{y}_{i-1} \\ &= \mathbf{S}^{\mathrm{H}}\mathbf{r} + (\mathbf{I} - \mathbf{S}^{\mathrm{H}}\mathbf{S})\mathbf{y}_{i-1}, \end{aligned} \quad (3.6)$$

where $\mathbf{y}_0 = \mathbf{0}$. From this recursion, we can express the soft output at the m^{th} stage as

$$\begin{aligned} \mathbf{y}_m &= \mathbf{S}^{\mathrm{H}}\mathbf{r} + (\mathbf{I} - \mathbf{S}^{\mathrm{H}}\mathbf{S})\mathbf{y}_{m-1} & (3.7) \\ &= \mathbf{S}^{\mathrm{H}}\mathbf{r} + (\mathbf{I} - \mathbf{S}^{\mathrm{H}}\mathbf{S})\mathbf{S}^{\mathrm{H}}\mathbf{r} + (\mathbf{I} - \mathbf{S}^{\mathrm{H}}\mathbf{S})^2\mathbf{y}_{m-2} & (3.8) \\ &= \cdots \quad \cdots & (3.9) \\ &= \sum_{i=1}^{m} (\mathbf{I} - \mathbf{S}^{\mathrm{H}}\mathbf{S})^{(i-1)}\mathbf{S}^{\mathrm{H}}\mathbf{r}. & (3.10) \end{aligned}$$

It is clear that the linear PIC scheme corresponds to linear matrix filtering that can be performed directly on the received chip-matched filtered signal vector, since the vector of decision statistics at the m^{th} stage can be obtained as

$$\mathbf{y}_m = \mathbf{G}_m^{\mathrm{H}}\mathbf{r} \quad (3.11)$$

where

$$\mathbf{G}_m^{\mathrm{H}} = \sum_{i=1}^{m} (\mathbf{I} - \mathbf{S}^{\mathrm{H}}\mathbf{S})^{(i-1)}\mathbf{S}^{\mathrm{H}}. \quad (3.12)$$

The matrix filter $\mathbf{G}_m^{\mathrm{H}}$ is then referred to as the equivalent one-shot cancellation filter[1] for an m-stage conventional linear PIC. Note also that

$$\mathbf{G}_1^{\mathrm{H}} = \mathbf{S}^{\mathrm{H}}, \tag{3.13}$$

therefore a 1-stage PIC corresponds to the conventional detector.

Since $\mathbf{G}_m$ is a linear filter, the noise term in $\mathbf{y}_m$ is still Gaussian but with a coloured correlation matrix

$$\mathrm{E}\left\{\mathbf{G}_m^{\mathrm{H}}\mathbf{n}\mathbf{n}^{\mathrm{H}}\mathbf{G}_m\right\} = \sigma^2\mathbf{G}_m^{\mathrm{H}}\mathbf{G}_m. \tag{3.14}$$

Given the spreading codes used, we can therefore analytically calculate the BER for user k at stage m using the same techniques as for the conventional matched filter detector [12] and the linear successive interference canceller [10]. Specifically for BPSK data modulation systems,

$$P_k^{(m)}(\sigma) = \frac{1}{2^{K-1}} \sum_{\substack{\mathbf{d}\in\{-1,1\}^K \\ d_k=1}} Q\left(\frac{\mathrm{Re}\{\mathbf{g}_{m,k}^{\mathrm{H}}\mathbf{A}\mathbf{d}\}}{\sigma\|\mathbf{g}_{m,k}\|}\right), \tag{3.15}$$

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

3.2 Convergence Issues

It is instructive to see how the linear conventional PIC behaves as the number of stages goes to infinity. Rewrite the expression for the equivalent filter here,

$$\mathbf{G}_m^{\mathrm{H}} = \sum_{i=1}^{m}(\mathbf{I} - \mathbf{S}^{\mathrm{H}}\mathbf{S})^{(i-1)}\mathbf{S}^{\mathrm{H}}. \tag{3.16}$$

Without losing generality, we assume that $N \geq K$. There exists a single value decomposition of $\mathbf{S}_{N\times K}$ as

$$\mathbf{S} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathrm{H}} \tag{3.17}$$

where $\mathbf{U}_{N\times N}$ and $\mathbf{V}_{K\times K}$ are unitary matrices and $\mathbf{\Sigma}$ is a $N \times K$ matrix of the form

$$\mathbf{\Sigma} = \begin{bmatrix} \sqrt{\mathbf{\Lambda}} \\ \mathbf{0} \end{bmatrix} \tag{3.18}$$

where $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_K)$ is a $K \times K$ diagonal matrix and the zero matrix is of appropriate dimensions. Since

$$\begin{aligned} \mathbf{S}^{\mathrm{H}}\mathbf{S} &= \mathbf{V}\mathbf{\Sigma}^{\mathrm{H}}\mathbf{U}^{\mathrm{H}}\mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathrm{H}} && (3.19)\\ &= \mathbf{V}\mathbf{\Sigma}^{\mathrm{H}}\mathbf{\Sigma}\mathbf{V}^{\mathrm{H}} && (3.20)\\ &= \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{\mathrm{H}}, && (3.21) \end{aligned}$$

[1] This appellation was first made for linear SIC in [19].

it is obvious that $\mathbf{\Lambda}$ is formed by the eigenvalues of $\mathbf{S}^{\mathrm{H}}\mathbf{S}$ which are all non-negative as defined in section 2.3. It follows then

$$\begin{aligned}
\mathbf{G}_m^{\mathrm{H}} &= \sum_{i=1}^{m}(\mathbf{I}-\mathbf{V}\mathbf{\Sigma}^{\mathrm{H}}\mathbf{U}^{\mathrm{H}}\mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathrm{H}})^{(i-1)}\mathbf{V}\mathbf{\Sigma}\mathbf{U}^{\mathrm{H}} && (3.22)\\
&= \sum_{i=1}^{m}\mathbf{V}(\mathbf{I}-\mathbf{\Sigma}^{\mathrm{H}}\mathbf{\Sigma})^{(i-1)}\mathbf{\Sigma}\mathbf{U}^{\mathrm{H}} && (3.23)\\
&= \mathbf{V}\begin{bmatrix}\sum_{i=1}^{m}(\mathbf{I}-\mathbf{\Lambda})^{(i-1)}\sqrt{\mathbf{\Lambda}}\\ \mathbf{0}\end{bmatrix}\mathbf{U}^{\mathrm{H}}. && (3.24)
\end{aligned}$$

A sufficient and necessary condition for $\mathbf{G}_m^{\mathrm{H}}$ to converge is then $\sum_{i=1}^{\infty}(\mathbf{I}-\mathbf{\Lambda})^{(i-1)}\sqrt{\mathbf{\Lambda}}$ converges, i.e., $\sum_{i=1}^{\infty}(1-\lambda_k)^{(i-1)}\sqrt{\lambda_k}$ converges for $k=1,2,\cdots,K$. The solution is obviously $0\le\lambda_k<2$, or in other words, all the eigenvalues of $\mathbf{S}^{\mathrm{H}}\mathbf{S}$ are less than 2.

Clearly, when the above eigenvalue constraint is satisfied,

$$\sum_{i=1}^{\infty}(1-\lambda_k)^{(i-1)}\sqrt{\lambda_k}=\begin{cases}\lambda_k^{-1}\sqrt{\lambda_k} & \text{if } \lambda_k>0,\\ 0 & \text{if } \lambda_k=0.\end{cases} \tag{3.25}$$

Hence,

$$\sum_{i=1}^{\infty}(\mathbf{I}-\mathbf{\Lambda})^{(i-1)}\sqrt{\mathbf{\Lambda}}=\begin{bmatrix}\sqrt{\mathbf{\Lambda}_1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}, \tag{3.26}$$

where $\mathbf{\Lambda}_1$ is a diagonal matrix contains all the non-zero eigenvalues of $\mathbf{S}^{\mathrm{H}}\mathbf{S}$. Therefore,

$$\begin{aligned}
\lim_{m\to\infty}\mathbf{G}_m^{\mathrm{H}} &= \mathbf{V}\begin{bmatrix}\mathbf{\Lambda}_1^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}\mathbf{V}^{\mathrm{H}}\mathbf{V}\begin{bmatrix}\sqrt{\mathbf{\Lambda}_1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}\mathbf{U}^{\mathrm{H}} && (3.27)\\
&= (\mathbf{S}^{\mathrm{H}}\mathbf{S})^{+}\mathbf{S}^{\mathrm{H}}. && (3.28)
\end{aligned}$$

In conclusion, the conventional PIC converges if and only if all eigenvalues of $\mathbf{S}^{\mathrm{H}}\mathbf{S}$ are less than 2. When this eigenvalue constraint is satisfied, the PIC detector converges to the decorrelating detector, i.e.,

$$\lim_{m\to\infty}\mathbf{G}_m^{\mathrm{H}}=\mathbf{G}_{\mathrm{dec}}^{\mathrm{H}}=(\mathbf{S}^{\mathrm{H}}\mathbf{S})^{+}\mathbf{S}^{\mathrm{H}}. \tag{3.29}$$

Consequently, the BER performance of a conventional PIC detector approaches that of the decorrelator given sufficient number of stages, provided that it does converge. Eqn. (3.12) and (3.29) also indicate that a linear PIC is realising the matrix inverse of a decorrelating detector indirectly through polynomial expansion (PE) [7].

However, the eigenvalue constraint is too strict to be satisfied in all but the simplest cases. This accounts for the unstable behaviour of the conventional PIC [27]. For instance, a simple numerical search indicates that the largest eigenvalue of a 15 user system with a processing gain of 31 is greater than 2 for more than 99% of all possible code-sets.

Furthermore, it is interesting to observe through simulations that a linear PIC detector often outperforms the decorrelator at some intermediate stages. Similar phenomenon has been reported for SIC in [19]. This implies that interference cancellation detector's overall capability is not limited to approximating the decorrelator. But this is yet to be exploited and thoroughly explained. Further study will be presented in the following chapters. The next section contains some numerical results.

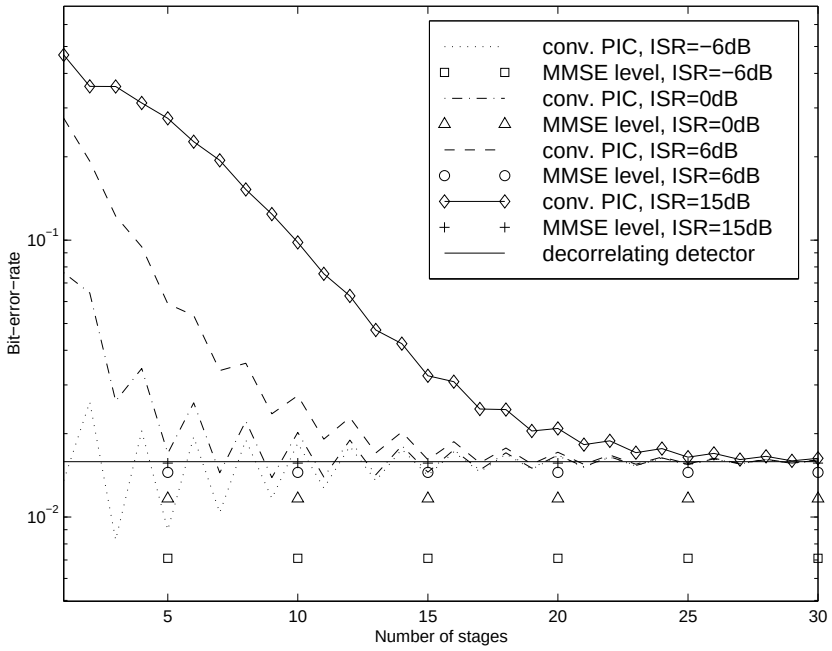

Figure 3.3: Stage-by-stage performance of the conventional PIC using short codes.

3.3 Numerical Results

The numerical results demonstrated in this section are based on a symbol-synchronous system assuming BPSK modulation and spreading formats. A processing gain of 31 is assumed in all instances. Short-code systems are considered in Fig. 3.3 through Fig. 3.6, while a long-code system is considered in Fig. 3.7.

Fig. 3.3 illustrates stage-by-stage BER performance of the first user of a 6-user conventional PIC detector. The signal-to-noise ratio (SNR) of the first user is 6 dB while the remaining 5 users are transmitting at the same power but ISR(dB)[2] different to user 1. ISRs of -6, 0, 6 and 15 dB are shown. The eigenvalues of $\mathbf{S}^{\mathrm{H}}\mathbf{S}$, the code correlation matrix, are (0.65805, 0.78619, 0.96682, 0.25276, 1.4664, 1.8698). Since they are all less than 2, the performance converges to that of the decorrelator as the stage number increases. In the meantime, the stronger the interference from other users is the slower it converges. However the BER can be much better than the decorrelating detector (albeit always worse than the MMSE detector) in some intermediate stages. It is also observed that BER performance often fluctuates like saw-teeth along the number of stages. This phenomenon is called ping-pong effect. The reason is that some of the eigenvalues are close to 2, which resembles an under-damped oscillation of an RLC circuit.

In Figure 3.4, a 4-user system is considered where the eigenvalues of the code correlation matrix are (0.41147, 0.45896, 0.96615, 2.1634). The BER performance of all 4 users is compared to that of the decorrelating detector. Since the largest eigenvalue is greater than 2, divergence is observed for the conventional PIC. All 4 users appear to follow the same pattern

[2]Here we simply use ISR as the short for interference-to-signal ratio.

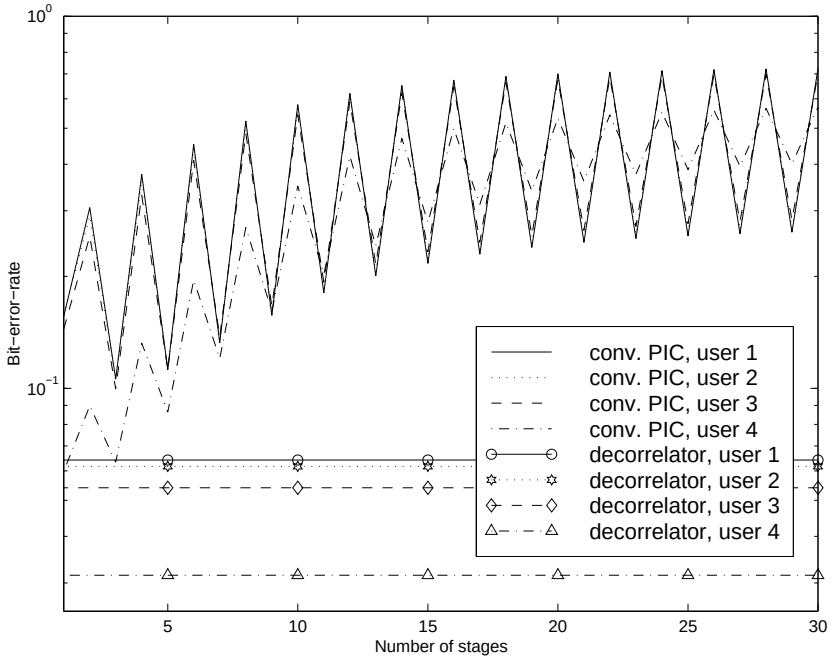

Figure 3.4: Divergence case of the conventional PIC.

of ping-pong fluctuation. It is rare for a 4-user system with a processing gain of 31 to have at least one eigenvalue greater than 2. A simple numerical search indicate that it occurs roughly 6 times in every 10 thousand random code-sets. Conversely for a 15-user system with the same processing gain, it happens for more than 99% of the code-sets, so that the conventional PIC hardly converges in this case.

Figure 3.5 demonstrates the BER performance of the conventional PIC detector versus the working SNR. Again a 6-user system is considered. The spreading codes used are the same as the system shown in Fig. 3.3. It is observed that the higher the SNR, the better the BER performance and the more number of PIC stages the closer the BER is to that of the MMSE detector. Furthermore, there is no error floor for a conventional PIC detector employing independent codes.

Figure 3.6 shows the BER performance of a conventional PIC detector under near-far environment, as compared with that of the conventional detector, the decorrelator and the MMSE detector. The same system as shown in Fig. 3.3 and Fig. 3.5 is considered. It is observed that the PIC generally performs better than the conventional detector but worse than the decorrelator, and the more number of PIC stages the better ability to combat the near-far effects. Since the PIC detector can never eliminate the MAI thoroughly as what the decorrelator does, the signal from the desired user may be totally buried under that of much stronger users, thus its near-far resistance is 0. This implies that no matter how large the number of PIC stages is, the BER performance will go up to 50% as the ISR becomes higher and higher. Note however that near-far resistance delineates only the performance of a CDMA receiver in the extreme situation where the interfering users are transmitting at infinitely higher SNRs. It is not a good indicator of the BER performance in most practical

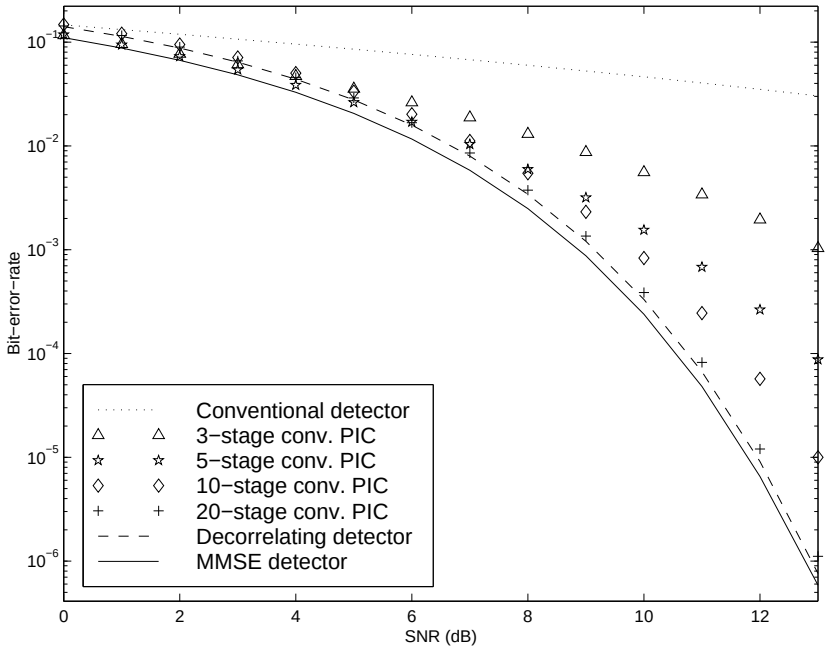

Figure 3.5: BER versus SNR of the conventional PIC.

cases.

In Figure 3.7 a system similar to the one in Figure 3.3 is considered but long codes are used instead. The stage-by-stage BER performance of the first user is shown. It shows that the PIC detector is not converging to the decorrelator but maintains a certain performance gap in this case. The reason is that a portion of all code-sets has at least one eigenvalue of the code correlation matrix greater than 2 and therefore suffers divergence. The overall performance is then an average of those that converge and those that do not. It can be inferred that the conventional PIC detector is useless for a system with the number of users comparable to the processing gain, since the chance of having an eigenvalue greater than 2 is so large that the detector suffers divergence for most of the codes.

3.4 Summary

In this chapter, the linear conventional PIC is analysed using matrix algebra. It is shown that the sufficient and necessary condition for a linear conventional PIC to converge is that all eigenvalues of the code correlation matrix are less than 2, which is not always true. Simulation results demonstrate that this detector suffers ping-pong effects in the BER performance as well as divergence. However, its favourable behaviour in the intermediate stages implies some unexploited merits to be studied further.

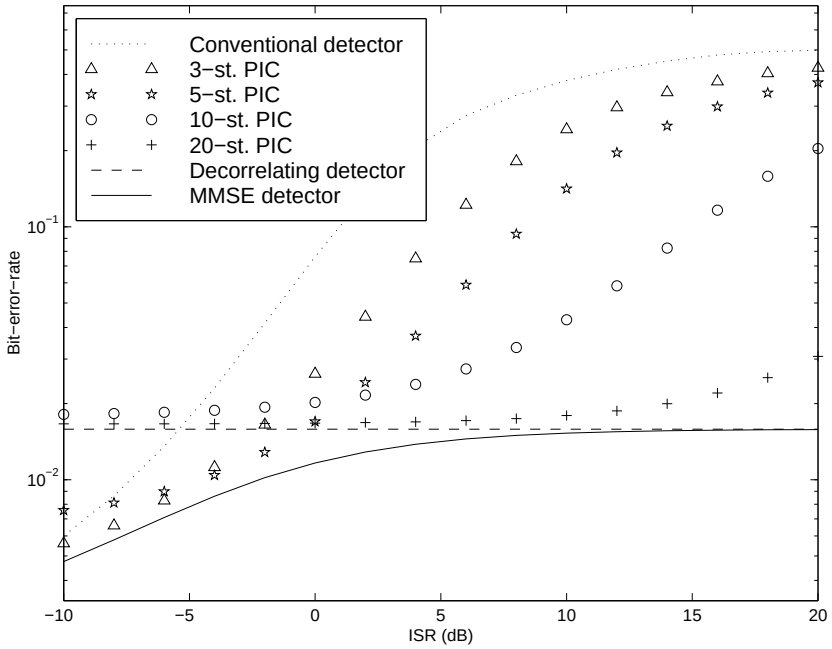

Figure 3.6: BER versus ISR of the conventional PIC.

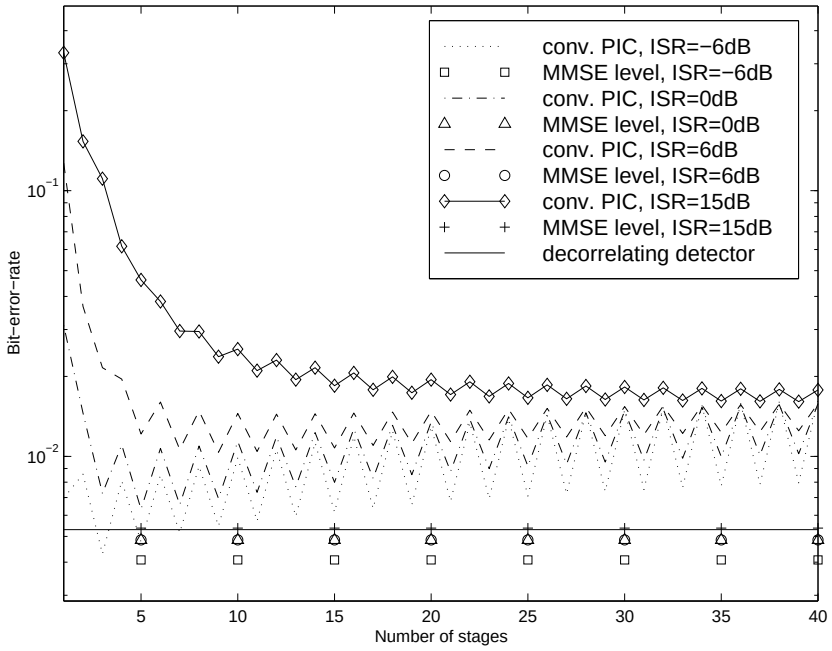

Figure 3.7: Stage-by-stage performance of the conventional PIC using long codes.

Chapter 4

Weighted Linear Parallel Interference Cancellation

In this chapter the weighted linear PIC scheme using short codes is mathematically described and analysed. It is shown in section 4.1 that linear PIC schemes, either conventional or weighted, correspond to linear matrix filtering that can be performed directly on the received chip-matched filtered signal vector.

