TECHNISCHE UNIVERSITÄT WIEN
VIENNA UNIVERSITY OF TECHNOLOGY

DISSERTATION

Low Complexity Multiuser Detectors for Randomly Spread CDMA Systems

ausgeführt zum Zwecke der Erlangung des akademischen Grades
eines Doktors der technischen Wissenschaften

eingereicht an der
Technischen Universität Wien
Fakultät für Elektrotechnik und Informationstechnik

durch

Dipl.-Ing. Laura Cottatellucci
Mommsengasse 4/14, A-1040 Wien
Matr. Nr. 0027671

Wien, im März 2006

Date of examination:
24 March 2006

Examiners:
Prof. Dr. Ralf R. Müller
NORWEGIAN UNIVERSITY OF SCIENCE AND TECHNOLOGY
TRONDHEIM, NORWAY

a. o. Prof. Dr. Franz Hlawatsch
TECHNISCHE UNIVERSITÄT WIEN
VIENNA, AUSTRIA

Prof. Dr. David N. C. Tse
UNIVERSITY OF CALIFORNIA AT BERKELEY
BERKELEY, CALIFORNIA, USA

To my mother
To my father

Contents

Kurzfassung

Mehrbenutzerdetektion ermöglicht eine relevante Steigerung der spektralen Effizienz von CDMA-Systemen, die jedoch mit einer stark zunehmenden Komplexität erkauft wird. Diese Arbeit befasst sich mit Entwurf und Analyse von linearen Mehrbenutzerempfängern *niedriger Komplexität* für CDMA-Systeme mit zufälligen Spreizcodes, großen Spreizfaktoren und vielen Benutzern, wenn die Komplexität linearer MMSE-Detektoren keine Echtzeitimplementierung zulässt.

Die vorgeschlagenen Mehrstufen-Detektoren haben eine modulare Struktur, die auf signalangepassten Filtern aufbaut; sie führen eine Projektion des beobachteten Signals auf einen Unterraum durch und filtern dieses anschließend. Die Komplexitätsreduktion ergibt sich, indem man die optimalen Filterkoeffizienten eines großen, aber endlichen Systems durch universelle Koeffizienten, d.h. durch die optimalen Koeffizienten eines Systems mit unendlicher Benutzerzahl und unendlichem Spreizfaktor approximiert. Der Entwurf universeller Gewichte geschieht mittels ausgefeilter Methoden der Theorie der Zufallsmatrizen und der freien Wahrscheinlichkeit. Dies ermöglicht lineare Mehrbenutzerdetektion mit einer Komplexität pro Bit die, wie bei einem Enfangsfilter, das nur auf einen Benutzer angepasst ist, nur linear mit der Anzahl der Benutzer wächst.

Die vorgeschlagenen Detektoren sind robust gegenüber Kanalnichtidealitäten wie frequenzselektivem Schwund oder Asynchronität; sie erreichen annähernd lineares MMSE-Verhalten mit einer Anzahl von Stufen, die sowohl viel kleiner als die Anzahl der Benutzer als auch von dieser unabhängig ist. Der vorgeschlagene Entwurf umfasst auch CDMA-Systeme mit mehreren Sende- und Empfangantennen und räumlich korrelierter Diversität im Sender und Empfänger.

Für asynchrone Systeme wird ein gleitendes Beobachtungsfenster vorgeschlagen, wodurch die Detektorgüte keinen Einbruch durch ein endliches Beobachtungsfenster erleidet und die Komplexität, im Gegensatz zu linearen MMSE-Detektoren, vergleichbar mit einem synchronen System bleibt. Dank diesem Ansatz können sogar lineare MMSE-Detektoren mit endlichem Fenster übertroffen werden.

Aufbauend auf Eigenschaften von Zufallsmatrizen, die im Verlauf dieser Arbeit entdeckt wurden, wird ein allgemeiner Ansatz zur asysmptotischen Analyse einer breiten Klasse von linearen Mehrbenutzer-Detektoren vorgestellt, inklusive der

vorgeschlagenen Mehrstufen-Detektoren.

Schließlich werden die Auswirkungen von Asynchronität, Chip-Wellenform und räumlicher Korrelation der Kanäle (bei mehreren Antennenelementen in Sender und Empfänger) untersucht.

Abstract

Multiuser detection can achieve a relevant increase in the spectral efficiency of CDMA systems at the cost of a considerable increase in complexity.

This work is focused on the design and analysis of *low complexity* linear multiuser receivers for CDMA systems with random spreading codes, large spreading factors and large number of users, when even linear MMSE detectors are computationally very intensive in real-time implementations.

The proposed multistage detectors have a modular structure based on matched filters and perform the projection of the observed signals onto a subspace and a successive filtering. The reduction in complexity is achieved by approximating the optimum filter coefficients of a large but finite system by universal weights, i.e. the optimum weights of a system with infinite users and spreading factor. The design of universal weights uses sophisticated tools of random matrix theory and free probability theory. Such a design enables linear multiuser detection with a complexity order per bit that scales linearly with the number of users as in a single user matched filter.

The proposed multistage detectors with universal weights efficiently cope with channel non-ideality such as frequency selective fading and asynchronism. They achieve near-linear MMSE performance with a number of stages much lower than the number of users and independent of it.

For asynchronous systems the proposed detectors include a sliding observation window, so that they do not suffer from performance degradation due to a finite observation window. They keep the same complexity as their counterpart for synchronous systems, in contrast to the linear MMSE detectors. With this approach they can even outperform a finite-window linear MMSE detector.

The design of multiuser detectors includes also CDMA systems with multiple transmitting and receiving antennas and spatial correlation of the channels.

Benefitting of properties of random matrices discovered in this work, a general framework for the asymptotic analysis of a wide class of linear multiuser detectors, including the proposed multistage detectors, is presented.

The effects of asynchronism, chip-pulse waveforms, and correlated spatial diversity are analyzed.

Acknowledgements

This thesis contains part of my research work in Forschungszentrum Telekommunikation Wien and it has been partially supported by the Kplus Programme of the Austrian Government.

First of all I would like to thank my committee members.

- Franz Hlawasch thought me most of the basic concepts of signal processing that I know in a very lucid and organized fashion. His thorough review of this manuscript and his detailed comments improved the readability and the editorial aspect of this thesis.

- Various works of David Tse have been a constant presence in my daily working life and the greatest source of inspiration for my research.

Particular thanks are to Alex Grant for allowing the final polishings on this manuscript and enabling the final defense of my Ph.D..

I would like to express my deepest gratefulness to all people that in different ways contributed to this work.

This Ph.D. thesis would not have been possible without the precious help I received from my advisor Ralf Müller. He created the conditions for the development of my research and allowed me to define my identity as researcher. He introduced me to random matrix theory and multiuser communications that played a major role in my Ph.D. thesis. His encouragements and his support have kept me motivated and determined even in the most difficult periods of my doctoral work.

The course of my graduate studies has been an important time to try to clarify and to define my idea of research and of research in engineering. This process received fundamental contributions from Giuseppe Caire, Mérouane Debbah, and Ralf Müller. In particular, the former helped me to retrieve the original reasons that motivated my choice of engineering studies.

Mérouane Debbah transmitted me energy and enthusiasm with interesting and fruitful discussions. His constant support has been fundamental to overcome difficult periods. In this respect the togetherness of Michele Casa, Giuseppe Razzano, Fabio Ricciato, Francesco Vacirca, and Danilo Valerio have been also helpful.

Eitan Altman, Tijani Chahed and Mérouane Debbah made the long waiting between the conclusion of this work and its defense smooth, stimulating, and fruitful.

Claudio Weidmann and Thomas Magesacher helped me for the German translation of the abstract. Additionally, some discussions with Claudio Weidmann were useful in the proof of Lemma 5. Gautam Gupta thoroughly read and discussed with me the proofs of Theorem 8 and of Theorem 9. Mérouane Debbah read the whole manuscript and provided useful comments.

Finally, I could not have done anything without the precious support of my family and of my friends. My mother, my sister Anna, my uncle Ezio, my close friend Settimia Michetti, who is practically part of my family, were always there for me. My friends Enrico Camilli, Egisto Catalani, Mérouane Debbah, Julia Gross, Emmanuele Grosicki, Liana Müller, and Roberto Pietrosanti have shown to be invaluable assets for my life.

Previously Published Work

Partial results of Chapter 3 appeared in the following papers:

- Laura Cottatellucci and Ralf R. Müller. Asymptotic analysis of multistage receivers for multipath fading channels. *Proc. of IEEE International Symposium on Information Theory (ISIT)*, page 243, Yokohama, Japan, June/July 2003.

- Laura Cottatellucci and Ralf R. Müller. Asymptotic design and analysis of full-multistage based receivers for multipath fading channels. *Proc. of Winter School on Coding and Information Theory*, Monte Verità, Switzerland, Feb 2003.

- Laura Cottatellucci and Ralf R. Müller. Multiuser interference mitigation with multistage detectors: Design and analysis for unequal powers. *Proc. of 36^{th} Annual Asilomar Conference on Signals, Systems, and Computers*, pp. 1948–1952, Pacific Grove, CA, U.S.A., Nov 2002.

- Laura Cottatellucci and Ralf R. Müller. Asymptotic design and analysis of multistage detectors with unequal powers. *Proc. of IEEE Information Theory Workshop (ITW)*, pp. 167–170, Bangalore, India, Oct 2002.

- Laura Cottatellucci and Ralf R. Müller. Performance assessment of polynomial detector in MIMO channels with measured data. *Proc. of IEE Technical Seminar on MIMO Communication Systems*, London, U.K., Dec 2001.

Chapter 3 appeared in its entirety as

- Laura Cottatellucci and Ralf R. Müller. A systematic approach to multistage detectors in multipath fading channels. *IEEE Transactions on Information Theory*, vol. 51, no. 9, pp. 3146–3158, Sep 2005.

The results on chip synchronous and symbol asynchronous CDMA systems presented in Section 4.3 appeared in

- Laura Cottatellucci, Ralf R. Müller and Mérouane Debbah. Efficient implementation of multiuser detectors for asynchronous CDMA. *Proc. 42nd Allerton Conf. on Communication, Control and Computing* , Monticello, Illinois, Sep/Oct 2004.

- Laura Cottatellucci, Ralf R. Müller and Mérouane Debbah. Asymptotic analysis of linear detectors for asynchronous CDMA systems. *Proc. of IEEE International Symposium on Information Theory (ISIT)*, Chicago, IL, U.S.A., June/July 2004.

A preliminary version of the analysis of asynchronous CDMA systems presented in Section 4.4 and Section 4.5 appeared in

- Laura Cottatellucci and Ralf R. Müller. Multistage detectors for asynchronous CDMA. Invited for *International Zurich Seminar on Communications (IZS)*, Zurich, Switzerland, Feb 2004.

The analysis of classical linear multiuser detectors in CDMA systems with correlated spatial diversity in Chapter 5 appeared in the following two papers:

- Laura Cottatellucci and Ralf R. Müller. CDMA Systems with correlated spatial diversity: A generalized resource pooling result. *IEEE Transactions on Information Theory*, submitted July 2004, to be revised.

- Laura Cottatellucci and Ralf R. Müller. A generalized resource pooling result for correlated antennas with applications to asynchronous CDMA. *Proc. of International Symposium on Information Theory and its Applications (ISITA)*, Parma, Oct 2004.

The design and analysis of multistage detectors with universal weights for systems with spatial diversity appear in

- Ralf R. Müller and Laura Cottatellucci. Joint antenna combining and multiuser detection. In *Smart antennas in Europe–State of the art*, T. Kaiser and A. Bourdoux and H. Boche and J. Fonollosa and J. B. Andersen and W. Utschik editors, EURASIP book series on Applied Signal Processing, 2005.

List of Figures

List of Tables

Glossary

Acronyms

BER	bit error rate
BPSK	binary phase shift keying
CDMA	code-division multiple-access
c.d.f.	cumulative distribution function
e.d.f.	empirical distribution function
GSO	Gram-Schmidt orthogonalization
i.i.d.	independent and identically distributed
l.h.s.	left hand side
MAI	multiple access interference
MAP	maximum a posteriori
MIMO	multiple input multiple output
MMSE	minimum mean square error
MSE	mean square error
MSWF	multistage Wiener filter
p.d.f.	probability density function
PIC	parallel interference cancelling
PDP	power delay profile
QPSK	quadrature phase shift keying
r.h.s.	right hand side
SER	symbol error rate
SINR	signal-to-interference and noise ratio
SNR	signal to noise ratio
SUMF	single user matched filter

Constants

$\boldsymbol{I}_n$	identity matrix of size $n \times n$
π	the number pi $\pi = 3.141592653589793\ldots$
e	Euler's number $\mathrm{e} = 2.7182818284459045\ldots$
j	imaginary unit

Functions

$1_{\mathcal{A}}(x)$	indicator function on the set $\mathcal{A}$ $1_{\mathcal{A}}(x) = \begin{cases} 1 & x \in \mathcal{A}, \\ 0 & \text{otherwise.} \end{cases}$
$1(x)$	indicator function on a right unbounded interval $1(x) = \begin{cases} 1 & x \geq 0, \\ 0 & \text{otherwise.} \end{cases}$
$1(x_1, x_2, \ldots, x_N)$	N-dimensional indicator function on right unbounded intervals $1(x_1, x_2 \ldots x_N) = \prod_{n=1}^{N} 1(x_n)$
$\delta(\lambda)$	Dirac's delta function
δ_{ij}	Kronecker symbol

Matrix Notation

$\boldsymbol{X} = (x_{ij})_{i=1,\ldots,n_1}^{j=1,\ldots,n_2}$	$n_1 \times n_2$ matrix whose (i,j)-element is the scalar x_{ij}
$\boldsymbol{X} = (\boldsymbol{X}_{ij})_{i=1,\ldots,n_1}^{j=1,\ldots,n_2}$	$n_1 q_1 \times n_2 q_2$ block matrix whose (i,j)-element is the $q_1 \times q_2$ matrix $\boldsymbol{X}_{ij}$
$\boldsymbol{X} > 0$	positive definite matrix

Operators

$.^H$	Hermitian operator
$.^T$	transposition operator
$.^*$	complex conjugate
$*$	convolution operator
argmin(.)	argument minimizing the expression in parentheses
E{.}	expectation
$\mathrm{Im}(\cdot)$	imaginary part
$\max(\cdot)$	maximum
$\min(\cdot)$	minimum
$a \bmod b$	remainder of the division of a by b
$\mathcal{O}$	Landau operator
$\Pr(\cdot)$	probability
$\mathrm{Re}(\cdot)$	real part
$\mathrm{span}(\cdot, \cdot, \ldots, \cdot)$	space spanned by the vector arguments
$\mathrm{tr}(\cdot)$	trace of the matrix argument
$\lVert \cdot \rVert_2$ or $\lVert \cdot \rVert$	Frobenius norm ($\lVert \boldsymbol{A} \rVert_2 = \sqrt{\mathrm{tr}(\boldsymbol{A}\boldsymbol{A}^H)}$)
$\lvert \cdot \rvert$	spectral norm ($\lvert \boldsymbol{A} \rvert = \max_{\boldsymbol{x}^H \boldsymbol{x} \leq 1} \lVert \boldsymbol{A}\boldsymbol{x} \rVert_2$)
$\otimes$	Kronecker product
$\lfloor \cdot \rfloor$	maximum integer not greater than the argument
$\lceil \cdot \rceil$	minimum integer not lower than the argument
$\wedge$	logical 'and'
$\vee$	logical 'or'

Relations

$\xrightarrow{a.s.}$, $\stackrel{a.s.}{=}$	convergence with probability 1 or almost sure convergence
$\xrightarrow{\mathcal{P}}$, $\stackrel{\mathcal{P}}{=}$	convergence in probability

Sets

$\mathbb{R}$	field of real numbers
$\mathbb{R}_0^+$	field of non-negative real numbers
$\mathbb{C}$	field of complex numbers
$\mathbb{Z}$	field of integer numbers
$\mathbb{Z}^+$	field of positive integer numbers

1 Introduction

In the course of time, engineers have had the challenging and exciting function of transforming utopia to dreams, and dreams to actual reality. So, Icarus' utopia of flying towards to the sun and Verne's dreams of walking on the moon and exploring the deep sea are concrete possibilities nowadays. A common dream of futurologists and visionaries, writers and children is the gift of ubiquity. As pointed out by Negroponte [1], the substitution of the heavy and slow atom with the bit, capable of travelling at the speed of light, allowed the dream to come true. Theoretical and applied digital communication engineers accepted the challenge: the intense, successful development of wireless communications at the late 1980's has been a unprecedented step forward to the achievement of ubiquity.

In contrast to the wireline medium characterized by reliability and large capacity, the wireless medium is unreliable and has low capacity due to path-loss, shadowing, and intersymbol interference. Additionally, it is intrinsically limited and can easily become a scarce resource when the service demand increases. Consequently, an efficient utilization of the available radio spectrum is a key requirement to make the technological reality as close as possible to the dream.

The last decades have experienced a deep rethinking and reformation of the conceptions about wireless multiuser communications, which have opened new ways and possibilities to exploit the wireless medium efficiently. New concepts such as multiuser detection, multiple antenna elements, and opportunistic communications are playing major roles in the field nowadays.

For some time, it was widely believed that the interference introduced by a large number of equal-power users was accurately approximated by a Gaussian random variable and consequently the single user matched filter was almost optimal for large systems. The near-far effect was considered detrimental and power control was the only available tool to combat them. In the early eighties, Verdú recognized the wrong assumptions that led to this misconception [2–5] and pointed out the large improvements in spectral efficiency achievable by taking into account the structure of the multiuser interference and by mitigating the cross-talk among users with an optimum multiuser detector. Multiuser detection techniques efficiently compensate for the near-far "problem" and open the way to the recent discovery in [6] that turned the near-far "problem" into an advantage for heavily loaded systems.

In the early nineties it was discovered that adding antennas in rich scattering environments increases proportionally the point-to-point data rates without extra transmitted power or bandwidth [7, 8]. These systems are referred to as multiple input multiple output (MIMO) systems. Antenna arrays provide spatial diversity and introduce new degrees of freedom in the received signal. The successive joint processing of the multiple received signals makes use of these further degrees of freedom to enhance the system capacity.

In the last lustrum, the resource pooling effect, discovered by Hanly and Tse [9], demonstrated that degrees of freedom in space and frequency are interchangeable. Moreover, the total number of degrees of freedom is the product of the degrees of freedom in space and frequency. A system with spreading factor N and L receive antennas is in many respects equivalent to a system with a single antenna and spreading factor NL. This suggests the idea to treat the two effects in the same way performing antenna array processing and multiuser detection jointly. Joint processing significantly outperforms techniques that exploit separately the degrees of freedom in space and in frequency.

Although these techniques promise large enhancements in spectral efficiency and are really appealing, their implementations in real systems is not straightforward and several issues are still open. The major problem is related to their complexity: the optimum maximum likelihood detector for CDMA systems has a complexity which is exponential in the number of users. This has fuelled the research on sub-optimum multiuser detectors with a substantially lower complexity in exchange for some tolerable performance degradation.

Our work is focused on the design of low complexity multiuser detectors for symbols transmitted on the uplink[1] mobile radio channel of a CDMA system. Our attention is concentrated especially on challenging scenarios with long random spreading sequences[2], large spreading factor, and many users. Systems with these characteristics are supported by current standards— an example is the FDD mode of UMTS— and the use of antenna arrays makes them more and more widespread. In these cases, even the quadratic complexity order per bit of a linear MMSE detector or a decorrelator can be computationally too intensive for real-time implementations.

This work focuses on linear multistage detectors with universal weights. They reach a very good compromise between performance and complexity, taking advantage exactly of what is considered to be deleterious in such scenarios: the numerous users and the long spreading sequences. In fact, multiuser communication systems can be modelled by random matrices whose entries are in general statistically dependent and whose size depends on the number of users, the spreading factor, the number of receiving and transmitting antennas, the observation window length, etc.

[1]The downlink channel can be regarded as a special case of uplink channel with channel coefficients equal for all users. Therefore, we can focus on the uplink without loss of generality.

[2]By long spreading sequences we mean spreading sequences that span more than a symbol interval.

The design of the universal weights benefits from the asymptotic self-averaging properties of these random matrices and reduces the computationally most demanding part of the detector to a computation of a polynomial depending on the statistical properties of the random matrices via few essential system parameters.

The statistical structure of random matrices considered in this work includes the typical non-ideality of a CDMA system. The effects of flat fading and frequency selective fading have been considered as well as asynchronism and the effects of the chip-pulse waveforms. Channel correlation in case of multiple antenna elements at the transmitters and the receiver has also been investigated.

We will now present an outline of the material and results contained in the individual chapters of this work.

In Chapter 2 we review the most relevant linear detectors and discuss the large system[3] performance analysis based on random matrix theory. The principles of random matrix theory useful for the following developments are also illustrated.

In Chapter 3 two families of linear multistage detectors with universal weights are introduced. The universal weights for asynchronous CDMA systems with flat fading and frequency selective fading channels are derived. A general framework for the performance analysis of a large class of linear multiuser detectors is proposed.

The multistage detectors perform a projection of the received signal onto a Krylov subspace and a successive filtering according to some optimality criterion. In case of detection of multiple users, as in uplink channels, a reduction in complexity requires both an appropriate choice of the bases of the projection subspaces and the use of universal weights. The bases of the projection subspaces should enable joint projection of the received signal for all K users of interest. In such a way most of the computations for the projections become identical and the complexity drops by a factor of K. The universal weighting is based on the approximation of the weights, optimum according to some optimality criterion, by weights optimum in the same sense for large CDMA systems.

The detectors "Type J-I" proposed in this work perform the joint projection of the received signal for all users and the filtering of the projections in a way that is asymptotically optimum in a MSE sense for each user. Thanks to the joint projection and the universal weighting, the Type J-I detector has near-linear MMSE performance with the same complexity order per bit as the *single user matched filter*.

The design of universal weights for detectors Type J-I is based on a self-averaging property of random matrices established in this work. The convergence of the empirical eigenvalue distribution of some random matrices is well known and widely utilized in multiuser communications (e.g., [9–12]). In this work we prove that also

[3]Throughout this work we refer to CDMA systems with number of users and spreading factor going to infinity with constant ratio as large CDMA systems.

the diagonal elements of powers of some random matrices converge to a deterministic limit depending on a small set of system parameters. The universal weights of Type J-I detectors are designed making use of these diagonal elements.

It is possible to design multistage detectors performing joint projection and universal weighting by utilizing the convergence of the empirical eigenvalue distribution at the cost of performance degradation. This yields the class of detectors referred to as detectors Type J-J. They have the same complexity order per bit as the *conventional* detectors and worse performance than detectors Type J-I.

Multistage detectors with universal weights based on the convergence of the eigenvalue distribution that are presented in parallel works [13, 14] utilize bases of the projection subspace that do not enable joint projection. In case of joint detection of multiple users they have the same large system performance as detectors Type J-I but keep the same complexity order per bit as the *linear MMSE* detector that is typically some order of magnitude higher than the complexity order per bit of a conventional detector.

The usefulness of the convergence of the diagonal elements is evident when we consider that the performance of linear multiuser detectors and especially multistage detectors is more naturally related to these diagonal elements than to the eigenvalue moments of some random matrices and only some optimum detector (linear MMSE detector and multistage Wiener filter) can be analyzed by the eigenvalue moments. Thanks to this property we develop a general framework for the performance analysis of a large class of linear detectors including the multistage Wiener filters, polynomial expansion detectors, and parallel interference cancelling detectors. For large systems, detectors Type J-I and Type J-J are equivalent to multistage Wiener filters and polynomial expansion detectors, respectively. The performance analysis disproves the widespread belief of the equivalence between multistage Wiener filters and polynomial expansion detectors. In general, the former outperform the latter and they are equivalent only with perfect power control.

In Chapter 4 the previous results are extended to asynchronous systems. In this scenario the multiple access interference is correlated from a symbol interval to the other. The acquirement of sufficient statistics requires an infinite observation window. A linear MMSE detector of practical use suffers from performance degradation due to the finite observation window. Additionally, the complexity increases with the observation window length. We propose a slightly modified version of the multistage detectors in Chapter 3 whose observation window expands with the number of stages. In contrast to the linear MMSE detector, detectors Type J-I for chip synchronous and symbol asynchronous CDMA systems are equivalent in performance and complexity to the corresponding multistage detectors for synchronous systems. The proposed multistage detectors employ a sliding observation window. Thanks to this feature detectors Type J-I achieve uniform multiuser efficiency for all users.

Since the polynomial expansion detectors and, thus, detectors Type J-J are intrinsically suboptimal for asynchronous systems, i.e. they have worse performance than detectors Type J-I also in case of perfect power control, the design is focused on detectors Type J-I.

Two relevant topics of investigation for asynchronous systems are addressed in the design of multistage detectors:

- The choice of a set of observables with the twofold aim of achieving low complexity and optimum or nearly optimum performance ;
- The effects of the chip pulse waveforms.

Processing the received signal by a lowpass filter and then sampling it at the Nyquist rate turned out to be the most convenient way to acquire the observables. It provides sufficient statistics and, in the meantime, enables joint processing and detection of all users. For a large class of chip-pulse waveforms, the universal weights for asynchronous systems take into account the chip-pulse effects via some coefficients that are very simply related to the power spectral density of the chip pulses.

A linear MMSE detector with given observation window optimizes the utilization of the available observables. The multistage detector with expanding observation window performs a suboptimal utilization of the observables but uses a wider set of statistics. The effects on the performance of these two approaches are investigated. For a linear MMSE detector we provide an algorithm to determine the large system SINR of any symbol whose spreading sequence is partially or completely received in the observation window. The performance of Type J-I detectors is also investigated. We show that the proposed multistage detector with a sufficiently large number of stages can outperform the linear MMSE detector.

The general framework for the performance analysis of a large class of detectors introduced in Chapter 3 is extended to asynchronous systems.

The joint effects of chip pulse waveforms and the distribution of the arrival time of the signals is analyzed. As long as B, the bandwidth of the chip-pulse waveform, is not greater than half the chip rate $\frac{1}{T_c}$, i.e. $B \leq \frac{1}{2T_c}$, the performance of asynchronous and synchronous systems is equivalent and independent of the arrival time distribution. As the bandwidth increases, the effects of the arrival time distribution becomes relevant and dependent on the chip pulse waveform. For square root Nyquist and raised cosine chip waveforms, the output SINR of detectors optimum in an MSE sense increases with the bandwidth for large asynchronous CDMA systems. In contrast, for large synchronous CDMA systems it keeps constant (square root Nyquist waveforms) or decreases (raised cosine waveforms) as the bandwidth increases.

The use of random matrix theory allows a concise and insightful description of the asynchronous large system behaviour. Few system parameters are sufficient to capture the system performance, namely the system load, i.e. the number of received symbols per chip, the variance of the additive Gaussian noise, the moments of the received power distribution and some coefficients related to the power spectral density of the chip-pulse waveforms.

In Chapter 5, CDMA systems with multiple antenna elements at the transmitters and the receivers (multiuser MIMO systems) and with possibly correlated channels are investigated. Linear multistage detectors Type J-I and Type J-J with universal weights are designed. The general framework proposed in Chapter 3 for the performance analysis of a large class of linear multiuser detectors is extended to this scenario with correlated spatial diversity.

The large system analyses of the linear MMSE detectors, the single user Baysian receivers and the single user matched filters are extended to the case of correlated channel gains proving rigourously and generalizing the results in [9].

Thanks to the random spreading, the performance of the investigated linear multiuser detectors are independent of the channel correlation at the transmitters and only the correlation at the receivers plays a relevant role.

The large system performance of the linear MMSE detector, the single user Bayesian filter and the single user matched filter are described by deterministic square matrices, $\boldsymbol{A}$, with dimensions equal to the number of receiving antenna elements. In contrast to the case of CDMA systems with single receiving antennas, the multiuser efficiency does not characterize univocally the systems. In fact, the multiuser efficiency depends on the direction of the user channel gain vector with respect to the eigenvectors of $\boldsymbol{A}$.

The conditions under which the resource pooling effect [9] arises, i.e., the interchangeability between degrees of freedom in space and frequency holds, are generalized.

Chapter 6 concludes the work discussing the contributions and drawing guidelines for future developments in this field.

2 Linear Multiuser Detection and Random Matrices

2.1 Introduction

For some time the development of spread-spectrum systems was driven by the belief that matched filter receivers were approximately optimum in large systems with equal powers since the multiple access interference could be modelled as Gaussian noise.

In its seminal work [2–4] Verdú discovered the enormous improvements in performance achievable by taking into account the structure of the multiple access interference instead of modelling it simply as a Gaussian noise.

The rationale behind this is that the output of a bank of filters matched to the spread waveforms of users provides sufficient decision statistics for the detection of all users [15]. In contrast, the output of a filter matched to the spread waveform of the user of interest is not a sufficient statistic for the detection of such a user.

This breakthrough opened the way to a flourishing technical and scientific production in multiuser detection.

The promises of multiuser detection in terms of increase in spectral efficiency could be fulfilled at the cost of a considerable increase in complexity. In fact, the optimum receiver investigated in [15] allows a dramatic improvement in performance in exchange for an increase in complexity, which is exponential in the number of the users. Therefore, there is a strong demand for algorithms that simplify the signal processing required for theoretically optimum communications.

The significant efforts devoted to the design of detectors for signals impaired by structured interference from other users yielded many suboptimal algorithms. An exhaustive overview on multiuser detection is beyond the scope of this work. The interested reader can refer to [16] and references therein.

In this chapter we focus on a large class of detectors, called, with an abuse in denomination, linear detectors. A linear detector consists of a linear filter followed by a set of threshold devices. They have been introduced with the goal of finding an acceptable compromise between performance and complexity. In fact, they yield a substantial improvement in performance compared to the conventional matched filter, while maintaining a lower complexity than the optimum detector investigated in [15].

Modelling of spreading matrices in CDMA systems by random matrices has been extremely fruitful for the theoretical analysis of systems with linear detectors. In this respect the interested reader can refer to the pioneering works in [11], [10], and [17]. In the large system limit, as both the transmitted signals K and the spreading factor N tend to infinity with a fixed ratio, the random matrices show self-averaging properties. These allow the description of the system in terms of few macroscopic system parameters and thus provide deep insights into the system behaviour.

Random matrix theory has proven to be a powerful tool not only from the theoretical perspective of performance analysis but also from the practical point of view of receiver design [18]. Since the low complexity detectors proposed in this work benefit particularly from such a tool, Chapter 2 illustrates also fundamental concepts of random matrix theory relevant in the development of the work.

2.2 Linear Multiuser Detection

The class of linear multiuser detectors consists of decision algorithms performing a linear transformation $T : \mathbb{C}^N \to \mathbb{C}^K$ on the decision statistics followed by a set of scalar quantizers.

Let $\mathbb{B}$ be the set of modulation symbols. The scalar quantizer is a nonlinear transformation

$$\mathrm{quant}_{\mathbb{B}} : \mathbb{C} \to \mathbb{B}.$$

$\mathrm{quant}_{\mathbb{B}}$ associates to a complex number the closest[1] element in the set $\mathbb{B}$. $\mathrm{quant}_{\mathbb{B}}^{(K)}$ denotes a nonlinear function of a K-dimensional complex vectors onto a K-dimensional vector in $\mathbb{B}$, i.e., $\mathrm{quant}_{\mathbb{B}}^{(K)} : \mathbb{C}^K \to \mathbb{B}^K$. The function $\mathrm{quant}_{\mathbb{B}}^{(K)}$ performs element-wise a $\mathrm{quant}_{\mathbb{B}}$ transformation. Then, the signals detected by linear multiuser detectors are given by

$$\boldsymbol{b}_{\text{det}} = \mathrm{quant}_{\mathbb{B}}^{(K)}(T(\boldsymbol{y})). \tag{2.1}$$

2.2.1 System Model

In this chapter a synchronous CDMA system on the uplink[2] of a flat fading mobile radio channel impaired by additive white Gaussian noise is considered.

[1] In this context we adopt the Euclidian distance as metric.

[2] The system model of the downlink channel can be regarded as a special case of the system model of an uplink channel with fading coefficients equal for all signals. Therefore, all results presented in this work can be specialized to the downlink channel. In the following, we will not consider the donwlink channel explicitly.

The use of such a system to illustrate the most relevant linear detectors enables us to keep the exposition simple and, in the meanwhile, to capture the features of linear detectors.

K users are active in the system and use a common spreading factor N. The system load, i.e., the number of transmitted symbols per chip, is defined as

$$\beta = \frac{K}{N}. \tag{2.2}$$

Each user and the base station are equipped with a single antenna. User k transmits a spread signal given, in the base-band domain, by

$$s_k(t) = \sum_{m=-\infty}^{+\infty} b_k[m] c_k^{(m)}(t). \tag{2.3}$$

Here, $b_k[m]$ is the m^{th} transmitted symbol belonging to the modulation symbol set $\mathbb{B}$;

$$c_k^{(m)}(t) = \sum_{u=0}^{N-1} s_{k,m}[u] \psi(t - mT_s - uT_c) \tag{2.4}$$

is the spreading waveform; and $s_{k,m}[u]$, $u \in [0, \dots, N-1]$, are elements of the signature sequence of user k in the m^{th} symbol interval. The spreading sequences are normalized to have unit energy, i.e. $\sum_{u=0}^{N-1} |s_{k,m}[u]|^2 = 1$, $\forall k, m$. T_s and T_c are the symbol and chip intervals, respectively. $\psi(t)$ is a square root Nyquist chip pulse waveform with unitary energy common to all users. The received signal is given by

$$y(t) = \sum_{k=1}^{K} a_{kk} s_k(t) + n(t), \tag{2.5}$$

where a_{kk} is a flat fading channel coefficient of user k and $n(t)$ is additive zero mean complex white Gaussian noise with two sided spectral density N_0. The received signal is processed by a filter matched to the chip-pulse waveform and sampled at the chip rate[3].