The concept of weighted linear PIC resembles the concept of the steepest descent method for updating adaptive filter weights to minimise the MSE. We demonstrate in section 4.2 the close relationship between the two and present a new PIC structure which is in fact a modified version of the structures suggested in [36] and [25].

The one-shot equivalent detector is derived in section 4.3. An analytical expression for the exact bit error rate is then obtained. In section 4.4 conditions on the eigenvalues of the correlation matrix and weights that ensure convergence are derived. It is shown that our structure may approach to the performance of the MMSE detector rather than the decorrelating detector which other PIC structures generally converge to.

Given the number of PIC stages, we derive an expression for the mean squared error of the one-shot cancellation filter in section 4.5 and devise techniques in section 4.6 for optimising the choice of weighting factors (or equivalently step sizes for the SDM) with respect to the MSE. It is shown that only K PIC stages are required for the equivalent one-shot 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. The effect of a parameter in the structure and the ordering of the weights are studied in section 4.7 and section 4.8 respectively.

Numerical results are presented in section 4.9 to support theoretical findings and section 4.10 summarises this chapter.

4.1 Structure of Weighted Linear PIC

In [24] Divsalar and Simon first suggested using weighting factors to make PIC more flexible. The general structure for an m-stage PIC is illustrated in Fig. 4.1, where the multi-stage PIC is seen to be a concatenation of m one-stage PICs. $\mu_1, \mu_2, \cdots, \mu_m$ denote the weighting

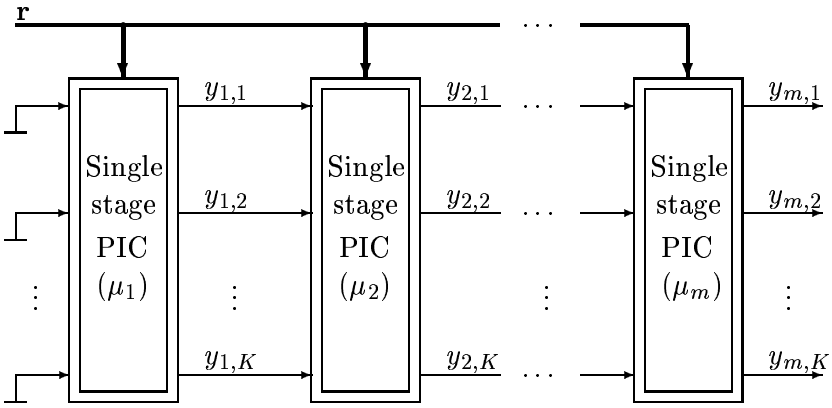

Figure 4.1: A general structure for a K-user, m-stage PIC.

factors, or weights[1], used accordingly to control the amount of interference cancellation at each stage. The weights in our structure can be complex.

The detailed structure of the i^{th} stage PIC is depicted in Fig. 4.2, where α is a non-negative parameter. The tentative decision devices are dropped in this figure since linear soft decisions are made in all the intermediate stages. Note that we have chosen to include the received signal amplitude in re-spreading and de-spreading (matched filtering), i.e., $\mathbf{a}_k$ is used rather than $\mathbf{s}_k$. This is different from most of previous proposals [17, 25, 36]. This is motivated by the relationship between PIC and the steepest descent method, to become clearer in section 4.2.

From the structure, we can get the relationship between the current decision statistic and the previous set of decision statistics. In general, the decision statistic for user k at stage i is a weighted sum of the statistics at stage $(i-1)$ and the statistic based on current cancellation, i.e.,

$$y_{i,k} = (1-\alpha\mu_i)y_{i-1,k} + \mu_i \mathbf{a}_k^{\mathrm{H}} \left(\mathbf{r} - \sum_{j=1}^{K} y_{i-1,j}\mathbf{a}_j \right). \tag{4.1}$$

In vector representation, we have the recursive formula of decision statistic as

$$\begin{aligned} \mathbf{y}_i &= (1-\alpha\mu_i)\mathbf{y}_i + \mu_i \mathbf{A}^{\mathrm{H}}(\mathbf{r} - \mathbf{A}\mathbf{y}_{i-1}) && (4.2)\\ &= [\mathbf{I} - \mu_i(\mathbf{A}^{\mathrm{H}}\mathbf{A} + \alpha\mathbf{I})]\,\mathbf{y}_{i-1} + \mu_i \mathbf{A}^{\mathrm{H}}\mathbf{r} && (4.3)\\ &= [\mathbf{I} - \mu_i(\mathbf{R} + \alpha\mathbf{I})]\,\mathbf{y}_{i-1} + \mu_i \mathbf{A}^{\mathrm{H}}\mathbf{r}. && (4.4) \end{aligned}$$

Substituting matrix $\mathbf{A}$ by $\mathbf{S}$, our structure embraces all the linear PIC schemes in [17, 25, 36] with the help of adjustable weighting factors. For $\mu_i = 1$, $\alpha = 0$ and $\mathbf{A}^{\mathrm{H}}$ replaced by $\mathbf{S}^{\mathrm{H}}$, Eqns. (3.6) and (4.4) are identical, hence the structure reduces to the conventional PIC in [17]. For a fixed parameter, $\mu_i = \mu$, and $\alpha = 0$, Eqn. (4.4) describes the structure in [25]. For varying μ_i and $\alpha = 0$, Eqn. (4.4) also describes the structure in [24].

[1] "Weight" and "weighting factor" are interchangeable terms throughout this thesis.

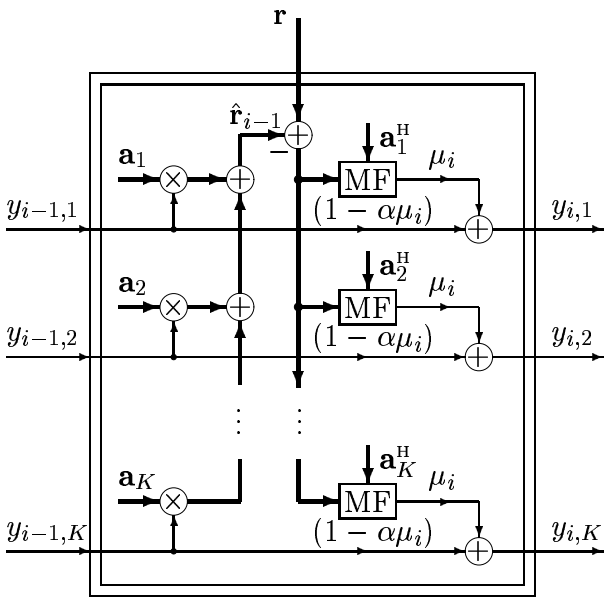

Figure 4.2: The i^{th} stage PIC, employing weight μ_i. MF — matched filtering.

4.2 Relationship to the Steepest Descent Method

In this section we show that a linear weighted PIC detector is a realisation of the steepest descent method using variable step sizes. The step sizes happen to be the weighting factors in the PIC structure. We start from deriving the filter updating equation of the steepest descent method, then we show its equivalence to the weighted PIC detector introduced in section 4.1.

A linear detector $\mathbf{G}$ is an $N \times K$ linear matrix filter

$$\mathbf{G} = [\mathbf{g}_1 \ \mathbf{g}_2 \ \cdots \ \mathbf{g}_K], \tag{4.5}$$

where $\mathbf{g}_1, \mathbf{g}_2, \cdots, \mathbf{g}_K$ are column vectors of length N. The filter output is then the following estimate of the transmitted data symbols,

$$\mathbf{y} = \mathbf{G}^{\mathrm{H}}\mathbf{r} = \mathbf{G}^{\mathrm{H}}(\mathbf{A}\mathbf{d} + \mathbf{n}). \tag{4.6}$$

The corresponding MSE is given by

$$J = \mathrm{E}\left\{\|\mathbf{y} - \mathbf{d}\|^2\right\} = \mathrm{E}\left\{\|\mathbf{G}^{\mathrm{H}}\mathbf{r} - \mathbf{d}\|^2\right\} = \sum_{k=1}^{K} \mathrm{E}\left\{|\mathbf{g}_k^H\mathbf{r} - d_k|^2\right\}. \tag{4.7}$$

Differentiating with respect to $\mathbf{g}_k$ yields

$$\begin{aligned}
\frac{\partial J}{\partial \mathbf{g}_k} &= \mathrm{E}\left\{\frac{\partial |\mathbf{g}_k^H\mathbf{r} - d_k|^2}{\partial \mathbf{g}_k}\right\} \\
&= \mathrm{E}\left\{\mathbf{r}^*(\mathbf{g}_k^H\mathbf{r} - d_k)\right\} \\
&= \mathrm{E}\left\{(\mathbf{r}\mathbf{r}^H)^*\right\}\mathbf{g}_k^* - \mathrm{E}\left\{d_k\mathbf{r}^*\right\} \\
&= [\mathbf{A}\mathbf{A}^H + \sigma^2\mathbf{I}]^*\mathbf{g}_k^* - \mathbf{a}_k^*.
\end{aligned}$$

The gradient with respect to $\mathbf{g}_k$ is then (see, e.g. [37])

$$\nabla_k J = 2\frac{\partial J}{\partial \mathbf{g}_k} = 2[(\mathbf{A}\mathbf{A}^H + \sigma^2\mathbf{I})^*\mathbf{g}_k^* - \mathbf{a}_k^*]. \tag{4.8}$$

The steepest descent method gives the following recursion for finding the minimum of MSE,

$$\begin{aligned}\mathbf{g}_k^*(i) &= \mathbf{g}_k^*(i-1) - \frac{1}{2}\mu_i \nabla_k J &(4.9)\\ &= \left[\mathbf{I} - \mu_i(\mathbf{A}\mathbf{A}^H + \sigma^2\mathbf{I})^*\right]\mathbf{g}_k^*(i-1) + \mu_i\mathbf{a}_k^*, &(4.10)\end{aligned}$$

where μ_i is a variable step size of the current stage. Treating the K filters as a filter bank, we have that

$$\mathbf{G}_i^{\mathrm{H}} = \mathbf{G}_{i-1}^{\mathrm{H}} - \mu_i\left[\mathbf{G}_{i-1}^{\mathrm{H}}(\mathbf{A}\mathbf{A}^{\mathrm{H}} + \sigma^2\mathbf{I}) - \mathbf{A}^{\mathrm{H}}\right], \tag{4.11}$$

where $\mathbf{G}_0 = \mathbf{0}$. The equivalent one-shot filter for an i-stage PIC detector in non-recursive form is then

$$\mathbf{G}_i^{\mathrm{H}} = \sum_{l=1}^{i}\mu_l\mathbf{A}^{\mathrm{H}}\prod_{j=l+1}^{i}(\mathbf{I} - \mu_j(\mathbf{A}\mathbf{A}^{\mathrm{H}} + \sigma^2\mathbf{I})). \tag{4.12}$$

Note that $\mathbf{G}_i^{\mathrm{H}}\mathbf{A}\mathbf{A}^{\mathrm{H}} = \mathbf{R}\mathbf{G}_i^{\mathrm{H}}$. Hence (4.11) can be expressed as

$$\mathbf{G}_i^{\mathrm{H}} = \mathbf{G}_{i-1}^{\mathrm{H}} - \mu_i\left[(\mathbf{R} + \sigma^2\mathbf{I})\mathbf{G}_{i-1}^{\mathrm{H}} - \mathbf{A}^{\mathrm{H}}\right]. \tag{4.13}$$

Post-multiplying with $\mathbf{r}$ gives

$$\mathbf{y}_i = [\mathbf{I} - \mu_i(\mathbf{R} + \sigma^2\mathbf{I})]\mathbf{y}_{i-1} + \mu_i\mathbf{A}^{\mathrm{H}}\mathbf{r}. \tag{4.14}$$

Note that this is equivalent to (4.4) with $\alpha = \sigma^2$. Therefore the linear PIC illustrated in Fig. 4.1 can be seen as a realisation of the steepest descent method for implementing the MMSE detector. When $\alpha = \sigma^2$, this structure implements the algorithm of Eqn. (4.14) exactly. The reason for introducing an arbitrary non-negative parameter α is because the exact noise level, σ^2, is usually not available at the receiver. This will become clear later on.

Also note that the SDM recursion inherently includes the received signal amplitude in cancellation processes. This motivates our special design of the PIC structure, which is different to previous proposals. The advantage of this design will become clearer later on.

4.3 One-shot Equivalent Filter

For $\alpha > 0$, $(\mathbf{R}+\alpha\mathbf{I})$ is always positive definite. For $\alpha = 0$, we assume that linearly independent spreading codes are used by all users so that $\mathbf{R}$ is positive definite and the decorrelating detector is valid for all users [13]. This assumption is almost always valid in a real scenario.

Subtracting $(\mathbf{R} + \alpha\mathbf{I})^{-1}\mathbf{A}^{\mathrm{H}}\mathbf{r}$ from both sides of (4.4),

$$\begin{aligned}\mathbf{y}_i - (\mathbf{R} + \alpha\mathbf{I})^{-1}\mathbf{A}^{\mathrm{H}}\mathbf{r} &= [\mathbf{I} - \mu_i(\mathbf{R} + \alpha\mathbf{I})]\,\mathbf{y}_{i-1} + \mu_i\mathbf{A}^{\mathrm{H}}\mathbf{r} - (\mathbf{R} + \alpha\mathbf{I})^{-1}\mathbf{A}^{\mathrm{H}}\mathbf{r}\\ &= [\mathbf{I} - \mu_i(\mathbf{R} + \alpha\mathbf{I})]\,[\mathbf{y}_{i-1} - (\mathbf{R} + \alpha\mathbf{I})^{-1}\mathbf{A}^{\mathrm{H}}\mathbf{r}]. &(4.15)\end{aligned}$$

Treating $[\mathbf{y}_i - (\mathbf{R} + \alpha\mathbf{I})^{-1}\mathbf{A}^{\mathrm{H}}\mathbf{r}]$ as a series of vectors and $[\mathbf{I} - \mu_i(\mathbf{R} + \alpha\mathbf{I})]$ as a series of matrices, i.e.,

$$\mathbf{x}_i = \mathbf{y}_i - (\mathbf{R} + \alpha\mathbf{I})^{-1}\mathbf{A}^{\mathrm{H}}\mathbf{r}, \quad i = 0, 1, \cdots, m, \tag{4.16}$$

and

$$\mathbf{F}_i = \mathbf{I} - \mu_i(\mathbf{R} + \alpha\mathbf{I}), \quad i = 1, \cdots, m, \tag{4.17}$$

we have a simple recursive equation for $i = 0, 1, \cdots, m-1$ as

$$\mathbf{x}_i = \mathbf{F}_i \mathbf{x}_{i-1}. \tag{4.18}$$

Clearly,

$$\mathbf{x}_m = \mathbf{F}_m \mathbf{F}_{m-1} \cdots \mathbf{F}_1 \mathbf{x}_0 \tag{4.19}$$

$$= \left[\prod_{i=1}^{m} \mathbf{F}_i\right] \mathbf{x}_0. \tag{4.20}$$

Substitute (4.16) and (4.17) back into (4.20),

$$\mathbf{y}_m - (\mathbf{R} + \alpha\mathbf{I})^{-1}\mathbf{A}^{\mathrm{H}}\mathbf{r} = \left[\prod_{i=1}^{m} (\mathbf{I} - \mu_i(\mathbf{R} + \alpha\mathbf{I}))\right] \left[\mathbf{y}_0 - (\mathbf{R} + \alpha\mathbf{I})^{-1}\mathbf{A}^{\mathrm{H}}\mathbf{r}\right]. \tag{4.21}$$

Setting $\mathbf{y}_0 = \mathbf{0}$, we have the non-recursive form expression for the set of decision statistics at the output of the m^{th} stage as

$$\mathbf{y}_m = \left[\mathbf{I} - \prod_{i=1}^{m} (\mathbf{I} - \mu_i(\mathbf{R} + \alpha\mathbf{I}))\right] (\mathbf{R} + \alpha\mathbf{I})^{-1}\mathbf{A}^{\mathrm{H}}\mathbf{r}. \tag{4.22}$$

This means that an m-stage PIC is equivalent to a linear matrix filter $\mathbf{G}_m$ described by

$$\mathbf{G}_m^{\mathrm{H}} = \left[\mathbf{I} - \prod_{i=1}^{m} (\mathbf{I} - \mu_i(\mathbf{R} + \alpha\mathbf{I}))\right] (\mathbf{R} + \alpha\mathbf{I})^{-1}\mathbf{A}^{\mathrm{H}}. \tag{4.23}$$

Again we can analytically calculate the BER for any user at any stage as mentioned in section 3.1. For a BPSK data modulation, (3.15) may also be used to compute the exact BER.

4.4 Convergence Issues

It is possible to rewrite Eqn. (4.22) in terms of a "steady-state" solution corrupted by some disturbance,

$$\mathbf{y}_m = \mathbf{y}_\infty - \mathbf{e}_m \tag{4.24}$$

where $\mathbf{y}_\infty$ is the "steady-state" filter output expressed by

$$\mathbf{y}_\infty = (\mathbf{R} + \alpha\mathbf{I})^{-1}\mathbf{A}^{\mathrm{H}}\mathbf{r}, \tag{4.25}$$

and $\mathbf{e}_m$ is some excess transient error related to the m^{th} stage,

$$\mathbf{e}_m = \left[\prod_{i=1}^{m} (\mathbf{I} - \mu_i(\mathbf{R} + \alpha\mathbf{I}))\right] (\mathbf{R} + \alpha\mathbf{I})^{-1}\mathbf{A}^{\mathrm{H}}\mathbf{r}. \tag{4.26}$$

It is obvious that we may force $\mathbf{e}_m$ to converge to $\mathbf{0}$ in order for $\mathbf{y}_m$ to converge. As mentioned in section 2.3, a Hermitian matrix $\mathbf{R}$ has an eigenvalue decomposition as $\mathbf{R} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\mathrm{H}}$ where

$\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, ..., \lambda_K)$. Clearly the term in the square brackets in Eqn. (4.26) is a polynomial in $\mathbf{R}$, which may be expressed as

$$\mathbf{U}\text{diag}(f(\lambda_1), f(\lambda_2), \cdots, f(\lambda_K))\mathbf{U}^{\text{H}}, \tag{4.27}$$

where

$$f(x) \triangleq \prod_{i=1}^{m} (1 - \mu_i(x + \alpha)) \,. \tag{4.28}$$

It is easy to get a sufficient (but not necessary) condition[2] for convergence (when m goes to ∞) as,

$$-1 < 1 - \mu_i(\lambda_k + \alpha) < 1, \tag{4.29}$$

for $k = 1, 2, \cdots, K$ and $m = 1, 2, \cdots$, or equivalently

$$0 < \mu_i < \frac{2}{\lambda_{\max} + \alpha} \tag{4.30}$$

where $\lambda_{\max}$ is the largest eigenvalue. The "steady-state" filter in this case would be:

$$\mathbf{G}_{\infty}^{\text{H}} = (\mathbf{R} + \alpha\mathbf{I})^{-1}\,\mathbf{A}^{\text{H}}, \tag{4.31}$$

which in case $\alpha = 0$ is the decorrelating filter and in case $\alpha = \sigma^2$ is the MMSE filter. For conventional PIC, $\alpha = 0$ and $\mu = 1$. In this case we have convergence if and only if $0 < \lambda_k < 2$ for $k = 1, 2, \cdots, K$, which collaborates the conclusion in section 3.2.

It is clear now that given knowledge of eigenvalues of the correlation matrix, weighting factors may be introduced to ensure convergence. Furthermore, the modified structure can be made to converge to the MMSE detector, which always has a better BER performance than the decorrelating detector. However, since it is obviously impossible to have unlimited number of stages in practice, the study of PIC with a pre-designated number of stages becomes of more interest and not surprisingly, reveals new aspects of the PIC scheme.

4.5 Mean Squared Error

Although bit error rate is the eventual criterion for a digital communication system, it is not unusual to see the mean squared error used as an indicator of the BER performance, since it is a quadratic function of the filter weights which results in a unique minimum. In this section we derive the MSE of an m-stage PIC detector.

At the m^{th} stage, the set of decision statistics is $\mathbf{y}_m = \mathbf{G}_m^{\text{H}}\mathbf{r}$ where $\mathbf{G}_m^{\text{H}}$ is described by (4.23). We may then express the MSE at the m^{th} stage as a function of the weighting factors, described by $\boldsymbol{\mu} = (\mu_1, \mu_2, ..., \mu_m)^{\top}$, and the parameter α,

$$J^{(m)}(\boldsymbol{\mu}, \alpha) = \text{E}\left\{\|\mathbf{G}_m^{\text{H}}\mathbf{r} - \mathbf{d}\|^2\right\} = \text{E}\left\{\mathbf{r}^{\text{H}}\mathbf{G}_m\mathbf{G}_m^{\text{H}}\mathbf{r} - 2\text{Re}\left\{\mathbf{d}^{\text{H}}\mathbf{G}_m^{\text{H}}\mathbf{r}\right\} + \mathbf{d}^{\text{H}}\mathbf{d}\right\} \tag{4.32}$$

[2]If an identical weight is used for all stages [25], i.e., $\mu_i = \mu$, it is also a necessary condition.