The discrete-time baseband signal at the receiver in the m^{th} symbol interval is given by

$$y[p] = \sum_{k=1}^{K} a_{kk} b_k[m] s_{km}[p - Nm] + n[p] \qquad p = Nm, \dots, N(m+1) - 1 \tag{2.6}$$

where $n[p]$ is white additive Gaussian noise with variance $\sigma^2 = N_0$.

[3]This approach will be further discussed and analyzed in Section 4.4.3.

In matrix notation, the system model is given by

$$\begin{aligned}\boldsymbol{y}(m) &= \boldsymbol{S}(m)\boldsymbol{A}\boldsymbol{b}(m) + \boldsymbol{n}(m)\\ &= \boldsymbol{H}(m)\boldsymbol{b}(m) + \boldsymbol{n}(m)\end{aligned} \tag{2.7}$$

where $\boldsymbol{y}(m)$ is the N-dimensional vector of received signal in the m^{th} symbol interval, i.e. $\boldsymbol{y}(m) = (y[mN], \dots y[N(m+1)-1])$, and $\boldsymbol{b}(m)$ is the K-dimensional vector of transmitted symbols. Hereinafter, we assume that the transmitted symbols are uncorrelated with zero mean[4] and, without loss of generality, with unit variance. Additionally, they are independent of the noise. $\boldsymbol{A}$ is a $K \times K$ complex diagonal matrix whose k^{th} diagonal element a_{kk} is the channel fading coefficient of user k. $\boldsymbol{S}(m)$ is the $N \times K$ matrix of spreading sequences whose k^{th} column is the spreading sequence of user k in the m^{th} symbol interval.

Throughout this chapter we focus on the detection of the symbols transmitted in the m^{th} symbol interval and we drop the index m in (2.7) without causing confusion.

2.2.2 Matched Filters

The single user matched filter, also called *conventional detector*, is the simplest strategy to demodulate CDMA signals and the optimal solution in single user systems.

In case of fading channels we consider a coherent matched filter and assume perfect knowledge of the fading coefficients. The soft detected symbol of user k is given by

$$\widehat{b}_{\mathrm{MF},k} = \int_{-\infty}^{+\infty} a_k c_k^*(t) y(t) \mathrm{d}t$$

in continuous time. In discrete time

$$\widehat{b}_{\mathrm{MF},k} = \boldsymbol{h}_k^H \boldsymbol{y},$$

where $\boldsymbol{h}_k$ is the k^{th} column of $\boldsymbol{H}$.

The matched filter is widely used in CDMA systems because of its low complexity. It is optimized for single user systems and does not take into account the effects of the structured multiple access interference. Therefore, its performance is very poor.

Interestingly, the output of a bank of filters matched to the spreading waveforms of all users provides a sufficient statistic [15] for multiuser detection. Thus, the matched filters are often used as the receiver front-end for the subsequent multiuser detection.

[4]The assumption on the mean is typically verified by the modulation constellations in use.

2.2.3 Linear MMSE Detection

The concept of linear MMSE detection originates from turning the problem of detection of transmitted symbols in a CDMA system into a problem of linear *estimation* [19] followed by a quantizer (see (2.1)) and from requiring the minimization of the mean square error between $\boldsymbol{b}$, the vector of transmitted symbols, and its linear *estimate* $\widehat{\boldsymbol{b}}$. Thus, the linear MMSE detector consists of

- A linear estimator defined by the matrix $\boldsymbol{T}_{\text{MMSE}}$ such that the linear estimate $\widehat{\boldsymbol{b}} = \boldsymbol{T}_{\text{MMSE}}\boldsymbol{y}$ minimizes the mean square error (MSE) $\mathrm{E}\{\|\widehat{\boldsymbol{b}} - \widehat{\boldsymbol{b}}\|^2\}$;
- A subsequent set of threshold devises.

The output of the linear MMSE detector is given by

$$\widehat{\boldsymbol{b}}_{\text{MMSE}} = \text{quant}_{\mathbb{B}}^{(K)}(\boldsymbol{T}_{\text{MMSE}}\boldsymbol{y}).$$

The linear transform $\boldsymbol{T}_{\text{MMSE}}$ that minimizes the MSE can be obtained by applying the following results of Bayesian estimation theory for general linear models (e.g. [20]) to system model (2.7).

Theorem 1 ***(Bayesian Gauss-Markov theorem)*** *Let the observed data be described by the Bayesian linear model form*

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{b} + \boldsymbol{n}, \tag{2.8}$$

where $\boldsymbol{y}$ is an $N \times 1$ observed data vector, $\boldsymbol{H}$ is a $N \times K$ observation matrix, $\boldsymbol{b}$ is a $K \times 1$ random vector of parameters whose realization is to be estimated and has mean $\mathrm{E}\{\boldsymbol{b}\}$ and covariance matrix $\mathbf{C}_{\boldsymbol{b}}$, and $\boldsymbol{n}$ is an $N \times 1$ random vector with zero mean and covariance matrix $\mathbf{C}_{\boldsymbol{n}}$. It is uncorrelated with $\boldsymbol{b}$ (the joint p.d.f. $f(\boldsymbol{n}, \boldsymbol{b})$ is otherwise arbitrary). Then, the linear MMSE estimator of $\boldsymbol{b}$ is

$$\begin{aligned}\widehat{\boldsymbol{b}} &= \mathrm{E}\{\boldsymbol{b}\} + \boldsymbol{C}_b\boldsymbol{H}^H(\boldsymbol{H}\boldsymbol{C}_b\boldsymbol{H}^H + \boldsymbol{C}_n)^{-1}(\boldsymbol{y} - \boldsymbol{H}\mathrm{E}\{\boldsymbol{b}\}) \\ &= \mathrm{E}\{\boldsymbol{b}\} + (\boldsymbol{C}_b^{-1} + \boldsymbol{H}^H\boldsymbol{C}_b^{-1}\boldsymbol{H})^{-1}\boldsymbol{H}^H\boldsymbol{C}_n^{-1}(\boldsymbol{y} - \boldsymbol{H}\mathrm{E}\{\boldsymbol{b}\}).\end{aligned} \tag{2.9}$$

The performance of the estimator is measured by the error $\boldsymbol{\epsilon} = \boldsymbol{b} - \widehat{\boldsymbol{b}}$ whose mean is zero and whose covariance matrix is

$$\begin{aligned}\boldsymbol{C}_\epsilon &= \mathrm{E}\{\boldsymbol{\epsilon}\boldsymbol{\epsilon}^H\} \\ &= \boldsymbol{C}_b - \boldsymbol{C}_b\boldsymbol{H}^H(\boldsymbol{H}\boldsymbol{C}_b\boldsymbol{H}^H + \boldsymbol{C}_n)^{-1}\boldsymbol{H}\boldsymbol{C}_b \\ &= (\boldsymbol{C}_b^{-1} + \boldsymbol{H}^H\boldsymbol{C}_n^{-1}\boldsymbol{H})^{-1}.\end{aligned} \tag{2.10}$$

The k^{th} diagonal element of the error covariance matrix coincides with the minimum Bayesian MSE for the estimation of b_k, the k^{th} element of $\boldsymbol{b}$.

This general result enables plenty of flexibility in the definition of the $\boldsymbol{H}$ matrix to model non-ideality of communication systems (e.g. asynchronism, frequency selectivity of the channel, multiple antennas at the receiver).

Applying (2.9) to system model (2.7) and taking into account the statistical constraints on the transmitted symbols and the noise yields

$$\widehat{\boldsymbol{b}} = (\boldsymbol{R} + \sigma^2 \mathrm{I})^{-1} \boldsymbol{H}^H \boldsymbol{y} \tag{2.11}$$

$$= \boldsymbol{H}^H (\boldsymbol{T} + \sigma^2 \mathrm{I})^{-1} \boldsymbol{y} \tag{2.12}$$

where $\boldsymbol{R} = \boldsymbol{H}^H \boldsymbol{H}$ and $\boldsymbol{T} = \boldsymbol{H}\boldsymbol{H}^H$.

The linear MMSE estimator maximizes the SINR [21]. This is a reasonable optimality criterion, especially when the multiuser receiver supplies soft decisions, rather than hard detected data, to an error control decoder.

The linear MMSE estimator in (2.9) is optimum in the general MMSE sense, without the constraint of linearity, for the estimation of a Gaussian complex parameter impaired by Gaussian noise. Its application to the detection of discrete symbols requires some attention depending on the modulation symbol set $\mathbb{B}$.

In case of binary symbols, the symbol set is better approximated by a one dimensional parameter space than by the two dimensional complex space, as noticed in [22]. This leads to the *widely linear MMSE detectors.* The interested reader is referred to [22] for further details on this topic.

If the power of modulated symbols is not constant the so-called bias problem of the linear MMSE detector arises. In this case, the *unbiased linear MMSE detector* should be adopted. Further details on it can be found in [17].

2.2.4 Parallel Interference Cancellation

Historically, the linear parallel interference cancellation (PIC) approach was proposed in [23, 24] to improve the performance of the nonlinear PIC techniques proposed first by Varanasi and Aazhang in [25].

Linear PIC detectors answer to the need of avoiding the matrix inversion required in the linear MMSE detector and perform substantially better than the matched filter.

They rely on simple processing elements and are constructed around the matched filter concept. The first estimate of the signal at the output of the matched filter is utilized to estimate the interference from all users. This estimation is subtracted from the original signal and an improved estimate of the transmitted symbols is performed. The PIC detectors are based on the iteration of this procedure. The linear PIC detectors differ from the original PIC detector in [25] in the fact that they benefit from the *soft estimates* of the interference while the latter estimates the interference making use of *detected* symbols.

Given $\widetilde{\boldsymbol{b}}_{i-1}$, the soft estimates of the symbols at the $(i-1)^{\text{th}}$ iteration, the i^{th} stage of a weighted linear PIC detector is described by

$$\widetilde{\boldsymbol{b}}_i = \boldsymbol{H}^H \boldsymbol{y} + \tau(\boldsymbol{I} - \boldsymbol{R})\widetilde{\boldsymbol{b}}_{i-1}$$

where τ is a scalar that can be conveniently designed to optimize the quantity of interference to be cancelled. In case of error on the estimates, a complete cancellation of the estimated interference obtained by assuming $\tau = 1$ can have a detrimental effect. The parameter τ can be optimized to minimize this effect.

The weighted linear PIC detector is given by

$$\begin{aligned} T_{\text{PIC}} &= \left(\sum_{i=0}^{M-1} \tau^i (\boldsymbol{I} - \boldsymbol{R})^i \right) \boldsymbol{H}^H \\ &= \sum_{m=1}^{M-1} \overline{w}_m \boldsymbol{R}^m \boldsymbol{H}^H \end{aligned} \tag{2.13}$$

with $M \in \mathbb{Z}^+$ and

$$\overline{w}_m = \sum_{\ell=m}^{M-1} \tau^\ell (-1)^m \binom{\ell}{m}. \tag{2.14}$$

The scalar coefficients $\overline{w}_k$ depends only on τ but do not depend on any system parameter, e.g. number of users, spreading factor, spreading sequences, channel gains. For $\tau = 1$ the weighted linear PIC detector reduces to the standard PIC detector.

2.2.5 Polynomial Expansion Detection

The polynomial expansion detectors approximate the inverse matrix $(\boldsymbol{R}+\sigma^2\boldsymbol{I})^{-1}$ in the linear MMSE detector (2.11) by a matrix polynomial in the correlation matrix $\boldsymbol{R}$ so that

$$\widehat{\boldsymbol{b}} = \sum_{k=0}^{M-1} w_k \boldsymbol{R}^k \boldsymbol{H}^H \boldsymbol{y} \tag{2.15}$$

with $M = 0, 1, \ldots K-1$ and w_k scalar coefficients designed according to some optimality criterion.

The rationale behind this approximation is that a $K \times K$ matrix $(\boldsymbol{R}+\sigma^2\boldsymbol{I})^{-1}$ can be expanded into a matrix polynomial of degree K. This is a direct consequence of the Cayley-Hamilton theorem. Let $\Pi(x) = \sum_{k=0}^{K} \alpha_k x^k$ be the characteristic polynomial of the matrix $\boldsymbol{R} + \sigma^2 \boldsymbol{I}$ with coefficients α_k, as well known, dependent of the eigenvalues of $\boldsymbol{R}+\sigma^2\boldsymbol{I}$ or, equivalently[5], of the eigenvalues of $\boldsymbol{R}$ and on the variance

[5] As well known, if λ_i, $i = 1, \ldots, K$, are the eigenvalues of $\boldsymbol{R}$, the eigenvalues of $\boldsymbol{R} + \sigma^2 \boldsymbol{I}$ are $\lambda_i + \sigma^2$, $i = 1, \ldots, K$.

σ^2. Thanks to the Cayley-Hamilton theorem $\boldsymbol{R} + \sigma^2 \boldsymbol{I}$ is a zero of $\Pi(x)$, i.e.

$$\sum_{k=0}^{K} \alpha_k (\boldsymbol{R} + \sigma^2 \boldsymbol{I})^k = \boldsymbol{0}. \tag{2.16}$$

Substituting $(\boldsymbol{R} + \sigma^2 \boldsymbol{I})^0 = (\boldsymbol{R} + \sigma^2 \boldsymbol{I})^{-1}(\boldsymbol{R} + \sigma^2 \boldsymbol{I})$ in (2.16) yields

$$\begin{aligned}
(\boldsymbol{R} + \sigma^2 \boldsymbol{I})^{-1} &= -\sum_{k=0}^{K-1} \frac{\alpha_{k+1}}{\alpha_0} (\boldsymbol{R} + \sigma^2 \boldsymbol{I})^k \\
&= -\sum_{k=0}^{K-1} \frac{\alpha_{k+1}}{\alpha_0} \sum_{\ell=0}^{k} \begin{pmatrix} k \\ \ell \end{pmatrix} \sigma^{2(k-\ell)} \boldsymbol{R}^\ell \\
&= -\sum_{\ell=0}^{K-1} \sum_{k=\ell}^{K-1} \left(\frac{\alpha_{k+1}}{\alpha_0} \begin{pmatrix} k \\ \ell \end{pmatrix} \sigma^{2(k-\ell)} \right) \boldsymbol{R}^\ell \\
&= \sum_{\ell=0}^{K-1} \widetilde{w}_\ell \boldsymbol{R}^\ell
\end{aligned} \tag{2.17}$$

with

$$\widetilde{w}_\ell = -\sum_{k=\ell}^{K-1} \frac{\alpha_{k+1}}{\alpha_0} \begin{pmatrix} k \\ \ell \end{pmatrix} \sigma^{2(k-\ell)}.$$

In contrast to the weighted linear PIC detector, the weights $\widetilde{w}_\ell$ of a polynomial expansion detector depend on the system through the characteristic polynomial coefficients α_k. Usually $M < K$ so that (2.15) is only an approximation of the linear MMSE detector. In [26] the weights w_k in (2.15) are designed by minimizing the average power of the soft output error between the full rank linear MMSE detector (2.11) and the polynomial expansion detector, i.e.

$$\boldsymbol{w} = \operatorname*{argmin}_{\boldsymbol{w}} \mathrm{E} \left\{ \left\| \left((\boldsymbol{R} + \sigma^2 \boldsymbol{I})^{-1} - \sum_{\ell=0}^{M-1} w_\ell \boldsymbol{R}^\ell \right) \boldsymbol{H}^H \boldsymbol{y} \right\|^2 \right\}$$

where $\boldsymbol{w} = (w_1, w_2, \ldots, w_M)^T$.

This criterion is equivalent to the joint minimization of the mean square error of all received signals:

$$\boldsymbol{w} = \operatorname*{argmin}_{\boldsymbol{w}} \mathrm{E} \left\{ \left\| \boldsymbol{b} - \sum_{\ell=0}^{M-1} w_\ell \boldsymbol{R}^\ell \boldsymbol{H}^H \boldsymbol{y} \right\|^2 \right\}.$$

The solution of the previous optimization problem is [26]

$$\boldsymbol{w} = \boldsymbol{\Phi}^{-1} \boldsymbol{\varphi} \tag{2.18}$$

where $\mathbf{\Phi}$ is an $M \times M$ matrix whose elements can be expressed in terms of the traces of powers of the autocorrelation matrix $\boldsymbol{R}$ as

$$\Phi_{ij} = \mathrm{tr}(\boldsymbol{R}^{i+j}) + \sigma^2 \mathrm{tr}(\boldsymbol{R}^{i+j-1}),$$

and $\boldsymbol{\varphi}$ is an M-dimensional vector with elements $\varphi_i = \mathrm{tr}(\boldsymbol{R}^i)$.

As in the case of linear PIC detectors, the matrix inversion required in the linear MMSE detector can be avoided. However, this approach requires the computation of $\mathrm{tr}(\boldsymbol{R}^i)$, $i = 1, \ldots, 2M$ (complexity $\mathcal{O}(K^3)$) and the inversion of an $M \times M$ matrix instead of a $K \times K$ matrix in order to compute the weights in (2.18). Therefore, the complexity order of polynomial expansion detection is determined by the complexity of the weight computation.

An alternative design criterion for the weights has been proposed in [18] for CDMA systems with random spreading and with perfect power control, i.e. $\boldsymbol{A} = \boldsymbol{I}$. The proposed weights minimize the signal-to-total-power-and-noise ratio of a large system with $K, N \to \infty$ and $\frac{K}{N} \to \beta$, where β is the system load of the actual finite system with finite number of users and finite spreading factor.

Thanks to the properties of random matrices, this asymptotic approximation of the weights makes them independent of the spreading sequences. If $M \ll K$, it enables a low complexity computation of the weights with negligible complexity by using the asymptotic spectral analysis of random matrices. With this design of the weights, the complexity order of polynomial expansion detection is determined by the complexity of the multiplications in (2.15). Asymptotically and for equal powers, the weight design in [18] is equivalent to the minimization of the MSE proposed in [26]. We will elaborate further on this equivalence in Chapter 3.

2.2.6 Multistage Wiener Filters

The reduced rank multistage Wiener filter has been proposed first in [27] as a byproduct of a multistage decomposition of the Wiener filter for the estimation of a scalar process when the observed signal is a vector process. The analytical description of such a decomposition is quite cumbersome and not directly necessary for the future developments in this work. Therefore, we omit it here and focus on some properties and equivalent representations useful for further studies. We refer the interested reader to [27–29] for analytical details on the original multistage Wiener filter representation.

The reduced-rank multistage Wiener filtering is concerned with the compression, or reduction in dimensionality, of the observed data prior to Wiener filtering. A "good" rank reduction aims to minimize the MSE between the output of a filter in the projection subspace and the output of a filter using all observables. In this respect, the projection subspace method proposed in [27] is very effective and outperforms other well known reduced rank techniques (e.g. principal components, cross-spectral methods [30]) as shown numerically in [27] and analytically in [28]. Because of the compression of the observed data prior to the Wiener filter the reduced-rank multistage Wiener filters provide only an approximation of the full rank Wiener filter.

Equivalent representations of the reduced-rank multistage Wiener filter are obtained by projecting the observables onto the same projection subspace proposed in [27] and then performing a linear MMSE filtering of the projection. Throughout this section, these equivalent representations are illustrated.

Let us consider the CDMA system in (2.7) and let us denote with $\chi_{M,k}(\boldsymbol{H})$ the M-dimensional projection subspace for the estimation of the k^{th} user symbol by the reduced rank multistage Wiener filtering in [27][6]. Let $\boldsymbol{B}_k$ be the matrix whose column vectors form a possibly nonorthogonal basis of $\chi_{M,k}(\boldsymbol{H})$. The projection of the observed signal onto $\chi_{M,k}(\boldsymbol{H})$ yields

$$\boldsymbol{y}' = (\boldsymbol{B}_k^H \boldsymbol{B}_k)^{-1} \boldsymbol{B}_k^H \boldsymbol{y}$$

where $\boldsymbol{y}'$ is an M-dimensional column vector.

The subsequent linear MMSE filter in $\chi_{M,k}(\boldsymbol{H})$ satisfies the Wiener-Hopf equation

$$\boldsymbol{w}'_k = (\mathrm{E}\{\boldsymbol{y}'\boldsymbol{y}'^H\})^{-1} \mathrm{E}\{\boldsymbol{y}' b_k\}$$

with

$$\mathrm{E}\{\boldsymbol{y}'\boldsymbol{y}'^H\} = (\boldsymbol{B}_k^H \boldsymbol{B}_k)^{-1} \boldsymbol{B}_k^H (\boldsymbol{T} + \sigma^2 \boldsymbol{I}) \boldsymbol{B}_k (\boldsymbol{B}_k^H \boldsymbol{B}_k)^{-1},$$

$\boldsymbol{T}$ defined in Section 2.2.3, and

$$\mathrm{E}\{\boldsymbol{y}' b_k\} = (\boldsymbol{B}_k^H \boldsymbol{B}_k)^{-1} \boldsymbol{B}_k^H \boldsymbol{h}_k.$$

For further studies, it is convenient to define

$$\begin{aligned} \boldsymbol{\Phi}_k &= \boldsymbol{B}_k^H (\boldsymbol{T} + \sigma^2 \boldsymbol{I}) \boldsymbol{B}_k \\ \boldsymbol{\varphi}_k &= \boldsymbol{B}_k^H \boldsymbol{h}_k \\ \boldsymbol{w}_k &= (\boldsymbol{\Phi}_k)^{-1} \boldsymbol{\varphi}_k. \end{aligned} \tag{2.19}$$

The soft estimate is given by

$$\widehat{b}_k = \boldsymbol{w}_k^H \boldsymbol{B}_k^H \boldsymbol{y}.$$

[6] To avoid to go into the details of the reduced rank multistage Wiener filter in [27], we intentionally do not specify the subspace $\chi_{M,k}(\boldsymbol{H})$ and its basis here. Later on in this section, we will present several possible bases of $\chi_{M,k}(\boldsymbol{H})$ of practical interest.

The reduced-rank Wiener filter computes the M statistics[7]

$$\boldsymbol{x}_k = \boldsymbol{B}_k^H \boldsymbol{y}.$$

Then, it filters $\boldsymbol{x}_k$ by $\boldsymbol{w}_k$.

Different choices of the basis of $\chi_{M,k}(\boldsymbol{H})$ yield equivalent representations of the reduced-rank multistage Wiener filter. Bases of $\chi_{M,k}(\boldsymbol{H})$ are given by [28]:

$$\chi_{M,k}(\boldsymbol{H}) = \operatorname{span}\{\boldsymbol{U}_k^m \boldsymbol{h}_k\}_{m=0}^{M-1}$$

with $\boldsymbol{U}_k = \boldsymbol{T} - \boldsymbol{h}\boldsymbol{h}_k^H + \sigma^2 \boldsymbol{I}$ or $\boldsymbol{U}_k = \boldsymbol{T} - \boldsymbol{h}_k \boldsymbol{h}_k^H$ and have been adopted in several works [13, 14, 28]. With this choice of the basis of $\chi_{M,k}(\boldsymbol{H})$ the reduced rank Wiener filter is given by

$$\widehat{b}_k = \sum_{m=0}^{M-1} (\boldsymbol{w}_k)_m \left(\boldsymbol{U}_k^m \boldsymbol{h}_k\right)^H \boldsymbol{y}.$$

Although filters using different bases are equivalent in the sense that they show equal performance, the choice of the basis of the projection subspace can have relevant effects on the complexity of the detector. We will elaborate further on this aspect in Section 3.3.

Hereinafter, we refer to the reduced rank multistage Wiener filters shortly as multistage Wiener filters or MSWF.

2.3 Performance Analysis

2.3.1 Performance Measures

Very useful performance measures in multiuser communication systems are the bit error rate (BER), i.e. the probability of decoding erroneously the transmitted bits, and the symbol error rate (SER), i.e. the probability of detecting erroneously the uncoded symbols. The analytical computation of BER and SER in a multiple access system is far from being trivial. They depend on the set of modulation symbols $\mathbb{B}$. The former depends also on the mapping. In order to simplify the computation of these performance measures we consider the simple case when all users transmit binary, equiprobable, antipodal symbols and the received signal is impaired by additive white Gaussian noise. If $\boldsymbol{C}$ is the linear transformation applied to the received vector $\boldsymbol{y}$ and $\boldsymbol{c}_k^H$ is its k^{th} row, so that the detected bit of user k is $b_{\text{det},k} = \text{quant}_{\mathbb{B}}(\boldsymbol{c}_k^H \boldsymbol{y})$, then the probability of detecting erroneously b_k, conditioned on $b_k = 1$ and on the

[7] Hereinafter, with an abuse of denomination, we refer to this computation as "projection".

vector of interfering bits $\boldsymbol{b}_I = (b_1, \dots b_{k-1}, b_{k+1}, \dots b_K)$, is given by

$$\begin{aligned}\mathcal{P}_{e,k}(b_k = 1, \boldsymbol{b}_I) &= \Pr(b_{\text{det},k} \neq b_k | b_k = 1, \boldsymbol{b}_I) \\ &= \mathrm{Q}\left(\frac{\boldsymbol{c}_k^H \boldsymbol{h}_k + \sum_{k \neq j} \boldsymbol{c}_i^H \boldsymbol{h}_j b_j}{\sqrt{\sigma^2 \boldsymbol{c}_k^H \boldsymbol{c}_k}}\right)\end{aligned}$$

where $\mathrm{Q}(x) = \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} \mathrm{e}^{-\frac{t^2}{2}} \mathrm{d}t$. The error probability is obtained by averaging over all sequences of interfering bits $\boldsymbol{b}_I$:

$$\overline{\mathcal{P}}_{e,k} = \frac{1}{2^{K-1}} \sum_{\boldsymbol{b}_I} \mathrm{Q}\left(\frac{\boldsymbol{c}_k^H \boldsymbol{h}_k + \sum_{k \neq j} \boldsymbol{c}_i^H \boldsymbol{h}_j b_j}{\sqrt{\sigma^2 \boldsymbol{c}_k^H \boldsymbol{c}_k}}\right). \tag{2.20}$$

We can recognize by inspection that the computation of $\overline{\mathcal{P}}_{e,k}$ has a complexity exponential in the number of interfering users. However, the average error probability can be accurately approximated by

$$\overline{\mathcal{P}}_{e,k} \approx \mathrm{Q}\left(\sqrt{\mathrm{SINR}_k}\right). \tag{2.21}$$

where SINR_k is the SINR of user k at the output of the detector $\boldsymbol{C}$. The reader is referred to Section 3.4 in [16] for the rationale behind this approximation. The accuracy of this approximation is supported by the results in [31]. Whereas at low signal-to-noise ratios the approximation (2.21) is generally good, for high signal-to-noise ratios it may be unreliable.

Equation (2.21) shows the relevance of the SINR at the output of a detector as performance measure. As it will be clear in the following, the output SINR of a linear detector can be computed with low complexity also for CDMA systems with a large number of users.

Denoting by P_k and P the useful power of user k and the total power at the output of the linear detector $\boldsymbol{C}$, respectively, the output SINR of $\boldsymbol{C}$ is given by

$$\begin{aligned}\mathrm{SINR}_k &= \frac{P_k}{P - P_k} \\ &= \frac{\mathrm{E}\{\|\boldsymbol{c}_k^H \boldsymbol{h}_k b_k\|^2\}}{\mathrm{E}\{\|\boldsymbol{c}_k^H (\boldsymbol{y} - \boldsymbol{h}_k b_k)\|^2\}} \\ &= \frac{\boldsymbol{c}_k^H \boldsymbol{h}_k \boldsymbol{h}_k^H \boldsymbol{c}_k}{\boldsymbol{c}_k^H (\boldsymbol{H}\boldsymbol{H}^H - \boldsymbol{h}_k \boldsymbol{h}_k^H + \sigma^2 \boldsymbol{I}) \boldsymbol{c}_k}.\end{aligned} \tag{2.22}$$

A performance measure widely used in multiuser communications is the multiuser efficiency. In order to introduce the concept of multiuser efficiency, we define the *effective energy* of user k to achieve a certain target bit error rate BER_0. It would be the energy that user k would require to achieve BER_0 in a single-user Gaussian channel with the same background noise level and maximum-a-posteriori (MAP) detection. Given BER_0, the multiuser efficiency of user k is the ratio between the effective energy and the actual energy required for user k to achieve BER_0 with the detector in use. This concept can be formalized in the following definition.

Definition 1 *Let* $\overline{\mathcal{P}}_{e,k}(\sigma^2, \boldsymbol{R}, \mathrm{det})$ *denote the average symbol error probability of user* k *after transmission through the multiple access channel with covariance matrix* $\boldsymbol{R}$ *and additive white Gaussian noise with variance* σ^2 *and after detection with detector* det*. The multiuser efficiency of user* k *with detector* det *is the number* η_k *such that*

$$\overline{\mathcal{P}}_{e,k}(\sigma^2, \boldsymbol{R}, \mathrm{det}) = \overline{\mathcal{P}}_{e,k}(\sigma^2/\eta_k, \boldsymbol{I}, \mathrm{MAP}). \tag{2.23}$$

The identity matrix in the r.h.s. of (2.23) guarantees that the channel is free of multiple access interference.

Since $\overline{\mathcal{P}}_{e,k}(\sigma^2/\eta_k, \boldsymbol{I}, \mathrm{MAP}) \approx \mathrm{Q}\left(\eta_k \frac{\boldsymbol{h}_k^H \boldsymbol{h}_k}{\sigma^2}\right)$, the substitution of (2.21) in (2.23) yields

$$\eta_k = \frac{\sigma^2}{\boldsymbol{h}_k^H \boldsymbol{h}_k} \mathrm{SINR}_{\mathrm{det},k} \tag{2.24}$$

where $\mathrm{SINR}_{\mathrm{det},k}$ is the SINR of user k at the output of detector dec. Since the effective energy is never greater than the actual energy required by user k to achieve the same BER, $\eta_k = [0, 1]$.

The concept of multiuser efficiency can be illustrated graphically. Let us plot, in logarithmic scale, the BER of user k at the output of detector det as a function of the SNR of user k in the following two cases:

(a) The channel of user k is not impaired by MAI and MAP detection is performed at the receiver.

(b) User k transmits its signal through the real communication system to be analyzed, i.e. a system affected by multiuser interference and equipped with a multiuser detector.

The curve describing scenario (a) is referred to as curve (a). Similarly, curve (b) is associated to scenario (b). For BER_0, a fixed value of BER, the multiuser efficiency in decibels is the negative shift that would bring the curve (b) to intersect the curve (a) in BER_0. Figure 2.1 illustrates this graphical interpretation.

The multiuser efficiency is a very interesting and useful performance measure in the performance analysis of large systems, i.e. when the number of users and the spreading factor of the system tend to infinity with their ratio converging to a constant. In fact, we will see in the following sections that, asymptotically, η_k is a

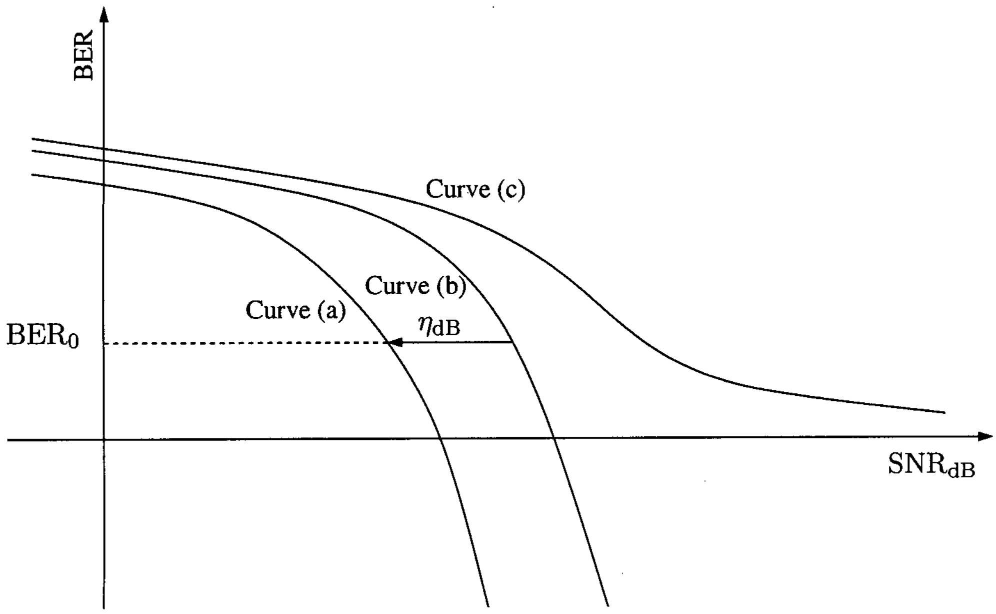


Figure 2.1: Graphical illustration of the concept of multiuser efficiency.

function of the SNR and of the system load independent of user k for some relevant detectors.

The *asymptotic multiuser efficiency* is defined as the multiuser efficiency for vanishing noise:

$$\eta_{0,k} = \lim_{\sigma^2 \to 0} \eta_k.$$

It can be shown that it measures the slope with which $\overline{\mathcal{P}}_{e,k}$ goes to zero (in logarithmic scale) in the high SNR region (see [16]). The graphical interpretation of the multiuser efficiency in Figure 2.1 shows that the asymptotic multiuser efficiency $\eta_{0,k}$ is a nonzero constant when curve (a) tends to be parallel to curve (b) in the high SINR region. When the BER does not vanish as the SNR goes to infinity, as for the curve (c), or it vanishes with slower rate than curve (a), the asymptotic multiuser efficiency vanishes, i.e. $\eta_{0,k} = 0$.

The *near-far resistance* of user k is defined as the asymptotic multiuser efficiency minimized over the received energies of all interfering users, i.e.

$$\overline{\eta}_{0,k} = \inf_{\forall a_{jj}, j \neq k} \eta_{0,k}.$$

being a_{jj} the received signal amplitudes introduced in Section 2.2.1.

2.3.2 Matched Filter

The SINR at the output of a matched filter is obtained from (2.22) for $\boldsymbol{c}_k = \boldsymbol{h}_k$

$$\mathrm{SINR}_k = \frac{|a_{kk}|^2}{\boldsymbol{s}_k^H \boldsymbol{H}\boldsymbol{H}^H \boldsymbol{s}_k - |a_{kk}|^2 + \sigma^2}.$$

If the spreading sequences are not orthogonal the error probability does not vanishes as σ^2 tends to zero [16]. This implies that the single user matched filters suffer from zero asymptotic multiuser efficiency, unless the spreading sequences are orthogonal:

$$\eta_{0,k} = \begin{cases} 1 & \boldsymbol{s}_i^H \boldsymbol{s}_k = 0 \; \forall i \neq k \\ 0 & \text{otherwise.} \end{cases}$$

The large system performance under the assumption of random spreading has been investigated in [10,11]. The asymptotic multiuser efficiency, as $N, K \to \infty$ with $\frac{K}{N} \to \beta$, converges in probability to a deterministic value that is independent of the spreading sequences and the users. It is a constant that characterizes unequivocally the detector and provides an insightful description of its performance.

Let us assume that the sequence of the empirical eigenvalue distributions of the matrix $\boldsymbol{A}^H\boldsymbol{A}$ converges to a limit c.d.f. $F_{|\boldsymbol{A}|^2}(\lambda)$ as the number of users tends to infinity. The limit multiuser efficiency as $K, N \to \infty$ with $\frac{K}{N} \to \beta$ is given by

$$\eta_{\mathrm{MF}} = \left(1 + \frac{\beta}{\sigma^2} \int \lambda \mathrm{d}F_{|\boldsymbol{A}|^2}(\lambda)\right)^{-1}.$$

With random spreading, the asymptotic performance is simply described by the system load β, the variance of the noise σ^2, and the mean of the received powers.

It is straightforward to verify that the asymptotic multiuser efficiency vanishes for large systems.

2.3.3 Linear MMSE Detection

The SINR at the output of a linear MMSE detector can be derived from (2.22) taking into account that $\boldsymbol{c}_k = \boldsymbol{h}_k^H(\boldsymbol{R} + \sigma^2\boldsymbol{I})^{-1}$. It results as

$$\mathrm{SINR}_k = \frac{\boldsymbol{h}_k^H(\boldsymbol{R} + \sigma^2\boldsymbol{I})^{-1}\boldsymbol{h}_k}{1 - \boldsymbol{h}_k^H(\boldsymbol{R} + \sigma^2\boldsymbol{I})^{-1}\boldsymbol{h}_k}.$$

Theorem 1 allows us to derive a simple relation between the SINR at the output of the linear MMSE detector and the MSE. In fact, (2.10) for $\boldsymbol{C}_b = \boldsymbol{I}$ yields

$$\begin{aligned} \mathrm{MSE}_k &= (\boldsymbol{C}_\epsilon)_{kk} \\ &= 1 - \boldsymbol{h}_k^H(\boldsymbol{R} + \sigma^2\boldsymbol{I})^{-1}\boldsymbol{h}_k \end{aligned} \tag{2.25}$$

and thus the following relation holds:

$$\mathrm{SINR}_k = \frac{1}{\mathrm{MSE}_k} - 1.$$

The linear MMSE detector is also the linear detector maximizing the SINRs of all users. This property can be verified by maximization of (2.22) with respect to $\boldsymbol{c}_k$.

The SINR at the output of the linear MMSE detector depends on the spreading sequences of all users, the channel realization, and the variance of the noise. Therefore, this performance measure does not provide general indications on the system behaviour. Better insight into the system performance is provided by a large system analysis. This asymptotic analysis investigates systems whose dimensions go to infinity with their ratio converging to a constant. It assumes random spreading sequences and benefits from the theory of random matrices. The asymptotic SINR of linear MMSE detectors in flat fading channels has been derived for spreading matrix $\boldsymbol{S}$ with i.i.d. entries in [10]. Systems with orthogonal spreading matrices modelled by random isometric matrices have been considered in [12]. The convergence rate has been investigated in [32].

The following theorem recapitulates the result for random spreading matrices with i.i.d. elements.

Theorem 2 *[10] Let us consider the system model (2.7). Let the $N \times K$ matrix $\boldsymbol{S}$ be a random matrix whose entries are centered i.i.d. with variance $\mathrm{E}\{|s_{ij}|^2\} = \frac{1}{N}$ and fourth moment such that[8] $\lim_{N\to\infty} N^2\mathrm{E}\{|s_{ij}|^4\} < +\infty$. Let the empirical distribution of the received powers $|a_{kk}|^2$ converges to a fixed distribution $F_{|\boldsymbol{A}|^2}(\lambda)$ as $K \to \infty$. Then, conditionally on $|a_{kk}|^2$, the received power of user k, the SINR of user k at the output of a linear MMSE detector converges in probability to the unique solution to the following fixed point equation, as $K, N \to \infty$ with $\frac{K}{N} \to \beta$:*

$$\mathrm{SINR}_k = \frac{|a_{kk}|^2}{\sigma^2 + \beta \mathrm{E}\left\{\frac{\lambda |a_{kk}|^2}{|a_{kk}|^2 + \lambda \mathrm{SINR}_k}\right\}} \tag{2.26}$$

where the expectation is taken over the limiting c.d.f. $F_{|\boldsymbol{A}|^2}(\lambda)$ of the received powers of the interferers.

The multiuser efficiency can be derived from (2.26) by substituting $\eta_k = \frac{\mathrm{SINR}_k \sigma^2}{|a_{kk}|^2}$. This yields

$$\eta_k = \frac{1}{1 + \beta \mathrm{E}\left\{\frac{\lambda}{\sigma^2 + \lambda \eta_k}\right\}}. \tag{2.27}$$

[8]Note that the constraints on the chips of the spreading sequences are verified in cases of practical interest (see e.g. binary spreading or Gaussian spreading). The constraint on the variance takes into account the usual normalization on the spreading sequences to have unit energy. The constraints on moments higher than 2 require that the higher moments of $\sqrt{N} s_{ij}$ are upper bounded. This is a mathematical constraint verified in physical systems.

The multiuser efficiency is independent of the user and characterizes unequivocally the system.

The spectral efficiency of CDMA systems with linear MMSE detectors has been investigated by Shamai and Verdú for AWGN channels [11] and flat fading channels [33].

2.3.4 Parallel Interference Cancellation

The weighted linear PIC detector (2.13) can be regarded as a polynomial expansion detector with nonoptimized weights (2.14). The filter of user k is

$$\boldsymbol{c}_k = \sum_{m=0}^{M-1} \overline{w}_m \boldsymbol{h}_k^H \boldsymbol{T}^m \tag{2.28}$$

with $\boldsymbol{T} = \boldsymbol{H}\boldsymbol{H}^H$.

The SINR of user k is obtained by substituting (2.28) in (2.22). For finite systems, SINR_k is a random variable depending on the channel realizations, the variance of the noise, the spreading sequences and the parameter τ.

The large system performance of a modified version of the weighted PIC detectors (2.13) is analyzed in [34][9] assuming equal received powers, i.e. $\boldsymbol{A} = \sqrt{P}\boldsymbol{I}$. The weighted PIC detector (2.13) is substituted by

$$\widetilde{T}_{\mathrm{PIC},k} = \sum_{m=0}^{M-1} \overline{w}_k P^{\frac{2m+1}{2}} \boldsymbol{s}_k^H (\boldsymbol{S}_{\sim k}\boldsymbol{S}_{\sim k}^H)^m \tag{2.29}$$

where $\boldsymbol{S}_{\sim k}$ is the matrix obtained from $\boldsymbol{S}$ by suppressing $\boldsymbol{s}_k$, the k^{th} column of $\boldsymbol{S}$. If the spreading matrix $\boldsymbol{S}$ satisfies the same assumptions as in Theorem 2, SINR_k converges to a deterministic value as $K, N \to \infty$ with $\frac{K}{N} \to \beta$,

$$\mathrm{SINR}_k = \frac{(\overline{\boldsymbol{w}}^T \boldsymbol{\varphi}_{\mathrm{MS}})^2}{\overline{\boldsymbol{w}}^T (\boldsymbol{\Phi}_{\mathrm{MS}} - \boldsymbol{\varphi}_{\mathrm{MS}}\boldsymbol{\varphi}_{\mathrm{MS}}^H)\overline{\boldsymbol{w}}} \tag{2.30}$$

with $\overline{\boldsymbol{w}} = (\overline{w}_0, \overline{w}_1, \ldots, \overline{w}_{M-1})^T$, $\boldsymbol{\varphi}_{\mathrm{MS}} = (P\varphi_1, \ldots, P^M\varphi_M)^T$,

$$\boldsymbol{\Phi}_{\mathrm{MS}} = \begin{pmatrix} P^2\varphi_2 + \sigma^2 P\varphi_1 & \cdots & P^{M+1}\varphi_{M+1} + \sigma^2 P^M \varphi_M \\ \cdots & \cdots & \cdots \\ P^{M+1}\varphi_{M+1} + \sigma^2 P^M\varphi_M & \cdots & P^{2M}\varphi_{2M} + \sigma^2 P^{2M-1}\varphi_{2M-1} \end{pmatrix}.$$

[9] In order to keep a uniform approach in presenting the system performance, the expression of SINR_k proposed in this work is equivalent but not identical to the the expression in [34].

φ_ℓ are the moments of the Marčenko-Pastur distribution [35]:

$$\varphi_\ell = \sum_{j=0}^{\ell-1} \frac{1}{j+1} \binom{\ell}{j} \binom{\ell+1}{j} \beta^{j+1}. \tag{2.31}$$

The following will be shown in Chapter 3:

- The complexity of the weighted PIC filter in (2.13) is lower than the complexity of the modified version (2.29) by a factor of K' if the detection of K' users is required.
- The performance of the weighted PIC detector in (2.13) coincides with the performance of the modified PIC detector (2.29) only in case of equal received powers.

A general framework that enables the performance analysis of standard PIC detectors is presented in Chapter 3. The results in Chapter 3 can be applied to the analysis of PIC detectors with unequal received powers and multipath fading channels. The performance of PIC detectors in asynchronous CDMA systems can be analyzed benefitting from the general results in Chapter 4. PIC detectors for CDMA systems with spatial diversity can be investigated using the results in Chapter 5.

2.3.5 Polynomial Expansion Detector

The asymptotic performance of polynomial expansion detection has been analyzed in [18] for the case of random spreading and matrices with i.i.d. entries and equal received powers. If the matrix $\boldsymbol{S}$ satisfies the conditions of Theorem 2, the limit SINR of user k, as $K, N \to \infty$ with $\frac{K}{N} \to \beta$, is deterministic and given by[10]:

$$\mathrm{SINR}_k = \frac{\boldsymbol{\varphi}_{\mathrm{MS}}^H \boldsymbol{\Phi}_{\mathrm{MS}}^{-1} \boldsymbol{\varphi}_{\mathrm{MS}}}{1 - \boldsymbol{\varphi}_{\mathrm{MS}}^H \boldsymbol{\Phi}_{\mathrm{MS}}^{-1} \boldsymbol{\varphi}_{\mathrm{MS}}} \tag{2.32}$$

with $\boldsymbol{\Phi}_{\mathrm{MS}}$ and $\boldsymbol{\varphi}_{\mathrm{MS}}$ defined in Section 2.3.4.

The performance of the polynomial expansion detectors for flat fading channels and multipath fading channels is investigated in Chapter 3. The performance analysis of asynchronous CDMA systems or systems with spatial diversity is based on the results presented in Chapter 4 and Chapter 5, respectively.

[10] As for the PIC detectors, we propose an expression for the SINR different but equivalent to the expression in [18] to keep a uniform approach.

2.3.6 Multistage Wiener Filtering

The asymptotic performance of the reduced rank multistage Wiener filter (MSWF) has been analyzed in [28,29]. For large[11] systems with perfect power control, i.e. equal received power for all users, the SINR converges to the deterministic value given in (2.32). In this scenario, MSWF and polynomial expansion detectors have the same performance. This supported the erroneous belief of their equivalence [36–38] in the general case. The equivalence of MSWF and polynomial expansion detectors is disproved in Chapter 3.

Based on the following theorem, the performance of MSWF can also be expressed as a continued fraction.

Theorem 3 *[28] Let $\boldsymbol{S}$ be as in Theorem 2 and $\boldsymbol{A} = P\boldsymbol{I}$. As $K = \beta N \to \infty$ the output SINR of the rank M MSWF converges in probability to the limit*

$$\mathrm{SINR}_M = \cfrac{P}{\sigma^2 + \beta \cfrac{P}{1 + \mathrm{SINR}_{M-1}}} \qquad \textit{for} \qquad M \geq 1$$

with $\mathrm{SINR}_0 = 0$. $\mathrm{SINR}_1 = \frac{P}{\sigma^2+\beta P}$ *is the large system limit of the output SINR for the matched filter.*

Thus, for example

$$\mathrm{SINR}_2 = \frac{P}{\sigma^2 + \beta \frac{P}{1+\frac{P}{\sigma^2+\beta P}}}.$$

In absence of background noise $\mathrm{SINR}_M = \sum_{m=1}^{M} \frac{1}{\beta^m}$. This relation shows two important characteristics of the MSWF that hold also for more general situations than in absence of noise and can be verified both analytically and numerically:

- The rank of the MSWF needed to achieve a target SINR within some small ϵ of the full rank SINR does not scale with the system size (K and N).
- As M increases, the limit SINR converges rapidly to the limit SINR of the full rank MMSE detector [28, 39]. Numerical results [28] show that the full-rank MMSE performance is essentially achieved with $M = 8$, for SNRs and system loads β of practical interest.

In analogy to the uniform power case, a continued fraction expression of the limit SINR is suggested in [28]. Let $F_{|\boldsymbol{A}|^2}(\lambda)$ be an arbitrary distribution of the received

[11] In this work we refer to systems with number of users and spreading factor that goes to infinity with ratio converging to the system load β as large CDMA systems.

powers. Then, the limit SINR at the output of a rank M MSWF is approximated by:

$$\mathrm{SINR}_{M+1} \approx \frac{P}{\sigma^2 + \beta \int_0^\infty \frac{P\lambda}{P + \lambda \mathrm{SINR}_M} \mathrm{d}F_{|\boldsymbol{A}|^2}(\lambda)}.$$

Recently, an exact continued fraction expression of the limit SINR of MSWF has been proposed in [38].

2.4 Useful Results from Random Matrices

In quantum mechanics, the energy levels of quanta are not directly observable, but can be characterized by the eigenvalues of matrices of observations. The empirical distribution of the eigenvalues of a random matrix is very complicated for large matrices. This fuelled the research on the limiting spectral analysis of large dimensional random matrices in the 1950s.

The application of random matrix theory to signal processing [40] and analysis of communication systems [8, 10, 17, 41–43] goes back to the 1990s. Since then, random matrix theory received much attention because random matrices appear in many applications of statistics and communication theory.

This section introduces some basic results on random matrices that will be useful in the remainder of this work. For general short introductions to random matrices and free probability the interested reader is referred to [44–46]. Monographs on these topics are due to Mehta [47], Girko [48], Voiculescu [49], and Hiai and Petz [50]. Applications of random matrices to communication theory are presented in [51]. This reference list does not claim to be exhaustive.

A random matrix of dimensions $N \times K$ consists of NK random elements and can be described by the joint distribution of its elements.

As an example, let us consider an $N \times K$ matrix $\boldsymbol{S}$ with i.i.d. zero mean entries of variance $\frac{1}{N}$. As N is finite, the eigenvalues of the matrix $\boldsymbol{T} = \boldsymbol{S}\boldsymbol{S}^H$ are random. Let λ_n, $n = 1, \ldots, N$ denote the eigenvalues of $\boldsymbol{T}$ and let $1(x)$ be the indicator function on a right unbounded interval, i.e. $1(x) = 1$ for $x \geq 0$ and zero elsewhere. The empirical distribution of the eigenvalues $F_{\boldsymbol{T}}^{(N)}(\lambda) = \frac{1}{N}\sum_{n=1}^{N} 1(\lambda - \lambda_n)$ is a random function depending on the realization of the matrix $\boldsymbol{S}$. Five empirical distributions of eigenvalues of a matrix $\boldsymbol{T}$ are plotted in Figure 2.2. They correspond to five independent realizations of the matrix $\boldsymbol{S}$ with constant aspect ratio[12] $\frac{K}{N} = \frac{1}{2}$ and various sizes of $\boldsymbol{S}$, namely $K = 4, 8, 64, 256, 1024$. For small matrices the empirical distribution of eigenvalues is completely arbitrary in a nonnegative support. However, as the matrix size becomes large the empirical eigenvalue distribution of $\boldsymbol{T}$ seems

[12]The aspect ratio is the ratio of the number of columns to the number of rows of a matrix.

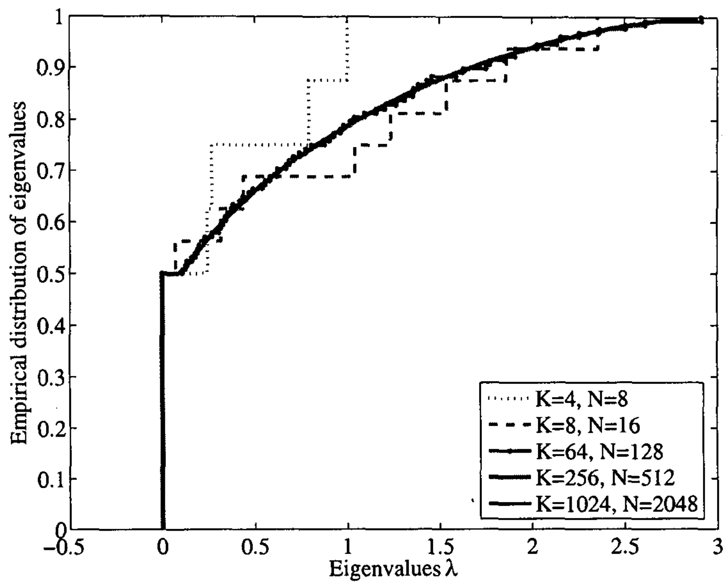


Figure 2.2: Empirical eigenvalue distributions of five independent realizations of the matrix $\boldsymbol{T} = \boldsymbol{S}\boldsymbol{S}^H$, with different size of $\boldsymbol{S}$ but equal aspect ratio $\beta = \frac{K}{N} = \frac{1}{2}$, namely $K = 4$, 8, 64, 256, 1024.

to converge to a deterministic distribution function. In Figure 2.2, the empirical eigenvalue distributions for $K = 256$ and $K = 1024$ are indistinguishable.

The convergence observed in Figure 2.2 is a general property. In fact, as $K, N \to \infty$ with ratio converging to a constant value β, the empirical eigenvalue distribution of $\boldsymbol{T}$ converges to a deterministic function known as Marčenko-Pastur distribution. The corresponding probability density function, for aspect ratio $\frac{K}{N} = \beta$, is

$$f_{\boldsymbol{T}}(\lambda) = \begin{cases} \frac{\sqrt{4\beta-(\lambda-1-\beta)^2}}{2\pi\lambda} & (1-\sqrt{\beta})^2 \leq \lambda \leq (1+\sqrt{\beta})^2 \\ (1-\beta)\delta(\lambda) & \text{elsewhere.} \end{cases}$$

Its moments have been already introduced in (2.31) and this result has been utilized in [11,17,34,52] to analyze the performance of multiuser detectors in terms of SINR, multiuser efficiency, or spectral efficiency. Since the asymptotic eigenvalue distribution is deterministic, the expectation of any function $g(\lambda)$ with respect to $F_{\boldsymbol{T}}(\lambda)$, i.e. $\mathrm{E}\{g(\lambda)\}$ is deterministic. In particular, the normalized traces $\frac{1}{N}\mathrm{tr}\boldsymbol{T}^\ell$, $\ell \in \mathbb{Z}^+$, converge to the deterministic eigenvalue moments $m_{\boldsymbol{T}}^\ell = \varphi_\ell$ in (2.31). Figure 2.3 illustrates this property. Let us consider random matrices $\boldsymbol{S}$ with different sizes, but identical aspect ratios. For each random matrix we generate 100 independent realizations. In Figure 2.3, $\widehat{m}_{\boldsymbol{T}}^1$ and $\widehat{m}_{\boldsymbol{T}}^2$, the normalized traces of $\boldsymbol{T}^1$ and $\boldsymbol{T}^2$, are shown

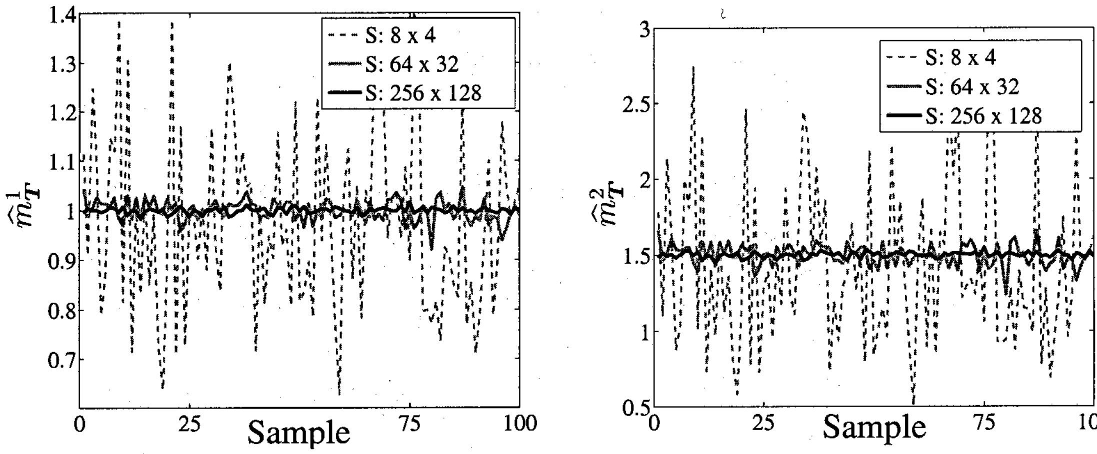


Figure 2.3: First two moments for 100 realizations of $\boldsymbol{T}$ with i.i.d. Gaussian entries.

for all realizations. It is apparent that $\widehat{m}_{\boldsymbol{T}}^1$ and $\widehat{m}_{\boldsymbol{T}}^2$ converge to 1 and $\frac{3}{2}$, respectively, as the size of $\boldsymbol{S}$ increases. The asymptotic moments $m_{\boldsymbol{T}}^\ell$ depend only on the aspect ratio β, as evident from (2.31).

The matrix $\boldsymbol{T}$ belongs to a wider class of random matrices whose limiting spectral analysis has been the object of thorough studies. In order to introduce the more general result, the Stieltjes transform is required. The Stieltjes transform was introduced by Stieltjes in 1894 [53] to address the problem of moments, i.e. to find an unknown probability distribution given its moments.

The Stieltjes transform $G(z)$ of a probability distribution function $F(x)$ is defined by

$$\begin{aligned} G(z) &= \int \frac{\mathrm{d}\, F(x)}{x - z} \\ &= -\frac{1}{z} \int \sum_{m=0}^{+\infty} \frac{x^m}{z^m} \mathrm{d}\, F(x) \end{aligned}$$

with $z \in \mathbb{C}$ and $\mathrm{Im}(z) > 0$. Given $G(z)$, the Stieltjes transform of $F(x)$, it is straightforward to derive the moments of $F(x)$ and the probability density function $f(x)$. In fact, the following relations hold (e.g. [50]):

$$-\lim_{z \to 0} \frac{\mathrm{d}^m}{\mathrm{d}z^m} \frac{G(z^{-1})}{z} = m! \int x^m \mathrm{d}F(x)$$

and

$$f(x) = \lim_{y \to 0^+} \frac{1}{\pi} \mathrm{Im}(G(x + jy)).$$

Let $\boldsymbol{U}$ and $\boldsymbol{V}$ be an $N \times N$ Hermitian matrix and a $K \times K$ diagonal matrix, respectively. Let $\boldsymbol{S}$ be defined as above, i.e. $\boldsymbol{S}$ is an $N \times K$ matrix with i.i.d. zero mean entries of variance $\frac{1}{N}$. The limiting spectral distribution of the random matrix $\boldsymbol{U} + \boldsymbol{S}\boldsymbol{V}\boldsymbol{S}^H$, with $\boldsymbol{S}$, $\boldsymbol{U}$, and $\boldsymbol{V}$ independent, has been investigated by Marčenko and Pastur [54] and Silverstein and Bai [55].

Theorem 4 *[55] Let the matrix $\boldsymbol{S}_N$, $N \in \mathbb{Z}^+$, be an $N \times K$ matrix with complex random i.i.d. entries $s_{ij}^{(N)}$ such that* $\mathrm{E}\{|s_{ij}^{(N)} - \mathrm{E}\{s_{ij}^{(N)}\}|^2\} = \frac{1}{N}$. *Let $\boldsymbol{V}_K$ be a $K \times K$ diagonal matrix with real entries whose empirical eigenvalue distribution converges almost surely in distribution to a probability distribution function $F_{\boldsymbol{V}}$, as $K \to \infty$. Finally, let $\boldsymbol{U}_N$ be a Hermitian $N \times N$ matrix whose empirical eigenvalue distribution converges almost surely to a nonrandom distribution function $F_{\boldsymbol{U}}$.*

If $\boldsymbol{S}_N$, $\boldsymbol{V}_K$, and $\boldsymbol{U}_N$ are independent and K and N tend to infinity with ratio $\frac{K}{N}$ converging to the constant β, then, almost surely, the empirical distribution of the eigenvalues of $\boldsymbol{U}_N + \boldsymbol{S}_N\boldsymbol{V}_N\boldsymbol{S}_N^H$ converges, as $N \to \infty$, to a nonrandom distribution F, whose Stieltjes transform $G(z)$ satisfies

$$G(z) = G_{\boldsymbol{U}}\left(z - \beta \int \frac{x \,\mathrm{d}G_{\boldsymbol{V}}(x)}{1 + xG(z)}\right)$$

where $G_{\boldsymbol{U}}$ and $G_{\boldsymbol{V}}$ are the Stieltjes transforms of $F_{\boldsymbol{U}}$ and $F_{\boldsymbol{V}}$, respectively.

Theorem 4 has been applied to derive Theorem 2 and Theorem 3[13]. Additionally, it has been utilized for the analysis of CDMA systems in [33, 57–60].

Girko analyzed the limiting spectral distribution of Gram random matrices[14] $\Xi\Xi^H$ with elements of Ξ statistically dependent and possibly not identically distributed. More precisely, Girko consider matrices Ξ of size $n_1 \times n_2$, with $n_1 = n_1(n)$, $n_2 = n_2(n)$, and $n \in \mathbb{Z}^+$, structured in blocks of size $q_1 \times q_2$ such that $n_1 = q_1 p_1(n)$ and $n_2 = q_2 p_2(n)$. $n_1(n)$ and $n_2(n)$ are increasing functions of n. The (i,j)-element of Ξ is denoted with ξ_{ij}; the (i,j)-block of Ξ is denoted by $\Xi_{ij}^{(n)}$. The matrix Ξ is of the

[13]The proofs of Theorem 2 in [10] and Theorem 3 in [56] clarify the existing relation between the SINR at the output of a linear MMSE detector and the output of a reduced rank MSWF for user k and the eigenvalues moments of the matrix $\boldsymbol{R} - \boldsymbol{h}_k\boldsymbol{h}_k$ with $\boldsymbol{h}_k$ and $\boldsymbol{R}$ defined in Section 2.2.2 and Section 2.2.3, respectively. Additional examples are in (2.30) and (2.32). Note that these relations hold only for special cases as the linear MMSE detectors and the reduced rank multistage Wiener filters. A more general relation between the SINR of a large class of linear detectors and the diagonal elements of the matrix $\boldsymbol{R}$ will be discussed in Chapter 3.

[14]The Gram matrix of the vectors $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \dots \boldsymbol{x}_K\}$ with respect to the inner product $< \cdot, \cdot >$ is the $K \times K$ matrix $\boldsymbol{G} = (g_{ij})_{i,j=1,\dots K}$ defined by $g_{ij} =< \boldsymbol{x}_j, \boldsymbol{x}_i >$. Defining $\boldsymbol{X} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \dots \boldsymbol{x}_K)$, $\boldsymbol{G} = \boldsymbol{X}^H\boldsymbol{X}$.

following form

$$\begin{pmatrix} \Xi_{1,1}^{(n)} & \Xi_{1,2}^{(n)} & \cdots & \Xi_{1,p_2}^{(n)} \\ \Xi_{2,1}^{(n)} & \Xi_{2,2}^{(n)} & \cdots & \Xi_{2,p_2}^{(n)} \\ \cdots & \cdots & \cdots & \cdots \\ \Xi_{p_1,1}^{(n)} & \Xi_{p_1,2}^{(n)} & \cdots & \Xi_{p_1,p_2}^{(n)} \end{pmatrix}$$

We adopt the notation $\Xi = (\xi_{ij})_{i=1,\ldots,p_1}^{j=1,\ldots,p_2}$ to represent the matrix Ξ by its elements ξ_{ij} and the notation $\Xi = (\Xi_{ij}^{(n)})_{i=1,\ldots,p_1}^{j=1,\ldots,p_2}$ to represent the matrix Ξ as a block matrix. Girko's results are focused on random matrices Ξ whose blocks are independent but, within a block $\Xi_{ij}^{(n)}$ the elements are possibly correlated and with nonzero mean. Girko's theorem shows that, under some conditions on the statistics of $\Xi_{ij}^{(n)}$, $i = 1, \ldots, p_1$ and $j = 1, \ldots, p_2$, the sequence of the empirical eigenvalue distribution of the matrix $\Xi\Xi^H$ converges almost surely to a deterministic probability distribution function as $n \to \infty$. This distribution function is characterized by its Stieltjes transform. The Stieltjes transform is obtained as solution of a system of equations called *canonical system of equation*. Let us denote by $\boldsymbol{A}_{ij}^{(n)}$ the mean of the matrix Ξ, i.e. $\boldsymbol{A}_{ij}^{(n)} = \mathrm{E}\{\Xi_{ij}^{(n)}\}$. The theorem requires that some conditions on $|\boldsymbol{A}_{ij}^{(n)}|$, the spectral norm of the matrix $\boldsymbol{A}_{ij}^{(n)}$, and on $\|\Xi_{ij}^{(n)} - \boldsymbol{A}_{ij}^{(n)}\|$, the Frobenius norm of the centered matrix $\Xi_{ij}^{(n)} - \boldsymbol{A}_{ij}^{(n)}$, are satisfied.

The following theorem summarizes Girko's results.

Theorem 5 [15] *[61] Let $\Xi_{n_1 \times n_2} = (\xi_{ij})_{i=1\ldots n_1}^{j=1\ldots n_2}$ be a random matrix composed of complex blocks $\Xi_{ij}^{(n)}$ of size $q_1 \times q_2$. We consider non symmetric block matrices $\widetilde{\Xi}_{p_1 \times p_2}$ of the form*

$$\widetilde{\Xi}_{p_1 \times p_2} = \Xi_{p_1 q_1 \times p_2 q_2} = (\Xi_{ij}^{(n)})_{i=1\ldots p_1}^{j=1\ldots p_2} \tag{2.33}$$

whose entries are the complex matrices $\Xi_{ij}^{(n)}$. p_1, p_2, q_1, q_2 are certain positive integers. p_1 and p_2 are increasing functions of $n \in \mathbb{Z}^+$. They go to infinity as $n \to \infty$. Let the random blocks $\Xi_{ij}^{(n)}$, $i = 1, \ldots, p_1$, $j = 1, \ldots, p_2$ be independent for every n, $\mathrm{E}\{\Xi_{ij}^{(n)}\} = \boldsymbol{A}_{ij}^{(n)}$, and

$$\lim_{n\to\infty}\left[\max_{i=1,\ldots,p_1}\sum_{j=1}^{p_2}\mathrm{E}\left\{\|\Xi_{ij}^{(n)} - \boldsymbol{A}_{ij}^{(n)}\|^2\right\} + \max_{i=1,\ldots,p_2}\sum_{j=1}^{p_1}\mathrm{E}\left\{\|\Xi_{ji}^{(n)} - \boldsymbol{A}_{ji}^{(n)}\|^2\right\}\right] < \infty \tag{2.34}$$

[15]The discrepancies in (2.37) and (2.38) between the statement of the theorem here and in [61] are due to typos in [61] discussed with the theorem's author in personal correspondence.