Since $J^{(m)}(\boldsymbol{\mu}, \alpha)$ is a scalar, we can apply the trace operator to the RHS directly. Also note that $\mathrm{tr}\{\cdot\}$ is a linear operation and satisfies $\mathrm{tr}\{\mathbf{AB}\} = \mathrm{tr}\{\mathbf{BA}\}$. Therefore,

$$\begin{aligned} J^{(m)}(\boldsymbol{\mu}, \alpha) &= \mathrm{tr}\left\{\mathrm{E}\left\{\mathbf{r}^{\mathrm{H}}\mathbf{G}_m\mathbf{G}_m^{\mathrm{H}}\mathbf{r} - 2\mathrm{Re}\left\{\mathbf{d}^{\mathrm{H}}\mathbf{G}_m^{\mathrm{H}}\mathbf{r}\right\} + \mathbf{d}^{\mathrm{H}}\mathbf{d}\right\}\right\} && (4.33)\\ &= \mathrm{tr}\left\{\mathbf{G}_m^{\mathrm{H}}\mathrm{E}\left\{\mathbf{r}\mathbf{r}^{\mathrm{H}}\right\}\mathbf{G}_m - 2\mathrm{Re}\left\{\mathbf{G}_m^{\mathrm{H}}\mathrm{E}\left\{\mathbf{r}\mathbf{d}^{\mathrm{H}}\right\}\right\} + \mathrm{E}\left\{\mathbf{d}\mathbf{d}^{\mathrm{H}}\right\}\right\} && (4.34)\\ &= \mathrm{tr}\left\{\mathbf{G}_m^{\mathrm{H}}(\mathbf{A}\mathbf{A}^{\mathrm{H}} + \sigma^2\mathbf{I})\mathbf{G}_m - 2\mathrm{Re}\left\{\mathbf{G}_m^{\mathrm{H}}\mathbf{A}\right\} + \mathbf{I}\right\} && (4.35) \end{aligned}$$

Let $\mathbf{G}_m^{\mathrm{H}} = \tilde{\mathbf{R}}\mathbf{A}^{\mathrm{H}}$ where

$$\tilde{\mathbf{R}} \triangleq \left[\mathbf{I} - \prod_{i=1}^{m}(\mathbf{I} - \mu_i(\mathbf{R} + \alpha\mathbf{I}))\right](\mathbf{R} + \alpha\mathbf{I})^{-1}. \tag{4.36}$$

We can therefore re-write (4.35) as

$$\begin{aligned} J^{(m)}(\boldsymbol{\mu}, \alpha) &= \mathrm{tr}\left\{\tilde{\mathbf{R}}\mathbf{A}^{\mathrm{H}}(\mathbf{A}\mathbf{A}^{\mathrm{H}} + \sigma^2\mathbf{I})\mathbf{A}\tilde{\mathbf{R}} - 2\mathrm{Re}\left\{\mathbf{R}\tilde{\mathbf{R}}\right\} + \mathbf{I}\right\} && (4.37)\\ &= \mathrm{tr}\left\{\tilde{\mathbf{R}}\mathbf{R}(\mathbf{R} + \sigma^2\mathbf{I})\tilde{\mathbf{R}} - 2\mathrm{Re}\left\{\mathbf{R}\tilde{\mathbf{R}}\right\} + \mathbf{I}\right\}. && (4.38) \end{aligned}$$

Since $\tilde{\mathbf{R}}$ is a polynomial in $\mathbf{R}$, it may be expressed using $\mathbf{R}$'s eigenvectors as (refer to section 2.3)

$$\tilde{\mathbf{R}} = \mathbf{U}\tilde{\boldsymbol{\Lambda}}\mathbf{U}^{\mathrm{H}} \tag{4.39}$$

where $\tilde{\boldsymbol{\Lambda}} = \mathrm{diag}(\tilde{\lambda}_1, \tilde{\lambda}_2, ..., \tilde{\lambda}_K)$ and for $k = 1, 2, \cdots, K$,

$$\tilde{\lambda}_k = \frac{1}{\lambda_k + \alpha}\left[1 - \prod_{i=1}^{m}(1 - \mu_i(\lambda_k + \alpha))\right]. \tag{4.40}$$

Therefore,

$$\begin{aligned} & J^{(m)}(\boldsymbol{\mu}, \alpha)\\ =\ & \mathrm{tr}\left\{\mathbf{U}\tilde{\boldsymbol{\Lambda}}\mathbf{U}^{\mathrm{H}}\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{\mathrm{H}}(\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{\mathrm{H}} + \sigma^2\mathbf{I})\mathbf{U}\tilde{\boldsymbol{\Lambda}}\mathbf{U}^{\mathrm{H}} - 2\mathrm{Re}\left\{\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{\mathrm{H}}\mathbf{U}\tilde{\boldsymbol{\Lambda}}\mathbf{U}^{\mathrm{H}}\right\} + \mathbf{I}\right\} && (4.41)\\ =\ & \mathrm{tr}\left\{\tilde{\boldsymbol{\Lambda}}\boldsymbol{\Lambda}(\boldsymbol{\Lambda} + \sigma^2\mathbf{I})\tilde{\boldsymbol{\Lambda}} - 2\mathrm{Re}\left\{\boldsymbol{\Lambda}\tilde{\boldsymbol{\Lambda}}\right\} + \mathbf{I}\right\} && (4.42)\\ =\ & \sum_{k=1}^{K}\left[\tilde{\lambda}_k\lambda_k(\lambda_k + \sigma^2)\tilde{\lambda}_k - 2\mathrm{Re}\left\{\lambda_k\tilde{\lambda}_k\right\} + 1\right] && (4.43)\\ =\ & \sum_{k=1}^{K}\frac{\sigma^2}{\lambda_k + \sigma^2} + \sum_{k=1}^{K}\left[\lambda_k(\lambda_k + \sigma^2)\left|\tilde{\lambda}_k - \frac{1}{\lambda_k + \sigma^2}\right|^2\right]. && (4.44) \end{aligned}$$

From (4.40), we have

$$\begin{aligned} \tilde{\lambda}_k - \frac{1}{\lambda_k + \sigma^2} &= \frac{1}{\lambda_k + \alpha}\left[1 - \prod_{i=1}^{m}(1 - \mu_i(\lambda_k + \alpha))\right] - \frac{1}{\lambda_k + \sigma^2} && (4.45)\\ &= \frac{1}{\lambda_k + \alpha} - \frac{1}{\lambda_k + \sigma^2} - \frac{1}{\lambda_k + \alpha}\prod_{i=1}^{m}(1 - \mu_i(\lambda_k + \alpha)) && (4.46)\\ &= \frac{1}{\lambda_k + \alpha}\left[\frac{\sigma^2 - \alpha}{\lambda_k + \sigma^2} - \prod_{i=1}^{m}(1 - \mu_i(\lambda_k + \alpha))\right]. && (4.47) \end{aligned}$$

The MSE is then expressed as

$$J^{(m)}(\boldsymbol{\mu},\alpha) = \sum_{k=1}^{K}\frac{\sigma^2}{\lambda_k+\sigma^2} + \sum_{k=1}^{K}\frac{\lambda_k(\lambda_k+\sigma^2)}{(\lambda_k+\alpha)^2}\left|\frac{\sigma^2-\alpha}{\lambda_k+\sigma^2} - \prod_{i=1}^{m}(1-\mu_i(\lambda_k+\alpha))\right|^2 . \quad (4.48)$$

The first term of (4.48) is nothing but the minimum MSE, which is independent of the weights and can be achieved only by the MMSE detector, i.e.,

$$J_{\text{MMSE}} = \mathrm{E}\left\{\|\mathbf{G}_{\text{MMSE}}^{\text{H}}\mathbf{r}-\mathbf{d}\|^2\right\} = \mathrm{E}\left\{\left\|\left(\mathbf{R}+\sigma^2\mathbf{I}\right)^{-1}\mathbf{A}^{\text{H}}\mathbf{r}-\mathbf{d}\right\|^2\right\} \quad (4.49)$$

$$= \sum_{k=1}^{K}\frac{\sigma^2}{\lambda_k+\sigma^2}, \quad (4.50)$$

while the second term then represents the degradation with respect to the MMSE filter, or, the excess MSE expressed as

$$J_{\text{ex}}^{(m)}(\boldsymbol{\mu},\alpha) = \sum_{k=1}^{K}\frac{\lambda_k(\lambda_k+\sigma^2)}{(\lambda_k+\alpha)^2}\left|\frac{\sigma^2-\alpha}{\lambda_k+\sigma^2} - \prod_{i=1}^{m}(1-\mu_i(\lambda_k+\alpha))\right|^2 . \quad (4.51)$$

In the following sections, we concentrate on exploiting the nature of the above equation and come up with a number of interesting properties of the PIC.

4.6 Optimisation of the Weights

In this section we consider the choice of optimal[3] weights with respect to minimising the MSE given that we have a limited number of stages ($m \ll \infty$).

First we assume that $\alpha = \sigma^2$ and show that linear PIC needs exactly K stages to converge to the MMSE detector. It is later shown that given fewer number of stages, $m \leq K$, and any $\alpha \geq 0$, a unique optimal set of weighting factors exists that will lead to the minimum achievable MSE.

For notational simplicity, we define $\phi_k = \lambda_k + \sigma^2$ and $\gamma_k = \lambda_k + \alpha$ for $k = 1, 2, \cdots, K$, so that

$$J_{\text{ex}}^{(m)}(\boldsymbol{\mu},\alpha) = \sum_{k=1}^{K}\frac{\lambda_k\phi_k}{\gamma_k^2}\left|\frac{\sigma^2-\alpha}{\phi_k} - \prod_{i=1}^{m}(1-\mu_i\gamma_k)\right|^2 . \quad (4.52)$$

4.6.1 An Identical Real Weighting Factor for All Stages

Considering the special case when an identical real weighting factor is used for all stages. From (4.52), we have the excess MSE expressed as

$$J_{\text{ex}}^{(m)}(\mu,\alpha) = \sum_{k=1}^{K}\frac{\lambda_k\phi_k}{\gamma_k^2}\left[\frac{\sigma^2-\alpha}{\phi_k} - (1-\mu\gamma_k)^m\right]^2 . \quad (4.53)$$

[3]The terms "optimal", "optimise" and "optimisation", whenever appear in this thesis, refer to the MSE criterion implicitly, unless otherwise stated. It is very close to (albeit not exactly) the best solution in the BER sense.

To locate the global minimum, equate its derivative with respect to μ to zero,

$$\frac{\mathrm{d}J_{\mathrm{ex}}^{(m)}(\mu,\alpha)}{\mathrm{d}\mu} = 2m\sum_{k=1}^{K}\frac{\lambda_k\phi_k}{\gamma_k}\left[\frac{\sigma^2-\alpha}{\phi_k}-(1-\mu\gamma_k)^m\right](1-\mu\gamma_k)^{m-1} = 0. \tag{4.54}$$

This is a real coefficient polynomial with a degree of $(2m-1)$ in μ, which has $(2m-1)$ roots and at least one of them is real. It is then possible to find out the global real minimum among all real roots through comparison. However, the result is not necessarily the global minimum of $J_{\mathrm{ex}}^{(m)}(\mu,\alpha)$, since we have discarded all complex minima.

Note that when $\alpha = 0$ is used here, our structure corresponds to the system considered in [25], which is the simplest weighted PIC structure. The above solution is then a supplement to that scheme as an analytical method for obtaining the optimal real weighting factor in the sense of the MSE.

However, if the noise level, σ^2, is available at the receiver, we can set $\alpha = \sigma^2$ to yield a even better result. In this case,

$$\frac{\mathrm{d}J_{\mathrm{ex}}^{(m)}(\mu,\alpha=\sigma^2)}{\mathrm{d}\mu} = \sum_{k=1}^{K}(-2m)\lambda_k(1-\mu\phi_k)^{2m-1} = 0 \tag{4.55}$$

has potentially $(2m-1)$ valid solutions. However, by considering the second derivative,

$$\frac{\mathrm{d}^2J_{\mathrm{ex}}^{(m)}(\mu,\alpha=\sigma^2)}{\mathrm{d}\mu^2} = \sum_{k=1}^{K}2m(2m-1)\lambda_k\phi_k(1-\mu\phi_k)^{2(m-1)} > 0, \tag{4.56}$$

we observe that it is always positive and therefore, Eqn. (4.54) has only one real root which is the optimal real weight for this m-stage PIC.

In a later section 4.7, Theorem 4.2 shows that the optimal global minimum for $J_{\mathrm{ex}}^{(m)}(\mu,\alpha=\sigma^2)$ must be real, which also implies that this unique real root of Eqn. (4.55) we have obtained must be the unique global minimum. It is then proved in Theorem 4.4 that the α-value has no contribution to the minimum achievable value of the MSE, hence assuming that only real weighting factor is valid, the MSE achieved by our modified structure with $\alpha = \sigma^2$ is always no larger than that achieved by using $\alpha = 0$.

Also in section 4.7, Theorem 4.3 gives lower and upper bounds on the value of the weights, which is also applicable for this case when the same parameter is used in all stages. Newton's method can then be employed to locate this unique real minimum within the valid region.

4.6.2 Special Case of $\alpha = \sigma^2$ and $m \geq K$

Consider Eqn. (4.51) with $\alpha = \sigma^2$. We then have $J_{\mathrm{ex}}^{(m)}(\boldsymbol{\mu}) \triangleq J_{\mathrm{ex}}^{(m)}(\boldsymbol{\mu},\alpha=\sigma^2)$ expressed as

$$J_{\mathrm{ex}}^{(m)}(\boldsymbol{\mu}) = \sum_{k=1}^{K}\frac{\lambda_k}{\phi_k}\prod_{i=1}^{m}\left|1-\mu_i\phi_k\right|^2. \tag{4.57}$$

Assuming that $m \geq K$, we can write out (4.57) as

$$\begin{aligned} J_{\text{ex}}^{(m)}(\boldsymbol{\mu}) &= \frac{\lambda_1}{\phi_1}|\underline{1-\mu_1\phi_1}|^2|1-\mu_2\phi_1|^2\cdots|1-\mu_K\phi_1|^2\cdots|1-\mu_m\phi_1|^2 \\ &\quad +\frac{\lambda_2}{\phi_2}|1-\mu_1\phi_2|^2|\underline{1-\mu_2\phi_2}|^2\cdots|1-\mu_K\phi_2|^2\cdots|1-\mu_m\phi_2|^2 \\ &\qquad\qquad \vdots \qquad\qquad \cdots \qquad\qquad \vdots \\ &\quad +\frac{\lambda_K}{\phi_K}|1-\mu_1\phi_K|^2|1-\mu_2\phi_K|^2\cdots|\underline{1-\mu_K\phi_K}|^2\cdots|1-\mu_m\phi_K|^2. \end{aligned} \tag{4.58}$$

Obviously we can now make $J_{\text{ex}}^{(m)}(\boldsymbol{\mu})$ zero by selecting the parameters in such a way that the underlined factors in (4.58) are zero. It is therefore clear that we can reach the MMSE solution if we choose the parameters as

$$\mu_i = \frac{1}{\lambda_i + \sigma^2}, \quad i = 1, 2, ...K. \tag{4.59}$$

So on condition that we set $\alpha = \sigma^2$, the linear PIC needs exactly K stages to implement the MMSE detector.[4]

4.6.3 General Case

In a practical system the number of PIC stages will be significantly less than K due to limitations on the overall receiver complexity, hence it is desirable to find the global minimum of the excessive MSE $J_{\text{ex}}^{(m)}(\boldsymbol{\mu}, \alpha)$ given such a constraint.

It is found that $J_{\text{ex}}^{(m)}(\boldsymbol{\mu}, \alpha)$ can be reduced to a quadratic function on $\mathbb{C}^m$ through a mapping defined as:

Definition 4.1. *Mapping* $T: \mathbb{C}^m \to \mathbb{C}^m$ *is given by* $\mathbf{x} = T(\boldsymbol{\mu})$, *where* $\mathbf{x} = (x_1, x_2, \cdots, x_m)^\top$, $\boldsymbol{\mu} = (\mu_1, \mu_2, \cdots, \mu_m)^\top$ *and for* $i = 1, 2, \cdots, m$,

$$x_i = (-1)^i \sum_{1 \leq j_1 < j_2 < \cdots < j_i \leq m} \mu_{j_1}\mu_{j_2}\cdots\mu_{j_i}. \tag{4.60}$$

Corollary 4.1. *Assuming that* $\mathbf{x} = T(\boldsymbol{\mu})$, *we have*

$$\prod_{i=1}^{m}(1-\mu_i\gamma) = 1 + \sum_{i=1}^{m} x_i\gamma^i. \tag{4.61}$$

Proof. An expansion of Eqn. (4.60) gives

$$\begin{aligned} x_1 &= (-1)(\mu_1 + \mu_2 + \cdots + \mu_m) \\ x_2 &= (-1)^2(\mu_1\mu_2 + \mu_1\mu_3 + \cdots + \mu_{m-1}\mu_m) \\ &\vdots \qquad\qquad \cdots \qquad\qquad \vdots \\ x_m &= (-1)^m\mu_1\mu_2\cdots\mu_m. \end{aligned} \tag{4.62}$$

[4] In a later section 4.7.2, Theorem 4.4 implies that given knowledge of the noise level, σ^2, the MMSE detector can always be achieved exactly by a K-stage PIC regardless of the value of α.

Clearly,

$$\prod_{i=1}^{m}(1-\mu_i\gamma) \tag{4.63}$$

$$= 1+(-1)(\mu_1+\mu_2+\cdots+\mu_m)\gamma+(-1)^2(\mu_1\mu_2+\mu_1\mu_3+\cdots+\mu_{m-1}\mu_m)\gamma^2 + \cdots+(-1)^m\mu_1\mu_2\cdots\mu_m\gamma^m \tag{4.64}$$

$$= 1+x_1\gamma+x_2\gamma^2+\cdots+x_m\gamma^m \tag{4.65}$$

$$= 1+\sum_{i=1}^{m}x_i\gamma^i. \tag{4.66}$$

□

Making use of Corollary 4.1, the individual product in (4.52) is then

$$\prod_{i=1}^{m}(1-\mu_i\gamma_k) = 1+\sum_{i=1}^{m}x_i\gamma_k^i = 1+\gamma_k\boldsymbol{\gamma}_k^\top\mathbf{x}, \tag{4.67}$$

where $\boldsymbol{\gamma}_k \triangleq \left(1,\gamma_k,...,\gamma_k^{m-1}\right)^\top$ for $k=1,2,\cdots,K$. The excess MSE, $J_{\text{ex}}^{(m)}(\boldsymbol{\mu},\alpha)$, can then be expressed concisely as a function of $\mathbf{x}$, and α,

$$J_{\text{ex}}^{(m)}(\mathbf{x},\alpha) = \sum_{k=1}^{K}\lambda_k\phi_k\left|\frac{1}{\phi_k}+\boldsymbol{\gamma}_k{}^T\mathbf{x}\right|^2. \tag{4.68}$$

Note that $J_{\text{ex}}^{(m)}(\mathbf{x},\alpha)$ is a quadratic function in $\mathbf{x}$. Differentiating with respect to $\mathbf{x}$ and equating to zero gives

$$\frac{\partial J_{\text{ex}}^{(m)}(\mathbf{x},\alpha)}{\partial\mathbf{x}} = 0 \tag{4.69}$$

$$\Rightarrow \sum_{k=1}^{K}\lambda_k\phi_k\left[\frac{1}{\phi_k}+\boldsymbol{\gamma}_k^\top\mathbf{x}^*\right]\boldsymbol{\gamma}_k = 0 \tag{4.70}$$

$$\Rightarrow \left[\sum_{k=1}^{K}\lambda_k\phi_k\boldsymbol{\gamma}_k\boldsymbol{\gamma}_k^\top\right]\mathbf{x}^* = -\sum_{k=1}^{K}\lambda_k\boldsymbol{\gamma}_k \tag{4.71}$$

$$\Rightarrow \mathbf{C}\mathbf{x}^* = -\mathbf{p}. \tag{4.72}$$

Therefore the minimum in $\mathbf{x}$ satisfies (4.72). Clearly $\mathbf{C}$ and $\mathbf{p}$ are real. It is shown in Appendix A that there always exists a real $\mathbf{x}$ that satisfies (4.72), which gives the global minimum of $J_{\text{ex}}^{(m)}(\mathbf{x},\alpha)$. It is proved that $\mathbf{C}$ is positive definite if and only if m or more of the eigenvalues of $\mathbf{R}$ are distinct ($\lambda_i \neq \lambda_j$), which results in a unique solution $\hat{\mathbf{x}} = -\mathbf{C}^{-1}\mathbf{p}$. $\hat{\mathbf{x}}$ is then the global minimum of $J_{\text{ex}}^{(m)}(\mathbf{x},\alpha)$. If $\mathbf{R}$ has $l<m$ distinct eigenvalues, denoted by $(\lambda_{k_1},\lambda_{k_2},\cdots,\lambda_{k_l})$, $1\le k_1<k_2<\cdots<k_l\le m$, then $\mathbf{C}$ is positive semi-definite and any solution to the following equation

$$(\boldsymbol{\gamma}_{k_1},\boldsymbol{\gamma}_{k_2},\cdots,\boldsymbol{\gamma}_{k_l})^\top\mathbf{x} = -(\frac{1}{\phi_{k_1}},\frac{1}{\phi_{k_2}},\cdots,\frac{1}{\phi_{k_l}})^\top \tag{4.73}$$

will give the minimum of $J_{\text{ex}}^{(m)}(\mathbf{x}, \alpha) = 0$ in $\mathbf{x}$, which is obviously the MMSE solution. This is an indefinite group equation with more unknowns than the number of equations, which has infinite number of solutions, including complex ones. Since all the solutions lead to the same minimum value of $J_{\text{ex}} = 0$, it is sufficient to choose an arbitrary real solution by convenience, which can be defined as "the global minimum" indiscriminately. Therefore, we hereafter always limit the selection of $\mathbf{x}$ in the region of $\mathbb{R}^m$, which simplifies our study of the problem.

Given $\mathbf{x} = (x_1, x_2, \cdots, x_m)^\top \in \mathbb{R}^m$ as the global minimum, we need to find an equivalent minimum in $\boldsymbol{\mu}$ which satisfies the mapping $\mathbf{x} = T(\boldsymbol{\mu})$. The solutions are not unique and they are given by the following theorem.

Theorem 4.1. *For any* $\mathbf{x} = (x_1, x_2, \cdots, x_m)^\top \in \mathbb{R}^m$, $\boldsymbol{\mu} = (\mu_1, \mu_2, \cdots, \mu_m)^\top \in \mathbb{C}^m$ *satisfies* $T(\boldsymbol{\mu}) = \mathbf{x}$ *if and only if* $\mu_1, \mu_2, \cdots, \mu_m$ *are the* m *roots of the following polynomial,*

$$p(\mu) = \mu^m + x_1\mu^{m-1} + x_2\mu^{m-2} + \cdots + x_m. \tag{4.74}$$

Proof. $\mu_1, \mu_2, \cdots, \mu_m$ are the m roots of (4.74) is equivalent to

$$p(\mu) = (\mu - \mu_1)(\mu - \mu_2)\cdots(\mu - \mu_m) \tag{4.75}$$

to hold for any $\mu \in \mathbb{C}$, i.e.,

$$\mu^m + x_1\mu^{m-1} + x_2\mu^{m-2} + \cdots + x_m = (\mu - \mu_1)(\mu - \mu_2)\cdots(\mu - \mu_m). \tag{4.76}$$

It follows then that the coefficients on both sides of (4.76) are equal, which corresponds to the m equalities expressed as Eqn. (4.62). This is equivalent to

$$\mathbf{x} = T(\boldsymbol{\mu}) \tag{4.77}$$

by definition of $T(\cdot)$. □

It is clear that the mapping T is not one-to-one since we have multiple solutions for $T(\boldsymbol{\mu}) = \mathbf{x}$, i.e., $\boldsymbol{\mu}$ satisfies $T(\boldsymbol{\mu}) = \mathbf{x}$ if and only if $\boldsymbol{\mu} = (\mu_{i_1}, \mu_{i_2}, \cdots, \mu_{i_m})^\top$, where $\mu_1, \mu_2, \cdots, \mu_m$ are the m roots of the polynomial (4.74) and $i_1, i_2, \cdots, i_m$ is any permutation of $1, 2, \cdots, m$. In general there are $m!$ minima in $\boldsymbol{\mu}$.

Regardless of the ordering, the m roots of (4.74) are exactly the m weighting factors that will give the minimum achievable MSE at the last stage of an m-stage PIC. The ordering of weighting factors, which significantly affects performance in the intermediate stages, will be discussed in section 4.8.

4.7 Effect of α

In this section, we show that the achievable MMSE does not depend on the choice of α. However, 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 may be complex.

4.7.1 Effect of α on Weighting Factors

It is obvious that the global minimum $\hat{\boldsymbol{\mu}}$ must be a function of α. Following theorem tells us that for $\alpha = \sigma^2$, it is always true that the corresponding weighting factors are real.