Let Lindeberg's condition be satisfied, i.e. for every $\tau > 0$,

$$\lim_{n\to\infty}\Bigg\{\max_{i=1,\dots,p_1}\sum_{j=1}^{p_2}\mathrm{E}\Big\{\|\Xi_{ij}^{(n)}-\boldsymbol{A}_{ij}^{(n)}\|^2 1_{\{\|\Xi_{ij}^{(n)}-\boldsymbol{A}_{ij}^{(n)}\|>\tau\}}(\Xi_{ij}^{(n)})\Big\}$$
$$+\max_{i=1,\dots,p_2}\sum_{j=1}^{p_1}\mathrm{E}\Big\{\|\Xi_{ji}^{(n)}-\boldsymbol{A}_{ji}^{(n)}\|^2 1_{\{\|\Xi_{ji}^{(n)}-\boldsymbol{A}_{ji}^{(n)}\|>\tau\}}(\Xi_{ji}^{(n)})\Big\}\Bigg\}=0, \tag{2.35}$$

where $1_{\{\|\Xi_{ji}^{(n)}-\boldsymbol{A}_{ji}^{(n)}\|>\tau\}}(\Xi_{ji}^{(n)})$ is the indicator function of a random event equal to 1 if the random argument $\Xi_{ji}^{(n)}$ satisfies the condition $\|\Xi_{ji}^{(n)}-\boldsymbol{A}_{ji}^{(n)}\|>\tau$ and zero otherwise. Additionally, let

$$\lim_{n\to\infty}\left[\max_{i=1,\dots,p_1}\sum_{j=1}^{p_2}|\boldsymbol{A}_{ij}^{(n)}|+\max_{i=1,\dots,p_2}\sum_{j=1}^{p_1}|\boldsymbol{A}_{ji}^{(n)}|\right]<\infty. \tag{2.36}$$

Then, with probability one, $\mu_n\{x,\Xi_{p_1q_1\times p_2q_2}\Xi_{p_1q_1\times p_2q_2}^H\}$, the empirical eigenvalue distribution of the matrix $\Xi_{p_1q_1\times p_2q_2}\Xi_{p_1q_1\times p_2q_2}^H$, satisfies

$$\lim_{n\to\infty}|\mu_n\{x,\Xi_{p_1q_1\times p_2q_2}\Xi_{p_1q_1\times p_2q_2}^H\}-F_n(x)|\overset{a.s.}{=}0,$$

where $F_n(x)$ is the distribution function whose Stieltjes transform is equal to

$$\int_0^\infty(x+z)^{-1}\mathrm{d}F_n(x)=\frac{1}{p_1q_1}\mathrm{Tr}\left(\widetilde{\boldsymbol{C}}_{p_1\times p_1}^{(1)}+\widetilde{\boldsymbol{A}}_{p_1\times p_2}\left(\widetilde{\boldsymbol{C}}_{p_2\times p_2}^{(2)}\right)^{-1}\widetilde{\boldsymbol{A}}_{p_1\times p_2}^H\right)^{-1},\quad \mathrm{Re}(z)>0.$$

The expectation is taken over ξ_{ij}, $i=1,\dots,n_1$ and $j=1,\dots,n_2$. $\widetilde{\boldsymbol{A}}_{p_1\times p_2}$ is a block matrix, mean of $\widetilde{\Xi}_{p_1\times p_2}$, i.e. $\widetilde{\boldsymbol{A}}_{p_1\times p_2}=(\boldsymbol{A}_{ij})_{i=1\dots p_1}^{j=1\dots p_2}$. $\widetilde{\boldsymbol{C}}_{p_1\times p_1}^{(1)}$ and $\widetilde{\boldsymbol{C}}_{p_2\times p_2}^{(2)}$ are diagonal block matrices, i.e.

$$\widetilde{\boldsymbol{C}}_{p_1\times p_1}^{(1)}=\mathrm{diag}\{\boldsymbol{C}_{kk}^{(1)}(z)\},\qquad \widetilde{\boldsymbol{C}}_{p_2\times p_2}^{(2)}=\mathrm{diag}\{\boldsymbol{C}_{ll}^{(2)}(z)\}$$

whose matrix elements $\boldsymbol{C}_{kk}^{(1)}(z)$ of size $q_1\times q_1$ and $\boldsymbol{C}_{kk}^{(2)}(z)$ of size $q_2\times q_2$ satisfy the canonical system of equations:

$$\boldsymbol{C}_{kk}^{(1)}=z\boldsymbol{I}_{q_1\times q_1}+\sum_{j=1}^{p_2}\mathrm{E}\left\{(\Xi_{kj}^{(n)}-\boldsymbol{A}_{kj}^{(n)})\left(\left[\widetilde{\boldsymbol{C}}_{p_2\times p_2}^{(2)}+\widetilde{\boldsymbol{A}}_{p_1\times p_2}^H[\widetilde{\boldsymbol{C}}_{p_1\times p_1}^{(1)}]^{-1}\widetilde{\boldsymbol{A}}_{p_1\times p_2}\right]^{-1}\right)_{jj}(\Xi_{kj}^{(n)}-\boldsymbol{A}_{kj}^{(n)})^H\right\},$$
$$k=1,\dots,p_1, \tag{2.37}$$

$$\boldsymbol{C}_{ll}^{(2)}=\boldsymbol{I}_{q_2\times q_2}+\sum_{j=1}^{p_1}\mathrm{E}\left\{(\Xi_{jl}^{(n)}-\boldsymbol{A}_{jl}^{(n)})^H\left(\left[\widetilde{\boldsymbol{C}}_{p_1\times p_1}^{(1)}+\widetilde{\boldsymbol{A}}_{p_1\times p_2}[\widetilde{\boldsymbol{C}}_{p_2\times p_2}^{(2)}]^{-1}\widetilde{\boldsymbol{A}}_{p_1\times p_2}^H\right]^{-1}\right)_{jj}(\Xi_{jl}^{(n)}-\boldsymbol{A}_{jl}^{(n)})\right\},$$
$$l=1,\dots,p_2, \tag{2.38}$$

There exists a unique solution $\boldsymbol{C}_{kk}^{(1)}$, $\boldsymbol{C}_{ll}^{(2)}$ to the previous canonical system of equations in the class of analytic functions $\boldsymbol{C}_{kk}^{(1)}(z)$ and $\boldsymbol{C}_{ll}^{(2)}(z)$, $k = 1, \dots, p_1$, $l = 1, \dots, p_2$ such that $\boldsymbol{C}_{kk}^{(1)}(z)$ and $\boldsymbol{C}_{ll}^{(2)}(z)$ are definite positive for $\mathrm{Re}(z) > 0$.

At first glance Theorem 5 seems to be of little practical use since the system of canonical equations (2.37) and (2.38) consists of infinite equations as $K, N \to \infty$. However, there are matrices of practical interest for which the system of canonical equations reduces to a finite number of equations, although $K, N \to \infty$.

This result has been applied to determine the capacity of MIMO Ricean channels in [62,63]. In such a case, the system of canonical equations reduces to two equations. In Chapter 5 it will be applied to the analysis of CDMA networks with multiple antennas at the receiving sites. The system of canonical equations consists of a number of equations equal to the square of the number of receiving antennas, in the most general case.

An intermediate result due to Bai and Silverstein [64] plays an important role in the following of this work. It analyzes the behaviour of a quadratic form $Q_N = \boldsymbol{x}_N^H \boldsymbol{C}_N \boldsymbol{x}_N$, with $\boldsymbol{C}_N$ an $N \times N$ complex matrix and $\boldsymbol{x}_N$ an N-dimensional random vector.

Lemma 1 *[64] Let $\boldsymbol{x}_N$ be an N-dimensional complex random vector with i.i.d. zero mean entries such that $\mathrm{E}\{|x_j|^2\} = \frac{1}{N}$ and let $\boldsymbol{C}_N$ be an $N \times N$ complex matrix independent of $\boldsymbol{x}_N$. Then, for any $p \geq 2$*

$$\mathrm{E}\left\{\left|\boldsymbol{x}_N^H \boldsymbol{C}_N \boldsymbol{x}_N - \frac{\operatorname{tr} \boldsymbol{C}_N}{N}\right|^p\right\} \leq K_p \left[(\mathrm{E}\{|x_j|^4\} \mathrm{tr}(\boldsymbol{C}_N \boldsymbol{C}_N^H))^{\frac{p}{2}} + \mathrm{E}\{|x_j|^{2p}\} \mathrm{tr}(\boldsymbol{C}_N \boldsymbol{C}_N^H)^{\frac{p}{2}}\right] \tag{2.39}$$

where K_p is a constant that does not depend on N, $\boldsymbol{C}_N$, or the distribution of x_j.

As a direct consequence of the previous lemma, for large matrices the quadratic form Q_N is well approximated by the normalized trace of $\boldsymbol{C}$, $\frac{\operatorname{tr} \boldsymbol{C}_N}{N}$ for large N. Since $\frac{\operatorname{tr} \boldsymbol{C}_N}{N} = \frac{\sum_{i=1}^N \lambda_i}{N}$, where λ_i denote the eigenvalues of $\boldsymbol{C}_N$, the quadratic form Q_N is well approximated also by average of the eigenvalues of $\boldsymbol{C}_N$. This is formalized in the following lemma.

Lemma 2 *Let $\boldsymbol{x}_N$ and $\boldsymbol{C}_N$ be as in Lemma 1. Additionally, let $\lim_{N\to\infty} N^3 \mathrm{E}\{|x_i|^6\} < +\infty$. Let $\{\boldsymbol{C}_N\}$ be a sequence of matrices $\boldsymbol{C}_N$ such that $F_{\boldsymbol{C}}^{(N)}$, the corresponding sequence of empirical eigenvalue distributions, converges to a nonrandom limit distribution $F_{\boldsymbol{C}}$ and $\int \lambda^p \mathrm{d}F_{\boldsymbol{C}}^{(N)}(\lambda) < +\infty$ for $p = \frac{3}{2}$ and 2 and $\forall N$[16]. Then, as $N \to \infty$, the quadratic form $Q_N = \boldsymbol{x}_N^H \boldsymbol{C}_N \boldsymbol{x}_N$ converges*

[16] This result is presented in the literature (e.g. [65]) under more restrictive conditions: $F_{\boldsymbol{C}}$ is required to have bounded support.

almost surely to the first moment of $F_{\boldsymbol{C}}$, i.e.

$$\lim_{N\to\infty} \boldsymbol{x}_N^H \boldsymbol{C}_N \boldsymbol{x}_N \overset{a.s.}{=} \lim_{N\to\infty} \frac{\mathrm{tr}\boldsymbol{C}_N}{N} = \int \lambda \mathrm{d}F_{\boldsymbol{C}}(\lambda). \tag{2.40}$$

The proof of this lemma is provided in Appendix A.

In the following lemma, a similar result is established for the form $\boldsymbol{x}_N^H \boldsymbol{C}_N \boldsymbol{y}_N$, where $\boldsymbol{x}_N$ and $\boldsymbol{y}_N$ are two independent random vectors.

Lemma 3 *Let $\boldsymbol{x}_N$ and $\boldsymbol{C}_N$ be as in Lemma 2. Let $\boldsymbol{y}_N$ be a vector similar to $\boldsymbol{x}_N$ and independent of it and of $\boldsymbol{C}_N$. Then,*

$$\lim_{N\to\infty} \boldsymbol{x}_N^H \boldsymbol{C}_N \boldsymbol{y}_N \overset{a.s.}{=} 0.$$

3 Efficient Multistage Detection for Synchronous Systems

3.1 Introduction

This chapter is focused on the design and analysis of low complexity multiuser detectors for synchronous CDMA systems with flat and frequency selective fading channels.

Making use of some properties of random matrices discovered in this work we introduce two multistage detectors with linear complexity order per bit, the same complexity order as the single user matched filter. A unified framework capable of describing large classes of multiuser detectors, such as the PIC detectors, the multistage Wiener filters and the polynomial expansion detectors, is adopted. A general result for the asymptotic performance of all detectors that fit into this framework is presented.

As shown in Chapter 2, the linear MMSE detector yields substantial improvements in performance, while maintaining a lower complexity than the optimum detector investigated in [15]. However, in systems with time-varying multiple access interference — due to, for example, long spreading sequences or fading channels — its computation in real time is very expensive. In fact, the linear MMSE detector requires the inversion of matrices that are at least of size $\min(K, N) \times \min(K, N)$, where K is the number of active users and N the spreading factor. When the system size is large, the complexity of a linear MMSE detector is prohibitive for real-time applications.

There is a large class of multiuser detectors that avoids matrix inversion: the linear multistage detectors. They are characterized by a modular structure and consist of a *projector* onto a subspace and a subsequent *filter*. The PIC, the MSWF filters, and the polynomial expansion detectors presented in Chapter 2 belong to such a class. All these detectors use the same Krylov subspace [36]. Hereafter, we refer to this subspace as *the projection subspace*. As already noticed in Section 2.3.5, the Krylov subspace has several useful properties:

- It need not be tracked.

- The subspace rank required to achieve a fixed level of performance does not scale with the system size [28].

- The multistage detector output SINR converges exponentially in the detector rank towards the linear MMSE detector output SINR [39] so that a low number of stages is sufficient to achieve near-linear MMSE performance.

The use of subspace methods does not allow for a significant reduction of the complexity order by default. In fact, the filter design, optimum in an MSE sense, has the same complexity order as the linear MMSE detector. A significant reduction in complexity can be obtained by approximating the optimum filter coefficients (also called weights) by asymptotic approximations [18, 52] at the cost of a slight degradation in performance [66] due to mismatch. The asymptotic multistage detectors, proposed first in [18, 52], take advantage of some asymptotic properties of random matrices such as the convergence of the eigenvalue moments to deterministic limits. These are independent of the spreading sequences and the channel realizations. Since these values can be expressed as a linear function of a small set of parameters, the asymptotic multistage weights can easily be computed off-line as a function of the eigenvalue moments. The complexity reduction promised by the use of asymptotic filter coefficients in [18, 52] inspired studies to design asymptotic weighting in different scenarios [37, 67–70]. Multistage detectors for systems with multipath fading channels have been considered only recently in parallel works for the downlink [13, 14, 70] and the uplink [13, 69]. Cottatellucci and Müller applied the multistage approach with asymptotic weights to both multiuser channel estimation for multipath fading and symbol detection [69]. The asymptotic performance of multistage detectors with no channel state information at the receiver is also analyzed in [69]. However, the application of this approach to channel estimation is beyond the scope of this work and it will not be considered further.

Thanks to the negligible computational complexity of the asymptotic filter design the complexity order of the detector is determined by the complexity of the projection onto the subspaces. Nevertheless, the projection complexity received little attention. In this work we consider jointly the projection complexity and the weighting complexity to significantly reduce the detector complexity. From the point of view of receiver complexity, it is desirable to perform the projection for all users jointly rather than using different projectors for each user if one wants to detect all users. In such a way most of the calculations of the projection become identical for all users and the complexity drops by a factor of K. This complexity reduction is possible only if the bases of the Krylov subspaces for all users can be chosen in an appropriate way to support the joint projection. Fortunately, such a set of bases does exists.

The low complexity of weight design and the asymptotic performance analysis of multistage detectors using such a set stem from the asymptotic convergence of the diagonal elements of random Gram matrices and their positive powers. This convergence is established in this work for the first time.

We design and analyze multistage detectors for CDMA systems in uplink with any kind of phase shift keying (PSK) symbol alphabets, random spreading, and multipath fading channels. We use subspace bases supporting the joint processing of all users so that all proposed multistage detectors have a linear complexity order per bit. From a conceptual point of view we focus on two asymptotic multistage detectors differing in the filter coefficients. Detector Type J-J uses a single set of weights satisfying the MMSE criterion jointly for all users. It is the counterpart with asymptotic weights of the polynomial expansion detector in Section 2.2.5. In detector Type J-I, the filter weights satisfy the MMSE criterion individually for each user. Detector Type J-I performs as well as the asymptotic multistage detectors in [13,37] but *its complexity is reduced by almost a factor of K* on the uplink CDMA channel. The detectors in [13, 37] will be referred to as detector Type I-I in the following.

Our analysis applies to a wider class of detectors than just Type J-J and J-I. It is applicable to any multistage detector using the same projection subspace bases, e.g. the linear "standard" partial parallel interference cancellation detectors. The asymptotic analysis can also be applied to the multistage Wiener filter. In fact, the asymptotic performance of the MSWF, the Type I-I detector, and the Type J-I detector are the same. This observation shows also the actual relation between the polynomial expansion detector and the MSWF. In the literature, the idea that those two detectors are equivalent is widely spread, explicitly claimed in [36] and implicitly assumed in [37, 38]. In contrast to this belief, we show that the two detectors differ and the MSWF outperforms the polynomial expansion detector in [26] for equal number of stages in general. The latter detector does not maximize the output SINR. This loss of optimality also affects the characteristics of its multiuser efficiency: In contrast to many of the other detectors analyzed in the literature, the multiuser efficiency of detector Type J-J depends on the received power of the user of interest.

The MSWF and the polynomial expansion detector coincide asymptotically in case of equal received powers for all users and, under these conditions, they are also equivalent to the multistage detector proposed in [18]. Therefore, polynomial expansion detectors can be efficiently utilized in scenarios with classical power control[1] whereas they degrade in performance, compared to MSWF, in CDMA systems with power imbalances.

[1]We refer to control mechanisms that force the received power of all users to be equal as classical power control. These mechanisms are adopted in systems that do not adopt multiuser detection and thus suffer from near-far effects. However, the current developments [6] show that heavily loaded systems can benefit from power imbalances.

Similarly to the asymptotic analysis of the linear MMSE detector in Section 2.3.3, the performance and the weighting of both detector Type J-J and detector Type J-I are independent of the spreading sequences and the fading channel realizations. They depend only on few macroscopic parameters, namely, the number of users per chip, the received power statistics, the noise variance, and the received power of the user of interest. The analysis proposed in this work provides deep insight into the system behaviour and clear guidelines for the design.

Recently, implementations of low complexity polynomial expansion detectors which do not benefit from the asymptotic approximation of the weights have been proposed in [71]. This approach utilizes an alternative basis of the projection subspace obtained by a Gram-Schmidt orthogonalization (GSO) [72]. With such a basis, it is possible to avoid the asymptotic weight design problem at the cost of the GSO, which can cause numerical problems for fixed-point arithmetic. Both the polynomial expansion detectors with universal weights and the polynomial expansion detectors in [71] require the multiplication of the received signal by the basis vectors. However, while the former perform the subsequent processing with negligible complexity, the latter requires the GSO, and then, the application of the Lanczos algorithm (see e.g. [72]) for the inversion of a symmetric Hessenberg matrix. As all other polynomial expansion detectors, the detectors in [71] also suffer from power imbalances. In the absence of a perfect power control they have worse performance than the MSWF and Type J-I detectors as apparent from Figure 3.4 and Figure 5.1.

These aspects of the polynomial expansion detector in [71] are magnified when the system is not synchronized. In fact, as it will be discussed in Chapter 4, the polynomial expansion detectors perform worse than the MSWF detectors also in case of perfect power control. Additionally, when applied to asynchronous systems, the finite approach in [71] has a complexity increasing with the length of the observation window, similar to the linear MMSE detector. In contrast, as it will be apparent in Chapter 4, multistage detectors with universal weights keep the same complexity as the equivalent detector for synchronous CDMA systems.

The linear multiuser detectors considered in this work assume perfect knowledge of the spreading sequences and the channel gains for both the desired users and the interfering users. Typically, this information is not available in the downlink. Then, an alternative class of detectors has to be utilized: the adaptive linear multiuser detectors. They approximate the ideal MMSE filter making use of adaptive algorithms. In this case training sequences for the estimation of the linear MMSE detector have to be transmitted or the spreading sequence of the user of interest has to be known. A complete overview on adaptive multiuser detectors is beyond the scope of this work. The interested reader is referred to the bibliographical notes in [16] and references therein. We emphasize here that adaptive implementations of MSWF are available [73–75] and their asymptotic performance has been analyzed in [29].

This chapter is structured as follows. Section 3.2 introduces the system model and the notation. In Section 3.3, we discuss criteria for the choice of the subspace bases and for filter optimization. We analyze their impact on performance, complexity, and design. The design of detectors Type J-J and Type J-I with universal asymptotic weights is illustrated in Section 3.4. Section 3.5 provides a performance analysis in asymptotic conditions. Section 3.6 presents numerical results and simulations assessing the degradation introduced by the asymptotic multistage detectors when used for finite systems and compares detector Type J-J and detector Type J-I in terms of performance. Conclusions on the analysis and design of low complexity multistage detectors for synchronous systems with flat and frequency selective fading are drawn in Section 3.7.

3.2 System Model

Let us consider a synchronous CDMA communication system with spreading factor N and K physical users, multipath fading, and additive noise at the receiver. Throughout this work the delay spread of the channel is small compared to the symbol interval so that the intersymbol interference can be neglected. Then, the equivalent baseband signals at the chip matched filter output are given by

$$\boldsymbol{y}(n) = \boldsymbol{H}(n)\boldsymbol{b}(n) + \boldsymbol{n}(n) \tag{3.1}$$

where $\boldsymbol{y}(n)$ is the N-dimensional received vector and $\boldsymbol{b}(n)$ is the K-dimensional column vector of transmitted symbols (one signal per each physical user) at the time instant n. The transmitted symbols belong to a finite alphabet in $\mathbb{C}$, they are zero mean and satisfy the relation $\mathrm{E}\{\mathbf{b}(n)\mathbf{b}(p)^H\} = \boldsymbol{I}\delta_{n,p}$. $\boldsymbol{n}(n)$ is the N-dimensional additive noise vector at the time instant n. The additive noise is circularly symmetric complex-valued white Gaussian with zero mean and variance σ^2.

The influence of spreading, transmission amplitudes, and fading is described by the $N \times K$ matrix [65]

$$\boldsymbol{H}(n) = \boldsymbol{S}(n)\boldsymbol{A}(n).$$

$\boldsymbol{A}(n)$ is the $KL \times K$ block diagonal matrix of received amplitudes taking into account the fading channel amplitudes and the transmitted powers. It consists of blocks of size $L \times 1$, assuming that the channels have impulse responses of lengths L with[2] $L \ll N$. $\mathbf{a}_k$ is the k^{th} block diagonal element of $\boldsymbol{A}$. The multipath channels are perfectly known at the receiver. In the asymptotic design and analysis carried out

[2] This last condition is implied by the assumption that the delay spread of the channel is small compared to the symbol interval.

in this work, we assume that the sequence of the joint empirical distributions of $\mathbf{a}_k$, $F_{\mathbf{a}}^{(K)}(a_1, a_2, \ldots, a_L) = \frac{1}{K}\sum_{k=1}^{K}\prod_{\ell=1}^{L} 1(a_\ell - (\mathbf{a}_k)_\ell)$, converges almost surely, as $K \to \infty$, to a non-random limit distribution function $F_{\mathbf{a}}(a_1, a_2, \ldots, a_L)$ with upper bounded support. The eigenvalues of the matrix $\boldsymbol{A}^H\boldsymbol{A}$ are given by $\lambda_k = \mathbf{a}_k^H\mathbf{a}_k$. Hereafter, we denote by $F_{|\boldsymbol{A}|^2}(\lambda)$ their asymptotic distribution. The matrix of ran-

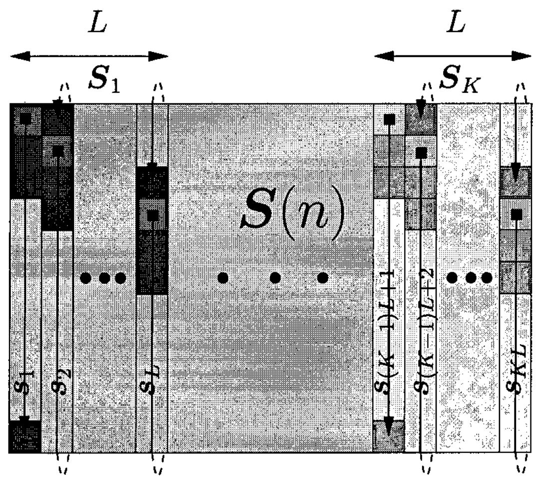


Figure 3.1: Structure of the spreading matrix $\boldsymbol{S}(n)$.

dom signature sequences $\boldsymbol{S}(n)$ is an $N \times KL$ random block matrix in $\mathbb{C}$ with blocks $\boldsymbol{S}_k = (\boldsymbol{s}_{(k-1)L+1}, \boldsymbol{s}_{(k-1)L+2}, \ldots, \boldsymbol{s}_{kL})$, $1 \leq k \leq K$, of size $N \times L$. Its structure is shown in Figure 3.1. The elements in a column vector $\boldsymbol{s}_{(k-1)L+1}$ are i.i.d. with zero mean and variance $\mathrm{E}\{|s_{j,(k-1)L+1}|^2\} = \frac{1}{N}$. Additionally, they are also i.i.d. from block to block. Within a block, the vector $\boldsymbol{s}_{(k-1)L+s}$ is $\boldsymbol{s}_{(k-1)L+1}$ cyclically shifted by $s-1$ positions. This downshift of the spreading sequence models the multipath fading. The cyclical downshift of the spreading sequence is a technical approximation justified by the assumption that $L \ll N$. These structures of the matrices $\boldsymbol{A}$ and $\boldsymbol{S}$ allow us to take into account the interchip interference due to multipath fading.

We adopt the following notation:

- $\beta = \frac{K}{N}$ for the system load;
- $\boldsymbol{h}_k(n)$ denotes the k^{th} column of $\boldsymbol{H}(n)$;
- $\boldsymbol{T}(n) = \boldsymbol{H}(n)\boldsymbol{H}(n)^H$;
- $\boldsymbol{R}(n) = \boldsymbol{H}(n)^H\boldsymbol{H}(n)$;
- $\boldsymbol{H}_{\sim k}(n)$ is the $N \times (K-1)$ matrix obtained from $\boldsymbol{H}(n)$ by removing the k^{th} column;

- $\boldsymbol{T}^m_{\sim k}(n) = (\boldsymbol{H}_{\sim k}(n)\boldsymbol{H}_{\sim k}(n)^H)^m$;
- $\boldsymbol{R}^m_{\sim k}(n) = (\boldsymbol{H}_{\sim k}(n)^H\boldsymbol{H}_{\sim k}(n))^m$.

By neglecting the intersymbol interference only quantities at the symbol-time index n appear in the system model. Therefore, the symbol-time index n will be omitted in what follows.

3.3 Multistage Detectors

3.3.1 Definitions

A linear multistage detector of order M for user k is a multiuser detector performing

1. a projection of the observed signal onto the Krylov subspace
$$\chi_{M,k}(\boldsymbol{H}) = \text{span}\{\boldsymbol{T}^m_{\sim k}\boldsymbol{h}_k\}_{m=0}^{M-1} \tag{3.2}$$
$$= \text{span}\{\boldsymbol{T}^m\boldsymbol{h}_k\}_{m=0}^{M-1}. \tag{3.3}$$
Note that, although other non-orthogonal bases slightly different have been proposed in literature too, these two[3] are capable to catch the main features of all non-orthogonal bases investigated in literature.
2. A subsequent processing of the projections by a filter designed according to an optimality criterion.

The choice of the Krylov subspace is motivated by two different observations. First, as shown in Section 2.2.3, the full-rank linear MMSE detector lies in $\chi_{K,k}(\boldsymbol{H})$, i.e. it is a linear combination of the basis vectors of $\chi_{K,k}(\boldsymbol{H})$. Second, the multistage filter output SINR converges exponentially in the filter rank M toward the full rank linear MMSE filter output SINR (see Section 2.3.6). Moroever, under the MMSE optimality criterion, the dimension M of the subspace needed to obtain a target SINR (e.g. within a small ϵ of the full rank SINR) does not scale with the system size (i.e. K and N) [28].

Both the projection and the filter design can be performed jointly for all users or individually for each user. This influences both the performance and the complexity of the resulting multistage detector. The joint projection is obtained using the vectors in (3.3) as a basis of $\chi_{M,k}(\boldsymbol{H})$. In this case, the projector[4] consists of a matched filter $\boldsymbol{H}^H$ and M stages each of them performing respreading — filtering by $\boldsymbol{H}$ — and successive matched filtering. The corresponding multistage detector is shown in Figure 3.2. Using the vectors in (3.2), no joint computation of the projections is known for $M > 2$ and K different projectors are required.

[3] About the identity of the subspaces spanned by the two bases in (3.2) and (3.3) see [28].
[4] We use here the word projection in the wide sense discussed in Section 2.2.6.

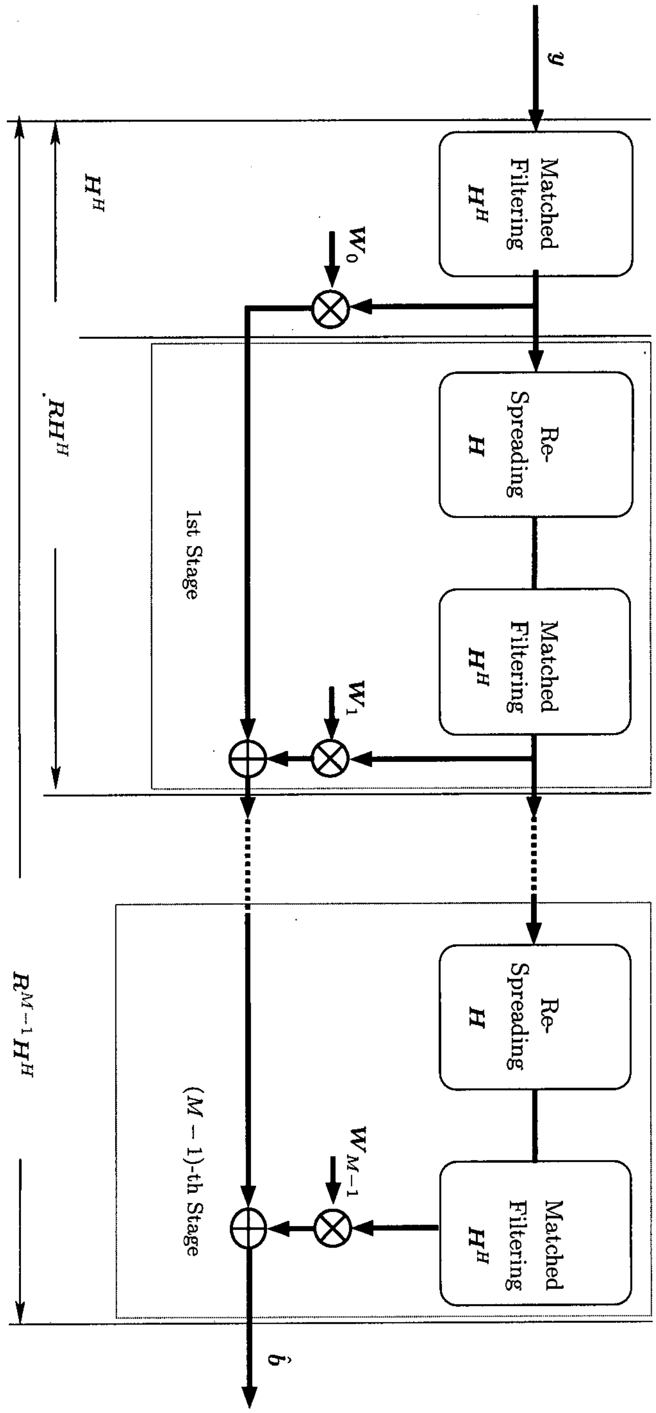


Figure 3.2: Type J-I detector for synchronous systems.

For the basis (3.3), the filter design can be performed jointly using the same filter coefficients for all users and choosing them, for example, by enforcing the minimization of the MSE averaged over all users [67]. Alternatively, we can design a different filter for each user minimizing the MSE individually. Table 3.1 shows the possible combinations and states the denominations.

	Joint Projection	**Individual Projection**
Joint Filtering	TYPE J-J	$\nexists$
Individual Filtering	TYPE J-I	TYPE I-I

Table 3.1: Multistage detector classification.

Detectors Type J-J are detectors with asymptotic weights approximating the polynomial expansion detectors in Section 2.2.5. Detectors Type I-I approximate the multistage Wiener filters in Section 2.2.6. Detectors Type J-I combine the advantages of detectors Type J-J in terms of complexity and of detectors Type I-I in terms of performance and are introduced in this work. Detectors Type J-I and Type I-I adopt the same optimality criterion in the same subspace and differ only in the choice of the subspace basis. Therefore, they have identical performance. However, they need, in general, different weights.

3.3.2 Complexity

Being a subspace methods does not imply that the multistage detectors have lower complexity order than the full rank linear MMSE detector. In fact, if we choose the minimization of the MSE as optimality criterion, the complexity of the filter coefficient design is identical to the complexity order of the linear MMSE detector. However, by approximating the optimum filter coefficients with the corresponding asymptotic limits in large systems, i.e. as $K, N \to \infty$ with $\frac{K}{N} \to \beta$, as proposed in [18, 52], the complexity of coefficient design becomes negligible with respect to the projection complexity. This justifies the efforts devoted to determine the asymptotic weighting in this work and, independently, in [13, 14, 37]. Referring to the denominations introduced in Table 3.1, the asymptotic weights of Type I-I detectors are designed in [13, 14] for the downlink and in [13, 37] for the uplink.

The complexity order per bit, driven by the projection complexity for detectors with asymptotic filter coefficients, is shown in Table 3.2. Table 3.2 distinguishes two cases: a single user is detected, typically in the downlink, and all users are detected, typically in the uplink. When the multiuser detection can be performed jointly for all

Detector	Detection of one user	Detection of all users
SUMF	$\mathcal{O}(K)$	$\mathcal{O}(K)$
Type J-J	$\mathcal{O}(K^2)$	$\mathcal{O}(K)$
Type J-I	$\mathcal{O}(K^2)$	$\mathcal{O}(K)$
Type I-I[5]	$\mathcal{O}(K^2)$	$\mathcal{O}(K^2)$
LMMSE	$\mathcal{O}(K^3)$	$\mathcal{O}(K^2)$

Table 3.2: Complexity order per *bit* ($K = \beta N$ is the number of active users).

users the complexity order per bit drops by a factor of K compared to the multiuser detection of a single user. In fact, for multistage detectors Type J-I and Type J-J the most of computations in the projection are common to all users, for the linear MMSE detector the matrix inversion is performed only once for the detection of all active users. Considering the advantages of the Type J-J and Type J-I detectors in terms of complexity with respect to Type I-I detectors and linear MMSE detectors, we focus on Type J-J and Type J-I detectors.

3.3.3 Individual Filtering: Type J-I Detectors

Projecting the received signal onto the subspaces $\chi_{M,k}(\boldsymbol{H})$ with $M < K$ we obtain an M-dimensional non-sufficient statistic of the received signal. We denote this statistic as $\boldsymbol{x}_k$

$$\boldsymbol{x}_k \triangleq \begin{bmatrix} \boldsymbol{h}_k^H \boldsymbol{y} \\ \boldsymbol{h}_k^H \boldsymbol{T} \boldsymbol{y} \\ \vdots \\ \boldsymbol{h}_k^H \boldsymbol{T}^{M-1} \boldsymbol{y} \end{bmatrix}. \tag{3.4}$$

The finite Type J-I detector for user k is defined as the linear transformation in $\chi_{M,k}(\boldsymbol{H})$[6]

$$\boldsymbol{m}_k^H = \sum_{m=0}^{M-1} (\boldsymbol{w}_k)_m \boldsymbol{h}_k^H \boldsymbol{T}^m \tag{3.5}$$

[5]For Type I-I detectors with one stage ($M = 2$) an implementation with complexity order $\mathcal{O}(K)$ is possible if all users are detected (e.g. uplink).

[6]Detectors Type J-J and Type I-I denote the detectors with asymptotic weights corresponding to polynomial expansion detectors and multistage Wiener filter, respectively. Thus, we keep the historical distinction between the finite optimum detectors and their asymptotic approximation. In contrast, detector Type J-I denotes both the finite optimum detector and the asymptotic approximation. In fact, the association of the Type J-I detector to the MSWF detector can be confused with a Type I-I detector with exact weights. It will be specified in the context which detector we refer to.

that satisfies the MMSE criterion, i.e. the weight vector $\boldsymbol{w}_k$ is given by

$$\begin{aligned} \boldsymbol{w}_k &= \arg\min_{\overline{\boldsymbol{w}}_k} \mathrm{E}\left\{\left\|\sum_{m=0}^{M-1}(\overline{\boldsymbol{w}}_k)_m \boldsymbol{h}_k^H \boldsymbol{T}^m \boldsymbol{y} - b_k\right\|^2\right\} \\ &= \arg\min_{\overline{\boldsymbol{w}}_k} \mathrm{E}\left\{\left\|\overline{\boldsymbol{w}}_k^H \boldsymbol{x}_k - b_k\right\|^2\right\}. \end{aligned} \tag{3.6}$$

From the second expression, the finite Type J-I detector reduces to scalar linear MMSE estimation on the non-sufficient statistic $\boldsymbol{x}_k$. Thus, the Wiener-Hopf theorem can be applied [20]:

$$\boldsymbol{w}_k = \boldsymbol{\Phi}_k^{-1}\boldsymbol{\varphi}_k, \tag{3.7}$$

where $\boldsymbol{\Phi}_k = \mathrm{E}\left\{\boldsymbol{x}_k\boldsymbol{x}_k^H\right\}$ and $\boldsymbol{\varphi}_k = \mathrm{E}\left\{b_k^*\boldsymbol{x}_k\right\}$. It is straightforward to verify that the following expressions hold

$$\boldsymbol{\Phi}_k = \begin{pmatrix} (\boldsymbol{R}^2)_{kk} + \sigma^2(\boldsymbol{R})_{kk} & \dots & (\boldsymbol{R}^{M+1})_{kk} + \sigma^2(\boldsymbol{R}^M)_{kk} \\ (\boldsymbol{R}^3)_{kk} + \sigma^2(\boldsymbol{R}^2)_{kk} & \dots & (\boldsymbol{R}^{M+2})_{kk} + \sigma^2(\boldsymbol{R}^{M+1})_{kk} \\ \dots & \dots & \dots \\ (\boldsymbol{R}^{M+1})_{kk} + \sigma^2(\boldsymbol{R}^M)_{kk} & \dots & (\boldsymbol{R}^{2M})_{kk} + \sigma^2(\boldsymbol{R}^{2M-1})_{kk} \end{pmatrix} \tag{3.8}$$

and

$$\boldsymbol{\varphi}_k = \left((\boldsymbol{R})_{kk}, (\boldsymbol{R}^2)_{kk}, \dots, (\boldsymbol{R}^M)_{kk}\right)^T, \tag{3.9}$$

where $(\boldsymbol{R}^s)_{kk}$ is the k^{th} diagonal element of the matrix $\boldsymbol{R}^s$.

The Type J-I detector is also the multistage detector in $\chi_{M,k}(\boldsymbol{H})$ that maximizes the signal-to-interference-and-noise-ratio SINR_k of user k at the detector output. This will be shown in Section 3.3.5.

The Type J-I detector for all users has structure

$$\boldsymbol{M} = \sum_{m=0}^{M-1} \boldsymbol{W}_m \boldsymbol{H}^H \boldsymbol{T}^m = \sum_{m=0}^{M-1} \boldsymbol{W}_m \boldsymbol{R}^m \boldsymbol{H}^H, \tag{3.10}$$

where $\boldsymbol{W}_m$ is the diagonal matrix whose k^{th} diagonal element is the m^{th} component of $\boldsymbol{w}_k$. It minimizes $\mathrm{E}\left\{\|\boldsymbol{M}\boldsymbol{y} - \boldsymbol{b}\|^2\right\}$.

3.3.4 Joint Filtering: Polynomial Expansion Detectors

In this section the concept of polynomial expansion detector presented in Section 2.2.5 is revisited from a completely different perspective. The polynomial expansion detectors are presented here as a subspace method in analogy to Type J-I detectors. The polynomial expansion detector is the linear transformation in[7]

[7]Note that in this case we consider the vector space of complex matrices of dimension $K \times N$. A basis in a vector subspace consists of elements of the vector space. In this case, the elements of the basis are $K \times N$ matrices.

$\chi_M(\boldsymbol{H}) = \mathrm{span}\{\boldsymbol{H}^H\boldsymbol{T}^m\}_{m=0}^{M-1}$,

$$\boldsymbol{L} = \sum_{m=0}^{M-1} w_m \boldsymbol{H}^H\boldsymbol{T}^m = \sum_{m=0}^{M-1} w_m \boldsymbol{R}^m\boldsymbol{H}^H \tag{3.11}$$

such that the scalar weights w_m minimize the mean square error $\mathrm{E}\left\{\|\boldsymbol{L}\boldsymbol{y} - \boldsymbol{b}\|^2\right\}$. Let us compare the polynomial expansion detector with the Type J-I detector. They differ in the weights: scalar weights characterize $\boldsymbol{L}$ while matrix weights appear in $\boldsymbol{M}$.

The weighting is

$$\boldsymbol{w} = \boldsymbol{\Phi}^{-1}\boldsymbol{\varphi} \tag{3.12}$$

where the elements of the M-dimensional vector $\boldsymbol{\varphi}$ and the elements of the $M \times M$ matrix $\boldsymbol{\Phi}$ can be expressed in terms of the traces of the powers of $\boldsymbol{R}$ as $(\boldsymbol{\Phi})_{ij} = \mathrm{tr}(\boldsymbol{R}^{i+j}) + \sigma^2\mathrm{tr}(\boldsymbol{R}^{i+j-1})$ and $(\boldsymbol{\varphi})_i = \mathrm{tr}(\boldsymbol{R}^i)$. This implies $\boldsymbol{\varphi} = \sum_{k=1}^K \boldsymbol{\varphi}_k$ and $\boldsymbol{\Phi} = \sum_{k=1}^K \boldsymbol{\Phi}_k$, with $\boldsymbol{\Phi}_k$ and $\boldsymbol{\varphi}_k$ defined in (3.8) and (3.9), respectively.

From the definition and the expression of polynomial expansion detectors in Section 2.2.5 it is evident that the polynomial expansion detectors minimize also the MSE between its own output and the output of a full rank linear MMSE detector.

3.3.5 Performance

For the full-rank linear MMSE detector, it is well known that the minimization of the MSE per user is equivalent to the minimization of the sum of the MSE of each user and also to the maximization of the SINR at the output of the filter for each physical user [76]. This is due to the fact that the detector is free to lie in the full space of the linear transformations that map $\mathbb{C}^N$ into $\mathbb{C}^K$. This property does not hold if the detector is forced to lie in a specific subspace as in the case of multistage detectors. Here, a difference between the joint minimization of the MSE (proposed in [26]) and the minimization of the MSE for each user (proposed in [28]) appears. The maximization of the SINR is achieved only in the latter case.

For any multistage detector in $\chi_{M,k}(\boldsymbol{H})$ with weight vector $\overline{\boldsymbol{w}}_k$, the MSE of user k is given by

$$\mathrm{MSE}_k = \overline{\boldsymbol{w}}_k^H\mathrm{E}\{\boldsymbol{x}_k\boldsymbol{x}_k^H\}\overline{\boldsymbol{w}}_k - 2\mathrm{Re}(\overline{\boldsymbol{w}}_k^H\mathrm{E}\{\boldsymbol{x}_k b_k^*\}) + 1. \tag{3.13}$$

Recalling that $\mathrm{E}\{\boldsymbol{x}_k\boldsymbol{x}_k^H\} = \boldsymbol{\Phi}_k$ and $\mathrm{E}\{\boldsymbol{x}_k b_k^*\} = \boldsymbol{\varphi}_k$, we obtain

$$\mathrm{MSE}_k \;=\; 1 - 2\,\mathrm{Re}\left(\boldsymbol{\varphi}_k^T\overline{\boldsymbol{w}}_k\right) + \overline{\boldsymbol{w}}_k^H\boldsymbol{\Phi}_k\overline{\boldsymbol{w}}_k. \tag{3.14}$$

The corresponding SINR for user k is given by

$$\mathrm{SINR}_k = \frac{P_k}{P - P_k} \tag{3.15}$$

where P_k is the useful power of user k at the detector output and P is the total power. We have

$$P = \mathrm{E}\{\overline{\boldsymbol{w}}_k^H \boldsymbol{x}_k \boldsymbol{x}_k^H \overline{\boldsymbol{w}}_k\} = \overline{\boldsymbol{w}}_k^H \boldsymbol{\Phi}_k \overline{\boldsymbol{w}}_k \tag{3.16}$$

and

$$P_k = \mathrm{E}\left\{\left|\sum_{m=0}^{M-1} (\overline{w}_k)_m \boldsymbol{h}_k^H \boldsymbol{T}^m \boldsymbol{H} \boldsymbol{e}_k b_k\right|^2\right\} = \overline{\boldsymbol{w}}_k^H \boldsymbol{\varphi}_k \boldsymbol{\varphi}_k^T \overline{\boldsymbol{w}}_k, \tag{3.17}$$

where $\boldsymbol{e}_k$ is a K-dimensional vector with all components equal to zero except the k^{th} that is equal to 1. Thus, (3.15) becomes

$$\mathrm{SINR}_k = \frac{\overline{\boldsymbol{w}}_k^H \boldsymbol{\varphi}_k \boldsymbol{\varphi}_k^T \overline{\boldsymbol{w}}_k}{\overline{\boldsymbol{w}}_k^H \left(\boldsymbol{\Phi}_k - \boldsymbol{\varphi}_k \boldsymbol{\varphi}_k^T\right) \overline{\boldsymbol{w}}_k}, \tag{3.18}$$

for any weight vector $\overline{\boldsymbol{w}}_k$. Specializing (3.14) and (3.18) to Type J-I detectors with $\overline{\boldsymbol{w}}_k = \boldsymbol{w}_k$ in (3.7), we obtain

$$\begin{aligned}
\mathrm{MSE}_{\mathrm{J-I},k} &= 1 - \boldsymbol{\varphi}_k^T \boldsymbol{\Phi}_k^{-1} \boldsymbol{\varphi}_k, &(3.19)\\
\mathrm{SINR}_{\mathrm{J-I},k} &= \frac{\boldsymbol{\varphi}_k^T \boldsymbol{\Phi}_k^{-1} \boldsymbol{\varphi}_k}{1 - \boldsymbol{\varphi}_k^T \boldsymbol{\Phi}_k^{-1} \boldsymbol{\varphi}_k} &(3.20)\\
&= \frac{1}{\mathrm{MSE}_{\mathrm{J-I},k}} - 1. &(3.21)
\end{aligned}$$

Calculating the gradient of SINR_k in (3.18) with respect to $\overline{\boldsymbol{w}}_k$ it is possible to verify that the Type J-I detector maximizes each SINR_k as already noticed in Section 3.3.3.

For the polynomial expansion detector, the performance becomes with $\overline{\boldsymbol{w}}_k = \boldsymbol{w}$ in (3.12)

$$\begin{aligned}
\mathrm{MSE}_{\mathrm{J-J},k} &= 1 - 2\boldsymbol{\varphi}_k^T \boldsymbol{\Phi}^{-1} \boldsymbol{\varphi} + \boldsymbol{\varphi}^T \boldsymbol{\Phi}^{-1} \boldsymbol{\Phi}_k \boldsymbol{\Phi}^{-1} \boldsymbol{\varphi}^T &(3.22)\\
\mathrm{SINR}_{\mathrm{J-J},k} &= \frac{1}{\dfrac{\boldsymbol{\varphi}^T \boldsymbol{\Phi}^{-1} \boldsymbol{\Phi}_k \boldsymbol{\Phi}^{-1} \boldsymbol{\varphi}}{\left(\boldsymbol{\varphi} \boldsymbol{\Phi}^{-1} \boldsymbol{\varphi}\right)^2} - 1} &(3.23)\\
&= \frac{\left(\boldsymbol{\varphi}_k^T \boldsymbol{\Phi}^{-1} \boldsymbol{\varphi}\right)^2}{\mathrm{MSE}_{\mathrm{J-J},k} - \left(\boldsymbol{\varphi}_k^T \boldsymbol{\Phi}^{-1} \boldsymbol{\varphi} - 1\right)^2}. &(3.24)
\end{aligned}$$

It can be shown that the polynomial expansion detector does not null the gradient of each SINR_k or the gradient of $\sum_{k=1}^{K} \mathrm{SINR}_k$. Therefore, the polynomial expansion detector does not maximize the SINR. Since the Type I-J detector with exact weights as well as the multistage Wiener filter do maximize the SINR, it also follows that the Type J-I detector outperforms the polynomial expansion detector in the same projection subspace.

Note that (3.21), the relation between SINR_k and MSE_k, also holds for the full-rank linear MMSE detector, while the equivalent relation (3.24) for the polynomial expansion detector is more involved. For $M = K$, both the polynomial expansion detector and the Type J-I detector coincide with the full-rank linear MMSE detector. This is a well known equivalence stated in [27] for the multistage Wiener filter and in [26] for the polynomial expansion detector.

In the following section we attack the problem of asymptotic weight design. The design of asymptotic weights entails the use of a stronger property of random matrices than the well known convergence of the eigenvalue distribution: the convergence of the diagonal elements of its positive integer powers. This property is established in Section 3.4.

3.4 Asymptotic Detector Design

The asymptotic multistage detectors are based on the idea of approximating the weights of the optimum multistage detectors with the corresponding weights of the detector for large systems. In fact, for finite K and N both $\mathrm{tr}(\boldsymbol{R}^m)$ and $(\boldsymbol{R}^m)_{kk}$, for $m \in \mathbb{Z}^+$ and $k = 1, \ldots, K$, are random variables because of the random assignment of the spreading sequences and of the channel gains. Their computation has complexity $\mathcal{O}(K^3)$. However, it is known that, as $K, N \to \infty$ with $\frac{K}{N} \to \beta$, $\mathrm{tr}(\boldsymbol{R}^m)$ tends to a deterministic value independent of the spreading sequences and depending only on the system load β and the limiting eigenvalue distribution $F_{\mathbf{a}}(a_1, a_2, \ldots, a_L)$. These asymptotic values can be computed at complexity $\mathcal{O}(1)$ [77] and need updating only when β and/or $F_{\mathbf{a}}(a_1, a_2, \ldots, a_L)$ change. We show that the same property holds also for the diagonal elements of the matrix $\boldsymbol{R}^m$. We will efficiently use this property for the design of Type J-I asymptotic weighting.

First, for the sake of simplicity, we derive the deterministic limit for flat fading channels and then we extend the results to multipath fading channels.

Theorem 6 *Let $\boldsymbol{A}$ be a $K \times K$ diagonal matrix in $\mathbb{C}$ with bounded elements and such that the sequence of the eigenvalue distribution of $\boldsymbol{A}^H\boldsymbol{A}$ converges almost surely, as $K \to \infty$, to a deterministic distribution function $F_{|\boldsymbol{A}|^2}(\lambda)$ with upper bounded support. Let $\boldsymbol{S} \in \mathbb{C}^{N \times K}$ have random i.i.d. zero mean entries with variance $\mathrm{E}\{|s_{ij}|^2\} = \frac{1}{N}$, and $\lim_{N\to\infty} \mathrm{E}\{N^3|s_{ij}|^6\} < +\infty$. Let $\boldsymbol{R} = \boldsymbol{A}^H\boldsymbol{S}^H\boldsymbol{S}\boldsymbol{A}$. Then, conditioned on a_{kk}, the k^{th} diagonal element of $\boldsymbol{A}$, $(\boldsymbol{R}^\ell)_{kk}$ converges almost surely, as $N, K \to \infty$ with $\frac{K}{N} \to \beta$, to the conditionally deterministic quantity*

$$R^\ell_{kk,\infty} = R^\ell_\infty(\lambda)|_{\lambda=|a_{kk}|^2} \tag{3.25}$$

with

$$R_{\infty}^{\ell}(\lambda) = \lambda \sum_{s=0}^{\ell-1} R_{\infty}^{s}(\lambda)\beta m_{\boldsymbol{R}}^{\ell-s-1}, \qquad \ell > 1 \tag{3.26}$$

for any $k, \ell \in \mathbb{Z}^+$. Here, $m_{\boldsymbol{R}}^s = \mathrm{E}\left\{\frac{1}{K}\mathrm{tr}\left(\boldsymbol{R}^s\right)\right\}$. The initial values of the recursion are $R_{kk,\infty}^0 = 1$ and $m_{\boldsymbol{R}}^0 = \beta^{-1}$. □

Theorem 6 is proven in Appendix B.1.

Note that the constraints on the chips are typically satisfied in practical systems with binary or Gaussian spreading. In fact, in order to normalize the spreading waveform to have unit energy the variance of the chips is typically $\frac{1}{N}$. The additional constraint $\lim_{N\to\infty} \mathrm{E}\{N^3|s_{ij}|^6\} < +\infty$ implies that $s'_{ij} = \sqrt{N}s_{ij}$, the chip re-scaled to have unit variance, has finite sixth moment. This property is also usually verified in physical systems.

A closed-form expression for the moments $m_{\boldsymbol{R}}^s$ can be found in [77]. Let us recall a fundamental property of linear algebra: *If $\widetilde{\lambda}_k$, $k = 1, \ldots K$ are the eigenvalues of the matrix $\boldsymbol{R}$, then $\frac{\sum_{k=1}^K \widetilde{\lambda}_k^\ell}{K} = \frac{\mathrm{tr}\boldsymbol{R}^\ell}{K}$.* Thus, an alternative recursive expression for the eigenvalue moments of $\boldsymbol{R}$ can be obtained noting that

$$\begin{aligned}
m_{\boldsymbol{R}}^{\ell} &= \mathrm{E}\{\widetilde{\lambda}\} \\
&= \lim_{K=\beta N\to\infty} \frac{\sum_{k=1}^K \widetilde{\lambda}_k^\ell}{K} \\
&= \lim_{K=\beta N\to\infty} \frac{\mathrm{tr}(\boldsymbol{R}^\ell)}{K} \\
&= \lim_{K=\beta N\to\infty} \frac{1}{K}\sum_{k=1}^{K} (\boldsymbol{R}^\ell)_{kk} \\
&= \int R_{\infty}^{\ell}(\lambda)\, \mathrm{d}F_{|\boldsymbol{A}|^2}(\lambda).
\end{aligned} \tag{3.27}$$

Substituting the right hand side of (3.26) in the right hand side of (3.27) we obtain the recursive expression of $m_{\boldsymbol{R}}^{\ell}$.

Corollary 1 *Let $\boldsymbol{A}$, $\boldsymbol{S}$ and $F_{|\boldsymbol{A}|^2}(\lambda)$ be as in Theorem 6. Then, the asymptotic eigenvalue moments of $\boldsymbol{R}$ are given by*

$$m_{\boldsymbol{R}}^{\ell} = \beta \sum_{s=0}^{\ell-1} m_{\boldsymbol{R}}^{\ell-s-1} \mathrm{E}\{\lambda R_{\infty}^{s}(\lambda)\} \tag{3.28}$$

where the expectation is taken over the c.d.f. $F_{|\boldsymbol{A}|^2}(\lambda)$. The initializing moment is $m_{\boldsymbol{R}}^0 = \beta^{-1}$. □

Note that, thanks to [55], (3.28) can be also used to calculate the eigenvalue moments for random matrices whose elements s_{ij} do not satisfy the constraint on the sixth moment. Theorem 6 and Corollary 1 suggest a simple algorithm to determine $R^{\ell}_{kk,\infty}$ and $m^{\ell}_{\boldsymbol{R}}$:

Algorithm 1

Initialization: *Let* $\rho_0(x) = 1$ *and* $\mu_0 = \beta^{-1}$.

ℓ^{th} step:

- *Define* $\rho_\ell(x) = \beta x \sum_{s=0}^{\ell-1} \rho_s(x)\mu_{\ell-1-s}$ *and write it as a polynomial in* x.
- *Assign* $\rho_\ell(|a_{kk}|^2)$ *to* $R^{\ell}_{kk,\infty}$. *Replace all monomials* $x, x^2, \dots, x^\ell$ *in the polynomial* $\rho_\ell(x)$ *by the moments* $m^1_{|\boldsymbol{A}|^2}, m^2_{|\boldsymbol{A}|^2}, \dots, m^\ell_{|\boldsymbol{A}|^2}$*, respectively and assign the result to* $m^{\ell}_{\boldsymbol{R}}$.

A closed-form expression for $R^{\ell}_{kk,\infty}$, $\ell \in \mathbb{Z}^+$, is provided in Appendix B.2. However, this expression requires an exhaustive search over the sum indices since they are not explicitly given. An exhaustive search is also required in the closed-form expression for the moments $m^{\ell}_{\boldsymbol{R}}$ in [77]. Therefore, the recursive approach is more practical.

The extension of the previous results to multipath fading is supported by the following theorem.

Theorem 7 *Let* $\boldsymbol{S}$ *be an* $N \times KL$ *random block matrix in* $\mathbb{C}$ *with blocks* $\boldsymbol{S}_k = (\boldsymbol{s}_{(k-1)L+1}, \boldsymbol{s}_{(k-1)L+2}, \dots \boldsymbol{s}_{kL})$, $1 \le k \le K$, *of size* $N \times L$. *The elements in a column vector* $\boldsymbol{s}_{(k-1)L+1}$ *are i.i.d. with zero mean, variance* $\frac{1}{N}$, *and* $\lim_{N\to\infty} \mathrm{E}\{N^3|s_{ij}|^6\} < +\infty$. *They are also i.i.d. from block to block. Within a block, the vector* $\boldsymbol{s}_{(k-1)L+s}$ *is* $\boldsymbol{s}_{(k-1)L+1}$ *cyclically down-shifted by* $s-1$ *positions. The empirical joint distribution of the received channel amplitudes* $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_K$ *converges to a limiting joint distribution with upper bounded support* $F_{\mathbf{a}}(a_1, a_2, \dots, a_L)$ *and the eigenvalue distribution of* $\boldsymbol{R} = \boldsymbol{A}^H\boldsymbol{S}^H\boldsymbol{S}\boldsymbol{A}$ *converges almost surely to a limiting distribution with bounded support*[8].

Additionally, let us assume that the joint probability density function $f_{\boldsymbol{a}}(a_1, a_2, \dots, a_L)$, *corresponding to the c.d.f.* $F_{\boldsymbol{a}}(a_1, a_2, \dots, a_L)$, *is an even function*[9] *of* $\mathrm{Re}(a_k)$ *and* $\mathrm{Im}(a_k)$ *for any* k *and for any value of the parameters* $(a_1, \dots, a_{k-1}, a_{k+1}, \dots, a_L)$. *Then, the following equivalences hold:*

[8] This assumption is of technical nature. Indeed, we conjecture that it follows from the nature of the support of $F_{\boldsymbol{a}}(a_1, \dots, a_L)$ and the statistics of the spreading sequences. Additionally, this condition can be substituted by a less restrictive condition. Applying Lemma 2 in its most general version, only the eigenvalue moments of $\boldsymbol{R}$ are required to be bounded.

[9] This condition is satisfied in the case of uncorrelated Rayleigh fading for example.

Equivalence-1 *The empirical eigenvalue distribution of $\boldsymbol{R}$ converges to the same limit as the eigenvalue distribution of the covariance matrix of a system with flat fading, received amplitude matrix $\widetilde{\boldsymbol{A}} = (\boldsymbol{A}^H\boldsymbol{A})^{\frac{1}{2}}$, and same system load β. The same property holds for the diagonal elements $R^{\ell}_{kk,\infty}$.*

Equivalence-2 *The empirical eigenvalue distribution of $\boldsymbol{R}$ converges to the same limit as the eigenvalue distribution of the covariance matrix $\overline{\boldsymbol{R}} = \boldsymbol{A}^H\overline{\boldsymbol{S}}^H\overline{\boldsymbol{S}}\boldsymbol{A}$ where $\overline{\boldsymbol{S}}$ is an $N \times LK$ matrix with all i.i.d. elements.*

The proof is in Appendix B.3.

Thanks to Equivalence 1 we can apply Algorithm 1 substituting $|a_{kk}|^2$ by $\boldsymbol{a}_k^H\boldsymbol{a}_k$.

An algorithm to compute the diagonal elements of $\boldsymbol{R}^{\ell}$, as $N, K \to \infty$ with $\frac{K}{N} \to \beta$, in the general case (no assumptions on $f_{\boldsymbol{a}}(a_1, \ldots, a_L)$) is formulated in Appendix B.3.

The conditions of Theorem 7 are verified when the channel gains are independent and Gaussian distributed. Therefore, Theorem 7 proves the conjecture of Equivalence-2 already in [65] for independent and Gaussian distributed channel gains. Additionally, the analysis in [65] can be extended to all multipath fading channels whose limit eigenvalue density functions satisfy the conditions of Theorem 7.

In order to derive the asymptotic weighting let us define the $M \times M$ matrices

$$\boldsymbol{\Phi}_k^{\infty} = \lim_{K=\beta N\to\infty} \boldsymbol{\Phi}_k \tag{3.29}$$

and

$$\boldsymbol{\Phi}^{\infty} = \lim_{K=\beta N\to\infty} \frac{\boldsymbol{\Phi}}{K}. \tag{3.30}$$

Their elements are $(\boldsymbol{\Phi}_k^{\infty})_{st} = R^{s+t}_{kk,\infty} + \sigma^2 R^{s+t-1}_{kk,\infty}$ and $(\boldsymbol{\Phi}^{\infty})_{st} = m_{\boldsymbol{R}}^{s+t} + \sigma^2 m_{\boldsymbol{R}}^{s+t-1}$ respectively. Additionally, let $\boldsymbol{\varphi}_k^{\infty}$ and $\boldsymbol{\varphi}^{\infty}$ be the M-dimensional vectors with respective elements $(\boldsymbol{\varphi}_k^{\infty})_s = R^{s}_{kk,\infty}$ and $(\boldsymbol{\varphi}^{\infty})_s = m_{\boldsymbol{R}}^{s}$.

The Type J-I detector with asymptotic weights is obtained using the weights that minimize (3.14) or, equivalently, maximize (3.18) as $K, N \to \infty$ with $\frac{K}{N} \to \infty$:

$$\boldsymbol{w}_k^{\infty} = (\boldsymbol{\Phi}_k^{\infty})^{-1}\boldsymbol{\varphi}_k^{\infty}, \tag{3.31}$$

where (3.7) was used.

The Type J-J detector is obtained as asymptotic approximation of the polynomial expansion detector with weights (3.11). The asymptotic weights of the Type J-J detector minimize the quantity

$$\begin{aligned}\lim_{K\to\infty}\sum_{k=1}^{K}\frac{\mathrm{MSE}_k}{K} &= \lim_{K=\beta N\to\infty}\frac{1}{K}\sum_{k=1}^{K}(\overline{\boldsymbol{w}}^T\boldsymbol{\Phi}\overline{\boldsymbol{w}}-2\mathrm{Re}(\boldsymbol{\varphi}_k^T\boldsymbol{w}_k)+1)\\ &= \boldsymbol{w}^T\boldsymbol{\Phi}^{\infty}\boldsymbol{w}-2\boldsymbol{w}^T\boldsymbol{\varphi}^{\infty}+1\end{aligned}$$

where (3.14) was used. This yields

$$\boldsymbol{w}^{\infty}=(\boldsymbol{\Phi}^{\infty})^{-1}\boldsymbol{\varphi}. \tag{3.32}$$

Let us consider the case when all received signals have the same power, i.e. $\boldsymbol{A}^H\boldsymbol{A}=\mathcal{P}\boldsymbol{I}$. Then, due to Theorem 6 and Corollary 1, $(\boldsymbol{R}^{\ell})_{kk}$ converges almost surely, as $K,N\to\infty$ with $\frac{K}{N}\to\beta$, to a value $R^{\ell}_{kk,infty}=R^{\ell}_{\infty}(\mathcal{P})$ that does not depend on the index k:

Corollary 2 *Let $\boldsymbol{S}$, $\boldsymbol{A}$, and $\boldsymbol{R}$ be as in Theorem 6. Additionally, let the matrix $\boldsymbol{A}$ be such that $\boldsymbol{A}^H\boldsymbol{A}=\mathcal{P}\boldsymbol{I}$. Then, for any $\ell,k\in\mathbb{Z}^+$, $(\boldsymbol{R}^{\ell})_{kk}$ converges almost surely, as $N,K\to\infty$ with $\frac{K}{N}$ constant, to the deterministic quantity*

$$(\boldsymbol{R}^{\ell})_{kk}\xrightarrow{a.s.}m^{\ell}_{\boldsymbol{R}}. \tag{3.33}$$

□

Corollary 2 ensures that $\boldsymbol{\Phi}_k^{\infty}=\boldsymbol{\Phi}^{\infty}$ and $\boldsymbol{\varphi}_k^{\infty}=\boldsymbol{\varphi}^{\infty}$ for $\boldsymbol{A}^H\boldsymbol{A}=\mathcal{P}\boldsymbol{I}$. Thus, Type J-J and Type J-I detectors coincide asymptotically in the equal power case. Additionally, in this case, the two asymptotic multistage detectors coincide also with the detector proposed in [18], which maximizes the ratio between the total useful power and the total noise and interference power at the detector output. For $\boldsymbol{A}^H\boldsymbol{A}=\mathcal{P}\boldsymbol{I}$, a closed-form expression of the eigenvalue moments is given by (2.31).

3.5 Asymptotic Performance Analysis

Let us consider a multistage detector for the k^{th} user using the basis (3.3) and weighting $\overline{\boldsymbol{w}}_k$. As $K,N\to\infty$ with $\frac{K}{N}\to\beta$, the MSE and the SINR are given by taking the limits of (3.14) and (3.18)

$$\mathrm{MSE}_k^{\infty} = 1-2\mathrm{Re}\left((\boldsymbol{\varphi}_k^{\infty})^T\overline{\boldsymbol{w}}_k\right)+\overline{\boldsymbol{w}}_k^T\boldsymbol{\Phi}_k^{\infty}\overline{\boldsymbol{w}}_k \tag{3.34}$$

$$\mathrm{SINR}_k^{\infty} = \frac{1}{\dfrac{\overline{\boldsymbol{w}}_k^H\boldsymbol{\Phi}_k^{\infty}\overline{\boldsymbol{w}}_k}{\overline{\boldsymbol{w}}_k^H\boldsymbol{\varphi}_k^{\infty}(\boldsymbol{\varphi}_k^{\infty})^T\overline{\boldsymbol{w}}_k}-1}. \tag{3.35}$$

Equation (3.35) can be immediately specialized to Type J-J and Type J-I detectors with (3.32) and (3.31) respectively:

$$\mathrm{SINR}^{\infty}_{\mathrm{J-J},k} = \frac{1}{\frac{(\boldsymbol{\varphi}^{\infty})^T(\boldsymbol{\Phi}^{\infty})^{-1}\boldsymbol{\Phi}^{\infty}_k(\boldsymbol{\Phi}^{\infty})^{-1}\boldsymbol{\varphi}^{\infty}}{\left((\boldsymbol{\varphi}^{\infty})^T(\boldsymbol{\Phi}^{\infty})^{-1}\boldsymbol{\varphi}^{\infty}_k\right)^2} - 1}, \tag{3.36}$$

$$\mathrm{SINR}^{\infty}_{\mathrm{J-I},k} = \frac{1}{\frac{1}{(\boldsymbol{\varphi}^{\infty}_k)^T(\boldsymbol{\Phi}^{\infty}_k)^{-1}\boldsymbol{\varphi}^{\infty}_k} - 1}. \tag{3.37}$$

They are the asymptotic limits of (3.23) and (3.20), respectively. In the asymptotic case, the performance depends only on the limiting distribution function $F_{\boldsymbol{a}}(a_1, \ldots, a_L)$, as well as on β and σ^2. If the conditions of Theorem 7 are fulfilled the performance depends on the eigenvalue moments of $\boldsymbol{A}^H\boldsymbol{A}$, $m_{|\boldsymbol{A}|^2}$, the received power of user k, $\mathcal{P}_k$, as well as on β and σ^2.