Theorem 4.2. *The global minimum of $J_{\text{ex}}^{(m)}(\boldsymbol{\mu}, \alpha = \sigma^2)$ in $\boldsymbol{\mu}$ must be real.*

Proof. Consider Eqn. (4.52) with $\alpha = \sigma^2$. We then have $J_{\text{ex}}^{(m)}(\boldsymbol{\mu}) \triangleq J_{\text{ex}}^{(m)}(\boldsymbol{\mu}, \alpha = \sigma^2)$ expressed as

$$J_{\text{ex}}^{(m)}(\boldsymbol{\mu}) = \sum_{k=1}^{K} \frac{\lambda_k}{\phi_k} \prod_{i=1}^{m} |1 - \mu_i \phi_k|^2 . \tag{4.78}$$

Let $\boldsymbol{\mu} = \text{Re}\{\boldsymbol{\mu}\} + j \cdot \text{Im}\{\boldsymbol{\mu}\} = \boldsymbol{\mu}^{(r)} + j \cdot \boldsymbol{\mu}^{(i)}$ be the global minimum of $J_{\text{ex}}^{(m)}(\boldsymbol{\mu})$ in $\boldsymbol{\mu}$, where $\boldsymbol{\mu}^{(r)}, \boldsymbol{\mu}^{(i)} \in \mathbb{R}^m$. It is clear that

$$\begin{aligned}
J_{\text{ex}}^{(m)}(\boldsymbol{\mu}) &= \sum_{k=1}^{K} \frac{\lambda_k}{\phi_k} \prod_{i=1}^{m} \left| 1 - \mu_i^{(r)} \phi_k - j \cdot \mu_i^{(i)} \phi_k \right|^2 && (4.79)\\
&= \sum_{k=1}^{K} \frac{\lambda_k}{\phi_k} \prod_{i=1}^{m} \left[\left(1 - \mu_i^{(r)} \phi_k \right)^2 + \left(\mu_i^{(i)} \phi_k \right)^2 \right] && (4.80)\\
&\geq \sum_{k=1}^{K} \frac{\lambda_k}{\phi_k} \prod_{i=1}^{m} \left(1 - \mu_i^{(r)} \phi_k \right)^2 && (4.81)\\
&= J_{\text{ex}}^{(m)}(\boldsymbol{\mu}^{(r)}). && (4.82)
\end{aligned}$$

This indicates that the excess MSE achieved by $\boldsymbol{\mu}$ is always no less than that achieved by $\text{Re}\{\boldsymbol{\mu}\}$. For $\boldsymbol{\mu}$ to be the global minimum, both side of (4.82) must be equal. It is obvious that $\mu_i^{(i)} = 0$, i.e., $\boldsymbol{\mu} \in \mathbb{R}^m$. □

Furthermore, the upper and lower bounds of the optimal weights are given in the following theorem.

Theorem 4.3. *The global minimum of $J_{\text{ex}}^{(m)}(\boldsymbol{\mu}, \alpha = \sigma^2)$ in $\boldsymbol{\mu} = (\mu_1, \mu_2, \cdots, \mu_m)^\top$ satisfies*

$$\mu_{\min} = \frac{1}{\lambda_{\max} + \sigma^2} \leq \mu_j \leq \frac{1}{\lambda_{\min} + \sigma^2} = \mu_{\max}, \quad j = 1, 2, \cdots, m, \tag{4.83}$$

where $\lambda_{\min}$ and $\lambda_{\max}$ are the smallest and the largest eigenvalue respectively.

Proof. Let $\boldsymbol{\mu}$ be the global minimum of $J_{\text{ex}}^{(m)}(\boldsymbol{\mu}) = J_{\text{ex}}^{(m)}(\boldsymbol{\mu}, \alpha = \sigma^2)$. From Theorem 4.2 we have $\boldsymbol{\mu} \in \mathbb{R}^m$, such that

$$J_{\text{ex}}^{(m)}(\boldsymbol{\mu}) = \sum_{k=1}^{K} \frac{\lambda_k}{\phi_k} \prod_{i=1}^{m} (1 - \mu_i \phi_k)^2 . \tag{4.84}$$

Consider its derivative with respect to μ_j,

$$\begin{aligned}
\frac{\partial J_{\text{ex}}^{(m)}(\boldsymbol{\mu})}{\partial \mu_j} &= -2 \sum_{k=1}^{K} \lambda_k (1 - \mu_j \phi_k) \prod_{\substack{i=1 \\ j \neq i}}^{m} (1 - \mu_i \phi_k)^2 && (4.85)\\
&= -2 \sum_{k=1}^{K} \lambda_k (1 - \mu_j \phi_k) C_{k,j}^2 && (4.86)
\end{aligned}$$

where

$$C_{k,j} \triangleq \prod_{\substack{i=1 \\ j\neq i}}^{m}(1-\mu_i\phi_k). \tag{4.87}$$

It is clear that for any $\mu < \mu_{\min}$,

$$1-\mu\phi_k > 1-\mu_{\min}\phi_k = 1-\frac{\lambda_k+\sigma^2}{\lambda_{\max}+\sigma^2} = \frac{\lambda_{\max}-\lambda_k}{\lambda_{\max}+\sigma^2} \geq 0, \tag{4.88}$$

so that

$$\left.\frac{\partial J_{\text{ex}}^{(m)}(\boldsymbol{\mu})}{\partial \mu_j}\right|_{\mu_j=\mu} = -2\sum_{k=1}^{K}\lambda_k(1-\mu\phi_k)C_{k,j}^2 \leq 0, \tag{4.89}$$

which implies that $J_{\text{ex}}^{(m)}(\boldsymbol{\mu})$ is non-increasing in the region $(-\infty, \mu_{\min})$. Similarly, it can be shown that $J_{\text{ex}}^{(m)}(\boldsymbol{\mu})$ is non-decreasing in the region $(\mu_{\max}, \infty)$. In consequence the global minimum of $J_{\text{ex}}^{(m)}(\boldsymbol{\mu})$ falls in the region of $[\mu_{\min}, \mu_{\max}]$. □

If $\alpha \neq \sigma^2$, the above approaches are not applicable. The optimal weighting factors can now be both negative and complex, and the real weighting factors may be out of the region bounded by $\mu_{\min}$ and $\mu_{\max}$. This is not a serious problem since the structure in Fig. 4.2 can easily accommodate any complex weighting factors, although complex weights introduce 2-fold complexity for BPSK modulation, or 4-fold complexity for any quadrature modulation schemes.

If the optimal weights are found to be complex numbers, but for implementation simplicity only real parameters are wanted, they can be located in the constrained set

$$\mathbb{B}_{\boldsymbol{\mu}} = \{\, \boldsymbol{\mu} = (\mu_1, \mu_2, \cdots, \mu_m) \mid \exists\, 1 \leq i < j \leq m,\ \mu_i = \mu_j,\ \boldsymbol{\mu} \in \mathbb{R}^m \,\}. \tag{4.90}$$

As shown in Appendix B this is the set of all $\boldsymbol{\mu}$ described by the limit where potential complex conjugate pairs become real. Since there is less freedom in selecting the weighting factors (a pair of them are forced to be equal), there will be some MSE degradation in comparison to allowing complex parameters.

Unfortunately it is not easy to locate such weighting factors. The weighting factors are the solution to a group of m multi-variate polynomials,

$$\frac{\partial J_{\text{ex}}^{(m)}(\boldsymbol{\mu},\alpha)}{\partial \mu_j} = 0, \quad j = 1, 2, \cdots, m, \tag{4.91}$$

each with a degree of $(2m-1)$. The mapping T can not help this time since now we are looking for a constrained minimum which is not necessarily the minimum of quadratic function $J_{\text{ex}}^{(m)}(\mathbf{x}, \alpha)$, i.e., the global unconstrained minimum. It is however possible to locate the minimum through various searching methods such as Newton's method, but this is too complicated to be practical.

Regardless of α, the optimal weighting factors, whether real or complex, are still dependent on σ^2. Usually σ^2, the noise level, is not available at the receiver, hence an estimate is used instead. The sensitivity of the performance of the PIC to a mismatch in σ^2 is investigated through simulations in section 4.9.

4.7.2 Effect of α on Achievable MMSE

It is obvious that the global minimum $\hat{\mathbf{x}}$, which satisfies $\partial J_{\text{ex}}^{(m)}(\mathbf{x},\alpha)/\partial x_i = 0$ for $1 \le i \le m$, must be a function of α. We can therefore express the minimum achievable excess MSE as a function of α,

$$J_{\text{ex,min}}^{(m)}(\alpha) = \inf_{\mathbf{x}\in\mathbb{C}^m} J_{\text{ex}}^{(m)}(\mathbf{x},\alpha). \tag{4.92}$$

The following theorem tells us how α will affect the achievable MSE.

Theorem 4.4. *The minimum of $J_{\text{ex}}^{(m)}(\mathbf{x},\alpha)$ is independent of the value of α.*

Proof. From section 4.6.3 it is clear that regardless of the value of α, the minimum of $J_{\text{ex}}^{(m)}(\mathbf{x},\alpha)$ in $\mathbf{x}$ is always real, therefore it is sufficient to consider the constraint problem of $\mathbf{x} \in \mathbb{R}^m$. With this confinement, we have

$$J_{\text{ex}}^{(m)}(\mathbf{x},\alpha) = \sum_{k=1}^{K} \lambda_k \phi_k \left[\frac{1}{\phi_k} + {\boldsymbol{\gamma}_k}^T \mathbf{x}\right]^2. \tag{4.93}$$

We can obtain the derivative of $J_{\text{ex}}^{(m)}(\mathbf{x},\alpha)$ with respect to α as

$$\begin{aligned}
\frac{\partial J_{\text{ex}}^{(m)}(\mathbf{x},\alpha)}{\partial \alpha} &= 2\sum_{k=1}^{K} \lambda_k \phi_k \left(\frac{1}{\phi_k} + \boldsymbol{\gamma}_k^\top \mathbf{x}\right) \frac{\partial \boldsymbol{\gamma}_k^\top}{\partial \alpha}\mathbf{x} && (4.94)\\
&= 2\sum_{i=1}^{m-1} i x_{i+1} \sum_{k=1}^{K} \lambda_k \phi_k \left(\frac{1}{\phi_k} + \boldsymbol{\gamma}_k^\top \mathbf{x}\right) \gamma_k^{i-1}. && (4.95)
\end{aligned}$$

On the other hand, the partial derivative of $J_{\text{ex}}^{(m)}(\mathbf{x},\alpha)$ with respect to x_i is expressed as

$$\begin{aligned}
\frac{\partial J_{\text{ex}}^{(m)}(\mathbf{x},\alpha)}{\partial x_i} &= 2\sum_{k=1}^{K} \lambda_k \phi_k \left(\frac{1}{\phi_k} + \boldsymbol{\gamma}_k^\top \mathbf{x}\right) \boldsymbol{\gamma}_k^\top \frac{\partial \mathbf{x}}{\partial \mathbf{x}_i} && (4.96)\\
&= 2\sum_{k=1}^{K} \lambda_k \phi_k \left(\frac{1}{\phi_k} + \boldsymbol{\gamma}_k^\top \mathbf{x}\right) \gamma_k^{i-1}. && (4.97)
\end{aligned}$$

It follows then

$$\frac{\partial J_{\text{ex}}^{(m)}(\mathbf{x},\alpha)}{\partial \alpha} = \sum_{i=1}^{m-1} i x_{i+1} \frac{\partial J_{\text{ex}}^{(m)}(\mathbf{x},\alpha)}{\partial x_i}. \tag{4.98}$$

At the minimum of $J_{\text{ex}}^{(m)}(\mathbf{x},\alpha)$, its partial derivative to any x_i is zero, hence both sides of Eqn. (4.98) are zero. Therefore, at the minimum,

$$\frac{\partial J_{\text{ex}}^{(m)}(\mathbf{x},\alpha)}{\partial \alpha} = 0, \tag{4.99}$$

i.e., the specific α-value has no influence on the minimum achievable MSE. □

This theorem implies that for any α we can find a corresponding $\boldsymbol{\mu}$ that will give us the achievable MMSE.

4.8 Ordering of the Weights

Section 4.6 tells us how to compute the m weights that lead to the minimum of $J_{\text{ex}}^{(m)}(\boldsymbol{\mu}, \alpha)$ for a PIC detector. There are then $m!$ different choices of $\boldsymbol{\mu}$ which are permutations of the m weighting factors. The order in which the m parameters are applied however, has a significant influence on the BER performance as well as the MSE performance at intermediate stages. Depending on the desired behaviour for intermediate stages, different criteria for the ordering can be adopted. In this thesis we have chosen to order the optimal step sizes according to a recursive minimisation of the current stage MSE.

Denote the optimal set of weighting factors as $V = \{\mu^{(1)}, \mu^{(2)}, \cdots, \mu^{(m)}\}$. Assuming that the weights for the first $(i-1)$ stages have been chosen from V as $\hat{\boldsymbol{\mu}}_{i-1} = (\hat{\mu}_1, \hat{\mu}_2, \cdots, \hat{\mu}_{i-1})^\top$, we then choose the i^{th} weighting factor $\hat{\mu}_i$ from the remaining set, $V^{(i)} = V - \{\hat{\mu}_1, \hat{\mu}_2, \cdots, \hat{\mu}_{i-1}\}$, which minimises the current stage MSE, i.e.,

$$\hat{\mu}_i = \arg \min_{\mu_i \in V^{(i)}} J_{\text{ex}}^{(i)}(\boldsymbol{\mu}_i, \alpha) \tag{4.100}$$

where $\boldsymbol{\mu}_i = (\hat{\mu}_1, \hat{\mu}_2, \cdots, \hat{\mu}_{i-1}, \mu_i)^\top$ and

$$J_{\text{ex}}^{(i)}(\boldsymbol{\mu}_i, \alpha) = \sum_{k=1}^{K} \frac{\lambda_k \phi_k}{\gamma_k^2} \left| \frac{\sigma^2 - \alpha}{\phi_k} - (1 - \mu_i \gamma_k) \prod_{j=1}^{i-1} (1 - \hat{\mu}_j \gamma_k) \right|^2 . \tag{4.101}$$

Since all elements of $\boldsymbol{\mu}_i$ but μ_i have already been determined, the current stage MSE may be expressed as a function of variable $\mu_i \in \mathbb{C}$,

$$J_{\text{ex}}^{(i)}(\mu_i, \alpha) = \sum_{k=1}^{K} \frac{\lambda_k \phi_k}{\gamma_k^2} \left| \frac{\sigma^2 - \alpha}{\phi_k} - (1 - \mu_i \gamma_k) \cdot t_k^{(i)} \right|^2 , \tag{4.102}$$

where we define

$$t_k^{(i)} = \prod_{j=1}^{i-1} (1 - \hat{\mu}_j \gamma_k). \tag{4.103}$$

$J_{\text{ex}}^{(i)}(\mu_i, \alpha)$ is a quadratic function of μ_i, hence it has one unique minimum in $\mathbb{C}$ as

$$\tilde{\mu}_i = \arg \min_{\mu_i \in \mathbb{C}} J_{\text{ex}}^{(i)}(\mu_i, \alpha). \tag{4.104}$$

Differentiating $J_{\text{ex}}^{(i)}(\mu_i, \alpha)$ with respect to μ_i, we have

$$\frac{\partial J_{\text{ex}}^{(i)}(\mu_i, \alpha)}{\partial \mu_i} = \sum_{k=1}^{K} \frac{\lambda_k \phi_k}{\gamma_k} \left[\frac{\sigma^2 - \alpha}{\phi_k} \cdot t_k^{(i)} - (1 - \mu_i \gamma_k) \cdot \left| t_k^{(i)} \right|^2 \right]. \tag{4.105}$$

Equate the above to 0 and solve the linear equation to yield the minimum $\tilde{\mu}_i$ as

$$\tilde{\mu}_i = \frac{\sum_{k=1}^{K} \frac{\lambda_k}{\gamma_k} \left[\phi_k \cdot \left| t_k^{(i)} \right|^2 - (\sigma^2 - \alpha) \cdot t_k^{(i)} \right]}{\sum_{k=1}^{K} \lambda_k \phi_k \cdot \left| t_k^{(i)} \right|^2}. \tag{4.106}$$

Clearly $J_{\text{ex}}^{(i)}(\mu_i, \alpha)$ is a parabolic function of the distance between μ_i and $\tilde{\mu}_i$, thus we can select, in the limited set of remaining weighting factors, $V^{(i)}$, the closest one to $\tilde{\mu}_i$ as the best for the current stage,

$$\hat{\mu}_i = \arg \min_{\mu_i \in V^{(i)}} |\mu_i - \tilde{\mu}_i|. \tag{4.107}$$

This procedure is repeated until the last weighting factor is determined.

The ordering algorithm described above attempts to force the MSE to drop as fast as possible in the early stages. Simulation results in the next section show that this scheme may force the BER to decrease monotonically when $\alpha = \sigma^2$, but not necessarily so for $\alpha \neq \sigma^2$. As a rule of thumb, whenever simplicity is the utmost concern, ordering the weighting factors in increasing order in amplitude will in general lead to moderate behaviour in the intermediate stages.

4.9 Numerical Results

The numerical examples considered in this section are based on a symbol-synchronous system with 15 users and a processing gain of 31. BPSK modulation and spreading formats are assumed. A randomly selected set of short codes is used in all instances 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).

Fig. 4.3 and Fig. 4.4 demonstrate the stage-by-stage performance of PIC with optimised weighting factors. Perfect power control is assumed in Fig. 4.3 and the SNR of all users is 7 dB. The performance of the conventional PIC is also shown in this figure. Since the largest eigenvalue is greater than 2, the condition for convergence for the conventional PIC is violated, so that divergence as well as the ping-pong effect is observed. In Fig. 4.4, the SNR of the first user is 7 dB while the remaining 14 users have the same SNR of 13 dB. For the right choice of weighting factors, it is clear that a 15-stage PIC can achieve exactly the MMSE performance regardless of the received power distribution and the value of parameter α in the structure. 3-stage PIC shows considerable improvement over the conventional detector[5]. In Fig. 4.3, 5-stage PIC gives very close to MMSE performance while in Fig. 4.4, where the interfering users are much stronger, the PIC needs 7 stages to get as close to the MMSE performance. It can be concluded for a PIC with the number of stages less than K that the stronger the signals from interfering users are, the farther its performance is from that of the MMSE detector. This is more clearly shown in Fig. 4.5 and will be discussed later.

The weighting factors here are ordered according to the criterion described in section 4.8, 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 when $\alpha \neq \sigma^2$. In case the weighting factors are sorted in decreasing order, the BER performance fluctuates around 50% for all but the first and the 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. 4.3. Obviously the ordering of the weighting factors is vital for performance in the intermediate stages.

Fig. 4.5 demonstrates PIC's ability to combat near-far effects. The SNR of the first user is 7 dB while the remaining 14 users have the same SNR but ISR(dB) different to user 1.

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

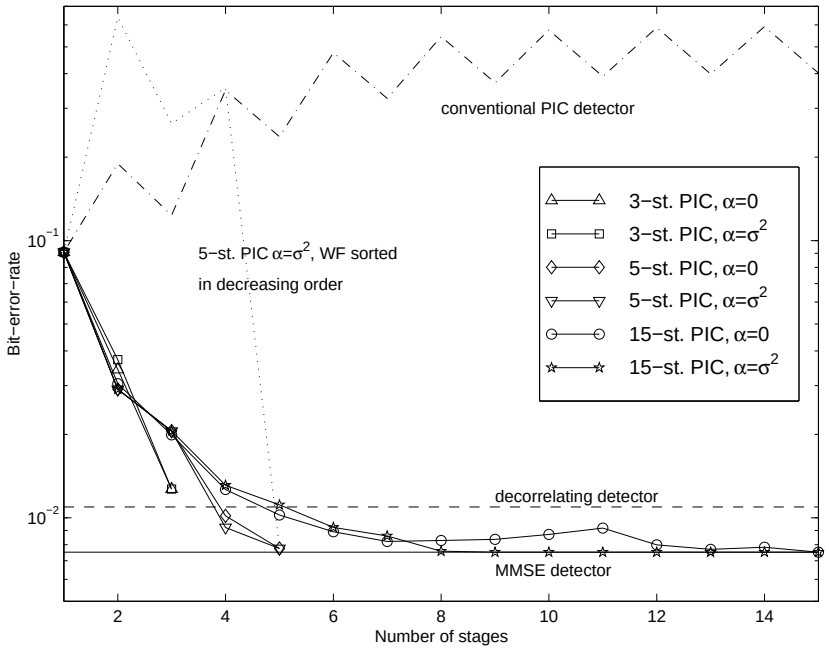

Figure 4.3: Stage-by-stage performance of PIC using short codes (perfect power control).

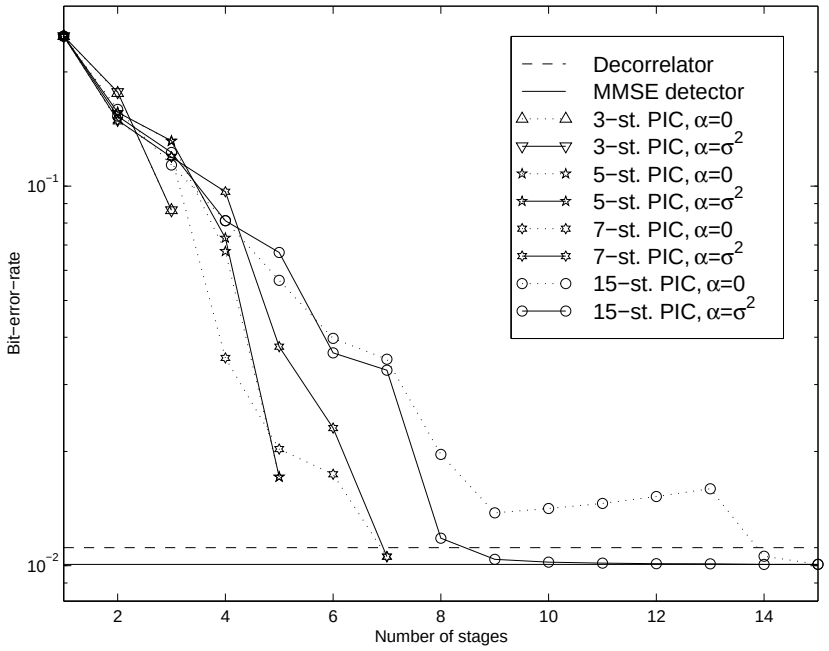

Figure 4.4: Stage-by-stage performance of PIC using short codes (ISR=6 dB).

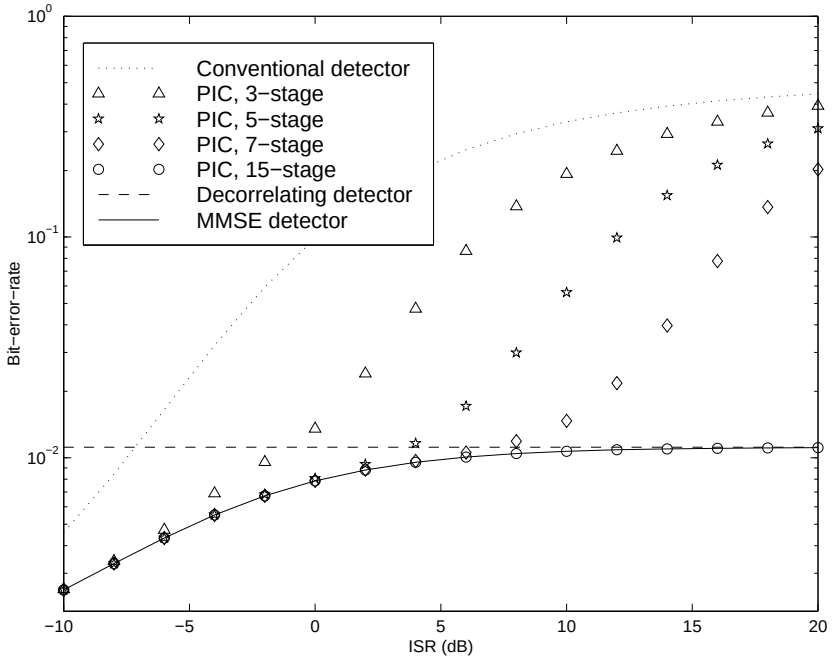

Figure 4.5: Near-far ability of the weighted linear PIC using short codes.

The weighted PIC is shown to behave better than the conventional detector in a near-far environment. The more number of stage deployed, the higher the MAI a PIC can resist. It is also instructive to compare Fig. 4.5 with Fig. 3.6. It is clear that given the same number of stages used, the weighted PIC is better than the conventional PIC in a near-far environment. Since a 15-stage weighted PIC achieves the MMSE detector exactly, it is at least as good as the decorrelating detector in combating near-far effect. However a PIC with fewer number of stages is not near-far resistant in the sense that when MAI becomes extremely strong, i.e., ISR $= \infty$, the detector fails to detect the signal from the desired user.

The weighting factors are determined based on a specific working SNR. The sensitivity of the detector performance at all SNRs ($0 \sim 14$ dB), to the estimation of the working SNR is illustrated in Fig. 4.6 and 4.7. Perfect power control is assumed. In Fig. 4.6 the weighting factors optimised for an SNR of 7 dB and $\alpha = 0.099763$ (which corresponds to a σ^2 at 7 dB) are used for various working 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. 4.7, where α is set to be 0. The system is observed to be practically insensitive 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

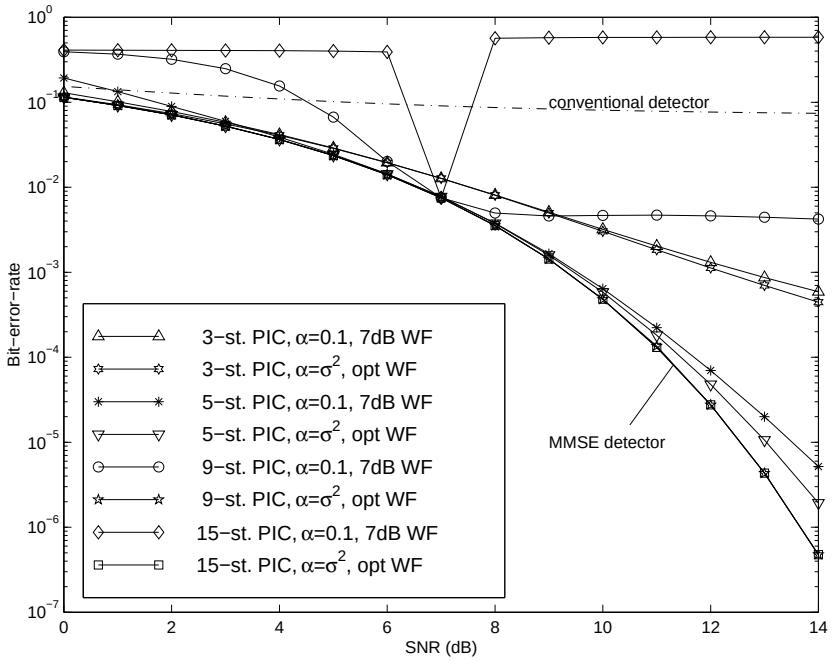

Figure 4.6: BER performance and sensitivity versus SNR using short codes. Weighting factors (WF) optimised for 7 dB and $\alpha = 0.099763$ (simplified to $\alpha = 0.1$ in the figure) are used for $0 \sim 14$ dB, in comparison to when weighting factors are optimised for each working SNR and corresponding α.

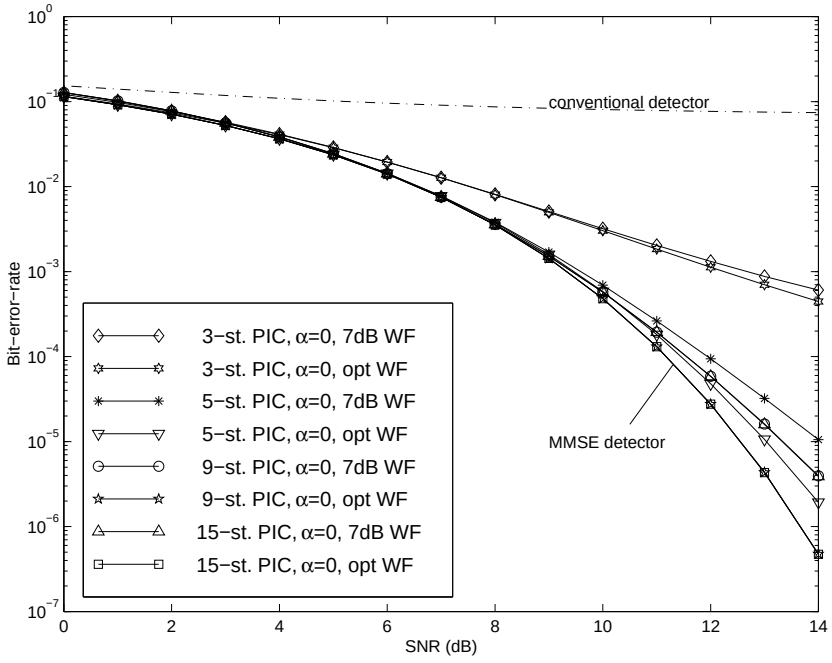

Figure 4.7: BER performance and sensitivity versus SNR using short codes. $\alpha = 0$ is assumed in all cases.Weighting factors (WF) for 7 dB are used for $0 \sim 14$ dB, in comparison to when weighting factors are optimised for each working SNR.

in general be complex.

4.10 Summary

In this chapter the weighted linear PIC scheme using short codes is mathematically described and analysed. The optimal set of weighting factors are found with respect to the minimum achievable MSE. Simulation shows significant improvement over the conventional detector as well as the conventional PIC.

Chapter 5

Extension to Long-code Systems

In long-code CDMA, the spreading codes as well as the corresponding correlation matrix will change for every symbol interval. Hence the optimal set of weighting factors given in Chapter 4, which depends on eigenvalues of the correlation matrix, have to be computed and updated symbol by symbol. As the solution to the optimal weighting factors involves eigenvalue decomposition, which has a complexity cubic to the number of users, such an approach is no better than implementing the MMSE detector directly. Instead we would like to consider using a fixed set of weighting factors that will work as well for random codes. This is possible because the correlation matrix of a CDMA channel is highly structured and centred around the identity matrix and as a consequence its eigenvalues are densely distributed. Furthermore, a rough estimation of the inverse of the correlation matrix can be sufficient in achieving good performance in multiuser detection. In other words, utilising knowledge of the statistical properties of the channel correlation matrix, we expect the PIC with a fixed set of compromised weights to give a fairly good estimate of the inverse of the matrices for a majority of all possible code-sets.

Different criteria in making the compromise over all code-sets can be adopted here and may reach different solutions. The author has chosen to minimise the ensemble average of the MSE over all possible channel matrices. The reason for this choice is that the MSE is always a quadratic function of filter taps, which has a unique global minimum. Again, this is not the best solution in the BER sense, but a reasonably close one. Following the approach in Chapter 4, we demonstrate in section 5.1 that a unique optimal set of fixed weights exists, which will give the minimum ensemble average of the MSE at the last stage of a multistage weighted PIC. For an m-stage PIC, the weights depend on the first $2m$ moments of the eigenvalues of channel correlation matrix.

Mathematical analysis of the eigenvalue distribution is conducted in section 5.2. Related work can be seen in [38] and [39], where the authors resorted to the asymptotic expression of the eigenvalue distribution as the size of the multiuser system goes to infinity. In this thesis we demonstrate a method for deriving the exact expressions for the moments of the eigenvalues, i.e., the moments of the correlation matrix, to use the terminology of [40]. The moments are found to be polynomials that may be easily evaluated given the processing gain, the number of active users and the received signal energies.

In section 5.3 the algorithm for computing the optimal weighting factors that achieve the minimum ensemble average of the MSE over random codes is summarised. The computational complexity of on-line updating these parameters is proportional to the number of users but

independent of the processing gain.

Numerical results are presented in section 5.4 and section 5.5 concludes this chapter.

5.1 Weighted Linear PIC for Long-code Systems

The structure of a weighted linear PIC working under long codes is no different to that working under short codes as shown in Fig. 4.1 and Fig. 4.2. For each symbol interval, the matched filtering waveforms represented by $\mathbf{a}_k$ in the structure is identical to the spreading waveforms used by the transmitter of the corresponding user.

5.1.1 Optimising the Weights

The excess MSE of an m-stage PIC has been derived in section 4.5 as

$$J_{\text{ex}}^{(m)}(\boldsymbol{\mu},\alpha) = \sum_{k=1}^{K} \frac{\lambda_k(\lambda_k+\sigma^2)}{(\lambda_k+\alpha)^2} \left| \frac{\sigma^2-\alpha}{\lambda_k+\sigma^2} - \prod_{i=1}^{m}(1-\mu_i(\lambda_k+\alpha)) \right|^2, \tag{5.1}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_K$ are the eigenvalues of a sample channel correlation matrix, $\mathbf{R} = \mathbf{A}^{\text{H}}\mathbf{A}$. The ensemble average of the excess MSE over random codes is defined as

$$\mathcal{J}_{\text{ex}}^{(m)}(\boldsymbol{\mu},\alpha) = \text{E}\left\{ J_{\text{ex}}^{(m)}(\boldsymbol{\mu},\alpha) \right\}. \tag{5.2}$$

where the expectation is taken over the probability density function (PDF) of the spreading code matrix $\mathbf{A}$. It is obviously a function of $\boldsymbol{\mu}$ and α and independent of the individual $\mathbf{A}$.

Define ϕ_k, γ_k and $\boldsymbol{\gamma}_k$ the same way as in section 4.6.3. It is clear that

$$\mathcal{J}_{\text{ex}}^{(m)}(\boldsymbol{\mu},\alpha) = \text{E}\left\{ \sum_{k=1}^{K} \frac{\lambda_k\phi_k}{\gamma_k^2} \left| \frac{\sigma^2-\alpha}{\phi_k} - \prod_{i=1}^{m}(1-\mu_i\gamma_k) \right|^2 \right\}. \tag{5.3}$$

Following the approach in section 4.6, we have an equivalent representation of the excess MSE in $\mathbf{x} = T(\boldsymbol{\mu})$ as

$$\mathcal{J}_{\text{ex}}^{(m)}(\mathbf{x},\alpha) = \text{E}\left\{ \sum_{k=1}^{K} \lambda_k\phi_k \left| \frac{1}{\phi_k} + \boldsymbol{\gamma}_k{}^T\mathbf{x} \right|^2 \right\}. \tag{5.4}$$

Since $\text{E}\{\cdot\}$ is a linear operator, it may exchange position with any other linear operators, such as trace, differentiation, and etc. All derivations of the unique minimum remains the same as for short codes only with all the coefficients replaced by their ensemble averaged values. Differentiating $\mathcal{J}_{\text{ex}}^{(m)}(\mathbf{x},\alpha)$ with respect to $\mathbf{x}$ and equating to zero yields the equation

$$\mathbf{C}\mathbf{x}^* = -\mathbf{p} \tag{5.5}$$

where

$$\mathbf{C} = \text{E}\left\{ \sum_{k=1}^{K} \lambda_k\phi_k\boldsymbol{\gamma}_k\boldsymbol{\gamma}_k^{\top} \right\} \tag{5.6}$$

and

$$\mathbf{p} = \mathrm{E}\left\{\sum_{k=1}^{K} \lambda_k \boldsymbol{\gamma}_k\right\}. \tag{5.7}$$

Obviously $\mathbf{C}$ is real and positive definite, hence it is clear that the minimum point of $\mathcal{J}_{\mathrm{ex}}^{(m)}(\mathbf{x}, \alpha)$ in $\mathbf{x}$ is determined by

$$\mathbf{x} = -\mathbf{C}^{-1}\mathbf{p} \quad \in \quad \mathbb{R}^m. \tag{5.8}$$

The optimal weighting factors are then the roots of an m^{th} order polynomial (same as Eqn. (4.74)) with $\mathbf{x}$ as its coefficients.

The problem now turns out to be the computation of all elements in $\mathbf{C}$ and $\mathbf{p}$. It is not difficult to obtain good numerical estimates based on Monte Carlo averaging over random codes. But such an approach entails very high computational complexity and therefore can only be carried out off-line. Furthermore, its dependence on the received energies as well as the number of active users renders such a method useless in a real dynamic system. It is then desirable to obtain analytical expressions for $\mathbf{C}$ and $\mathbf{p}$ so that they can be evaluated with moderate computational requirements.

5.1.2 Moment of Correlation Matrix

It is helpful here to define the moment of a correlation matrix. It is some kind of measurement of the distribution of the eigenvalues of sample correlation matrices.

Definition 5.1. *The r^{th} order moment of the correlation matrix is defined as*

$$M_r = \mathrm{E}\left\{\frac{1}{K}\sum_{k=1}^{K} (\lambda_k)^r\right\} \tag{5.9}$$

where $\lambda_1, \lambda_2, \cdots, \lambda_K$ are eigenvalues of a sample correlation matrix.

Expanding (5.8) yields

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = -\begin{bmatrix} C_2 & C_3 & \cdots & C_{m+1} \\ C_3 & C_4 & \cdots & C_{m+2} \\ \vdots & \vdots & & \vdots \\ C_{m+1} & C_{m+2} & \cdots & C_{2m} \end{bmatrix}^{-1} \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_m \end{bmatrix}, \tag{5.10}$$

where for $i = 2, 3, \cdots, 2m$ and $j = 1, 2, \cdots, m$,

$$C_i = \frac{1}{K}\mathrm{E}\left\{\sum_{k=1}^{K} \lambda_k \phi_k \gamma_k^{i-2}\right\} = \sum_{l=2}^{i} \binom{i-2}{l-2} \alpha^{i-l}\left(M_l + \sigma^2 M_{l-1}\right), \tag{5.11}$$

$$p_j = \frac{1}{K}\mathrm{E}\left\{\sum_{k=1}^{K} \lambda_k \gamma_k^{j-1}\right\} = \sum_{l=1}^{j} \binom{j-1}{l-1} \alpha^{j-l} M_l. \tag{5.12}$$

Obviously the coefficients of $\mathbf{C}$ and $\mathbf{p}$ are determined by the first $2m$ moments of the correlation matrix.

5.1.3 Ordering the Weights

The ordering of weighting factors is similar to the procedure described in section 4.8.

The ideal minimum at the i^{th} stage, $\tilde{\mu}_i$, can be expressed as

$$\tilde{\mu}_i = \frac{\mathrm{E}\left\{\sum_{k=1}^{K} \frac{\lambda_k}{\gamma_k}\left[\phi_k \cdot \left|t_k^{(i)}\right|^2 - (\sigma^2-\alpha)\cdot t_k^{(i)}\right]\right\}}{\mathrm{E}\left\{\sum_{k=1}^{K} \lambda_k\phi_k \cdot \left|t_k^{(i)}\right|^2\right\}} \tag{5.13}$$

where $t_k^{(i)}$ is given by

$$t_k^{(i)} = \prod_{j=1}^{i-1}(1-\hat{\mu}_j\gamma_k) = 1 + \sum_{j=1}^{i-1} x_j\gamma_k^j = 1 + \gamma_k \mathbf{x}_{i-1}^{\top}\boldsymbol{\gamma}_k^{(i-1)}, \tag{5.14}$$

in which $\mathbf{x}_{i-1} \triangleq T\left(\hat{\boldsymbol{\mu}}_{i-1}\right)$ and $\boldsymbol{\gamma}_k^{(i-1)} \triangleq \left(1,\gamma_k,...,\gamma_k^{i-2}\right)^{\top}$.

In the following we briefly show that $\tilde{\mu}_i$ can also be evaluated based on the moments of the correlation matrix.

Clearly $t_k^{(i)}$ is a polynomial in λ_k, and so is $\left|t_k^{(i)}\right|^2$. They have a degree of $(i-1)$ and $(2i-2)$ respectively. The denominator in Eqn. (5.13) is then an expectation of a polynomial in λ_k, which depends only on the first $2i$ moments of the correlation matrix.

Since $\sigma^2-\alpha = \phi_k - \gamma_k$, the numerator in Eqn. (5.13) is the expectation of

$$\sum_{k=1}^{K}\frac{\lambda_k}{\gamma_k}\left[\phi_k\left|t_k^{(i)}\right|^2 - (\phi_k-\gamma_k)\cdot t_k^{(i)}\right] = \sum_{k=1}^{K}\lambda_k\left[\phi_k\cdot\frac{\left|t_k^{(i)}\right|^2 - t_k^{(i)}}{\gamma_k} + t_k^{(i)}\right]. \tag{5.15}$$

Note that

$$\begin{aligned}
\left|t_k^{(i)}\right|^2 - t_k^{(i)} &= \left|1+\gamma_k\mathbf{x}_{i-1}^{\top}\boldsymbol{\gamma}_k^{(i-1)}\right|^2 - (1+\gamma_k\mathbf{x}_{i-1}^{\top}\boldsymbol{\gamma}_k^{(i-1)}) && (5.16)\\
&= 1 + 2\gamma_k\mathrm{Re}\left\{\mathbf{x}_{i-1}^{\top}\right\}\boldsymbol{\gamma}_k^{(i-1)} + \gamma_k^2\left|\mathbf{x}_{i-1}^{\top}\boldsymbol{\gamma}_k^{(i-1)}\right|^2 - 1 - \gamma_k\mathbf{x}_{i-1}^{\top}\boldsymbol{\gamma}_k^{(i-1)} && (5.17)\\
&= \gamma_k\left[\left(\mathbf{x}_{i-1}^{\top}\boldsymbol{\gamma}_k^{(i-1)}\right)^* + \gamma_k\left|\mathbf{x}_{i-1}^{\top}\boldsymbol{\gamma}_k^{(i-1)}\right|^2\right], && (5.18)
\end{aligned}$$

which has a factor of γ_k, hence (5.15) is also a polynomial in λ_k. The numerator in Eqn. (5.13) is then the expectation of a polynomial in λ_k with a degree of $2i$, which can also be evaluated based on the first $2i$ moments.

In conclusion, at the i^{th} stage $\tilde{\mu}_i$ can be determined by the first $2i$ moments of the correlation matrix.

5.2 Derivation of Moments

In the previous section we introduced a set of weighting factors for a linear PIC working under long codes. They are optimal in the sense of the ensemble average of the MSE. The weighting factors are found to be determined by moments of the correlation matrix. In this section we derive analytical expressions for the moments so that they can be evaluated without resorting to Monte Carlo averaging.

Consider the n^{th} chip of user k's spreading waveform as a random variable, denoted by S_{nk}. The corresponding chip waveform observed at the receiver may be expressed as $A_{nk} = \sqrt{w_k} S_{nk}$. For a long-code system all the chips A_{nk}, $n = 1, 2, \cdots, N$, $k = 1, 2, \cdots, K$, are mutually independent zero-mean random variables, i.e.,

$$\begin{aligned} \mathrm{E}\left\{A^*_{n_1 k_1} A_{n_2 k_2}\right\} &= \frac{\sqrt{w_{k_1} w_{k_2}}}{N} \delta(n_1 - n_2) \cdot \delta(k_1 - k_2) \\ &= \begin{cases} w_{k_1}/N & \text{if } n_2 = n_1 \text{ and } k_2 = k_1, \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \tag{5.19}$$

The cross-correlation between user i and user j is also a random variable

$$R_{ij} = \mathbf{a}_i^{\mathrm{H}} \mathbf{a}_j = \sum_{n=1}^{N} A^*_{ni} A_{nj}, \tag{5.20}$$

which is also the element in the i^{th} row and the j^{th} column of the correlation matrix $\mathbf{R}$. A_{nk} and R_{ij} are in capital form because they are random variables.

Starting from Definition (5.1),

$$M_r = \mathrm{E}\left\{\frac{1}{K}\sum_{k=1}^{K}(\lambda_k)^r\right\} = \frac{1}{K}\mathrm{E}\left\{\mathrm{tr}\left\{\mathbf{\Lambda}^r\right\}\right\} = \frac{1}{K}\mathrm{E}\left\{\mathrm{tr}\left\{\mathbf{R}^r\right\}\right\}, \tag{5.21}$$

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

$$\mathrm{tr}\left\{\mathbf{R}^r\right\} = \sum_{k_1=1}^{K} (\mathbf{R}^r)_{k_1 k_1} \tag{5.22}$$

$$= \sum_{k_1=1}^{K} (\mathbf{R} \cdot \mathbf{R}^{r-1})_{k_1 k_1} \tag{5.23}$$

$$= \sum_{k_1=1}^{K} \sum_{k_2=1}^{K} R_{k_1 k_2} (\mathbf{R}^{r-1})_{k_2 k_1} \tag{5.24}$$

$$= \cdots \quad \cdots \tag{5.25}$$

$$= \sum_{k_1=1}^{K} \sum_{k_2=1}^{K} \cdots \sum_{k_r=1}^{K} R_{k_1 k_2} R_{k_2 k_3} \cdots R_{k_{r-1} k_r} R_{k_r k_1}. \tag{5.26}$$

Therefore,

$$M_r = \frac{1}{K} \sum_{k_1=1}^{K} \sum_{k_2=1}^{K} \cdots \sum_{k_r=1}^{K} \mathrm{E}\left\{R_{k_1 k_2} R_{k_2 k_3} \cdots R_{k_{r-1} k_r} R_{k_r k_1}\right\} \tag{5.27}$$

$$= \frac{1}{K} \sum_{k_1=1}^{K} \sum_{k_2=1}^{K} \cdots \sum_{k_r=1}^{K} \cdot \sum_{n_1=1}^{N} \sum_{n_2=1}^{N} \cdots \sum_{n_r=1}^{N} \mathrm{E}\left\{A_{n_1 k_1} A^*_{n_1 k_2} A_{n_2 k_2} A^*_{n_2 k_3} \cdots A_{n_r k_r} A^*_{n_r k_1}\right\}. \tag{5.28}$$

Here A_{nk} are independent random variable selected from a q-ary PSK[1] constellation with equal probability, thus only terms containing all complex conjugate pairs of the variables A_{nk} are relevant. The expectation taken over all code-sets implies that M_r has nothing to do with the particular spreading codes used in each symbol interval, rather, it depends on N, K, and the received signal energies. It is then possible to obtain M_r through evaluation of all combinations of indices.

Note, however,

$$\mathrm{E}\left\{(A_{nk})^r\right\} = \left(\frac{w_k}{N}\right)^{\frac{r}{2}} \sum_{m=0}^{q-1} e^{2\pi m r/q} = \begin{cases} \left(\frac{w_k}{N}\right)^{\frac{r}{2}} & \text{if } r/q \in \mathbb{Z}, \\ 0 & \text{otherwise.} \end{cases} \tag{5.29}$$

so that $\mathrm{E}\left\{A_{nk}^2\right\}$ would be w_k/N for BPSK but 0 for QPSK spreading, hence M_r would in general be different for different spreading schemes.