As shown in [28], the output SINRs of detector Type J-I and of detector Type I-I are proportional to the received power of user k, $\mathcal{P}_k$. Therefore, the multiuser efficiency[10]

$$\eta_k = \frac{\sigma^2}{\mathcal{P}_k}\mathrm{SINR}_k \tag{3.38}$$

is independent of $\mathcal{P}_k$ like for the linear MMSE detector (see Section 2.3.3). This is due to the fact that the filter coefficients of each user are optimum in the projection subspace. In contrast to many detectors analyzed in the literature, the SINR of detector Type J-J depends on $\mathcal{P}_k$ by a non-linear relation as can easily be verified by inspection. Therefore, the multiuser efficiency of this latter detector does depend on $\mathcal{P}_k$. This reflects the fact that the filter coefficients are optimized with respect to an ideal average user. The farther the users are away from the average the more sub-optimal is their detection. The system suffers from a sort of near-far effect that results in poorer performance for users with higher or lower received powers than the average.

A straightforward implication of the fact that the Type J-I detector outperforms the Type J-J detector (or equivalently the multistage Wiener filter outperforms the polynomial expansion detector) is that, for any user, the constant multiuser efficiency of detector Type J-I is an upper-bound for the multiuser efficiency of detector Type J-J (see Figure 3.4 in Section 3.6). In Section 3.6, this behavior is verified numerically.

Whereas Equations (3.34) and (3.35) provide asymptotic performance of multistage detectors, Equations (3.14) and (3.18) allow the performance evaluation of detectors Type J-J and Type J-I when they are used in real scenarios with finite

[10] We implicitly assume that the remaining noise and interference are Gaussian.

system size by setting $\overline{w}_k = \Phi^\infty \varphi^\infty$ and $\overline{w}_k = \Phi_k^\infty \varphi_k^\infty$, respectively. However, the performance for finite systems depends on the specific realizations of $\boldsymbol{A}$ and $\boldsymbol{S}$.

3.6 Numerical Results

Numerical results and simulations presented in this section were obtained using for each user a channel with an exponentially decaying power-delay-profile (PDP) with a decrease of 30 dB within the channel length $L = 15$ and block Rayleigh fading. Denoting the variances of the L taps of the PDP with $p_0, p_1, \dots p_{L-1}$, the ℓ^{th} tap of each channel is complex Gaussian distributed with variance p_ℓ. Then, the characteristic function of the eigenvalues of $\boldsymbol{A}^H\boldsymbol{A}$ is given by

$$\Phi_{|\boldsymbol{A}|^2}(j\omega) = \prod_{\ell=0}^{L-1} \frac{1}{1 - 2jp_\ell\omega}. \tag{3.39}$$

We calculate the eigenvalue moments from the relation $j^n m^n_{|\boldsymbol{A}|^2} = \left.\frac{\mathrm{d}^n \Phi_{|\boldsymbol{A}|^2}}{\mathrm{d}\omega^n}\right|_{\omega=0}$, the SINRs of Type J-J and Type J-I detectors in asymptotic conditions by (3.36) and (3.37), and the multiuser efficiency for the two detectors by (3.38). In Figure 3.3 the families of the curves $\eta^\infty_{\text{J-I}}$ versus $\frac{E_s}{N_0}$ parameterized by the system load β are plotted for $M = 2$ in dashed lines and for $M = 4$ in solid lines. The improvement in η obtained by increasing the number of stages is negligible for low $\frac{E_s}{N_0}$ and becomes more and more relevant for increasing $\frac{E_s}{N_0}$.

In Figure 3.4, the large system multiuser efficiency of Type J-J detectors (solid lines) is plotted as a function of P_k, the received power of the user of interest for different level of the background noise. In Figure 3.4, the multiuser efficiency of Type J-J detectors is also compared to multiuser efficiency of Type J-I detectors (dotted lines). As already mentioned, in contrast to many other detectors analyzed in the literature and in contrast to Type J-I detectors, the multiuser efficiency of Type J-J detectors depends on P_k. This dependence is stronger for low system loads and high SNR while it tends to vanish for systems heavily loaded and at low SNR. As shown analytically in Section 3.5, the constant multiuser efficiency of detector Type J-I provides an upper bound for the multiuser efficiency of detector Type J-J.

The performance degradation of both Type J-J and Type J-I detectors with asymptotic weighting compared to the polynomial expansion detector, the MSWF, and the full-rank linear MMSE detector were assessed by simulations. The simulations were performed using QPSK modulation, in the presence of multipath fading, and assuming perfect knowledge of the channel. Figure 3.5 shows the BER versus[11]

[11]Since we use a QPSK modulation and we focus on uncoded transmission, we have $\frac{E_b}{N_0} = \frac{E_s}{2N_0}$.

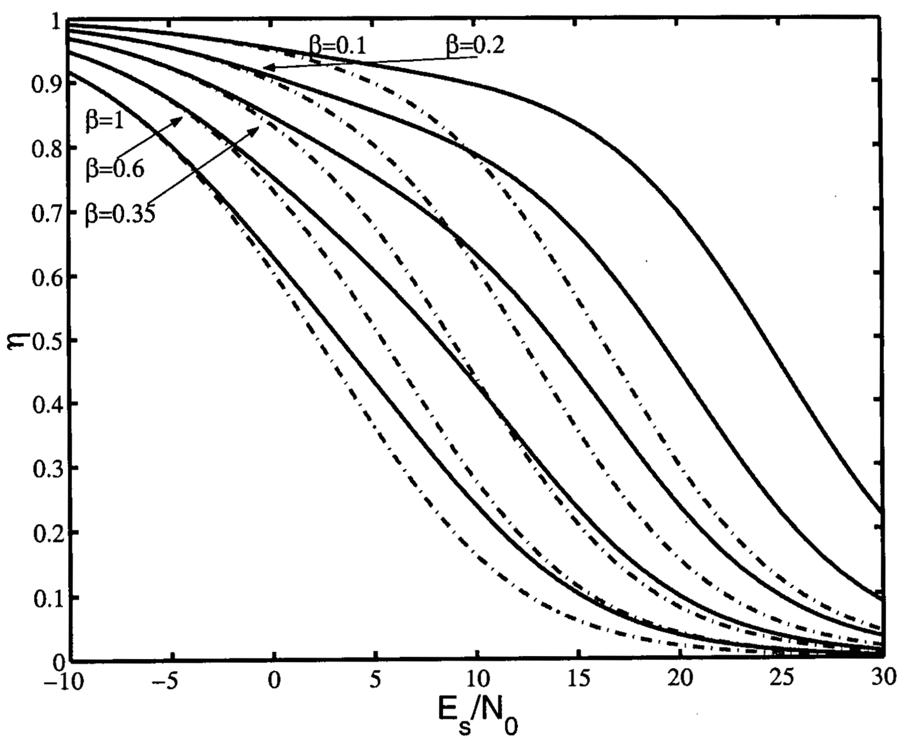


Figure 3.3: Multiuser efficiency η versus SNR for Type J-I detector with $M = 4$ (solid line) and $M = 2$ (dashed line). Frequency selective fading with exponentially decaying PDP and channel length $L = 15$.

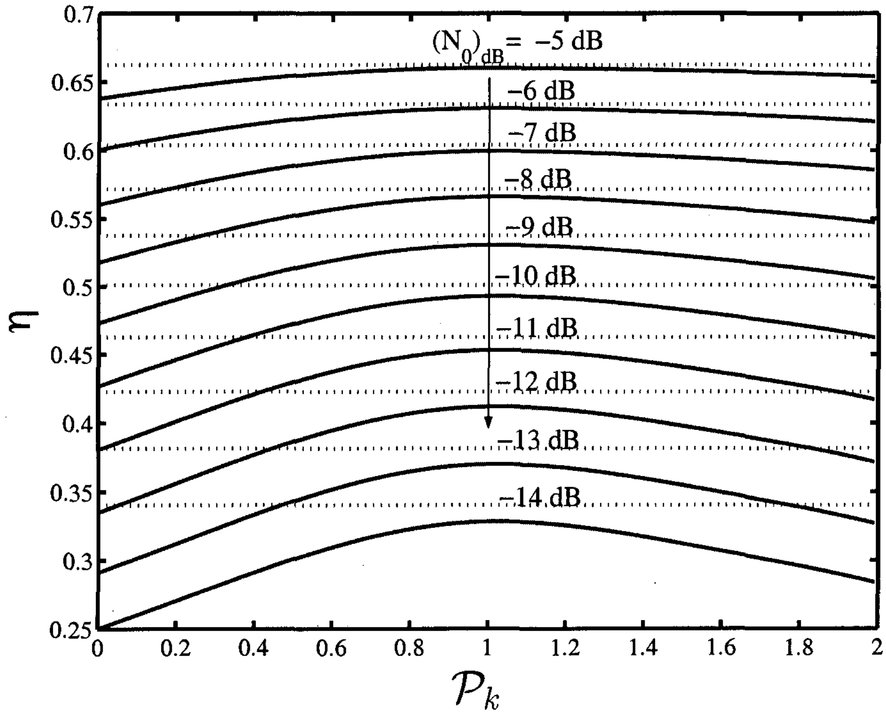


Figure 3.4: Multiuser efficiency versus power of the user of interest $\mathcal{P}_k$ for Type J-J detector (solid line) and Type J-I detector (dotted line). Frequency selective fading with exponentially decaying PDP and $L = 15$. Parameter setting: $M = 4$, $\beta = 0.5$, and $\mathrm{E}\{\mathcal{P}_k\} = 1$.

$\frac{E_b}{N_0}$ for multistage detectors with $M = 4$, $\beta = 0.5$, and $N = 128$[12]. The performance degradation due to the asymptotic approximation of weights is negligible in situations of practical interest. In fact, for $N = 128$ the curves of Type J-J and Type J-I detectors with asymptotic weights almost match the corresponding detectors with exact weighting. Using the same exponentially decaying PDP for all users, the performance degradation of the Type J-J detector with respect to the Type J-I detector is negligible for small $\frac{E_b}{N_0}$ and becomes relevant at larger $\frac{E_b}{N_0}$, as expected from the theoretic performance in Figure 3.4.

Figure 3.6 shows how the performance of the Type J-I detector with asymptotic weights increases for an increasing number of stages.

The effects of the mismatch between asymptotic and exact weights has been analyzed assuming flat fading channels, QPSK modulation, and $\beta = \frac{1}{2}$. Figure 3.7 compares $\mathrm{BER}_{\mathrm{J-I}}$, the BER at the output of a Type J-I detector, with $\mathrm{BER}_{\mathrm{MSWF}}$,

[12]This value of the spreading factor is in use in the UMTS FDD mode. This has motivated its choice.

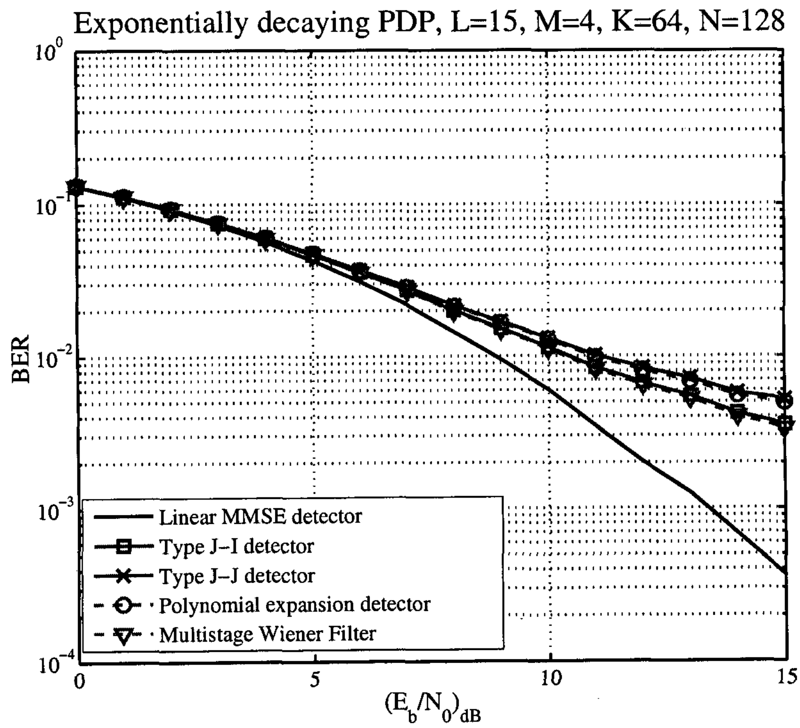


Figure 3.5: BER versus $\frac{E_b}{N_0}$ for $\beta = 0.5$.

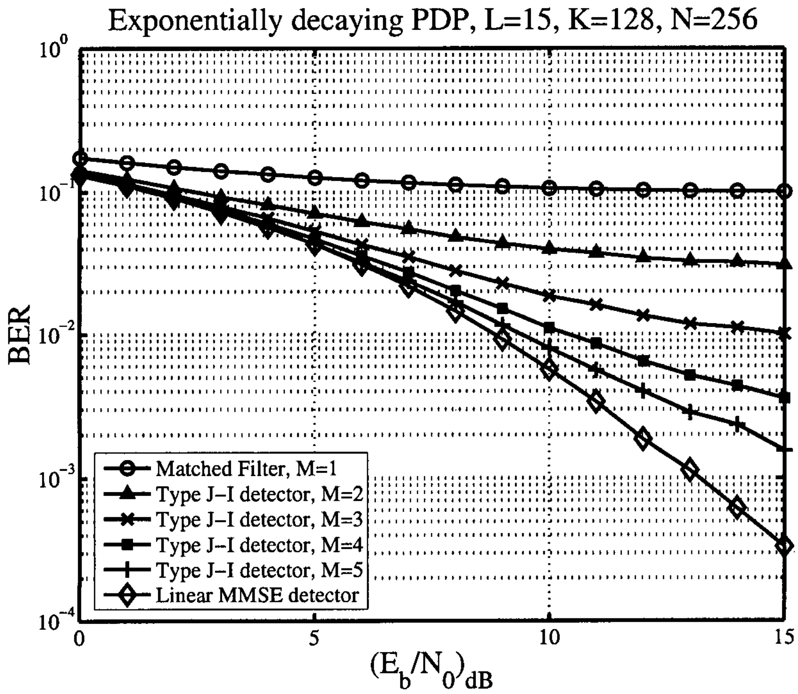


Figure 3.6: BER versus $\frac{E_b}{N_0}$ for $\beta = 0.5$ and varying number of stages.

the BER at the output of a MSWF, for $M = 4$ stages and increasing spreading factors. Performance degradation becomes visible when $\frac{E_b}{N_0}$ increases. However, this mismatch at high $\frac{E_b}{N_0}$ becomes rapidly irrelevant when the spreading factor is increased. Figure 3.8 plots $\epsilon = \frac{\mathrm{BER}_{\mathrm{J-I}} - \mathrm{BER}_{\mathrm{MSWF}}}{\mathrm{BER}_{\mathrm{MSWF}}}$, the relative mismatch of the BER at the output of the two compared detectors, as a function of $\frac{E_b}{N_0}$. The logarithmic scale allows to make visible also the mismatch at low $\frac{E_b}{N_0}$, otherwise not visible.

3.7 Conclusions

In this chapter we identified a general framework that is able to catch the main features of multistage detectors with asymptotic weights in terms of performance and complexity. Both the projection onto the Krylov subspace and the filtering can be performed jointly for all users or individually for each single user. The type of projection affects essentially the complexity while the type of filtering has an impact on the performance.

Considerations on the projection showed that only a joint projection can decrease the complexity order per bit from quadratic to linear.

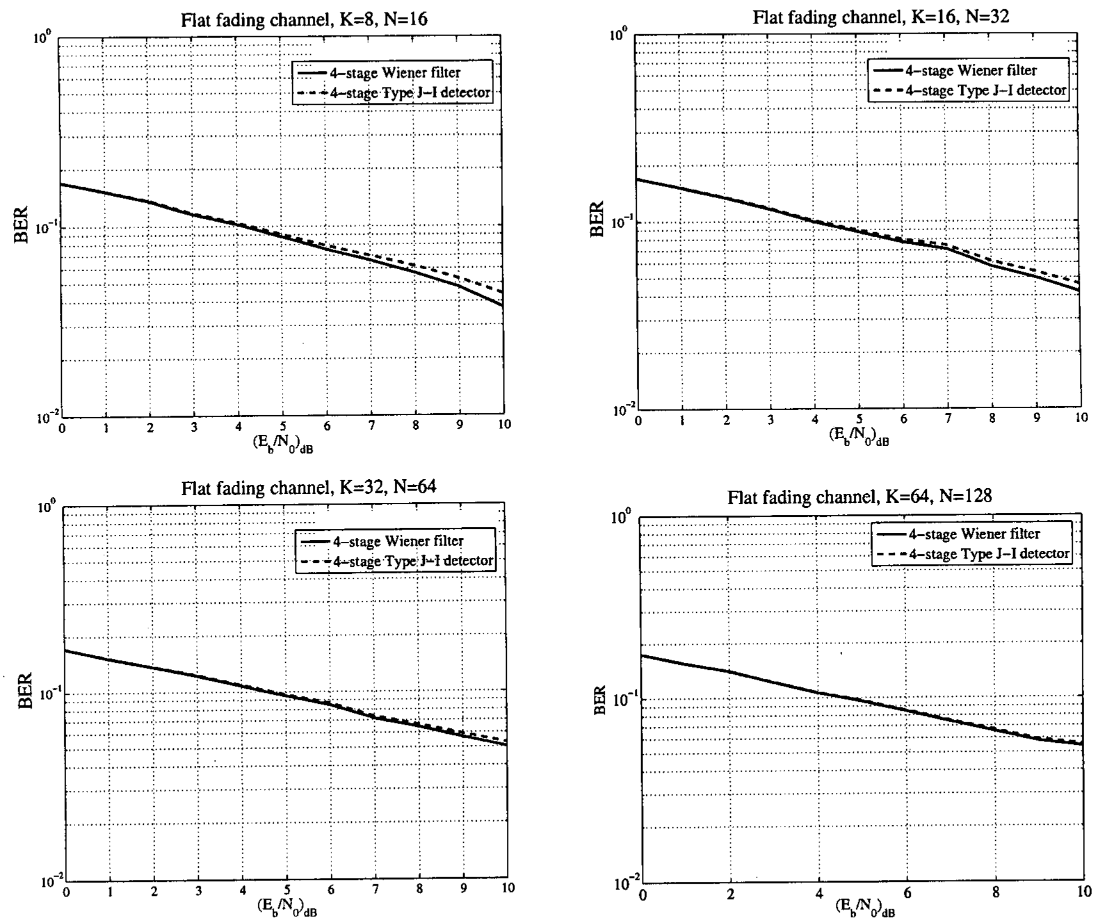


Figure 3.7: Comparison between $\mathrm{BER}_{\mathrm{J-I}}$ and $\mathrm{BER}_{\mathrm{MSWF}}$ for $M = 4$ stages, increasing spreading factors, $N = 16$, 32, 64, 128, and constant system load $\beta = \frac{1}{2}$.

of a single user is required.

For low complexity multiuser detection with Type J-I detectors we use statistics (A).

The general result for the asymptotic analysis and design of multiuser detectors provided in this chapter can be useful for the optimization of chip-pulse waveforms. However, this application is beyond the scope of this work.

Chapter 4 is organized in six sections. Section 4.2 introduces the general system model for asynchronous systems. Section 4.3 is focused on the analysis of a wide class of linear detectors and the design of low complexity multistage detectors for chip synchronous but symbol asynchronous systems. A detector structure with sliding observation window that does not suffer from windowing effects is proposed. Chip asynchronous but symbol quasi synchronous CDMA systems are studied in Section 4.4, and the effects of the chip pulse waveforms and the time delay distribution on the system performance are investigated. In Section 4.5, the results of Section 4.3 and Section 4.4 are applied to the design and analysis of totally asynchronous CDMA systems. Some conclusions are drawn in Section 4.6.

4.2 General System Model

Let us consider an asynchronous CDMA system with K users in the uplink channel. Each user and the base station are equipped with a single antenna. The channel is flat fading and impaired by additive white Gaussian noise. Then, the signal received at the base station, in complex base-band notation, is given by

$$y(t) = \sum_{k=1}^{K} a_{kk} s_k(t-\tau_k) + n(t) \qquad t \in [-\infty, +\infty].$$

Here, a_{kk} is the received signal amplitude of user k, which takes into account the transmitted amplitude, the effects of the flat fading, and the carrier phase offset; τ_k is the time delay of user k; $n(t)$ is a zero mean white, complex Gaussian process with two-sided power spectral density N_0; and $s_k(t)$ is the spread signal of user k. We have

$$s_k(t) = \sum_{m=-\infty}^{+\infty} b_k[m] c_k^{(m)}(t),$$

where $b_k[m]$ is the m^{th} transmitted symbol of user k and

$$c_k^{(m)}(t) = \sum_{u=0}^{N-1} s_{k,m}[u] \psi(t - mT_s - uT_c)$$

is its spreading waveform at time m. Here, $s_{k,m}$, is the spreading sequence of user k in the m^{th} symbol interval[9] with elements $s_{k,m}[u]$, $u = 0, \ldots, N-1$, T_s and $T_c = \frac{T_s}{N}$ are the symbol and chip periods, respectively.

The users' symbols $b_k[m]$ are uncorrelated and identically distributed random variables with $\mathrm{E}\{|b_k[m]|^2\} = 1$ and $\mathrm{E}\{b_k[m]\} = 0$. The elements of the spreading sequences $s_{k,m}[u]$ are assumed to be i.i.d. random variables with $\mathrm{E}\{|s_{k,m}[u]|^2\} = \frac{1}{N}$ and $\mathrm{E}\{s_{k,m}[u]\} = 0$. This assumption properly models the spreading sequences of some CDMA systems currently in use, such as the long spreading codes of the FDD (Frequency Division Duplex) mode in a UMTS uplink channel.

The chip waveform $\psi(t)$ is bandlimited with bandwidth B and energy $E_\psi = \int_{-\infty}^{+\infty} |\psi(t)|^2 \mathrm{d}t$. Because of the the constraint on the variance of the chips, i.e., $\mathrm{E}\{|s_{k,m}[k]|^2\} = \frac{1}{N}$, the mean energy of the signature waveform satisfies $\mathrm{E}\left\{\int_{-\infty}^{+\infty} |c_k^{(m)}(t)|^2 \mathrm{d}t\right\} = E_\psi$. Without loss of generality we can assume (i) user 1 as reference user so that $\tau_1 = 0$, (ii) the time delay to be, at most, one symbol interval so that $\tau_k \in [0, T_s)$.[10]

The front-end of the multiuser detector performs:

- A lowpass filtering with lowpass band $|f| \leq \frac{r}{2T_c}$ where $r \in \mathbb{Z}^+$ satisfies the constraint $B \leq \frac{r}{2T_c}$ so that condition of the sampling theorem is satisfied. The filter is normalized to obtain an overall amplification factor for the information bearing signal equal to one, i.e., the frequency response of the lowpass filter is

$$G(f) = \begin{cases} \frac{1}{\sqrt{E_\psi}} & |f| \leq \frac{r}{2T_c} \\ 0 & |f| > \frac{r}{2T_c}. \end{cases}$$

- A subsequent continuous-discrete time conversion by conventional sampling at rate $\frac{r}{T_c}$.

With this choice of the front-end, the conditions of the sampling theorem are satisfied so that the sampled signal provides sufficient statistics and the chip rate is a multiple of the sampling rate. Additionally, the discrete-time noise is still white with zero mean and variance $\sigma^2 = \frac{N_0 r}{E_\psi T_c}$.

The discrete-time signal at the front-end output is given by

$$y[p] = \sum_{k=1}^{K} a_k \sum_{m=-\infty}^{+\infty} b_k[m] \tilde{c}_k^{(m)} \left(\frac{p}{r} T_c - \tau_k\right) + n[p] \tag{4.1}$$

[9] The spreading sequence of user k possibly varies from symbol to symbol. This model is general and enables a proper description also of the spreading sequences of some CDMA systems currently in use such as the long spreading codes of the FDD (Frequency Division Duplex) mode in a UMTS uplink channel.

[10] For a thorough discussion on this assumption the reader can refer to [16].

with $p \in \mathbb{Z}$ and

$$\widetilde{c}_k^{(m)} = \sum_{u=0}^{N-1} s_{k,m}[u]\widetilde{\psi}(t-(u+mN)T_c)\,.$$

Here, $n[p]$ is discrete-time, complex-valued, zero mean, white Gaussian noise with variance $\sigma^2 = \frac{N_0 r}{E_\psi T_c}$ and $\widetilde{\psi}(t)$ is the pulse shape $\psi(t)$ normalized to have unit energy, i.e. $\widetilde{\psi}(t) = \frac{\psi(t)}{\sqrt{E_\psi}}$.

4.3 Symbol Asynchronous but Chip Synchronous CDMA Systems

4.3.1 System Model

In this section we focus on the analysis of symbol asynchronous but chip synchronous systems, i.e. we assume that the time delays τ_k are integer multiples of T_c, as in [81] and [80]. Without loss of generality we assume that the users are ordered according to increasing time delay with respect to the reference user. Let $r_k \triangleq \frac{\tau_k}{T_c} \in [0, 1, \ldots, N-1]$ be the time delay normalized to the chip period. Under the additional assumptions of chip synchronicity, use of sinc chip pulses, and sampling at the chip rate[11], $\widetilde{c}_k^{(m)}(qT_c - \tau_k)$, $q \in \mathbb{Z}$ simplifies to

$$\begin{aligned}
\widetilde{c}_k^{(m)}(qT_c-\tau_k) &= \sum_{u=0}^{N-1} s_{k,m}[u]\mathrm{sinc}(q-u-mN-r_k)\\
&= \sum_{u=0}^{N-1} s_{k,m}[u]\delta(q-u-mN-r_k)\\
&= \begin{cases} s_{k,m}[q-mN-r_k] & \text{for } mN+r_k \leq q \leq (m+1)N+r_k-1,\\ 0 & \text{otherwise.}\end{cases}
\end{aligned}$$

Then, the received signal from user k at the time instant $qT_c = (p+mN)T_c$ with $p \in 0, \ldots, N-1$, is given by

$$\sum_{m=-\infty}^{+\infty} b_k[m]\widetilde{c}_k^{(m)}((p+mN)T_c-\tau_k) = \begin{cases} b_k[m-1]s_{k,m-1}[N+p-r_k] & \text{if } p < r_k,\\ b_k[m]s_{k,m}[p-r_k] & \text{if } p \geq r_k.\end{cases}$$

[11] The results presented in this section hold also for CDMA systems transmitting square root Nyquist pulses and using as front-end an analog filter matched to the chip pulse followed by a sampler at the chip rate. The constraint on the sinc chip pulse is imposed here only to be consistent with the general system model (4.1) and the choice of the front-end made in Section 4.2.

Finally, the general system model (4.1) simplifies to

$$y[p+mN] = \sum_{k=1}^{K} a_{kk} d_{k,m}[p] + n[p+mN], \qquad p = 0, \ldots, N-1, \ m \in \mathbb{Z} \tag{4.2}$$

where $y[p+mN]$ is the sampled received signal at time instants $(p+mN)T_c$, and $n[p+mN]$ is the complex-valued white Gaussian noise with zero mean and variance $\sigma^2 = \frac{N_0}{E_s T_c}$. $d_{k,m}[p]$ is the received signal at time instant $p+mN$ from user k and it is given by

$$d_{k,m}[p] = \begin{cases} b_k[m-1] s_{k,m-1}[N+p-r_k] & \text{if } p < r_k \\ b_k[m] s_{k,m}[p-r_k] & \text{if } p \geq r_k \end{cases} \quad \text{and} \quad p \in [0, \ldots, N-1].$$

This definition reflects the fact that, if $p < r_k$ the receiver is still receiving the $(m-1)^{\text{st}}$ symbol of user k and $d_{k,m}[p]$ depends on $b_k[m-1]$ and the spreading sequence $\boldsymbol{s}_{k,m-1}$ at time instant $m-1$. Otherwise $d_{k,m}[p]$ depends on $b_k[m]$ and the spreading sequence $\boldsymbol{s}_{k,m}$.

For the following developments it is convenient to rewrite the system model (4.2) in matrix notation. Let $\boldsymbol{y}(m) = (y[mN], y[1+mN], \ldots, y[(m+1)N-1])^T$ and $\boldsymbol{b}(m) = (b_1[m], b_2[m], \ldots, b_K[m])^T$ be the vector of the observed signal and the vector of the transmitted symbols in the m^{th} symbol interval, respectively. We denote by $\boldsymbol{S}_u(m)$ the $N \times K$ matrix containing columnwise and appropriately shifted the parts of the spreading sequences $\boldsymbol{s}_{k,m}$ received in the m^{th} symbol interval. Similarly, $\boldsymbol{S}_d(m)$ is the $N \times K$ matrix containing columnwise and appropriately shifted the parts of the spreading sequences $\boldsymbol{s}_{k,m}$ received in the $(m+1)^{\text{st}}$ symbol interval. More specifically, the elements of $\boldsymbol{S}_u(m)$ and $\boldsymbol{S}_d(m)$ are given by

$$(\boldsymbol{S}_u(m))_{ij} = \begin{cases} 0 & i-1 < r_j, \\ s_{j,m}[i-1-r_j] & i-1 \geq r_j, \end{cases}$$

$$(\boldsymbol{S}_d(m))_{ij} = \begin{cases} s_{j,m}[N+i-1-r_j] & i \leq r_j, \\ 0 & i > r_j. \end{cases}$$

In this chapter, with a slight abuse of notation, we denote the matrix of the spreading sequences for asynchronous systems with $\boldsymbol{S}(m) = \begin{bmatrix} \boldsymbol{S}_u(m) \\ \boldsymbol{S}_d(m) \end{bmatrix}$ and the matrix of spreading sequences for synchronous systems[12] with $\boldsymbol{S}'(m) \in \mathbb{C}^{N \times K}$. More intuitively the matrix $\boldsymbol{S}(m) \in \mathbb{C}^{2N \times K}$ can be obtained by a vertical concatenation of

[12]In the previous chapter $\boldsymbol{S}$ was used to denote the matrix of the spreading sequences for synchronous systems.

$\boldsymbol{S}'(m)$ with an $N \times K$ matrix of zeros and by a subsequent cyclic down-shift of the k^{th} column by r_k positions $\forall k$. $\boldsymbol{S}_u(m)$ and $\boldsymbol{S}_d(m)$ are the upper and lower block of the matrix $\boldsymbol{S}(m)$, respectively. Figure 4.1 illustrates the structure of the matrix $\boldsymbol{S}(m)$.

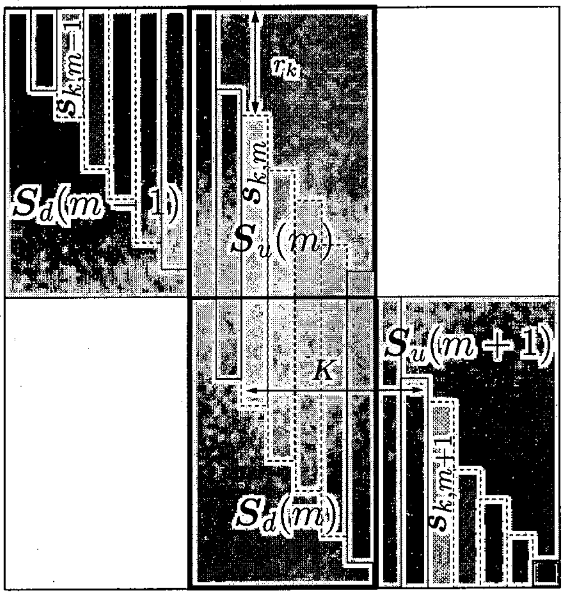


Figure 4.1: Graphical representation of the matrix $\boldsymbol{S}(m)$ (bold frame) and its blocks $\boldsymbol{S}_u(m)$ and $\boldsymbol{S}_d(m)$. The vertical bars represent the spreading sequences $\boldsymbol{s}_{k,m}$ shifted by r_k elements. The down-shift of the k^{th} spreading sequence, r_k, equals the time delay of user k normalized by the chip period T_c. The element of the matrix not covered by bars are zero.

Let $\boldsymbol{A} = \text{diag}(a_{1,1}, a_{2,2}, \ldots a_{K,K})$ denote the $K \times K$ diagonal matrix of complex received amplitudes. Furthermore, let $\boldsymbol{H}(m) = \boldsymbol{S}(m)\boldsymbol{A}$ and, consistently with the definitions of $\boldsymbol{S}_u(m)$ and $\boldsymbol{S}_d(m)$, $\boldsymbol{H}_u(m) = \boldsymbol{S}_u(m)\boldsymbol{A}$ and $\boldsymbol{H}_d(m) = \boldsymbol{S}_d(m)\boldsymbol{A}$. Then, the baseband discrete-time asynchronous system in the uplink is described by

$$\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{B}} + \boldsymbol{\mathcal{N}} \tag{4.3}$$

where $\boldsymbol{\mathcal{Y}} = [\ldots, \boldsymbol{y}^T(m-1), \boldsymbol{y}^T(m), \boldsymbol{y}^T(m+1)\ldots]^T$ and $\boldsymbol{\mathcal{B}} = [\ldots, \boldsymbol{b}^T(m-1), \boldsymbol{b}^T(m), \boldsymbol{b}^T(m+1)\ldots]^T$ are the infinite-length vectors of received and transmitted symbols respectively; $\boldsymbol{\mathcal{N}}$ is an infinite-length vector of white Gaussian noise with variance σ^2; and $\boldsymbol{\mathcal{H}}$ is a bi-diagonal block matrix with infinite block rows and block columns that is given by

$$\mathcal{H} = \begin{bmatrix} \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \cdots & \mathbf{0} & \boldsymbol{H}_d(m-2) & \boldsymbol{H}_u(m-1) & \mathbf{0} & \cdots & \cdots & \cdots \\ \cdots & \cdots & \mathbf{0} & \boldsymbol{H}_d(m-1) & \boldsymbol{H}_u(m) & \mathbf{0} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \mathbf{0} & \boldsymbol{H}_d(m) & \boldsymbol{H}_u(m+1) & \mathbf{0} & \cdots \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix} \tag{4.4}$$

for $m \in \mathbb{Z}$. This matrix models the effects of the spreading sequences, of the received amplitudes, and of time delays.