It is very hard (if not impossible) to reach a general expression for M_r, $r = 1, 2, \cdots$. However, given r as a determined positive integer, we start from (5.28) and discuss all combinations of k_1 to k_r as well as n_1 to n_r. After some elementary albeit very tedious manipulations, we may get an expression for M_r. This procedure involves a "pass-through" of a tree-like structure.

5.2.1 Tree-like Structure

For implementing an m-stage PIC, we need to derive expressions for M_1, M_2, $\cdots$, M_{2m} respectively. In other words, given any $1 \leq r \leq 2m$, we would like to have an expression for M_r, where K, N, and the received signal energies are treated as variables.

Consider Eqn. (5.28), where $1 \leq k_1, k_2, \cdots, k_r \leq K$ are the indices involved in derivation of M_r. Since K is a variable, it is not easy to exhaust all combinations of indices. However, for any realisation of the indices, $1 \leq k_1, k_2, \cdots, k_r \leq K$, we can classify the r indices into $s \leq r$ equivalence classes as $x_1, x_2, \cdots, x_s$ assuming numerical equality as the equivalence relation. Hence all indices in each class are equal, while indices from different classes are different. It is then possible to exhaust all combinations of equivalence classes since r is not a variable, but a known integer. The moment is now a summation of many terms where each term has the property that the equality pattern of all indices is determined, i.e., for any $1 \leq i < j \leq r$, either k_i and k_j (n_i and n_j) are equal or not, should be determined.

This classification is better illustrated in Fig. 5.1 by a tree-like structure. A 4-level summation is expanded into a hierarchical structure. Each node in layer i represents a summation over k_i, on condition of certain equivalence relations among the first i indices. The summand in the parentheses is then further divided into a summation of its child-nodes in layer $(i+1)$. These child-nodes exhaust all possible values of k_{i+1}. Their summation equates the summand represented by their parent-node. The number above each node denotes the number of equivalence classes constituted by the indices up-to the layer of that particular node. It can also be recognised as the degree of freedom of these indices. Clearly, a node with a degree of j will have $(j+1)$ children, among which the first j children will inherit a freedom of j while the last child will have a freedom of $(j+1)$.

[1]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.

Take for example the leaf connected to the root via 3 thick-lined branches in the figure. The summand represented by this particular leaf may be expressed as

$$\sum_{k_1=1}^{K} \sum_{\substack{k_2=1 \\ k_2 \neq k_1}}^{K} \sum_{\substack{k_3=1 \\ k_3 \neq k_1 \\ k_3 \neq k_2}}^{K} \sum_{k_4=k_2}^{k_2} \mathrm{E}\left\{R_{k_1 k_2} R_{k_2 k_3} R_{k_3 k_4} R_{k_4 k_1}\right\}. \tag{5.30}$$

Here k_1, k_2, k_3 and k_4 have been divided into 3 equivalence classes, i.e., $x_1 = \{k_1\}$, $x_2 = \{k_2, k_4\}$ and $x_3 = \{k_3\}$. In consequence this leaf has a degree of 3. This leaf, together with its 3 brothers, constitutes its parent-node, which has no restriction on the value of k_4.

This tree-like structure can be easily extended to any order. All the leaves of the tree then exhaust all combinations of the equivalence classes. It is then straightforward to derive moments of any order through exploiting this tree-like structure.

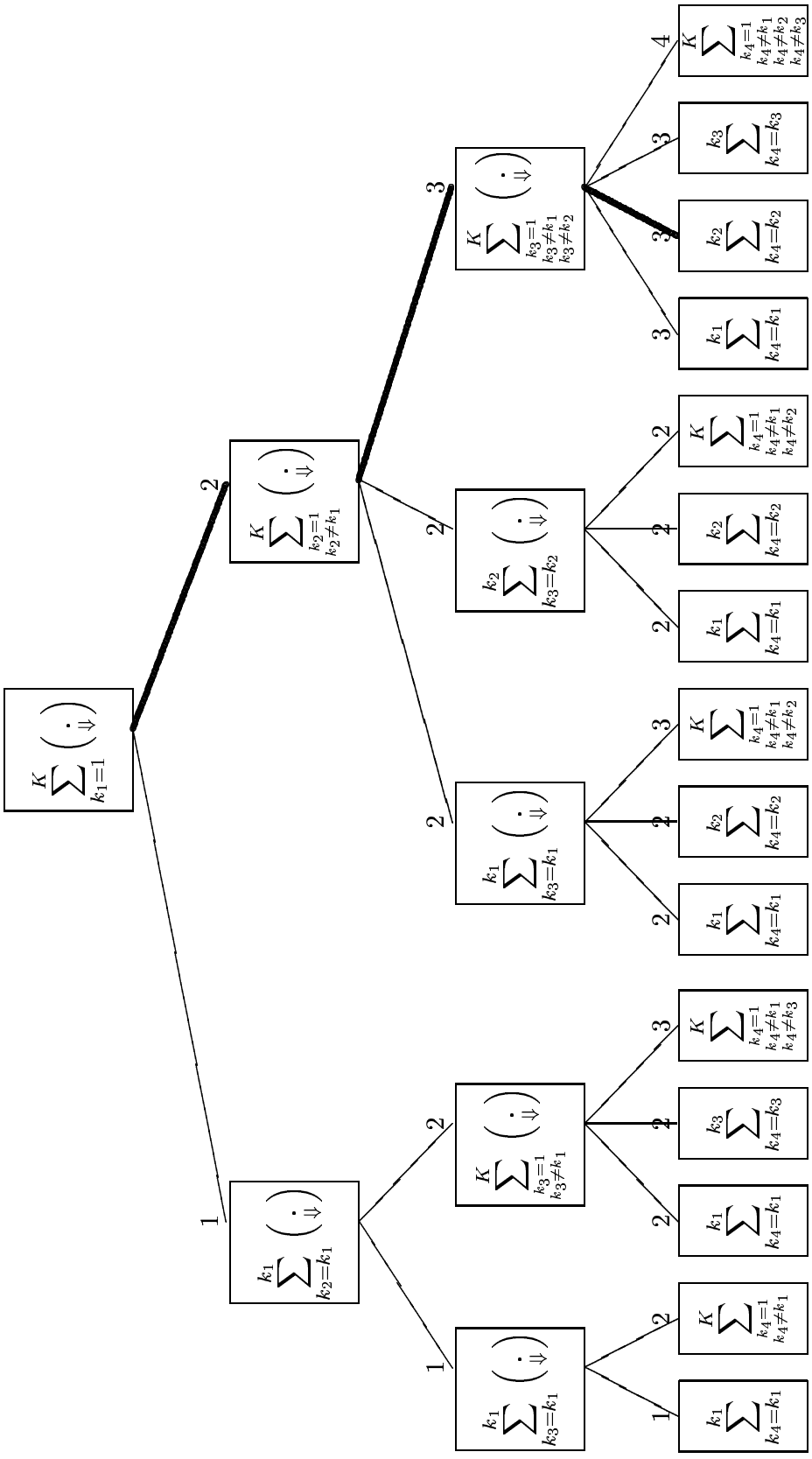

Figure 5.1: Expansion of multi-level summation as a tree-like structure.

5.2.2 Symbolic Manipulations

Before proceeding, we define the r^{th} order moment of the received signal energies.

Definition 5.2. *The r^{th} order moment of the received signal energies is defined as*

$$\mathcal{E}_r = \sum_{k=1}^{K} (w_k)^r. \tag{5.31}$$

Note that for the simplest case of perfect power control, i.e., $w_1 = w_2 = \cdots = w_K = 1$, we have $\mathcal{E}_r = K$ for $r = 1, 2, \cdots$.

In this subsection we introduce an example to show the derivation of an expression for the second order moment (M_2) by hand. Following the same principles it is possible to obtain analytical expressions for moments of any order. In fact the r^{th} order moment M_r is shown to be a polynomial in N (the processing gain), K (the number of active users), as well as the first r moments of the received signal energies.

Starting from (5.28), we can express the 2nd order moment of the correlation matrix as

$$M_2 = \frac{1}{K} \sum_{k_1=1}^{K} \sum_{k_2=1}^{K} \mathrm{E}\{R_{k_1k_2} R_{k_2k_1}\} \tag{5.32}$$

$$= \frac{1}{K} \left[\sum_{k_1=1}^{K} \sum_{k_2=k_1}^{k_1} \mathrm{E}\{R_{k_1k_2} R_{k_2k_1}\} + \sum_{k_1=1}^{K} \sum_{\substack{k_2=1 \\ k_2 \neq k_1}}^{K} \mathrm{E}\{R_{k_1k_2} R_{k_2k_1}\} \right]. \tag{5.33}$$

M_2 is now a sum of 2 terms, where in the first term $k_2 = k_1$, while in the second term $k_2 \neq k_1$. The expectation in each of the 2 terms can then be determined by the relation of the indices involved but regardless of their individual value. The first expectation term is simply

$$\mathrm{E}_{k_2=k_1}\{R_{k_1k_2} R_{k_2k_1}\} = \mathrm{E}\left\{R_{k_1k_1}^2\right\} = \mathrm{E}\left\{ \left(\sum_{n=1}^{N} |A_{nk_1}|^2 \right)^2 \right\}$$

$$= \left(\sum_{n=1}^{N} \frac{1}{N} w_{k_1} \right)^2 = w_{k_1}^2 \tag{5.34}$$

and the second one is

$$\mathrm{E}_{k_2 \neq k_1}\{R_{k_1k_2} R_{k_2k_1}\} = \sum_{n_1=1}^{N} \sum_{n_2=1}^{N} \mathrm{E}\left\{A_{n_1k_1}^* A_{n_1k_2} A_{n_2k_2}^* A_{n_2k_1}\right\}. \tag{5.35}$$

Note that here $R_{k_1k_2} R_{k_2k_1}$ is further expressed by a 2-level summation over n_1 and n_2, which can be split into 2 terms according to the "equality pattern" of the indices n_1 and n_2, i.e.,

$$\mathrm{E}_{k_2 \neq k_1}\{R_{k_1k_2} R_{k_2k_1}\} = \left[\sum_{n_1=1}^{N} \sum_{n_2=n_1}^{n_1} + \sum_{n_1=1}^{N} \sum_{\substack{n_2=1 \\ n_2 \neq n_1}}^{N} \right] \mathrm{E}\left\{A_{n_1k_1}^* A_{n_1k_2} A_{n_2k_2}^* A_{n_2k_1}\right\} \tag{5.36}$$

The second summand in (5.36) is obviously zero since it is the expectation of a product of 4 zero-mean independent random variables. Therefore

$$\mathrm{E}_{k_2\neq k_1}\left\{R_{k_1k_2}R_{k_2k_1}\right\} = \sum_{n_1=1}^{N}\sum_{n_2=n_1}^{n_1}\mathrm{E}\left\{A^*_{n_1k_1}A_{n_1k_2}A^*_{n_2k_2}A_{n_2k_1}\right\} \tag{5.37}$$

$$= \sum_{n_1=1}^{N}\mathrm{E}\left\{|A_{n_1k_1}|^2|A_{n_1k_2}|^2\right\} \tag{5.38}$$

$$= \sum_{n_1=1}^{N}\frac{1}{N^2}w_{k_1}w_{k_2} \tag{5.39}$$

$$= \frac{1}{N}w_{k_1}w_{k_2}. \tag{5.40}$$

Substituting (5.34) and (5.40) back into (5.33), we have

$$M_2 = \frac{1}{K}\left[\sum_{k_1=1}^{K}w_{k_1}^2 + \sum_{k_1=1}^{K}\sum_{\substack{k_2=1\\k_2\neq k_1}}^{K}\frac{1}{N}w_{k_1}w_{k_2}\right] \tag{5.41}$$

$$= \frac{1}{K}\left[\mathcal{E}_2 + \frac{1}{N}\sum_{k_1=1}^{K}w_{k_1}\left(\sum_{k_2=1}^{K}w_{k_2} - w_{k_1}\right)\right] \tag{5.42}$$

$$= \frac{1}{K}\left[\mathcal{E}_2 + \frac{1}{N}(\mathcal{E}_1^2 - \mathcal{E}_2)\right] \tag{5.43}$$

$$= \frac{1}{KN}[\mathcal{E}_1^2 + (N-1)\mathcal{E}_2]. \tag{5.44}$$

Using the same techniques it is possible to derive moments of any order. The derivation of moments becomes however more and more involved as the order increases. The reason is that the number of terms in the expansion of M_r, denoted by $\mathcal{N}(r)$, increases too fast with r. As stated in section 5.2.1, the evolution of the tree-like structure follows this rule: A node of degree j fathers $(j+1)$ child-nodes, among which j nodes inherit a degree of j while the remaining one has a degree of $(j+1)$. Although a general expression for the number of nodes in each layer may be too complicated to be useful here, it is not difficult to calculate those for the first few layers, which are 1, 2, 5, 15, 52, 203, 877, 4140, 21147, 115975, $\cdots$ respectively. To show the trend of the series, we plot the number of terms of up to the 20th layer in Fig. 5.2. From the curve in Fig. 5.2(a), it is observed that $\mathcal{N}(r)$ increases faster than exponentially. Consider the solution to $x^x = \mathcal{N}(r)$ as shown in Fig. 5.2(b), it may be concluded that $\mathcal{N}(r)$ is approximately of the order of $O((\kappa\cdot r)^{(\kappa\cdot r)})$ where κ stands for the slope of the curve. Not surprisingly it is not feasible to derive high order moments by hand.[2]

However, since all combinations of indices can be exhausted systematically, we can let a computer do the symbolic manipulations. With the help of a computer program written in C++ the first 10 moments have been obtained assuming BPSK spreading and they are

[2]It is possible to reduce the number of terms by exploiting identical terms and saving as many intermediate results as possible. Nevertheless, the author believe that the problem of exhausting all the terms is NP-hard, i.e., cannot be reduced to the level of $O(r^a)$.

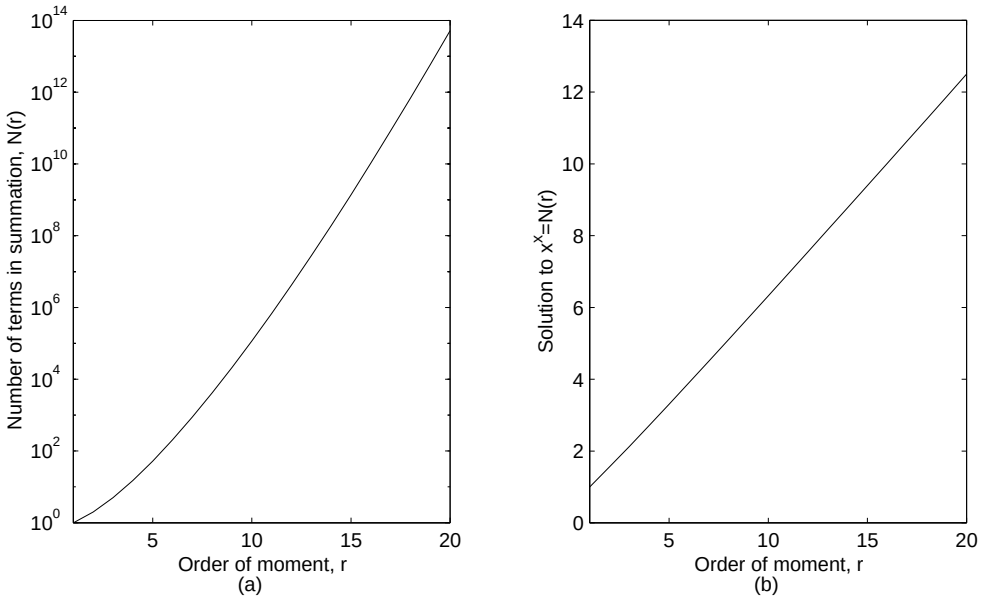

Figure 5.2: Number of terms in summation versus the order of moment.

included in Appendix C. Based on these expressions, it is straightforward to compute the weighting factors of a 5-stage PIC for long-code systems. Knowledge of received signal energies are necessities, but no power control is assumed, i.e., the SNRs of all users need not to be the same. It would be manageable for a computer to get analytical expression for several more moments. Nevertheless, deriving high order moments would be a difficult mathematical and computer science problem by itself.

There is no known general expression for M_r even for the simplest case of perfect power control. Mathematicians however have studied this problem and suggested some insightful approximations.

5.2.3 Dag Jonsson's Approximations for Perfect Power Control

The problem of moments of covariance matrix has been studied by Dag Jonsson, who gave the following theorem in 1982 [40].

Theorem 5.1. *(Dag Jonsson's Theorem) Assume* S_{nk}, $n = 1, 2, \cdots, N$, $k = 1, 2, \cdots, K$, *to be independent real random variables that satisfy*

$$\mathrm{E}\{S_{nk}\} = 0, \quad \mathrm{Var}\{S_{nk}\} = 1, \quad \mathrm{E}\left\{S_{nk}^4\right\} < \infty. \tag{5.45}$$

Let $R_{kl} = (1/N)\sum_{n=1}^{N} S_{nk}S_{nl}$ *denote elements of a* $K \times K$ *scaled sample covariance matrix* $\mathbf{R}$ *and let* $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_K$ *be the corresponding eigenvalues of* $\mathbf{R}$. *Define the* r^{th} *order moment of the covariance matrix the same as in Definition 5.1:*

$$M_r = \frac{1}{K}\mathrm{E}\left\{\sum_{k=1}^{K}(\lambda_k)^r\right\}. \tag{5.46}$$

The moment is then a polynomial in N and K with a total degree of r divided by N^r, which can be expressed as

$$M_r = \frac{1}{N^r}\left\{\sum_{j=0}^{r-1} a_{r,j} N^{r-j} K^j + (\textit{terms in } N \textit{ and } K \textit{ with total degree}) < r\right\} \tag{5.47}$$

where

$$a_{r,j} = \frac{1}{j+1}\binom{r}{j}\binom{r-1}{j}, \; j = 0, 1, \ldots, r-1. \tag{5.48}$$

Corollary 5.1. *Suppose that K and N depend on a variable l, so that when $l \to \infty$,*

$$N(l) \textit{ and } K(l) \to \infty, \; 0 < \frac{K(l)}{N(l)} < y, \; y < \infty. \tag{5.49}$$

It is clear then

$$\lim_{\substack{l\to\infty \\ K/N<\infty}} M_r = \sum_{j=0}^{r-1} a_{r,j} \left(\frac{K}{N}\right)^j, \; n = 1, 2, \cdots. \tag{5.50}$$

This corollary highlights the most significant terms (terms with a total degree of r) in the summation in (5.47) and may be used as a coarse approximation to M_r. For example, the above corollary gives the following approximations that agree with the exact expressions given by Eqn. (C.1) through (C.4),

$$M_1 = 1, \tag{5.51}$$

$$M_2 \approx 1 + \frac{K}{N}, \tag{5.52}$$

$$M_3 \approx 1 + 3\frac{K}{N} + \left(\frac{K}{N}\right)^2, \tag{5.53}$$

$$M_4 \approx 1 + 6\frac{K}{N} + 6\left(\frac{K}{N}\right)^2 + \left(\frac{K}{N}\right)^3. \tag{5.54}$$

Assuming that r is small, N is quite large and $K \leq N$, these approximations are found to be fairly accurate. Take M_6 as an example, the relative error is 18% for $N = 31$ and $K = 15$, and 5% for $N = 128$ and $K = 64$. This could be sufficient for computing weighting factors for a PIC detector, since the mean squared error of the detector output is not very sensitive to the estimation error of moments.

Nonetheless, rigorous expressions for M_1, M_2, $\cdots$, M_{2m} should be derived before implementing a practical system. In fact this can be carried out once and for all, which is by no means a difficult task compared to the realisation of the whole interference canceller structure. Hence the purpose of this subsection is to refer to known results in the literature that corroborate the exact expressions listed in the appendix.

5.3 On-line Computational Complexity

In this section, we discuss the computational complexity involved for updating the weighting factors when the received signal energies change, or some users join or leave the system. Consider an m-stage PIC with K active users in a CDMA system with a processing gain of N. The receiver is supposed to have perfect knowledge of the received signal energies and the noise variance. Assuming that analytical expressions for M_r, $r = 1, 2, \cdots, 2m$ have been derived and are ready for use, the procedure of updating the optimal set of weighting factors can be summarised as 6 steps as shown in Table 5.1 together with the approximate FLOPS[3] needed by each step.

OPERATION	FLOPS
1. Evaluate $\mathcal{E}_1, \mathcal{E}_2, \cdots, \mathcal{E}_{2m}$ using Definition 5.2.	$4mK$
2. Evaluate $M_1, M_2, \cdots, M_{2m}$ using expressions in Appendix C.	$2m^3$
3. Evaluate all elements of $\mathbf{C}$ and $\mathbf{p}$ using Eqn. (5.11) and (5.12).	$16m^2 + 3m^2 = 19m^2$
4. Solve $\hat{\mathbf{x}}$ from Eqn. (5.5).	m^3
5. Obtain the m weights as roots of Eqn. (4.74).	m^3
6. Order the weights as described in section 5.1.3.	$2m^3$
Total:	$4mK + 6m^3 + 19m^2$

Table 5.1: Step-by-step weight updating procedure and its complexity.

It takes $4mK$ FLOPS to evaluate the first $2m$ moments of the received signal energies in step 1 and $2m^3$ for the first $2m$ moments of the correlation matrix in step 2. In step 3, evaluating all elements of $\mathbf{C}$ and $\mathbf{p}$ takes $16m^2$ and $3m^2$ FLOPS respectively. In step 4, $\hat{\mathbf{x}}$ is obtained by solving the group equations described by (5.5), which takes no more than m^3 FLOPS. Substituting $\hat{\mathbf{x}}$ into Eqn. (4.74), we get the desired weighting factors by solving the m^{th} order polynomial, which is equivalent to computing the eigenvalues of its companion matrix. Therefore step 5 also needs m^3 FLOPS. The ordering of weighting factors is also based on the moments but independent of K. For the i^{th} stage it takes no more than $6i^2$ to evaluate (5.13). Hence the complexity of step 5 is roughly $2m^3$.

In all, the total complexity for updating the weights is no more than $(4mK + 6m^3 + 19m^2)$, which is independent of the processing gain. Regardless of m, which may be a very small number as compared to K, we may say that the total complexity is linear in the number of active users. Hence the complexity of updating the weights is probably 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 conspicuously increase the overall system complexity.

5.4 Numerical Results

[3]Consider 1 summation, multiplication, division or power function as one floating-point operation (FLOP).

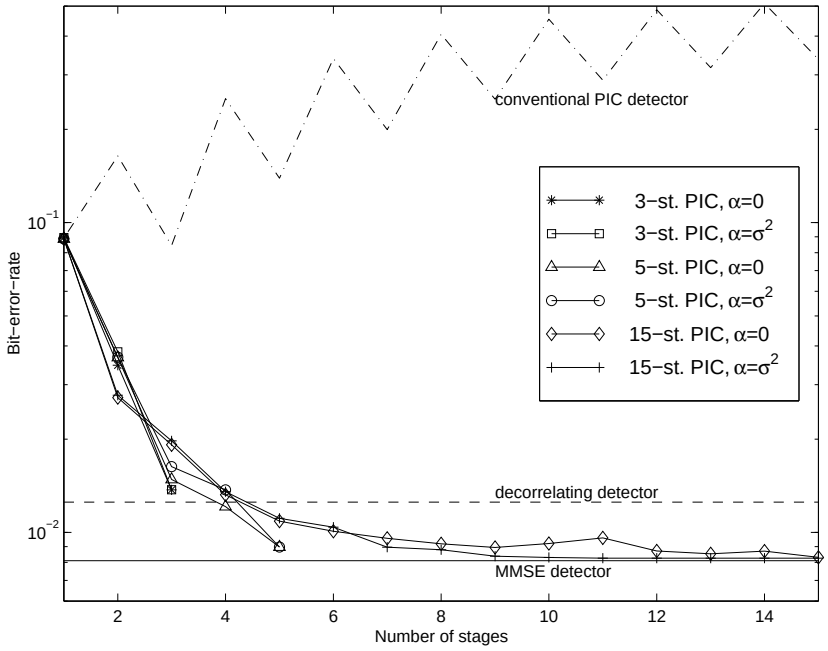

Figure 5.3: Stage-by-stage performance of the weighted linear PIC using long codes.

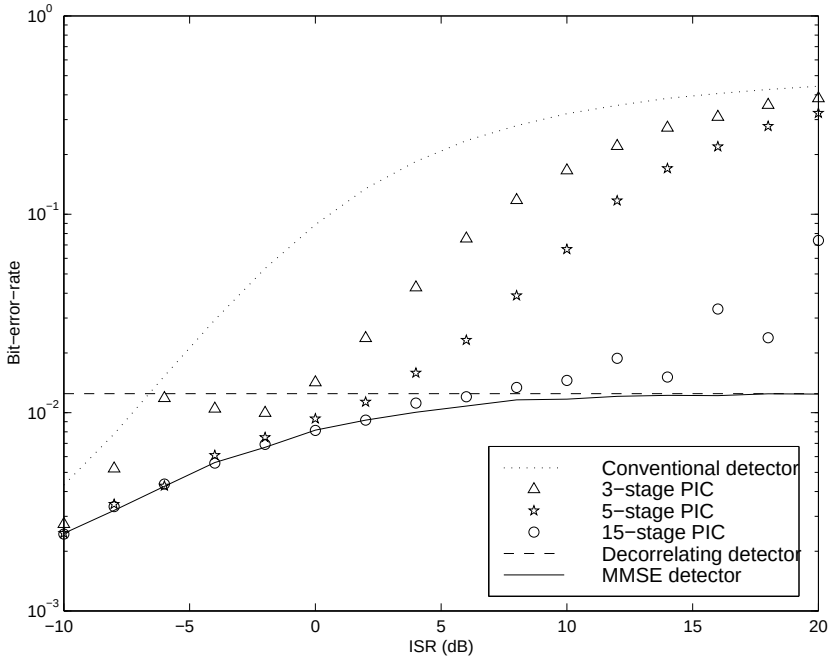

Figure 5.4: Near-far ability of the weighted linear PIC for a long-code system.

Fig. 5.3 shows the stage-by-stage BER performance of PIC using long codes. An SNR of 7 dB is assumed for all 15 users. We observe that the conventional PIC scheme diverges since the largest eigenvalue for a $K = 15$, $N = 31$ system, is almost always greater than 2. A numerical search indicate that this is the case for 99 % of all the possible code-sets. Divergence and the ping-pong effect can be overcome by a proper choice of weighting factors. The weighting factors are optimised for 7 dB and ordered as described in section 5.1.1. Significant improvement over the conventional detector can be achieved using merely 3 stages where the weighting factors are (0.63431, 0.39633, 1.6468) for $\alpha = 0.099763$ (which corresponds to σ^2 at 7 dB) and (0.79515, 0.40435, 1.2876) for $\alpha = 0$. A 5-stage detector can give close to average MMSE performance, where the weighting factors are (0.60861, 0.42511, 1.0799, 0.34456, 2.5765) for $\alpha = 0.099763$ and $(0.51180 - 0.019155i,\ 0.51180 + 0.019155i,\ 1.4383 - 0.78476i,\ 0.35226,\ 1.4383 + 0.78476i)$ for $\alpha = 0$. A 15-stage PIC gives very close to MMSE performance, however exact MMSE performance is not achieved. This is no surprise since exact MMSE performance will be achieved only if all relevant weighting factors for all possible spreading codes are included in the set, i.e., a prohibitively large number of stages are required.

Fig. 5.4 shows the BER performance of a weighted PIC detector for a long-code system under near-far environment, as compared with that of the conventional detector, the decorrelator and the MMSE detector. A 15-user system with a processing gain of 31 is considered. The SNR of the first user is 7 dB while the remaining 14 users have the same SNR but ISR (dB) different to user 1. The curves show the BER of the first user only. It is observed that the PIC generally performs better than the conventional detector but worse than the MMSE detector, and the larger number of PIC stages the better its ability to combat the near-far effect. In comparison to Fig. 4.5, the near-far ability of a weighted PIC working under long codes is not as good as that under short codes. The reason is the set of weights used in a long-code system has to compromise all possible code-sets, hence results in some performance loss. A 15-stage PIC does not achieve the MMSE detector under random codes and is not near-far resistant. In fact the near-far resistance of a PIC detector for long-code system is always 0. The fluctuation of the performance curve of a 15-stage PIC is due to numerical effects in computing the weights. Note that the BER curve for a 3-stage PIC is not monotonically increasing, i.e., in certain circumstances stronger interfering users help user 1 to achieve better BER performance. The reason is that the moments of the correlation matrix are not monotonically increasing functions of the ISR, and we know that smaller moments indicate that the correlation matrix is "closer" to the identity matrix, a property which makes the PIC perform better.

Since the optimal set of weights are dependent on the received signal energies, or equivalently, the signal-to-noise ratios, which are to be estimated at the receiver, we would like to see whether the estimation error severely undermines detector performance. The sensitivity of the BER to the choice of the working SNR is illustrated in Fig. 5.5. Again equal power users are assumed. The weighting factors optimised for an SNR of 7 dB together with the corresponding α 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. In other words, the curves show how the performance can be degraded if the working SNR is different from what we have estimated (7 dB). Note that for long codes it still holds that the achievable MMSE is independent of the selection of α. The performance is therefore virtually the same for $\alpha = 0$ and $\alpha = \sigma^2$. Increasing the number of stages increases the sensitivity to working noise level slightly. It is promising since for 3-stage, 5-stage or even 9-stage PIC, a set of weighting factors optimised for 7 dB would be working well under a wide

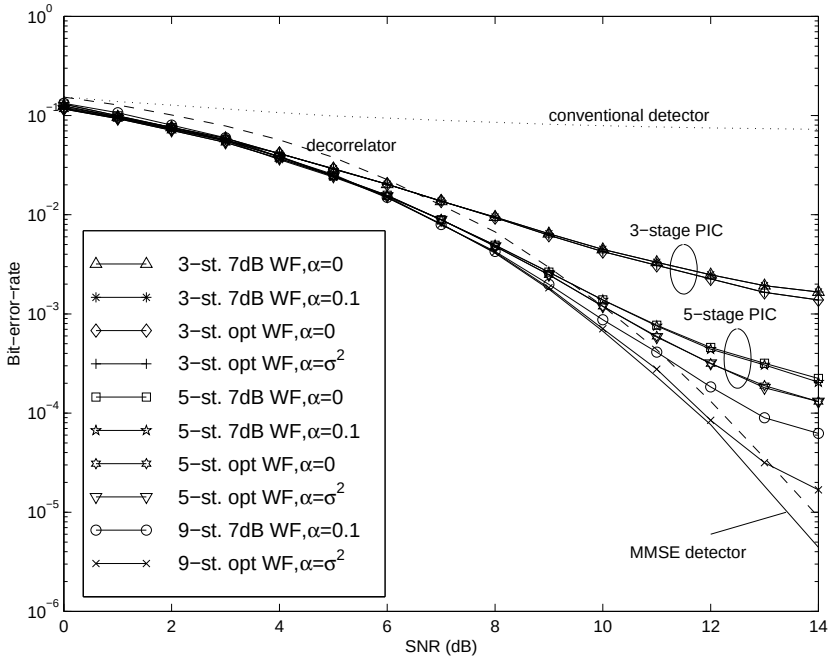

Figure 5.5: BER performance and sensitivity versus SNR using long codes. Weighting factors (WF) optimised for 7 dB are used for $0 \sim 14$ dB, in comparison to when weighting factors are optimised for each working SNR. Both $\alpha \sim 7$ dB and $\alpha = 0$ are shown.

range of noise level. 9-stage PIC is seen to be very close to MMSE detector in a wide range of SNRs.

5.5 Summary

In this chapter, we have investigated the case of using a fixed set of weighting factors that minimises the ensemble averaged value of the MSE over random codes, while the received signal power distribution is allowed to be changing slowly. The complexity for updating the weighting factors on-line in case that the received signal power distribution varies or the number of active users changes is linear to the number of users, which only increase overall system complexity very slightly.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

In this thesis, we have developed a mathematical approach to linear PIC. It is shown that the linear PIC, whether conventional or weighted, is equivalent to a one-shot linear matrix filter. It follows that the exact bit error rate can be calculated analytically. Conditions for the PIC to converge is studied and a modified weighted PIC structure is suggested which converges to the MMSE detector.

Based on analysis of the MSE, it is shown that for a K-user short-code system only K stages are necessary to achieve the MMSE detector exactly. For fewer stages, an analytical approach is derived for finding the optimal choice of weighting factors that will lead to the minimum achievable MSE. 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.

For long codes, an algorithm is proposed for finding a fixed set of weighting factors that will give the minimum ensemble average of the MSE over random codes. The coefficients however, depends on the received signal energies as well as the noise level. The computation complexity is linear to the number of users but independent of the processing gain, which enables updating the coefficients on-line each time the received signal energies or the number of active users change. Simulation shows that for moderate near-far situations, significant performance improvements are observed even for a small number of stages. For as few as 3 stages, it is possible for a PIC to get close to the MMSE performance. In this case, the detector is relatively insensitive to any mismatches in the working SNR for both $\alpha = 0$ and $\alpha = \sigma^2$.

In summary, this thesis presents an MSE-oriented study of linear parallel interference cancellation. It helps to improve the understanding of interference cancellation techniques and suggests analytical methods for the optimisation of the linear PIC, both for short-code and long-code systems.

6.2 Future Work

For the fulfilment of a practical multi-user receiver for CDMA based on interference cancellation techniques, the following areas are of great interest.

- The Conjugate Gradient Method

  Recent work done by Elders-Boll *et al.* [18] reveals that all existing linear interference cancellation structures belong to the family of iterative methods for computing matrix inversion by means of matrix splitting. The PIC scheme is equivalent to a Jacobi iteration whereas the SIC corresponds to a Gauß-Seidel iteration [32]. The linear interference cancellation concept is further enriched by many other well-developed methods. Among them the Hestenes-Stiefel conjugate gradient method [32, Chap. 10] is especially promising. It resembles the steepest descent method but tends to adapt its searching path according to the residual error and therefore shows superior performance. Its complexity is slightly higher than the PIC detector.

  A new type of interference canceller can be expected based on the conjugate gradient method. Its properties are yet to be studied.

- Study of Asynchronous Scenario

  An asynchronous model is more realistic for the up-link channel of a cellular mobile system. In this case the MAI includes also interference from the previous and the following symbols from other users. This has made the channel an IIR rather than an FIR filter. A suitable weighted structure for an interference canceller coping with asynchronous scenario is yet to be developed and analysed.

- Non-linear Interference Cancellation

  Simulations show that a non-linear interference canceller which makes use of non-linear decision functions such as hyperbolic tangent, clip function or even hard-decision in the intermediate stages may perform much better than a linear one. This attractive phenomenon is yet to be explained and necessary techniques to be introduced to achieve satisfactory receiver performance.

Appendix A

Solution to x

In section 4.6 the minimum of J_{ex} in $\mathbf{x}$ is found to satisfy (4.72), rewritten as follows,

$$\left[\sum_{k=1}^{K} \lambda_k \phi_k \boldsymbol{\gamma}_k \boldsymbol{\gamma}_k^\top\right] \mathbf{x}^* = -\sum_{k=1}^{K} \lambda_k \boldsymbol{\gamma}_k \quad \Rightarrow \quad \mathbf{C}\mathbf{x}^* = -\mathbf{p}. \tag{A.1}$$

There exists a unique solution if $\mathbf{C}$ is non-singular. $\mathbf{C}$ is positive semi-definite, since for any real vector $\mathbf{v}$,

$$\mathbf{v}^\top \mathbf{C} \mathbf{v} = \sum_{k=1}^{K} \lambda_k \phi_k (\boldsymbol{\gamma}_k^\top \mathbf{v})^2 \geq 0. \tag{A.2}$$

Consequently $\mathbf{C}$ is singular if and only if there exists a $\mathbf{v} \neq \mathbf{0}$ such that $\boldsymbol{\gamma}_k^\top \mathbf{v} = 0$ for all $1 \leq k \leq K$, i.e.,

$$(\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, ..., \boldsymbol{\gamma}_K)^\top \mathbf{v} = \mathbf{0}. \tag{A.3}$$

This constitutes K equations in m unknowns. If m or more of the K equations are linearly independent, there is no non-zero solutions to (A.3), which results in a non-singular $\mathbf{C}$. This corresponds to that there exists at least m vectors in $(\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, ..., \boldsymbol{\gamma}_K)^\top$ that are independent, or equivalently

$$\exists\, 1 \leq k_1 < k_2 < \cdots < k_m \leq K \text{ so that } \det(\boldsymbol{\gamma}_{k_1}, \boldsymbol{\gamma}_{k_2}, \cdots, \boldsymbol{\gamma}_{k_m}) \neq 0. \tag{A.4}$$

Note that $(\boldsymbol{\gamma}_{k_1}, \boldsymbol{\gamma}_{k_2}, \cdots, \boldsymbol{\gamma}_{k_m})$ is a Vandermonde system [32], thus

$$\begin{aligned} \det(\boldsymbol{\gamma}_{k_1}, \boldsymbol{\gamma}_{k_2}, \cdots, \boldsymbol{\gamma}_{k_m}) &= \begin{vmatrix} 1 & 1 & \cdots & 1 \\ \gamma_{k_1} & \gamma_{k_2} & \cdots & \gamma_{k_m} \\ & \cdots & \cdots & \\ \gamma_{k_1}^{(m-1)} & \gamma_{k_2}^{(m-1)} & \cdots & \gamma_{k_m}^{(m-1)} \end{vmatrix} && \text{(A.5)} \\ &= \prod_{1 \leq j < i \leq m} (\gamma_{k_i} - \gamma_{k_j}) && \text{(A.6)} \\ &= \prod_{1 \leq j < i \leq m} (\lambda_{k_i} - \lambda_{k_j}). && \text{(A.7)} \end{aligned}$$

It then follows that at least m eigenvalues of $\mathbf{R}$ must be distinct for $\mathbf{C}$ to be non-singular, so that the solution to (A.1) is $\hat{\mathbf{x}} = -\mathbf{C}^{-1}\mathbf{p}$.

If this is not the case, $\mathbf{C}$ is singular and its inverse does not exist. Assuming there are l distinct eigenvalues, $l < m$, and the eigenvalue λ_{k_i}, $1 \leq i \leq l$, is h_{k_i}-fold. (A.1) is then reduced to

$$\sum_{i=1}^{l} h_{k_i} \lambda_{k_i} \phi_{k_i} \boldsymbol{\gamma}_{k_i} \boldsymbol{\gamma}_{k_i}^{\top} \mathbf{x} = -\sum_{i=1}^{l} h_{k_i} \lambda_{k_i} \boldsymbol{\gamma}_{k_i}. \tag{A.8}$$

An obvious solution is that $\boldsymbol{\gamma}_{k_i}^{\top} \mathbf{x} = -1/\phi_{k_i}$ for $i = 1, 2, \cdots, l$, i.e.,

$$(\boldsymbol{\gamma}_{k_1}, \boldsymbol{\gamma}_{k_2}, \cdots, \boldsymbol{\gamma}_{k_l})^{\top} \mathbf{x} = -(\frac{1}{\phi_{k_1}}, \frac{1}{\phi_{k_2}}, \cdots, \frac{1}{\phi_{k_2}})^{\top}. \tag{A.9}$$

At this time $J_{\text{ex}}^{(m)}(\mathbf{x}, \alpha) = 0$ from (4.68). (A.9) is an indefinite group equation since there are m unknowns, where the coefficient matrix has rank $l < m$. There are infinite number of solutions, including complex ones. It is sufficient to choose an arbitrary real solution to (A.9), so that $J_{\text{ex}}^{(m)}(\mathbf{x}, \alpha)$ achieves its minimum 0.

In summary, if $\mathbf{R}$ has m or more distinct eigenvalues, $\mathbf{C}$ is non-singular and $\hat{\mathbf{x}} = -\mathbf{C}^{-1}\mathbf{p}$ is the unique solution. If $\mathbf{R}$ has $l < m$ distinct eigenvalues, any solution to (A.9) will satisfy $\mathbf{C}\mathbf{x} = -\mathbf{p}$ and lead to $J_{\text{ex}}^{(m)}(\mathbf{x}, \alpha) = 0$.

Appendix B

Locating the Real Step Sizes

Based on Definition 4.1, we define a constrained version of mapping T with real result region.

Definition B.1. *Mapping* $T_R:\ \mathbb{G}_{\boldsymbol{\mu}} \to \mathbb{R}^m$ *is given by* $\mathbf{x} = T(\boldsymbol{\mu})$, *where* $\mathbb{G}_{\boldsymbol{\mu}} = \{\, \boldsymbol{\mu} \mid\ T(\boldsymbol{\mu}) \in \mathbb{R}^m \}$ *and mapping* T *is given by Definition 4.1 in section 4.6.3.*

Theorem B.1. *Define* $\mathbb{D}_{\mathbf{x}} = \{\, \mathbf{x} \mid\ \exists\, \boldsymbol{\mu} \in \mathbb{R}^m,\ \mathbf{x} = T_R(\boldsymbol{\mu}) \} \subseteq \mathbb{R}^m$. $\mathbb{B}_{\mathbf{x}}$, *the boundary set of* $\mathbb{D}_{\mathbf{x}}$ *in* $\mathbb{R}^m$, *is then determined by* $\mathbb{B}_{\mathbf{x}} = T_R(\mathbb{B}_{\boldsymbol{\mu}})$, *where*

$$\mathbb{B}_{\boldsymbol{\mu}} = \{\, \boldsymbol{\mu} = (\mu_1, \mu_2, \cdots, \mu_m) \mid\ \exists\, 1 \le i < j \le m,\ \mu_i = \mu_j,\ \boldsymbol{\mu} \in \mathbb{R}^m \}\,. \tag{B.1}$$

Proof. Since $\mathbb{B}_{\mathbf{x}}$ is the boundary set of $\mathbb{D}_{\mathbf{x}}$,

$$\forall \mathbf{x} \in \mathbb{B}_{\mathbf{x}},\ \forall \delta > 0,\ \exists \mathbf{x}^{(1)} \in \mathbb{D}_{\mathbf{x}},\ \mathbf{x}^{(2)} \notin \mathbb{B}_{\mathbf{x}},\ \text{so that } |\mathbf{x}^{(1)} - \mathbf{x}| < \delta,\ |\mathbf{x}^{(2)} - \mathbf{x}| < \delta. \tag{B.2}$$

Therefore,

$$\exists \left\{ \mathbf{x}_n^{(1)} \in \mathbb{D}_{\mathbf{x}} \right\},\ \lim_{n\to\infty} \mathbf{x}_n^{(1)} = \mathbf{x}; \tag{B.3}$$

$$\exists \left\{ \mathbf{x}_n^{(2)} \notin \mathbb{D}_{\mathbf{x}} \right\},\ \lim_{n\to\infty} \mathbf{x}_n^{(2)} = \mathbf{x}. \tag{B.4}$$

Making use of Theorem 4.1, we can find corresponding series in $\mathbb{G}_{\boldsymbol{\mu}}$,

$$\exists \left\{ \boldsymbol{\mu}_n^{(1)} \in \mathbb{R}^m \right\} \subseteq \mathbb{G}_{\boldsymbol{\mu}},\ \lim_{n\to\infty} T_R(\boldsymbol{\mu}_n^{(1)}) = \mathbf{x}; \tag{B.5}$$

$$\exists \left\{ \boldsymbol{\mu}_n^{(2)} \notin \mathbb{R}^m \right\} \subseteq \mathbb{G}_{\boldsymbol{\mu}},\ \lim_{n\to\infty} T_R(\boldsymbol{\mu}_n^{(2)}) = \mathbf{x}. \tag{B.6}$$

It is obvious that $\left\{ \boldsymbol{\mu}_n^{(1)} \right\}$ and $\left\{ \boldsymbol{\mu}_n^{(2)} \right\}$ are bounded series, otherwise $T_R(\boldsymbol{\mu}_n^{(1)})$ and $T_R(\boldsymbol{\mu}_n^{(2)})$ will not converge. From theory of limitation, bounded series must have sub-series that converge, therefore,

$$\exists \left\{ \boldsymbol{\mu}_{n_k}^{(1)} \in \mathbb{R}^m \right\},\ \lim_{k\to\infty} \boldsymbol{\mu}_{n_k}^{(1)} = \boldsymbol{\mu}^{(1)}; \tag{B.7}$$

$$\exists \left\{ \boldsymbol{\mu}_{n_k}^{(2)} \notin \mathbb{R}^m \right\},\ \lim_{k\to\infty} \boldsymbol{\mu}_{n_k}^{(2)} = \boldsymbol{\mu}^{(2)}. \tag{B.8}$$

Since real vector series will converge only to a real vector, thus $\boldsymbol{\mu}^{(1)} \in \mathbb{R}^m$, which has corresponding $\mathbf{x} = T_R(\boldsymbol{\mu}^{(1)}) \in \mathbb{D}_{\mathbf{x}}$. Based on Theorem 4.1, $\boldsymbol{\mu}^{(2)}$ is a solution to $\mathbf{x} = T_R(\boldsymbol{\mu}^{(2)})$, hence $\boldsymbol{\mu}^{(2)} \in \mathbb{R}^m$.

So far it is proved that $\mathbb{B}_{\mathbf{x}} \subseteq \mathbb{D}_{\mathbf{x}}$. In the following we will prove that $\mathbb{B}_{\mathbf{x}} = T_R(\mathbb{B}_{\boldsymbol{\mu}})$.