We will also consider the system corresponding to a finite observation window of length W symbols centered at the m^{th} transmitted symbol of the reference user. In order to keep the notation simple we assume W to be an odd integer. However, the results presented in the following hold for any rational number W such that $WN \in \mathbb{Z}$. The system model has the following form:

$$\mathcal{Y}_{N,W}(m) = \mathcal{H}_{N,W}(m)\mathcal{B}_{N,W}(m) + \mathcal{N}_{N,W}(m) \tag{4.5}$$

with $\mathcal{Y}_{N,W}(m)=\begin{bmatrix} \boldsymbol{y}(m-\frac{W-1}{2}) \\ \vdots \\ \boldsymbol{y}(m) \\ \vdots \\ \boldsymbol{y}(m+\frac{W-1}{2}) \end{bmatrix}$; $\mathcal{B}_{N,W}(m)=\begin{bmatrix} \boldsymbol{b}(m-\frac{W+1}{2}) \\ \vdots \\ \boldsymbol{b}(m) \\ \vdots \\ \boldsymbol{b}(m+\frac{W-1}{2}) \end{bmatrix}$; $\mathcal{N}_{N,W}(m)=\begin{bmatrix} \boldsymbol{n}(m-\frac{W-1}{2}) \\ \vdots \\ \boldsymbol{n}(m) \\ \vdots \\ \boldsymbol{n}(m+\frac{W-1}{2}) \end{bmatrix}$;

and

$$\mathcal{H}_{N,W}(m) = \begin{bmatrix} \boldsymbol{H}_d(m-\frac{W+1}{2}) & \boldsymbol{H}_u(m-\frac{W-1}{2}) & \mathbf{0} & \cdots \\ \mathbf{0} & \boldsymbol{H}_d(m-\frac{W-1}{2}) & \boldsymbol{H}_u(m-\frac{W+1}{2}) & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ \cdots & \mathbf{0} & \boldsymbol{H}_d(m+\frac{W-3}{2}) & \boldsymbol{H}_u(m+\frac{W-1}{2}) \end{bmatrix}.$$

Throughout this section we adopt the following notation:

- $\boldsymbol{\hbar}(k,m)$ is the column of $\mathcal{H}$ corresponding to user k at the symbol interval m, i.e., it is the infinite-length column of $\mathcal{H}$ containing the k^{th} column of the matrix $\boldsymbol{H}(m)$;
- $\boldsymbol{\hbar}(s:r,m)$ denotes the matrix made up of the $s-r+1$ columns of $\mathcal{H}$, $\boldsymbol{\hbar}(s,m), \boldsymbol{\hbar}(s+1,m), \ldots \boldsymbol{\hbar}(r,m)$;
- $\mathcal{T} = \mathcal{H}\mathcal{H}^H$;
- $\mathcal{R} = \mathcal{H}^H\mathcal{H}$;

- $\mathcal{T}_{N,W}(m) = \mathcal{H}_{N,W}(m)\mathcal{H}_{N,W}^{H}(m)$;
- $\mathcal{R}_{N,W}(m) = \mathcal{H}_{N,W}^{H}(m)\mathcal{H}_{N,W}(m)$.

For $K \to \infty$, the sequence of the empirical eigenvalue distributions of $\boldsymbol{A}\boldsymbol{A}^H$ converges almost surely to $F_{|\boldsymbol{A}|^2}(\lambda)$, a non-random distribution function with upper bounded support. Let $\rho^{(N)} = \frac{\tau_k}{T_s} = \frac{\tau_k}{N}$ denote the random variable of the time delays normalized by T_s. For a given N, $F_P^{(N)}(\rho)$ denotes the empirical distribution function of ρ. We assume that, as $N \to \infty$, the sequence $\{F_P^{(N)}(\rho)\}$ converges to the probability distribution function $F_P(\rho)$. For the sake of generality, we do not assume that the random variables of the received powers λ and time delays ρ are independent; their joint probability distribution function will be denoted by $F_{|\boldsymbol{A}|^2,P}(\lambda,\rho)$. This model is appropriate for the design of coherent detectors since both the time delays τ_k, $k = 1, \dots, K$ and the received powers $|a_{k,k}|^2$, $k = 1, \dots, K$ are known, i.e. are deterministic. In this case, given the system load $\beta = \frac{K}{N}$,

$$F_{|\boldsymbol{A}|^2,P}(\lambda,\rho) = \frac{1}{K}\sum_{k=1}^{K} 1(\lambda - |a_{k,k}|^2, \rho - \frac{\tau_k}{T_s})$$

where $1(x, y)$ is the bi-dimensional indicator function on right unbounded intervals.

4.3.2 Linear MMSE Detection

For a given observation window of length W and symmetric around the m^{th} symbol interval the linear MMSE estimator of $\mathcal{B}_{N,W}(m)$ is given by

$$\begin{aligned}\widehat{\mathcal{B}}_{N,W}(m) &= \mathcal{H}_{N,W}^{H}(m)[\mathcal{T}_{N,W}(m) + \sigma^2 \boldsymbol{I}_{N,W}]^{-1}\mathcal{Y}_{N,W}(m)\\ &= [\mathcal{R}_{N,W}(m) + \sigma^2\boldsymbol{I}_{N,W}]^{-1}\mathcal{H}_{N,W}^{H}(m)\mathcal{Y}_{N,W}(m).\end{aligned}$$

By applying the multistage decomposition of the Wiener filter proposed in [27] and discussed in Section 2.2.6 the linear MMSE estimator can be rewritten as follows

$$\widehat{\mathcal{B}}_{N,W}(m) = \sum_{\ell=0}^{K(W+1)-1} \mathcal{W}_{N,W,\ell}^{(K(W+1)-1)}(m)\mathcal{R}_{N,W}^{\ell}(m)\mathcal{H}_{N,W}^{H}(m)\mathcal{Y}_{N,W}(m). \qquad (4.6)$$

Note that (4.6) coincides with the Type J-I detector with $K(W+1)-1$ stages introduced in Section 3.3.3 (see (3.10)). $\mathcal{W}_{N,W,\ell}^{(K(W+1)-1)}(m)$ are diagonal matrices whose j^{th} diagonal elements are obtained as solution of the Yule-Walker system of equations:

$$\boldsymbol{w}_j(m) = (\boldsymbol{\Phi}_{N,W,j}^{(K(W+1)-1)}(m))^{-1}\boldsymbol{\varphi}_{N,W,j}^{(K(W+1)-1)}(m)$$

and $\boldsymbol{w}_j(m) = ((\mathcal{W}_{N,W,0}^{(K(W+1)-1)}(m))_j, (\mathcal{W}_{N,W,1}^{(K(W+1)-1)}(m))_j, \dots, (\mathcal{W}_{N,W,K(W+1)-1}^{(K(W+1)-1)}(m))_j)$. For $s \in \mathbb{Z}^+$, $\boldsymbol{\Phi}_{N,W,j}^{(s)}(m)$ is an $s \times s$ real matrix with elements $(\boldsymbol{\Phi}_{N,W,j}^{(s)}(m))_{u,v} =$

$(\boldsymbol{\mathcal{R}}_{N,W}^{u+v}(m))_{jj} + \sigma^2(\boldsymbol{\mathcal{R}}_{N,W}^{u+v-1}(m))_{jj}$, and $\boldsymbol{\varphi}_{N,W,j}^{(s)}(m)$ is an s-dimensional column vector with elements $(\boldsymbol{\varphi}_{N,W}^{(s)}(m))_u = (\boldsymbol{\mathcal{R}}_{N,W}^{u}(m))_{jj}$. Let us introduce the quantities

$$\mathrm{SINR}_k^{(s)} = \frac{(\boldsymbol{\varphi}_{N,W,k}^{(s)}(m))^H(\boldsymbol{\Phi}_{N,W,k}^{(s)}(m))^{-1}\boldsymbol{\varphi}_{N,W,k}^{(s)}(m)}{1-((\boldsymbol{\varphi}_{N,W,k}^{(s)}(m))^H(\boldsymbol{\Phi}_{N,W,k}^{(s)}(m))^{-1}\boldsymbol{\varphi}_{N,W,k}^{(s)}(m))} \qquad s \in \mathbb{Z}^+. \tag{4.7}$$

For $s \geq K(W+1)-1$, $\mathrm{SINR}_k^{(K(W+1)-1)}$ is the SINR of the k^{th} transmitted symbol in $\boldsymbol{\mathcal{B}}_{N,W}(m)$ at the output of the linear MMSE detector for asynchronous systems[13]. For $s < K(W+1)-1$, $\mathrm{SINR}_k^{(s)}$ is the SINR of the k^{th} transmitted symbol at the output of an s-stage Wiener filter. Therefore, $\mathrm{SINR}_k^{(s)}$, $s = 1, 2, \ldots$, provides a family of lower bounds for the SINR of the k^{th} transmitted symbol in $\boldsymbol{\mathcal{B}}_{N,W}(m)$ at the output of the full rank linear MMSE detector. For synchronous CDMA systems the output SINR of reduced rank multistage filters converges exponentially in the filter rank toward the output SINR of a full rank linear MMSE filter [39]. Throughout this section we utilize this property and we analyze the performance of multistage detectors with finite observation windows and increasing number of stages. In Section 4.3.3 we verify numerically that, also for asynchronous systems, for $s \geq 8$ the lower bounds $\mathrm{SINR}_k^{(s)}$ are so close to the supremum to be indistinguishable from it.

Making use of (4.7), the problem of determining the family of lower bounds reduces to determining the diagonal elements of the matrix $\boldsymbol{\mathcal{R}}_{N,W}^{\ell}(m)$, $\ell = 1, \ldots, 2s$, as $K = \beta N \to \infty$. Recursive algorithms to determine $\lim_{K=\beta N \to \infty}(\boldsymbol{\mathcal{R}}_{N,W}^{\ell}(m))_{jj}$, $\ell \in \mathbb{Z}^+$, are provided in Theorem 8 and Theorem 9 for equal and unequal received powers, respectively.

In Theorem 8 we consider an asynchronous system with finite observation window length W and equal received powers. Thus, without loss of generality, we can assume that $\boldsymbol{A} = \boldsymbol{I}$ in (4.5). The assumptions in Theorem 8 summarize and formalize the characteristics of the system model (4.5) introduced in Section 4.3.1.

Theorem 8 *Let $\{\rho_j^{(N)}\}$ be a non decreasing sequence of elements in $\{0, \frac{1}{N}, \ldots, \frac{N-1}{N}, \}$, obtained by sorting K independent realizations of a random variable and let $F_{P^{(N)}}(\rho^{(N)})$ be its empirical distribution function (e.d.f.)[14]. Let $\widehat{h}_{ij}^{(N)}(\ell)$, $i, j = 1, 2, \ldots$ and $\ell = 1, \ldots, W+1$ be random variables[15] in $\mathbb{C}$. Let the matrix $\boldsymbol{H}^{(N)}(\ell) \in \mathbb{C}^{2N \times K}$, $\ell = 1, \ldots, W+1$ have elements $h_{ij}^{(N)}(\ell) = \widehat{h}_{ij}^{(N)}(\ell)$, for $\rho_j^{(N)}N + 1 \leq i \leq \rho_j^{(N)}N + N$, $1 \leq j \leq K$, and the remaining elements equal to*

[13]The equivalence of a linear MMSE detector and a full-rank multistage Wiener filter has been thoroughly discussed in Chapter 3

[14]$\rho_j^{(N)}$ models the time delays, normalized to the chip interval, of user j. The order reflects the ordering of the users assumed in the system models (4.3) and (4.5).

[15]The random variable $\widehat{h}_{ij}^{(N)}(\ell)$ models an element of the spreading sequence of user j at the symbol period ℓ.

zero[16]. *Let* $\boldsymbol{H}_d^{(N)}(\ell), \boldsymbol{H}_u^{(N)}(\ell) \in \mathbb{C}^{N\times K}$ *be, respectively, the lower and upper block of the matrix* $\boldsymbol{H}(\ell)^{(N)}$, *i.e.* $\boldsymbol{H}^{(N)}(\ell) = [(\boldsymbol{H}_u^{(N)}(\ell))^T, (\boldsymbol{H}_d^{(N)}(\ell))^T]^T$. *Finally, let* $\boldsymbol{\mathcal{H}}_{N,W}$ *be a* $WN \times (W+1)K$ *random bi-diagonal block matrix structured as follows:*

$$\boldsymbol{\mathcal{H}}_{N,W}=\begin{bmatrix} \boldsymbol{H}_d^{(N)}(1) & \boldsymbol{H}_u^{(N)}(2) & \mathbf{0} & \cdots & \cdots & \cdots \\ \mathbf{0} & \boldsymbol{H}_d^{(N)}(2) & \boldsymbol{H}_u^{(N)}(3) & \mathbf{0} & \cdots & \cdots \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \cdots & \cdots & \mathbf{0} & \boldsymbol{H}_d^{(N)}(W{-}1) & \boldsymbol{H}_u^{(N)}(W) & \mathbf{0} \\ \cdots & \cdots & \cdots & \mathbf{0} & \boldsymbol{H}_d^{(N)}(W) & \boldsymbol{H}_u^{(N)}(W{+}1) \end{bmatrix}. \tag{4.8}$$

Let us assume that:

(a) $\widehat{h}_{ij}^{(N)}(\ell)$, $i,j = 1,2,\ldots$ *and* $\ell = 1,\ldots,W+1$, *are i.i.d. with* $\mathrm{E}\{\widehat{h}_{ij}^{(N)}(\ell)\} = 0$, $\mathrm{E}\{|\widehat{h}_{ij}^{(N)}(\ell)|^2\} = \frac{1}{N}$, *and* $\lim_{N\to\infty} \mathrm{E}\{N^3|\widehat{h}_{ij}^{(N)}(\ell)|^6\} < +\infty$.

(b) As $N \to \infty$, *the sequence of e.d.f.* $\{F_{P^{(N)}}(\rho)\}$ *converges almost surely to a cumulative distribution function (c.d.f.)* $F_P(\rho)$[17].

(c) The spectral radius of the matrix $\boldsymbol{\mathcal{H}}_{N,W}$ *is almost surely upper bounded*[18].

(d) K *is a function of* N, *i.e.* $K = K(N)$, *satisfying* $\lim_{N\to\infty} \frac{K(N)}{N} = \beta$.

Then, the diagonal elements of the Gram matrices $\boldsymbol{\mathcal{T}}_{N,W}^m = (\boldsymbol{\mathcal{H}}_{N,W}\boldsymbol{\mathcal{H}}_{N,W}^H)^m$ *and* $\boldsymbol{\mathcal{R}}_{N,W}^m = (\boldsymbol{\mathcal{H}}_{N,W}^H\boldsymbol{\mathcal{H}}_{N,W})^m$ *converge with probability one to a deterministic value*

$$\lim_{K=\beta N\to\infty} \left(\boldsymbol{\mathcal{T}}_{N,W}^m\right)_{nn} \overset{a.s.}{=} \mathcal{T}_W^m(x) \tag{4.9}$$

$$\lim_{K=\beta N\to\infty} \left(\boldsymbol{\mathcal{R}}_{N,W}^m\right)_{kk} \overset{a.s.}{=} \mathcal{R}_W^m(y) \tag{4.10}$$

where $x = \lim_{N\to\infty} \frac{n}{N}$, *being* n *a function of* N *with values in* $\{1,\ldots,WN\}$, *i.e.* $n = n(N) : \mathbb{Z}^+ \to \{1,\ldots,WN\}$; $y = \lim_{N\to\infty} \frac{k}{N}$, *being* k *a function of* N *with values in* $\{1,\ldots,(W+1)K\}$, *i.e.* $k = k(N) : \mathbb{Z}^+ \to \{1,\ldots,(W+1)K\}$; *and* $\mathcal{R}_W^m(y)$

---

[16] The matrix $\boldsymbol{H}^{(N)}(\ell)$ models the spreading sequence matrix of transmitted symbols at time instant ℓ taking into account the time delays. It corresponds to the matrix $\boldsymbol{S}(\ell)$ or equivalently to the matrix $\boldsymbol{H}(\ell)$, since in this theorem we focus on the case $\boldsymbol{A} = \boldsymbol{I}$.

[17] Note that by assuming the $\rho_j^{(N)}$ independent with identical distribution function $F_P(\rho)$ the Glivenko-Cantelli theorem (see e.g. [89]) guarantees the almost sure convergence of $\{F_{P^{(N)}}(\rho)\}$ to $F_P(\rho)$. Therefore, condition (b) is redundant and stated explicitly in the theorem only for the sake of clarity.

[18] Theorem 8 holds also under the less restrictive condition that the eigenvalue moments of the matrix $\boldsymbol{\mathcal{H}}_{N,W}^H\boldsymbol{\mathcal{H}}_{N,W}$ are upper bounded.

and $\mathcal{T}_W^m(x)$ *determined by the following recursion:*

$$\mathcal{T}_W^{n+1}(x) = \sum_{s=0}^{n} \mathcal{T}_W^s(x) f(\mathcal{R}_W^{n-s}, x) \qquad 0 \le x \le W \tag{4.11}$$

$$\mathcal{R}_W^{n+1}(y) = \sum_{s=0}^{n} \mathcal{R}_W^s(y) g(\mathcal{T}_W^{n-s}, y) \qquad 0 \le y \le (W+1)\beta \tag{4.12}$$

where

$$f(\mathcal{R}_W^n, x) \triangleq \int_{r(x)}^{r(x)+\beta} \mathcal{R}_W^n(y)\,\mathrm{d}y \qquad 0 \le x \le W$$

$$g(\mathcal{T}_W^n, y) \triangleq \int_{\max(0,c(y)-1)}^{\min(W,c(y))} \mathcal{T}_W^n(x)\mathrm{d}x \qquad 0 \le y \le (W+1)\beta \tag{4.13}$$

with

$$r(x) \triangleq \beta\left[F_P(x - \lfloor x \rfloor) + \lfloor x \rfloor\right] \qquad 0 \le x \le W, \tag{4.14}$$

$$c(y) \triangleq \lfloor \frac{y}{\beta} \rfloor + F_P^{-1}\left(\frac{y}{\beta} - \lfloor \frac{y}{\beta} \rfloor\right) \qquad 0 \le y \le (W+1)\beta. \tag{4.15}$$

The recursion is initialized by $\mathcal{T}_W^0(x) = 1$, $\mathcal{R}_W^0(y) = 1$, $f(\mathcal{R}_W^0, x) = \beta$, *and* $g(\mathcal{T}_W^0, y) = l(y)$ *with*

$$l(y) \triangleq \begin{cases} F_P^{-1}\left(\frac{y}{\beta}\right) & 0 \le y \le \beta \\ 1 & \beta < y < W\beta \\ 1 - F_P^{-1}\left(\frac{y}{\beta} - W\right) & \beta W \le y \le \beta(W+1) \end{cases}.$$

Theorem 8 is proven in Appendix C.

The assumption (c), i.e., that the spectrum of the matrix $\boldsymbol{\mathcal{R}}_{N,W}$ is upper bounded, is of technical nature. Indeed, we conjecture that it follows from the assumptions on $h_{ij}(k)$ since we verified this property by extensive computer simulations. For the matrix $\boldsymbol{H}'$ for synchronous systems, the fact that the spectral radius is bounded was verified by computer simulations in [40] and it was proven in [64]. However, no analogous result for the matrix $\boldsymbol{\mathcal{R}}_{N,W}$ is known to the author.

In the following we give an interpretation of the quantities that appear in the recursion of Theorem 8. Figure 4.2 illustrates the structure of the matrix $\boldsymbol{\mathcal{H}}_{NW}$ as $N, K \to \infty$ and the meaning of the functions $r(x)$, and $l(y)$. The shaded region of the matrix corresponds to random nonzero elements while the remaining region corresponds to zero elements. The function $l(y)$ is the "height" of the shaded region in position y; the function $r(x)$ is the "width" of the zero region on the left of the

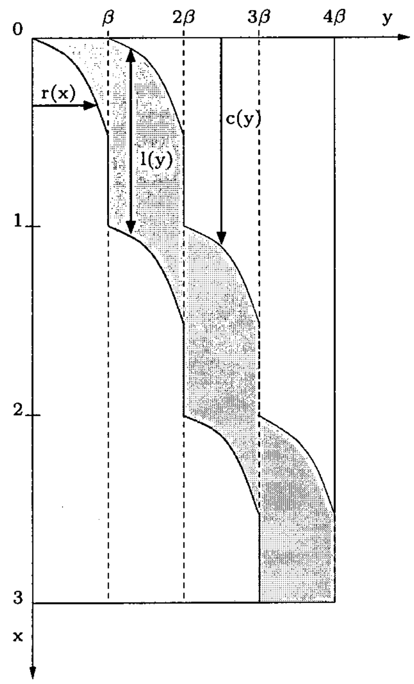


Figure 4.2: Graphical representation of the structure of the matrix $\mathcal{H}_{N,W}$ as $K, N \to \infty$ with $\frac{K}{N}$ and the functions $r(x)$, $c(y)$, and $l(y)$.

shaded region in position x; finally $c(y)$ is the "height" of the zero region on the right of the shaded region in position y. The functions $\mathcal{R}_W^n(y)$, $\mathcal{T}_W^n(x)$, $f(\mathcal{R}_W^n, x)$, and $g(\mathcal{T}_W^n, y)$ admit interesting interpretations. By definition $\mathcal{R}_W^n(y)$ and $\mathcal{T}_W^n(x)$ are the asymptotic deterministic values of the diagonal elements of the matrices $\boldsymbol{\mathcal{R}}_{N,W}^n$ and $\boldsymbol{\mathcal{T}}_{N,W}^n$, respectively, as $\frac{K}{N} \to \infty$ with ratio converging to a constant β. Being $f(\mathcal{R}_W^n, x)$ the integral of $\mathcal{R}_W^n(y)$ over the interval $[y_1, y_2]$ with $y_1 = r(x)$ and $y_2 = r(x) + \beta$, $f(\mathcal{R}_W^n, x)$ can be interpreted as the trace, normalized by N, of a submatrix of $\boldsymbol{\mathcal{R}}_W^n$ including all rows and columns of $\boldsymbol{\mathcal{R}}_W^n$ whose indices, normalized by N, are in the interval $[y_1, y_2]$. Similarly, since $g(\mathcal{T}_W^n, y)$ is the integral of $\mathcal{T}_W^n(x)$ over the interval $[x_1, x_2]$ with $x_1 = \max(0, c(y) - 1)$ and $x_2 = \min(W, c(y))$, $g(\mathcal{T}_W^n, y)$ can be interpreted as the normalized trace of a submatrix of $\boldsymbol{\mathcal{T}}_W^n$ including all rows and columns of $\boldsymbol{\mathcal{T}}_W^n$ whose indices, normalized by N, are in the interval $[x_1, x_2]$.

The following example explains the use of the theorem. Let us assume $W = 3$ and the time delay uniformly distributed in the interval $[0, T_s]$, then the limit distribution of the normalized time delay ρ is $F_P(\rho) = \rho$, with $\rho = [0, 1]$, $r(x) = \beta x$ with $x \in [0, 3]$,

$$l(y) = \begin{cases} \frac{y}{\beta} & 0 \le y \le \beta \\ 1 & \beta \le y \le 3\beta \\ 4 - \frac{y}{\beta} & 3\beta \le y \le 4\beta \end{cases} . \tag{4.16}$$

Therefore, $\mathcal{T}_3^1(x) = \beta$ and $\mathcal{R}_3^1(y) = l(y)$,

$$f(\mathcal{R}_3^1, x) = \begin{cases} \int\limits_{\beta x}^{\beta} \frac{y}{\beta} \mathrm{d}y + \int\limits_{\beta}^{\beta x+\beta} \mathrm{d}y & 0 \le x \le 1 \\ \int\limits_{\beta x}^{\beta x+\beta} \mathrm{d}y & 1 \le x \le 2 \\ \int\limits_{\beta x}^{3\beta} \mathrm{d}y + \int\limits_{3\beta}^{\beta x+\beta} \left(4 - \frac{y}{\beta}\right) \mathrm{d}y & 2 \le x \le 3 \end{cases} \tag{4.17}$$

and $g(\mathcal{W}_3^1, y) = \beta$, $0 \le y \le 4\beta$. Then, we can apply (4.11) and (4.12) to determine $\mathcal{T}_3^2(x)$ and $\mathcal{R}_3^2(y)$ and proceed recursively.

In Figure 4.3 the asymptotic values $\mathcal{R}_3^m(y)$ for $m = 1 \dots 6$ are compared to the values $(\boldsymbol{\mathcal{R}}_{N,3}^m)_{kk}$ of a single realization of $\boldsymbol{\mathcal{H}}_{N,3}$, for $N = 2048$ and $\beta = \frac{1}{2}$. Simulations with various distributions of the elements h_{ij} show that the diagonal elements of finite large matrices match very well the asymptotic values $\mathcal{R}_3^m(y)$ determined by the recursion in Theorem 8.

At first glance the recursion in Theorem 8 seems to be useful for the asymptotic analysis of linear detectors whereas it appears too complex for the design of low complexity multistage detectors similar to the detectors proposed for synchronous CDMA systems. However, a more careful analysis demonstrates very useful properties of the asymptotic values $\mathcal{R}_W^m(y)$ for the design of low complexity multistage detectors. With the goal of reducing the complexity of multiuser detector for asynchronous CDMA systems, we consider a symbol asynchronous but chip synchronous CDMA system with an observation window length $W = 6$ and $\beta = \frac{1}{2}$. The asymptotic values of the diagonal elements of $\boldsymbol{\mathcal{R}}_{N,6}^m$, $\mathcal{R}_6^m(y)$, $m = 1, 2, 3, 4, 5$ are plotted in Figure 4.4. The solid lines show the shape of $\mathcal{R}_6^m(y)$ while the dashed lines show the corresponding values of $\mathcal{R}_6^m(y)$ for a completely synchronous system. $\mathcal{R}_6^1(y)$ and $\mathcal{R}_6^2(y)$ coincide with the corresponding values of the synchronous system in the interval $y \in [\beta, \dots, 6\beta]$. In the interval $[0, \beta]$ and $[6\beta, 7\beta]$ the values of $\mathcal{R}_7^1(y)$ and $\mathcal{R}_7^2(y)$ for synchronous and asynchronous systems differ. From a mathematical point of view, this is due to the "tails" of the length of the spreading sequence $l(y)$ in the intervals $[0, \beta]$ and $[6\beta, 7\beta]$ (see Figure 4.2, $l(y) = 1$ for any value of x except in the intervals $[0, \beta]$ and $[W\beta, (W+1)\beta]$. We refer to these intervals as the "tails" of $l(y)$.). From a physical perspective, this behaviour stems from the fact that the symbols at the border of the observation window are observed only partially, i.e. only a subset of the chips of the whole spreading sequence is observed. For increasing powers of $\boldsymbol{\mathcal{R}}_{6,N}$, the effects of the "tails" start propagating inside the observation window. $\mathcal{R}_6^3(y)$ and $\mathcal{R}_6^4(y)$ for asynchronous systems coincide with the corresponding values for synchronous systems only in the interval $[2\beta, 5\beta]$. $\mathcal{R}_6^5(y)$ for synchronous and asynchronous systems is equal in the interval $[3\beta, 4\beta]$.

This behaviour is completely general and shows that, for fixed $s \in \mathbb{Z}^+$, there exists an observation window sufficiently large such that, in the center of the observation window, the values $\mathcal{R}_W^m(y)$, $m = 1, \ldots, s$ for synchronous and asynchronous systems are equal. Then, the values of $\mathcal{R}_W^m(y)$ for synchronous systems are very simple to compute using Algorithm 1 in Section 3.4 and the design of low complexity multiuser detectors for symbol asynchronous but chip synchronous systems can benefit from this "local equivalence".

In the following this property is established rigorously for CDMA systems with unequal received powers. First, Theorem 8 is generalized to the case of unequal received powers in the following Theorem 9. Then, the "local equivalence" is derived from Theorem 9 in Corollary 3. The assumptions of Theorem 9 summarize and formalize the characteristics of the system model (4.3) in case of power imbalances.

Theorem 9 *Let $\{(\lambda_i^{(N)}, \rho_i^{(N)})\}$, for $i = 1, \ldots, K$, be a sequence of K pairs in $\mathbb{R}^+ \times \{0, \frac{1}{N}, \ldots \frac{N-1}{N}\}$ sorted so that[19] $\rho_1^{(N)} \leq \rho_2^{(N)} \leq \ldots \leq \rho_K^{(N)}$. Let $F_{\Lambda^{(N)}, P^{(N)}}(\lambda^{(N)}, \rho^{(N)}) = \frac{1}{K}\sum_{i=1}^{K} 1(\lambda^{(N)} - \lambda_i^{(N)}) 1(\rho^{(N)} - \rho_i^{(N)})$ be the correspondent empirical joint distribution. Let $\widehat{s}_{ij}^{(N)}(\ell)$, $i, j = 1, 2, \ldots$ and $\ell = 1, \ldots, W+1$, be complex random variables. Let the matrix $\boldsymbol{S}^{(N)}(\ell) \in \mathbb{C}^{2N \times K}$, $\ell = 1, \ldots, W+1$ have elements $s_{ij}^{(N)}(\ell) = \widehat{s}_{ij}^{(N)}(\ell)$ for $\rho_j^{(N)} N + 1 \leq i \leq \rho_j^{(N)} N + N$, $1 \leq j \leq K$, and the remaining elements equal to zero[20]. Furthermore, let $\boldsymbol{S}_d^{(N)}(\ell), \boldsymbol{S}_u^{(N)}(\ell) \in \mathbb{C}^{N \times K}$ be, respectively, the lower and upper block of the matrix $\boldsymbol{S}(\ell)^{(N)}$, i.e. $\boldsymbol{S}^{(N)}(\ell) = [(\boldsymbol{S}_u^{(N)}(\ell))^T, (\boldsymbol{S}_d^{(N)}(\ell))]^T$. Let $\boldsymbol{\mathcal{S}}_{N,W}$ be a $WN \times (W+1)K$ random bi-diagonal block matrix structured as follows:*

$$\boldsymbol{\mathcal{S}}_{N,W} = \begin{bmatrix} \boldsymbol{S}_d^{(N)}(1) & \boldsymbol{S}_u^{(N)}(2) & \mathbf{0} & \cdots & \cdots & \cdots \\ \mathbf{0} & \boldsymbol{S}_d^{(N)}(2) & \boldsymbol{S}_u^{(N)}(3) & \mathbf{0} & \cdots & \cdots \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \cdots & \cdots & \mathbf{0} & \boldsymbol{S}_d^{(N)}(W-1) & \boldsymbol{S}_u^{(N)}(W) & \mathbf{0} \\ \cdots & \cdots & \cdots & \mathbf{0} & \boldsymbol{S}_d^{(N)}(W) & \boldsymbol{S}_u^{(N)}(W+1) \end{bmatrix}. \tag{4.18}$$

Additionally, let $\boldsymbol{\mathcal{A}}_K = \mathrm{diag}\{\underbrace{\boldsymbol{A}_K, \ldots, \boldsymbol{A}_K}_{W+1 \text{ times}}\}$ be a $(W+1)K \times (W+1)K$ block diagonal matrix with $(W+1)$ blocks equal to $\boldsymbol{A}_K$. Furthermore[21], $(\boldsymbol{A}_K \boldsymbol{A}_K^H)_{ii} = \lambda_i^{(N)}$. Finally, let $\boldsymbol{\mathcal{H}}_{N,W}$ be the $WN \times (W+1)K$ matrix given by $\boldsymbol{\mathcal{H}}_{N,W} = \boldsymbol{\mathcal{S}}_{N,W} \boldsymbol{\mathcal{A}}_K$.

[19] In a pair $\{(\lambda_i^{(N)}, \rho_i^{(N)})\}$, $\lambda_i^{(N)}$ models the received power $|a_{ii}|^2$ of user i and $\rho_i^{(N)}$ models the time delay of user i normalized by the symbol period T_s. The order reflects the ordering of the users assumed in the system models (4.3) and (4.5).

[20] $\boldsymbol{S}^{(N)}(\ell)$ models the spreading sequence matrix of transmitted symbols at time instant ℓ.

[21] The matrix $\boldsymbol{A}_K$ models the matrix of received amplitudes in the system models (4.3) and (4.5). Additionally, λ_i is the received power from user i, consistently with the definition of $\lambda_i^{(N)}$ in the pair $(\lambda_i^{(N)}, \rho_i^{(N)})$

Let us assume that:

(a) $\widehat{s}_{ij}^{(N)}(\ell)$, $i,j = 1,2,\ldots$ *and* $\ell = 1,\ldots,W+1$, *are i.i.d. with* $\mathrm{E}\{\widehat{s}_{ij}^{(N)}(\ell)\} = 0$, $\mathrm{E}\{|\widehat{s}_{ij}^{(N)}(\ell)|^2\} = \frac{1}{N}$, *and* $\lim_{N\to\infty} \mathrm{E}\{N^3|\widehat{s}_{ij}^{(N)}(\ell)|^6\} < +\infty$.

(b) As $N \to \infty$, *the sequence of e.d.f.* $\{F_{\Lambda^{(N)}P^{(N)}}(\lambda,\rho)\}$ *converges almost surely to a differentiable c.d.f.* $F_{\Lambda P}(\lambda,\rho)$.

(c) λ *and* ρ *are deterministically related, i.e.* $F_{\Lambda|P}(\lambda|\rho=\rho_0) = 1(\lambda - \lambda(\rho_0))$. *Here,* $\lambda = \lambda(\rho)$ *denotes* ρ *as a deterministic function of* λ. $\lambda(\rho)$ *is upper bounded.*

(d) The spectral radius of the matrix $\boldsymbol{\mathcal{H}}_{N,W}$ *is almost surely upper bounded*[22].