Rewrite $\boldsymbol{\mu}_{n_k}^{(2)}$ as $\boldsymbol{\nu}_k = \left(\nu_1^{(k)}, \nu_2^{(k)}, \cdots, \nu_m^{(k)}\right)^\top$ and $\boldsymbol{\mu}^{(2)}$ as $\boldsymbol{\nu}$, then

$$\forall\, k > 0,\ \boldsymbol{\nu}^{(k)} \notin \mathbb{R}^m,\ \boldsymbol{\nu} \in \mathbb{R}^m,\ \text{ and } \lim_{k\to\infty} \boldsymbol{\nu}^{(k)} = \boldsymbol{\nu}. \tag{B.9}$$

Since $\boldsymbol{\nu}^{(k)} \notin \mathbb{R}^m$ but $T_R(\boldsymbol{\nu}^{(k)}) \in \mathbb{R}^m$, there must have at least one complex conjugate pair among $\nu_1^{(n)}, \nu_2^{(n)}, \cdots, \nu_m^{(n)}$, which are the m roots of an m^{th} order real coefficient polynomial. It is not difficult to deduce that there exist i and j, $1 \le i < j \le m$, so that a sub-series of $\boldsymbol{\nu}^{(k)}$, denoted as $\left\{\boldsymbol{\xi}^{(l)} = \boldsymbol{\nu}^{(k_l)} \mid l = 1, 2, \cdots\right\}$, satisfies $\mathrm{Re}(\xi_i^{(l)}) = \mathrm{Re}(\xi_j^{(l)})$ for any l. In the meantime,

$$\lim_{l\to\infty} \xi_i^{(l)} = \nu_i \in \mathbb{R}; \tag{B.10}$$

$$\lim_{l\to\infty} \xi_j^{(l)} = \nu_j \in \mathbb{R}. \tag{B.11}$$

Therefore

$$\nu_i = \lim_{l\to\infty} \mathrm{Re}\{\xi_i^{(l)}\} = \lim_{l\to\infty} \mathrm{Re}\{\xi_j^{(l)}\} = \nu_j. \tag{B.12}$$

Equivalently, $\mu_i^{(2)} = \mu_j^{(2)}$, i.e., $\boldsymbol{\mu}^{(2)} \in \mathbb{B}_{\boldsymbol{\mu}}$. In conclusion, $\mathbb{B}_{\mathbf{x}} = T(\mathbb{B}_{\boldsymbol{\mu}})$. □

Appendix C

Results of Computer Aided Symbolic Manipulations

The exact expression for the first 10 moments of the correlation matrix assuming BPSK spreading have been obtained with the help of a computer program written in C++. They are listed below. They are found to be polynomials of the processing gain, the number of active users, and the moments of the received signal energies. Based on these expressions, the optimal set of weights for up to a 5-stage PIC using random-codes can be computed efficiently.

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

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

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

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

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

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

$$
\begin{aligned}
M_7 &= \frac{1}{KN^6}[\mathcal{E}_1^7 + \mathcal{E}_1^5\mathcal{E}_2(21N-21) + \mathcal{E}_1^4\mathcal{E}_3(55N^2-125N+70) \\
&+ \mathcal{E}_1^3\mathcal{E}_2^2(50N^2-85N+35) + \mathcal{E}_1^3\mathcal{E}_4(80N^3-305N^2+365N-140) \\
&+ \mathcal{E}_1^2\mathcal{E}_2\mathcal{E}_3(84N^3-233N^2+137N+12) \\
&+ \mathcal{E}_1^2\mathcal{E}_5(59N^4-331N^3+612N^2-412N+72) \\
&+ \mathcal{E}_1\mathcal{E}_2^3(11N^3-43N^2+51N-19) \\
&+ \mathcal{E}_1\mathcal{E}_2\mathcal{E}_4(29N^4-171N^3+219N^2-25N-52) \\
&+ \mathcal{E}_1\mathcal{E}_3^2(14N^4-40N^3-2N^2+74N-46) \\
&+ \mathcal{E}_1\mathcal{E}_6(17N^5-152N^4+418N^3-376N^2+11N+82) \\
&+ \mathcal{E}_2^2\mathcal{E}_3(3N^4-39N^3+81N^2-85N+40) \\
&+ \mathcal{E}_2\mathcal{E}_5(2N^5-42N^4+118N^3-87N^2+41N-32) \\
&+ \mathcal{E}_3\mathcal{E}_4(2N^5-16N^4+17N^3+15N^2+32N-50) \\
&+ \mathcal{E}_7(N^6-21N^5+105N^4-147N^3+14N^2+48)] \qquad \text{(C.7)}
\end{aligned}
$$

$$
\begin{aligned}
M_8 &= \frac{1}{KN^7}[\mathcal{E}_1^8 + \mathcal{E}_1^6\mathcal{E}_2(28N-28) + \mathcal{E}_1^5\mathcal{E}_3(91N^2-203N+112) \\
&+ \mathcal{E}_1^4\mathcal{E}_2^2(105N^2-175N+70) + \mathcal{E}_1^4\mathcal{E}_4(175N^3-645N^2+750N-280) \\
&+ \mathcal{E}_1^3\mathcal{E}_2\mathcal{E}_3(259N^3-663N^2+362N+42) \\
&+ \mathcal{E}_1^3\mathcal{E}_5(191N^4-991N^3+1717N^2-1099N+182) \\
&+ \mathcal{E}_1^2\mathcal{E}_2^3(56N^3-176N^2+190N-70) \\
&+ \mathcal{E}_1^2\mathcal{E}_2\mathcal{E}_4(171N^4-771N^3+856N^2-58N-198) \\
&+ \mathcal{E}_1^2\mathcal{E}_3^2(78N^4-187N^3-70N^2+375N-196) \\
&+ \mathcal{E}_1^2\mathcal{E}_6(108N^5-767N^4+1761N^3-1313N^2-113N+324) \\
&+ \mathcal{E}_1\mathcal{E}_2^2\mathcal{E}_3(49N^4-318N^3+585N^2-642N+326) \\
&+ \mathcal{E}_1\mathcal{E}_2\mathcal{E}_5(42N^5-407N^4+941N^3-747N^2+533N-362) \\
&+ \mathcal{E}_1\mathcal{E}_3\mathcal{E}_4(40N^5-172N^4-10N^3+395N^2+155N-408) \\
&+ \mathcal{E}_1\mathcal{E}_7(23N^6-268N^5+906N^4-865N^3-164N^2-124N+492) \\
&+ \mathcal{E}_2^4(N^4-20N^3+42N^2-42N+19) \\
&+ \mathcal{E}_2^2\mathcal{E}_4(3N^5-81N^4+236N^3-427N^2+645N-376) \\
&+ \mathcal{E}_2\mathcal{E}_3^2(3N^5-57N^4+202N^3-396N^2+546N-298) \\
&+ \mathcal{E}_2\mathcal{E}_6(2N^6-69N^5+308N^4-437N^3+768N^2-1510N+938) \\
&+ \mathcal{E}_3\mathcal{E}_5(2N^6-30N^5+68N^4-181N^3+918N^2-1701N+924) \\
&+ \mathcal{E}_4^2(N^6-11N^5+20N^4+17N^3+244N^2-689N+418) \\
&+ \mathcal{E}_8(N^7-28N^6+182N^5-308N^4+133N^3-1120N^2+2772N-1632)] \quad \text{(C.8)}
\end{aligned}
$$

$$
\begin{aligned}
M_9 &= \frac{1}{KN^8}[\mathcal{E}_1^9 + \mathcal{E}_1^7\mathcal{E}_2(36N-36) + \mathcal{E}_1^6\mathcal{E}_3(140N^2-308N+168) \\
&+ \mathcal{E}_1^5\mathcal{E}_2^2(196N^2-322N+126) + \mathcal{E}_1^5\mathcal{E}_4(336N^3-1211N^2+1379N-504) \\
&+ \mathcal{E}_1^4\mathcal{E}_2\mathcal{E}_3(644N^3-1568N^2+812N+112) \\
&+ \mathcal{E}_1^4\mathcal{E}_5(488N^4-2419N^3+4017N^2-2478N+392) \\
&+ \mathcal{E}_1^3\mathcal{E}_2^3(196N^3-539N^2+539N-196) \\
&+ \mathcal{E}_1^3\mathcal{E}_2\mathcal{E}_4(648N^4-2540N^3+2540N^2-72N-576) \\
&+ \mathcal{E}_1^3\mathcal{E}_3^2(281N^4-601N^3-367N^2+1295N-608) \\
&+ \mathcal{E}_1^3\mathcal{E}_6(419N^5-2653N^4+5436N^3-3467N^2-695N+960) \\
&+ \mathcal{E}_1^2\mathcal{E}_2^2\mathcal{E}_3(329N^4-1491N^3+2423N^2-2739N+1478) \\
&+ \mathcal{E}_1^2\mathcal{E}_2\mathcal{E}_5(326N^5-2171N^4+4186N^3-3313N^2+2954N-1982) \\
&+ \mathcal{E}_1^2\mathcal{E}_3\mathcal{E}_4(288N^5-853N^4-766N^3+2805N^2+448N-1922) \\
&+ \mathcal{E}_1^2\mathcal{E}_7(186N^6-1587N^5+4027N^4-2202N^3-2000N^2-1066N+2642) \\
&+ \mathcal{E}_1\mathcal{E}_2^4(18N^4-176N^3+363N^2-380N+175) \\
&+ \mathcal{E}_1\mathcal{E}_2^2\mathcal{E}_4(70N^5-860N^4+2315N^3-4214N^2+5971N-3282) \\
&+ \mathcal{E}_1\mathcal{E}_2\mathcal{E}_3^2(69N^5-548N^4+1553N^3-3428N^2+4944N-2590) \\
&+ \mathcal{E}_1\mathcal{E}_2\mathcal{E}_6(58N^6-849N^5+3024N^4-4822N^3+9387N^2-15016N+8218) \\
&+ \mathcal{E}_1\mathcal{E}_3\mathcal{E}_5(55N^6-388N^5+318N^4-806N^3+8013N^2-14878N+7686) \\
&+ \mathcal{E}_1\mathcal{E}_4^2(27N^6-134N^5-142N^4+502N^3+2530N^2-6013N+3230) \\
&+ \mathcal{E}_1\mathcal{E}_8(30N^7-443N^6+1757N^5-1578N^4+669N^3 \\
&\quad -12875N^2+25940N-13500) \\
&+ \mathcal{E}_2^3\mathcal{E}_3(4N^5-142N^4+455N^3-869N^2+1096N-544) \\
&+ \mathcal{E}_2^2\mathcal{E}_5(3N^6-141N^5+706N^4-1977N^3+4903N^2-6280N+2786) \\
&+ \mathcal{E}_2\mathcal{E}_3\mathcal{E}_4(6N^6-196N^5+891N^4-2623N^3+6671N^2-7841N+3092) \\
&+ \mathcal{E}_2\mathcal{E}_7(2N^7-103N^6+664N^5-1540N^4+4927N^3 \\
&\quad -14818N^2+17802N-6934) \\
&+ \mathcal{E}_3^3(N^6-23N^5+152N^4-558N^3+1144N^2-1106N+390) \\
&+ \mathcal{E}_3\mathcal{E}_6(2N^7-50N^6+199N^5-997N^4+5736N^3-14251N^2+13881N-4520) \\
&+ \mathcal{E}_4\mathcal{E}_5(2N^7-34N^6+110N^5-259N^4+3386N^3-10256N^2+9601N-2550) \\
&+ \mathcal{E}_9(N^8-36N^7+294N^6-588N^5+861N^4-9360N^3 \\
&\quad +28044N^2-27504N+8288)] \qquad \text{(C.9)}
\end{aligned}
$$

$$
\begin{aligned}
& M_{10} \\
= \; & \frac{1}{KN^9} \cdot [\mathcal{E}_1^{10} + \mathcal{E}_1^8\mathcal{E}_2(45N-45) + \mathcal{E}_1^7\mathcal{E}_3(204N^2-444N+240) \\
+ \; & \mathcal{E}_1^6\mathcal{E}_2^2(336N^2-546N+210) + \mathcal{E}_1^6\mathcal{E}_4(588N^3-2086N^2+2338N-840) \\
+ \; & \mathcal{E}_1^5\mathcal{E}_2\mathcal{E}_3(1386N^3-3262N^2+1624N+252) \\
+ \; & \mathcal{E}_1^5\mathcal{E}_5(1069N^4-5144N^3+8292N^2-4973N+756) \\
+ \; & \mathcal{E}_1^4\mathcal{E}_2^3(546N^3-1372N^2+1288N-462) \\
+ \; & \mathcal{E}_1^4\mathcal{E}_2\mathcal{E}_4(1894N^4-6820N^3+6299N^2+41N-1414) \\
+ \; & \mathcal{E}_1^4\mathcal{E}_3^2(791N^4-1563N^3-1223N^2+3543N-1548) \\
+ \; & \mathcal{E}_1^4\mathcal{E}_6(1235N^5-7307N^4+13861N^3-7773N^2-2390N+2374) \\
+ \; & \mathcal{E}_1^3\mathcal{E}_2^2\mathcal{E}_3(1400N^4-5135N^3+7500N^2-8721N+4956) \\
+ \; & \mathcal{E}_1^3\mathcal{E}_2\mathcal{E}_5(1498N^5-8184N^4+13681N^3-10656N^2+11207N-7546) \\
+ \; & \mathcal{E}_1^3\mathcal{E}_3\mathcal{E}_4(1240N^5-2876N^4-4639N^3+11841N^2+1088N-6654) \\
+ \; & \mathcal{E}_1^3\mathcal{E}_7(860N^6-6219N^5+12803N^4-2615N^3-9378N^2-5415N+9964) \\
+ \; & \mathcal{E}_1^2\mathcal{E}_2^4(138N^4-883N^3+1752N^2-1890N+883) \\
+ \; & \mathcal{E}_1^2\mathcal{E}_2^2\mathcal{E}_4(628N^5-4758N^4+11497N^3-21425N^2+30094N-16036) \\
+ \; & \mathcal{E}_1^2\mathcal{E}_2\mathcal{E}_3^2(589N^5-2772N^4+6497N^3-16772N^2+25232N-12774) \\
+ \; & \mathcal{E}_1^2\mathcal{E}_2\mathcal{E}_6(581N^6-5281N^5+14999N^4-25034N^3+55078N^2-80771N+40428) \\
+ \; & \mathcal{E}_1^2\mathcal{E}_3\mathcal{E}_5(510N^6-2119N^5-951N^4-413N^3+41175N^2-74292N+36090) \\
+ \; & \mathcal{E}_1^2\mathcal{E}_4^2(244N^6-622N^5-2331N^4+3941N^3+13575N^2-28503N+13696) \\
+ \; & \mathcal{E}_1^2\mathcal{E}_8(304N^7-3028N^6+7826N^5-274N^4+2232N^3-80267N^2+135809N-62602) \\
+ \; & \mathcal{E}_1\mathcal{E}_2^3\mathcal{E}_3(101N^5-1627N^4+4867N^3-9180N^2+10755N-4916) \\
+ \; & \mathcal{E}_1\mathcal{E}_2^2\mathcal{E}_5(96N^6-2025N^5+8634N^4-22534N^3+49507N^2-57582N+23904) \\
+ \; & \mathcal{E}_1\mathcal{E}_2\mathcal{E}_3\mathcal{E}_4(188N^6-2455N^5+8426N^4-27906N^3+69400N^2-72649N+24996) \\
+ \; & \mathcal{E}_1\mathcal{E}_2\mathcal{E}_7(77N^7-1601N^6+7800N^5-19224N^4+62562N^3-154400N^2+160188N-55402) \\
+ \; & \mathcal{E}_1\mathcal{E}_3^3(31N^6-268N^5+1300N^4-5490N^3+11843N^2-10776N+3360) \\
+ \; & \mathcal{E}_1\mathcal{E}_3\mathcal{E}_6(73N^7-802N^6+1730N^5-8427N^4+58772N^3-139370N^2+122738N-34714) \\
+ \; & \mathcal{E}_1\mathcal{E}_4\mathcal{E}_5(71N^7-509N^6-372N^5-1611N^4+39454N^3-96169N^2+69196N-10060) \\
+ \; & \mathcal{E}_1\mathcal{E}_9(38N^8-696N^7+3151N^6-2478N^5+10794N^4 \\
& \quad -118147N^3+282550N^2-228124N+52912) \\
+ \; & \mathcal{E}_2^5(N^5-50N^4+156N^3-325N^2+407N-189) \\
+ \; & \mathcal{E}_2^3\mathcal{E}_4(4N^6-244N^5+1442N^4-4512N^3+9144N^2-8482N+2648) \\
+ \; & \mathcal{E}_2^2\mathcal{E}_3^2(6N^6-312N^5+1598N^4-4602N^3+9836N^2-9420N+2894) \\
+ \; & \mathcal{E}_2^2\mathcal{E}_6(3N^7-219N^6+1806N^5-7878N^4+27173N^3-51099N^2+39744N-9530) \\
+ \; & \mathcal{E}_2\mathcal{E}_3\mathcal{E}_5(6N^7-311N^6+2115N^5-9781N^4+33007N^3-53474N^2+29416N-978) \\
+ \; & \mathcal{E}_2\mathcal{E}_4^2(3N^7-143N^6+806N^5-2711N^4+13698N^3-25815N^2+12778N+1384) \\
+ \; & \mathcal{E}_2\mathcal{E}_8(2N^8-144N^7+1259N^6-4224N^5+21605N^4 \\
& \quad -87885N^3+143886N^2-78405N+3906) \\
+ \; & \mathcal{E}_3^2\mathcal{E}_4(3N^7-107N^6+909N^5-4384N^4+13282N^3-19556N^2+11371N-1518) \\
+ \; & \mathcal{E}_3\mathcal{E}_7(2N^8-76N^7+459N^6-3393N^5+23257N^4-71490N^3+89216N^2-34315N-3660) \\
+ \; & \mathcal{E}_4\mathcal{E}_6(2N^8-52N^7+261N^6-1574N^5+14347N^4-49750N^3+52008N^2+6620N-21862) \\
+ \; & \mathcal{E}_5^2(N^8-22N^7+120N^6-331N^5+6390N^4-20618N^3+13580N^2+16384N-15504) \\
+ \; & \mathcal{E}_{10}(N^9-45N^8+450N^7-1050N^6+3633N^5-45741N^4 \\
& \quad +157980N^3-173420N^2+15792N+42400)] \qquad \text{(C.10)}
\end{aligned}
$$

List of Publications

[1] 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.

[2] 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.

[3] D. Guo, L. K. Rasmussen, S. Sun, and T. J. Lim, "A matrix-algebraic approach to linear parallel interference cancellation in CDMA," *submitted to IEEE Trans. Commun.*, Dec. 1997.

[4] D. Guo, L. K. Rasmussen, and T. J. Lim, "Linear parallel interference cancellation in random-code CDMA multiuser detection," *submitted to the IEEE J. Selected Areas Commun.*, Sept. 1998.

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

Bibliography

[1] P. W. Baier, P. Jung, and A. Klein, "Taking the challenge of multiple access for third–generation cellular mobile radio systems – a European view," *IEEE Communication Magazine*, vol. 34, pp. 82–89, Feb. 1996.

[2] R. L. Pickholtz, L. B. Milstein, and D. L. Schilling, "Spread spectrum for mobile communications," *IEEE Trans. Veh. Technol.*, May 1991.

[3] A. J. Viterbi, *CDMA – Principles of Spread Spectrum Communication.* Addison–Wesley Publishing Company, 1995.

[4] 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.

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

[6] "Mobile station/base station compatibility standard for dual-mode wideband spread spectrum cellular system." TIA/EIA Interim Standard IS-95, July 1993.

[7] S. Moshavi, "Multi-user detection for DS-CDMA communications," *IEEE Communication Magazine*, vol. 34, pp. 124–136, 1996.

[8] A. Duel-Hallen, J. Holtzman, and Z. Zvonar, "Multiuser detection for CDMA systems," *IEEE Personal Communications*, April 1995.

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

[10] L. K. Rasmussen, T. J. Lim, and T. M. Aulin, "Breadth-first maximum-likelihood detection in multiuser CDMA," *IEEE Trans. Commun.*, vol. 45, pp. 1176–1178, Oct. 1997.

[11] A. Duel-Hallen, "Decorrelating decision–feedback multiuser detector for synchronous code–division multiple–access channel," *IEEE Trans. Commun.*, vol. 41, pp. 285–290, Feb. 1993.

[12] 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.

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

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

[15] T. J. Lim and L. K. Rasmussen, "Adaptive symbol and parameter estimation in asynchronous multiuser CDMA detectors," *IEEE Trans. Commun.*, vol. 45, pp. 213–220, Feb. 1997.

[16] 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.

[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] H. Elders-Boll, H. D. Schotten, and A. Busboom, "Efficient implementation of linear multiuser detectors for asynchronous CDMA systems by linear interference cancellation," *European Transactions on Telecommunications*, no. 4, 1998.

[19] L. K. Rasmussen, T. J. Lim, and A.-L. Johansson, "A matrix-algebraic approach to successive interference cancellation in CDMA," *submitted to IEEE Trans. Commun.*, May 1997. revised July 1998.

[20] S. Sun, L. K. Rasmussen, T. J. Lim, and H. Sugimoto, "A matrix-algebraic approach to linear hybrid interference cancellation in CDMA," in *Proceedings of IEEE 1998 International Conference on Universal Personal Communications*, pp. 1219–1223, Florence, Italy, Oct. 1998.

[21] S. Sun, L. K. Rasmussen, H. Sugimoto, and T. J. Lim, "A hybrid interference canceller in CDMA," in *Proceedings IEEE Fifth International Symposium on Spread Spectrum Techniques and Applications*, vol. 1, pp. 150–154, Sun City, South Africa, Sept. 1998.

[22] M. K. Varanasi, "Group detection for synchronous Gaussian code-division multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 41, pp. 1083–1096, July 1995.

[23] 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.

[24] D. Divsalar and M. Simon, "Improved CDMA performance using parallel interference cancellation," Tech. Rep. 95-21, Jet Propulsion Lab., California Inst. of Tech., Oct. 1995.

[25] 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).

[26] 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.

[27] 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.

[28] 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.

[29] D. Guo, L. K. Rasmussen, S. Sun, and T. J. Lim, "A matrix-algebraic approach to linear parallel interference cancellation in CDMA," *submitted to IEEE Trans. Commun.*, Dec. 1997.

[30] D. Guo, L. K. Rasmussen, and T. J. Lim, "Linear parallel interference cancellation in random-code CDMA multiuser detection," *submitted to the IEEE J. Selected Areas Commun.*, Sept. 1998.

[31] J. G. Proakis, *Digital Communications*. Prentice Hall, 1993.

[32] G. H. Golub and C. F. van Loan, *Matrix Computations*. The Johns Hopkins University Press, 2nd ed., 1989.

[33] S. Verdú, *Multiuser detection*. Cambridge University Press, 1998.

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

[35] L. B. Nelson and H. V. Poor, "Iterative multiuser receivers for CDMA channels: An EM–based approach," *IEEE Trans. Commun.*, vol. 44, pp. 1700–1710, Dec. 1996.

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

[37] S. Haykin, *Adaptive Filter Theory*. Prentice Hall, 3rd ed., 1996.

[38] 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.

[39] D. Tse and S. Hanly, "Linear multiuser receivers: effective interference, effective bandwidth and capacity," *submitted to the IEEE Trans. Inform. Theory*, 1998.

[40] D. Jonsson, "Some limit theorems for the eigenvalues of a sample covariance matrix," *Journal of Multivariate Analysis*, Dec. 1982.

[41] J. S. Li and J. G. Zha, *Linear Algebra*. Press of the University of Science & Technology of China, 1989. (in Chinese).

[42] P. D. Alexander, *Coded multiuser CDMA*. PhD thesis, University of South Australia, 1996.

[43] L. Wei, *Single-user and multi-user modulation and detection methods for mobile communications*. PhD thesis, University of South Australia, Dec. 1994.

[44] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical recipes in C — the art of scientific computing.* Cambridge University Press, 1988.