(e) K *is a function of* N, *i.e.* $K = K(N)$, *satisfying* $\lim_{N\to\infty} \frac{K(N)}{N} = \beta$.

Then, the diagonal elements of the Gram matrix $\boldsymbol{\mathcal{T}}_{N,W}^m = (\boldsymbol{\mathcal{H}}_{N,W}\boldsymbol{\mathcal{H}}_{N,W}^H)^m$ *converge with probability one to a deterministic value*

$$\lim_{K=\beta N\to\infty} \left(\boldsymbol{\mathcal{T}}_{N,W}^m\right)_{nn} \overset{a.s.}{=} \mathcal{T}_W^m(x)$$

where $x = \lim_{N\to\infty} \frac{n}{N}$, *being* n *a function of* N *with values in* $\{1,\ldots,WN\}$, *i.e.* $n = n(N) : \mathbb{Z}^+ \to \{1,\ldots,WN\}$.

Conditioned on $\lambda\left(y - \lfloor\frac{y}{\beta}\rfloor\beta\right)$, *the diagonal elements of the matrix* $\boldsymbol{\mathcal{R}}_{N,W}^m = (\boldsymbol{\mathcal{H}}_{N,W}^H\boldsymbol{\mathcal{H}}_{N,W})^m$ *converge with probability one to a deterministic value*

$$\lim_{K=\beta N\to\infty} \left(\boldsymbol{\mathcal{R}}_{N,W}^m\right)_{kk} \overset{a.s.}{=} \mathcal{R}_W^m(y)$$

where $y = \lim_{N\to\infty} \frac{k}{N}$, *being* k *a function of* N *with values in* $\{1,\ldots,(W+1)K\}$, *i.e.* $k = k(N) : \mathbb{Z}^+ \to \{1,\ldots,(W+1)N\}$. *The limits* $\mathcal{R}_W^m(y)$ *and* $\mathcal{T}_w^m(x)$ *are determined by the following recursion:*

$$\mathcal{T}_W^{n+1}(x) = \sum_{s=0}^{n} \mathcal{T}_W^s(x) f(\mathcal{R}_W^{n-s}, x) \qquad 0 \le x \le W \tag{4.19}$$

$$\mathcal{R}_W^{n+1}(y) = \sum_{s=0}^{n} \mathcal{R}_W^s(y) g(\mathcal{T}_W^{n-s}, y) \qquad 0 \le y \le (W+1)\beta \tag{4.20}$$

where

$$f(\mathcal{R}_W^n, x) \triangleq \beta \int_{r(x)}^{r(x)+\beta} \mathcal{R}_W^n(y)\lambda\left(y - \lfloor\frac{y}{\beta}\rfloor\beta\right) \mathrm{d}F_P(y) \qquad 0 \le x \le W \tag{4.21}$$

$$g(\mathcal{T}_W^n, y) \triangleq \lambda\left(y - \lfloor\frac{y}{\beta}\rfloor\beta\right) \int_{\max(0,c(y)-1)}^{\min(W,c(y))} \mathcal{T}_W^n(x)\mathrm{d}x \qquad 0 \le y \le (W+1)\beta \tag{4.22}$$

[22] On this condition similar considerations as in Theorem 8 hold.

with

$$r(x) \triangleq \beta\left[F_P(x - \lfloor x \rfloor) + \lfloor x \rfloor\right] \qquad 0 \leq x \leq W, \tag{4.23}$$

$$c(y) \triangleq \left\lfloor \frac{y}{\beta} \right\rfloor + F_P^{-1}\left(\frac{y}{\beta} - \left\lfloor \frac{y}{\beta} \right\rfloor\right) \qquad 0 \leq y \leq (W+1)\beta. \tag{4.24}$$

The recursion is initialized by $\mathcal{T}_W^0(x) = 1$, $\mathcal{R}_W^0(y) = 1$, $f(\mathcal{R}_W^0, x) = \beta \mathrm{E}_{|\boldsymbol{A}|^2}\{\lambda\}$ *and* $g(\boldsymbol{\mathcal{T}}_W^0, y) = \lambda\left(y - \lfloor\frac{y}{\beta}\rfloor\beta\right) l(y)$ *with*

$$l(y) \triangleq \begin{cases} F_P^{-1}\left(\frac{y}{\beta}\right) & 0 \leq y \leq \beta \\ 1 & \beta < y < W\beta \\ 1 - F_P^{-1}\left(\frac{y}{\beta} - W\right) & \beta W \leq y \leq \beta(W+1) \end{cases} . \tag{4.25}$$

Theorem 9 is proven in Appendix C Section C.2.

Although the recursion proposed in Theorem 9 is perhaps too complex for being of great practical use in the design of low complexity multiuser detectors, the following useful corollary stems from it.

Corollary 3 *Let the assumptions of Theorem 9 hold. Let* $\mathcal{R}_W^m(y)$ *and* $\mathcal{R}_W'^m(y)$ *be the asymptotic deterministic values of the diagonal elements of* $\boldsymbol{\mathcal{R}}_{N,W}^m$ *for asynchronous and synchronous systems, respectively, as* $K, N \rightarrow \infty$ *with* $\frac{K}{N} \rightarrow \beta$. *Assume* $\lceil m \rceil < \frac{W+1}{2}$. *Then, for* $\ell = 0, 1, \ldots, 2m$

$$\mathcal{R}_W^\ell(y) = \mathcal{R}_W'^\ell(y)$$

for $y \in \left[\beta\lceil\frac{\ell}{2}\rceil, \beta(W + 1 - \lceil\frac{\ell}{2}\rceil)\right]$.

Let us notice that from (4.5), the system model for symbol asynchronous but chip synchronous systems, a synchronous system is obtained when equal time delays for all users are assumed. The matrix $\boldsymbol{\mathcal{R}}_{N,W}^m$ then becomes a block diagonal matrix. Additionally, each block of $\boldsymbol{\mathcal{R}}_{N,W}^m$ for a synchronous system is equal to the matrix $\boldsymbol{R}^m$ of the single-symbol system analyzed in Chapter 3. Corollary 3 follows from this observation and by induction from Theorem 9. Corollary 3 allows us to calculate the diagonal elements $\mathcal{R}_W^\ell(y)$ by means of Algorithm 1 presented in Section 3.4. This result is quite useful from a practical point of view, as it will be apparent in Section 4.3.3. In fact, the weight design for asynchronous systems reduces to the one for synchronous systems already solved in Chapter 3.

4.3.3 Multistage Detection

In this section we extend the concept of multistage detectors with asymptotic weights to asynchronous CDMA systems.

In Chapter 3 we pointed out that polynomial expansion detectors are suitable in scenarios with approximately equal received powers for all users while multistage Wiener filters outperform polynomial expansion detectors in scenarios with power imbalances. To give an intuition of this we can compare (3.7) and (3.12). The polynomial expansion detectors replace the optimum values $(\boldsymbol{R}^s)_{kk}$ in (3.8) and (3.9) by $\mathrm{tr}\boldsymbol{R}^s$. If the values of $(\boldsymbol{R}^s)_{kk}$ are closed to $\mathrm{tr}\boldsymbol{R}^s$, as in the case of synchronous CDMA systems with almost equal received powers, the performance degradation of polynomial expansion detectors is negligible. However, if the values of $(\boldsymbol{R}^s)_{kk}$ have large variance the approximation of $(\boldsymbol{R}^s)_{kk}$ with $\mathrm{tr}\boldsymbol{R}^s$ can be severely suboptimal. In an asynchronous system, $(\boldsymbol{\mathcal{R}}^s_{N,W})_{kk}$ varies considerably within the observation window even for equal received powers, as is apparent from Figure 4.3 and Figure 4.4. Therefore, in asynchronous systems, polynomial expansion detectors suffer more from their sub-optimality than in synchronous systems. In the following we focus on detectors Type J-I.

The design and analysis of detectors Type J-I for asynchronous systems benefit mainly from the following two observations:

- From Corollary 3 it is apparent that, for an observation window sufficiently large, the diagonal elements of the matrices[23] $\boldsymbol{\mathcal{R}}^s_W$, $s = 1, \ldots, m$, in the center of the observation window, coincide with the diagonal elements of the matrices $\boldsymbol{R}^s$, $s = 1, \ldots, m$, of the corresponding synchronous system (see Section 3.2). Therefore, the design of the asymptotic weights for the detection of the transmitted symbols in the center of the observation window reduces to the low complexity design of the weights for synchronous systems detailed in Chapter 3.[24]

- The band structure of the matrix $\boldsymbol{\mathcal{H}}$ in (4.4) enables an implementation of multistage detectors with infinite sliding observation window but finite delay.

The joint use of these two properties has the following implications:

[23] Given m, the observation window length W has to be chosen to satisfy the assumptions of Corollary 3.

[24] This observation suggests the possibility of defining a Type J-J detector for symbol asynchronous but chip synchronous systems simply replacing the asymptotic weights of a Type J-I detector with the weights of a Type J-J detector for synchronous systems. Note that this weighting would not correspond to any practical polynomial expansion detector with finite observation window since it implies an infinite observation window. Additionally, its asymptotic performance differs from the performance of a polynomial expansion detector, as $K, N \to \infty$ with constant ratio, for any finite choice of the observation window W of the polynomial expansion detector.

- The design of the asymptotic weights for symbol asynchronous but chip synchronous systems coincides with the design of the asymptotic weights for synchronous systems.

- Multistage detectors for asynchronous systems can be implemented with a sliding observation window. As a consequence, the multiuser efficiency of all received symbols from all users is identical in the large system limit.

- The complexity order per bit of multistage detectors for symbol asynchronous but chip synchronous systems is the same as the one of equivalent detectors for synchronous systems.

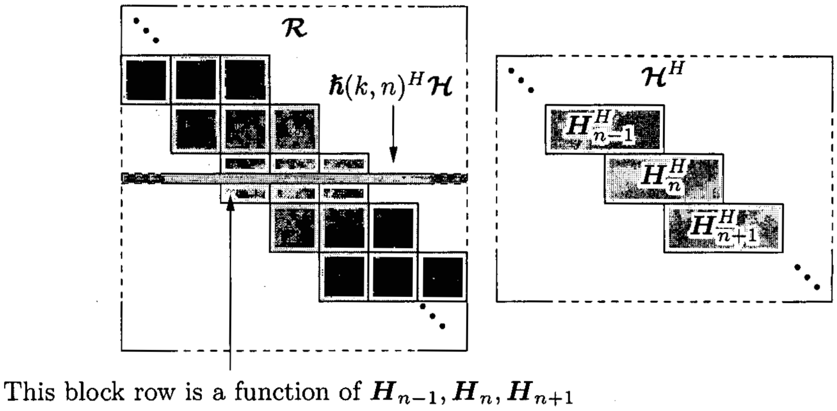
Choice of the basis of the projection subspace

As already discussed in Section 3.3, a linear multistage detector of order M performs a projection of the observed signal onto an M-dimensional Krylov subspace and a subsequent processing of the projections by a filter designed according to an optimality criterion.

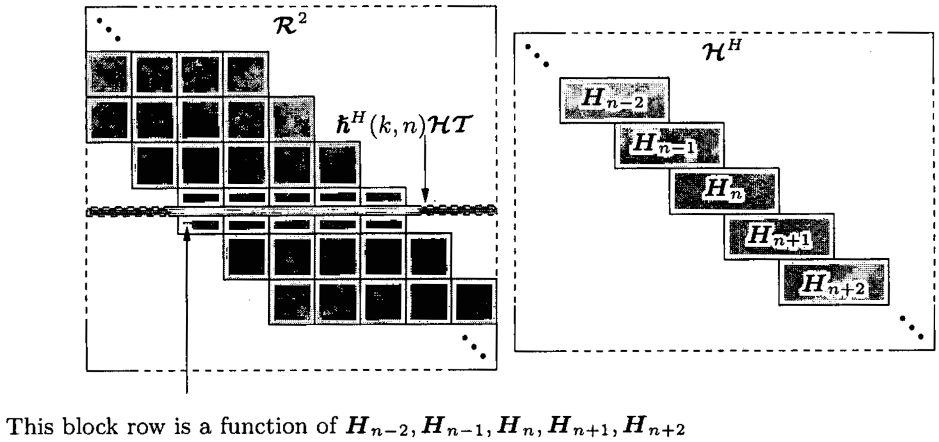
A straightforward extension of multistage detectors to symbol asynchronous but chip synchronous systems would replace the matrices $\boldsymbol{H}$ and $\boldsymbol{T}$ and the vector $\boldsymbol{h}_k$ in (3.3) with the finite matrices $\boldsymbol{\mathcal{H}}_{N,W}(n)$ and $\boldsymbol{\mathcal{T}}_{N,W}(n)$ and the k^{th} column vector extracted from vector $\boldsymbol{\mathcal{H}}_{N,W}(n)$, $\boldsymbol{\hbar}_{N,W}(k,n)$, respectively. By this straightforward extension two kinds of performance degradation would occur. One for using a subspace method instead of a full rank approach and one due to windowing. However, an implementation of multistage detectors with finite delay is possible while avoiding windowing effects. Let us consider the unlimited system model (4.3) and let us use the subspace

$$\chi_{M,k,n}(\boldsymbol{\mathcal{H}}) = \text{span}\left\{\boldsymbol{\hbar}(k,n)^H \boldsymbol{\mathcal{T}}^m\right\}_{m=0}^{M},$$

where $\boldsymbol{\mathcal{T}} = \boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{H}}^H$ and $\boldsymbol{\hbar}(k,n)$ is the column of $\boldsymbol{\mathcal{H}}$ corresponding to user k at the symbol interval n. Because of the bi-diagonal block structure of $\boldsymbol{\mathcal{H}}$, the matrix $\boldsymbol{\mathcal{T}}$ is a tri-diagonal block matrix and its power $\boldsymbol{\mathcal{T}}^m$ is a $(2m+1)$-diagonal matrix (see Figure 4.5). Therefore, the vector $\boldsymbol{\hbar}^H(k,n)\boldsymbol{\mathcal{T}}^m$ has, at most, $(2m+1)N$ nonzero elements and the M-stage detector for the unlimited system model can be implemented with a finite delay equal to MT_s. This property is illustrated in Figure 4.5. Figure 4.5.a shows $\boldsymbol{\hbar}^H(k,n)\boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{H}}^H = \boldsymbol{\hbar}^H(k,n)\boldsymbol{\mathcal{T}}$. The unlimited row vector $\boldsymbol{\hbar}^H(k,n)\boldsymbol{\mathcal{T}}$ depends only on $\boldsymbol{H}(n-1), \boldsymbol{H}(n), \boldsymbol{H}(n+1)$ and has nonzero elements only in the symbol intervals $n-1, n, n+1$. Therefore, the statistic $\boldsymbol{\hbar}^H(k,n)\boldsymbol{\mathcal{T}}\boldsymbol{\mathcal{Y}}$ depends only on $\boldsymbol{y}(n-1), \boldsymbol{y}(n), \boldsymbol{y}(n+1)$. Figure 4.5.b shows that this property extends to the other elements of the basis of the Krylov subspace. Namely, the statistic $\boldsymbol{\hbar}^H(k,n)\boldsymbol{\mathcal{T}}^2\boldsymbol{\mathcal{Y}}$ depends only on $\boldsymbol{y}(n-2), \boldsymbol{y}(n-1), \boldsymbol{y}(n), \boldsymbol{y}(n+1), \boldsymbol{y}(n+2)$.



(a)



(b)

Figure 4.5: Vectors of the basis of the Krylov subspace for asynchronous CDMA systems. (a) Decomposition of the basis vector $\hbar^H(k,n)\mathcal{T}$ into the product of the row vector $\hbar^H(k,n)\mathcal{H}$ by the matrix $\mathcal{H}^H$. The vector $\hbar^H(k,n)\mathcal{T}$ depends only on $H(n-1), H(n), H(n+1)$. (b) Decomposition of the basis vector $\hbar^H(k,n)\mathcal{T}^2$ into the product of the row vector $\hbar^H(k,n)\mathcal{H}\mathcal{T}$ by the matrix $\mathcal{H}^H$. The vector $\hbar^H(k,n)\mathcal{T}^2$ depends only on $H(n-2), H(n-1), H(n), H(n+1), H(n+2)$.

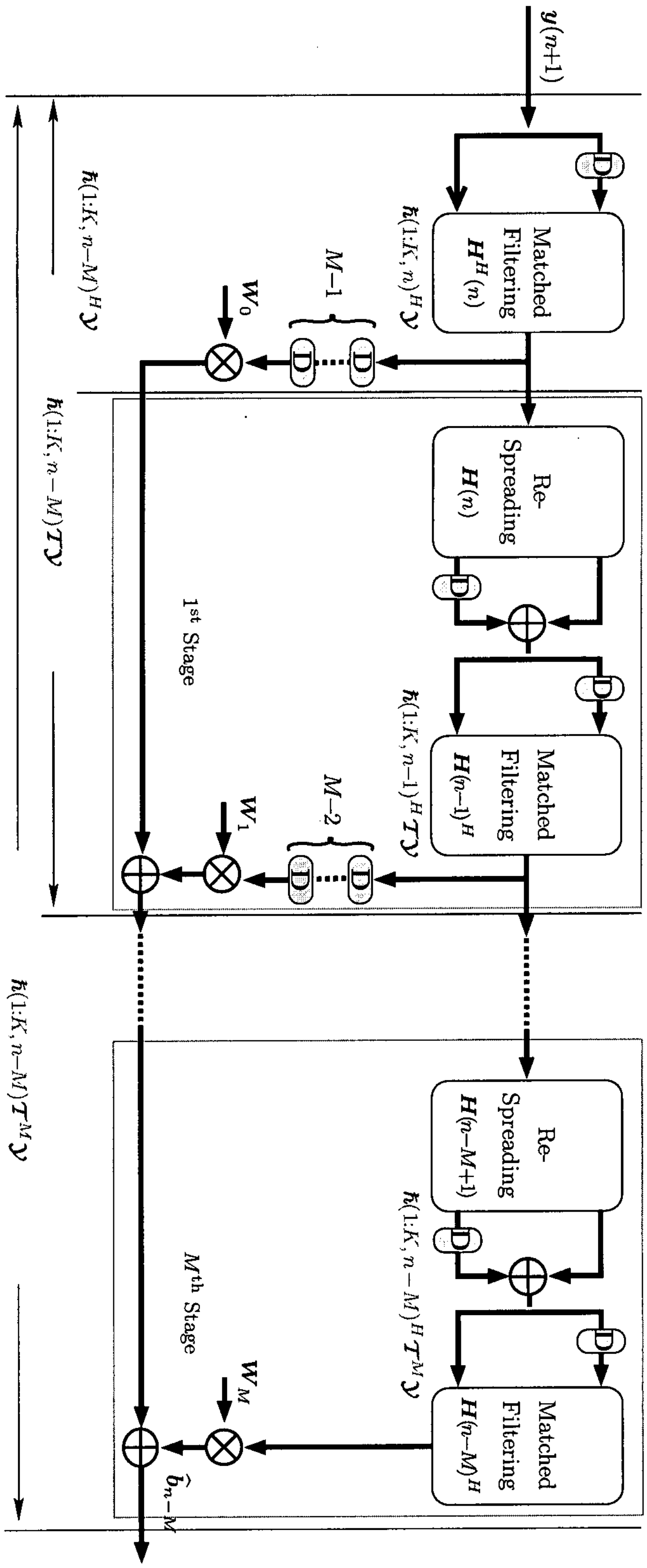


Figure 4.6: Multistage detector Type J-I for chip synchronous and symbol asynchronous CDMA systems.

The structure of a multistage detector for symbol-asynchronous but chip-synchronous systems with a sliding observation window expanding with the number of stages is shown in Figure 4.6.

Each stage of the multistage detector consists of a re-spreading block and a subsequent matched filter, as well as a stage of a multistage detector for synchronous systems (see Figure 3.2). However, the spreading and re-spreading blocks process a signal received during two symbol periods instead of a signal received in a single symbol period, as the signal processed by a multistage detector for synchronous systems (for symbol asynchronous but chip synchronous systems $\boldsymbol{H}(m) \in \mathbb{C}^{2N\times K}$, whereas for synchronous systems $\boldsymbol{H}'(m) \in \mathbb{C}^{N\times K}$). Furthermore, between the re-spreading and the subsequent matched filter the signal is properly delayed and combined. The j^{th} block receives as input the K-dimensional vector $\boldsymbol{\hbar}^H(1\!:\!K, n-j+1)\mathcal{T}^{j-1}\mathcal{Y}$, where $\boldsymbol{\hbar}(s\!:\!r,n)$ denotes the matrix $[\boldsymbol{\hbar}(s,n), \boldsymbol{\hbar}(s+1,n), \ldots, \boldsymbol{\hbar}(r,n)]$ with unbounded number of rows and $r-s+1$ columns, and provides as output the K-dimensional vector $\boldsymbol{\hbar}^H(1\!:\!K, n-j)\mathcal{T}^{j}\mathcal{Y}$. The vector $\boldsymbol{\hbar}^H(1\!:\!K, n-j+1)\mathcal{T}^{j-1}\mathcal{Y}$ is multiplied or re-spread by the matrix $\boldsymbol{H}(n-j+1)$. The re-spreading block provides two output vectors, the upper part vector $\boldsymbol{v}_u(n,j) = \boldsymbol{H}_u(n-j+1)\boldsymbol{\hbar}^H(1\!:\!K, n-j+1)\mathcal{T}^{j-1}\mathcal{Y}$ and the vector $\boldsymbol{v}_d(n,j) = \boldsymbol{H}_d(n-j+1)\boldsymbol{\hbar}^H(1\!:\!K, n-j+1)\mathcal{T}^{j-1}\mathcal{Y}$. The vectors $\boldsymbol{v}_d(n-1,j)$ and $\boldsymbol{v}_u(n-1,j)+\boldsymbol{v}_d(n-2,j)$ are memorized in the delay blocks. The input to the subsequent matched filter is given by

$$\begin{bmatrix} \boldsymbol{v}_u(n-1,j)+\boldsymbol{v}_d(n-2,j) \\ \boldsymbol{v}_u(n,j)+\boldsymbol{v}_d(n-1,j) \end{bmatrix} = \begin{bmatrix} (\boldsymbol{H}_u(n-j)\boldsymbol{\hbar}^H(1\!:\!K,n-j)+\boldsymbol{H}_d(n-j-1)\boldsymbol{\hbar}^H(1\!:\!K,n-j-1))\mathcal{T}^{j-1}\mathcal{Y} \\ (\boldsymbol{H}_u(n-j+1)\boldsymbol{\hbar}^H(1\!:\!K,n-j+1)+\boldsymbol{H}_d(n-j)\boldsymbol{\hbar}^H(1\!:\!K,n-j))\mathcal{T}^{j-1}\mathcal{Y} \end{bmatrix}. \tag{4.26}$$

The output of the j^{th} stage is delayed by $(M-j)T_s$ before being used as input of the filters defined by the matrix weights $\boldsymbol{W}_0(n-M), \boldsymbol{W}_1(n-M), \ldots, \boldsymbol{W}_M(n-M)$ to provide $\widehat{\boldsymbol{b}}(n-M)$, the soft estimate of $\boldsymbol{b}(n-M)$.

The following considerations provide further insight into the structure of the j^{th} stage. The j^{th} stage calculates $\boldsymbol{\hbar}^H(1\!:\!K, n-j)\mathcal{T}^{j}\mathcal{Y}$ from a partial knowledge of the vector $\mathcal{H}^H\mathcal{T}^{j-1}\mathcal{Y}$. This knowledge is limited to the K-dimensional vectors $\boldsymbol{\hbar}^H(1\!:\!K, n-j-s)\mathcal{T}^{j-1}\mathcal{Y}$, with $s=-1,0,\ldots$. Thanks to the fact $\boldsymbol{\hbar}(1\!:\!K, n-j)$ is nonzero only corresponding to the symbol periods $n-j$ and $n-j+1$, the knowledge of the vector $\mathcal{T}^{j}\mathcal{Y}$ corresponding to those symbol periods is sufficient to compute $\boldsymbol{\hbar}^H(1\!:\!K, n-j)\mathcal{T}^{j}\mathcal{Y}$. It is straightforward to verify that $\mathcal{T}^{j}\mathcal{Y}$ in the symbol interval $n-j$ is given by

$$[\boldsymbol{H}_u(n-j)\boldsymbol{\hbar}^H(1\!:\!K,n-j)+\boldsymbol{H}_d(n-j-1)\boldsymbol{\hbar}^H(1\!:\!K,n-j-1)]\mathcal{T}^{j-1}\mathcal{Y}$$

and in the symbol interval $n-j+1$ is given by

$$[\boldsymbol{H}_u(n-j+1)\boldsymbol{\hbar}^H(1\!:\!K,n-j+1)+\boldsymbol{H}_d(n-j)\boldsymbol{\hbar}^H(1\!:\!K,n-j)]\mathcal{T}^{j-1}\mathcal{Y},$$

i.e. the vectors provided as input of the matched filter in the j^{th} stage (cfr. (4.26)). Therefore, the signals at the output of the re-spreading block are delayed and combined to compute these two vectors and to provide the required input to the subsequent matched filter.

Asymptotic Design and Analysis

Let $\boldsymbol{W}_m(n)$ and $\boldsymbol{\mathcal{W}}_m$ be a $K \times K$ and an unlimited diagonal matrix, respectively. Let us denote by $w_{m\ell}(n)$ the ℓ^{th} diagonal element of the matrix $\boldsymbol{W}_m(n)$, then $\boldsymbol{W}_m(n) = \mathrm{diag}(w_{m1}(n), w_{m2}(n), \ldots, w_{mK}(n))$ and $\boldsymbol{\mathcal{W}}_m = \mathrm{diag}\{\ldots, \boldsymbol{W}_m(n), \boldsymbol{W}_m(n+1), \ldots\}$. The multistage detector Type J-I for asynchronous systems is the linear operator $\boldsymbol{\mathcal{M}} = \sum_{m=0}^{M-1} \boldsymbol{\mathcal{W}}_m \boldsymbol{\mathcal{H}}^H \boldsymbol{\mathcal{T}}^m$ where the $\boldsymbol{\mathcal{W}}_m$ (i.e. the matrix weight of the filter in the Krylow subspace) are chosen such that $\mathrm{E}\{\|\boldsymbol{\mathcal{M}}\boldsymbol{\mathcal{Y}} - \boldsymbol{\mathcal{B}}\|^2\}$ is minimum. This is equivalent to the minimization of the mean square error for each component $(\boldsymbol{b}(n))_k$ of $\boldsymbol{\mathcal{B}}$ in the corresponding subspace $\chi_{M,k,n}(\boldsymbol{\mathcal{H}})$. Applying again the Wiener-Hopf theorem as in Chapter 3, the weight matrices $\boldsymbol{W}_m(n)$ can be derived by the following equation:

$$\boldsymbol{w}_k(n) = (\boldsymbol{\Phi}_k(n))^{-1}\boldsymbol{\varphi}_k(n)$$

where $(\boldsymbol{w}_k(n))_m = (\boldsymbol{W}_m(n))_{kk}$, $\boldsymbol{\varphi}_k(n)$ is an M-dimensional vector , $\boldsymbol{\Phi}_k(n) \in \mathbb{R}^{M\times M}$, $(\boldsymbol{\varphi}_k(n))_m = (\boldsymbol{\mathcal{R}}^{m+1}(n))_{kk}$, $(\boldsymbol{\Phi}_k(n))_{lm} = (\boldsymbol{\mathcal{R}}^{l+m}(n))_{kk} + \sigma^2(\boldsymbol{\mathcal{R}}^{l+m-1}(n))_{kk}$, $\boldsymbol{\mathcal{R}} = \boldsymbol{\mathcal{H}}^H\boldsymbol{\mathcal{H}}$, and $(\boldsymbol{\mathcal{R}}^m(n))_{kk} = \boldsymbol{\hbar}(k,n)^H\boldsymbol{\mathcal{T}}^{m-1}\boldsymbol{\hbar}(k,n)$ denotes the diagonal element of the matrix $\boldsymbol{\mathcal{R}}^m$ corresponding to the user k at time instant n. The output SINR of user k is given by [67]

$$\mathrm{SINR}_k(n) = \frac{\boldsymbol{\varphi}_k^T(n)(\boldsymbol{\Phi}_k(n))^{-1}\boldsymbol{\varphi}_k(n)}{1 - \boldsymbol{\varphi}_k^T(n)(\boldsymbol{\Phi}_k(n))^{-1}\boldsymbol{\varphi}_k(n)}. \tag{4.27}$$

In the asymptotic case, as $N, K \to \infty$ with $\frac{K}{N} = \beta$, we can apply Corollary 3 to obtain

$$\lim_{K=\beta N\to\infty} (\boldsymbol{\mathcal{R}}^m(n))_{kk} = R^m_{kk,\infty} \qquad \forall 1 \le m \le 2M, \tag{4.28}$$

where $R^m_{kk,\infty}$ are the asymptotic (deterministic) diagonal elements of the matrix $\boldsymbol{R}^m$ for synchronous systems introduced in Chapter 3. Recursive and closed form expressions for $R^m_{k,\infty}$ can be found in Theorem 6 or Algorithm 1 (see Section 3.4) and in Theorem 15 in Section B.2, respectively.

Numerical Results

Throughout this section, we consider linear MMSE detectors with observation window $W = 3$ and equal received powers within a chip synchronous but symbol asynchronous CDMA system. Figure 4.7 shows $\mathrm{SINR}_{\mathrm{LMMSE}}$, the output SINR of a linear

MMSE detector, for a system with $\beta = \frac{1}{2}$ and $\frac{E_b}{N_0} = 7$ dB. As for the synchronous case, the convergence of lower bounds $\mathrm{SINR}_k^{(s)}$ in (4.7), as $s \to \infty$, toward the SINR of a linear MMSE detector, $\mathrm{SINR}_{\mathrm{LMMSE}}$ is very fast and the lower bound corresponding to $M = 7$ (line with markers in Figure 4.7) is undistinguishable from the one obtained for $M = 8$. The SINR reaches its maximum for the transmitted symbol centered in the observation window and decreases smoothly for the transmitted symbols whose spread signal is still completely observed ($y \in [\beta, 3\beta]$). The performance degrades rapidly for symbols only partially included in the observation window. In contrast to the synchronous case, in the asynchronous case the multistage detectors with M sufficiently large can outperform the full rank linear MMSE detector with finite observation window W. This is due to the fact that both detectors use only a subset of the sufficient statistics, but, with the proposed subspace basis, multistage detectors intrinsically use a wider and wider subset as the number of stages increases, whereas the full rank linear MMSE detector exploits always the same statistic (the use of a wider statistic would require a longer observation window and thus a larger complexity[25]).

These theoretical results are completely supported by simulations. We considered a CDMA system with 64 users, Gaussian random spreading, frequency flat Rayleigh fading, spreading factor 128 and QPSK modulation under two different conditions, namely, for synchronous received symbols and for chip synchronous but symbol asynchronous received symbols. We compare the BER of the multistage detectors described in Chapter 3 for synchronous transmitted symbols with the BER of detectors Type J-I for chip synchronous but symbol asynchronous systems introduced in this chapter. Figure 4.8 shows the BER versus $\frac{E_b}{N_0}$ for a varying number of stages. The BER for chip synchronous but symbol asynchronous systems matches completely the BER for synchronous systems.

4.4 Chip Asynchronous and Symbol Quasi-Synchronous CDMA Systems

4.4.1 System Model

In the following, CDMA systems such that the time delays of the received signals τ_k, $k = 1, \ldots, K$ are smaller than the chip delay, T_c, are referred to as *symbol quasi-synchronous but chip asynchronous CDMA systems.*

In this section we consider symbol quasi-synchronous but chip asynchronous systems and we assume that the time delays τ_k satisfy the constraints $\tau_k \leq T_c$,

[25]Note that complexity grows cubic with the window size, but linearly with the number of stages.

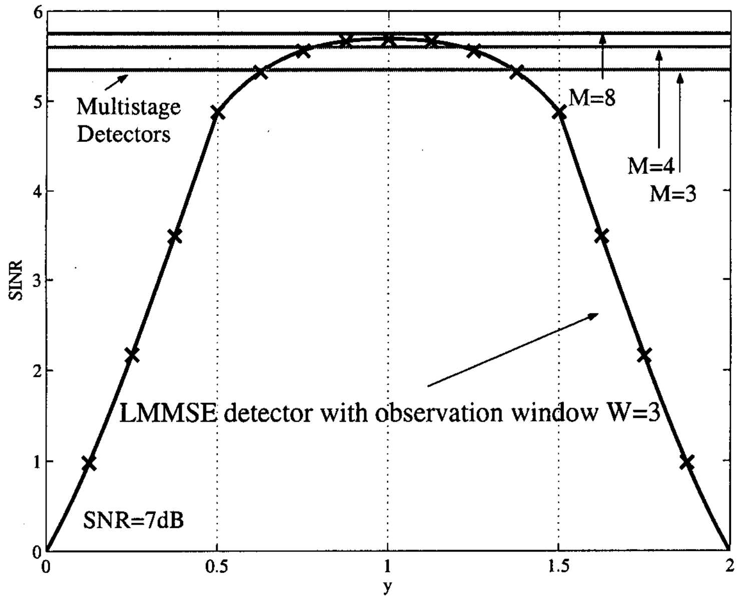


Figure 4.7: Chip synchronous but symbol asynchronous CDMA system with equal received powers, $\beta = \frac{1}{2}$ and $\frac{E_b}{N_0} = 7$ dB. *Asymptotic* $\mathrm{SINR}_{\mathrm{LMMSE}}$ *for* $W = 3$ *and multistage detector* SINR *for varying* M *versus the position of the detected symbol in the observation window.*

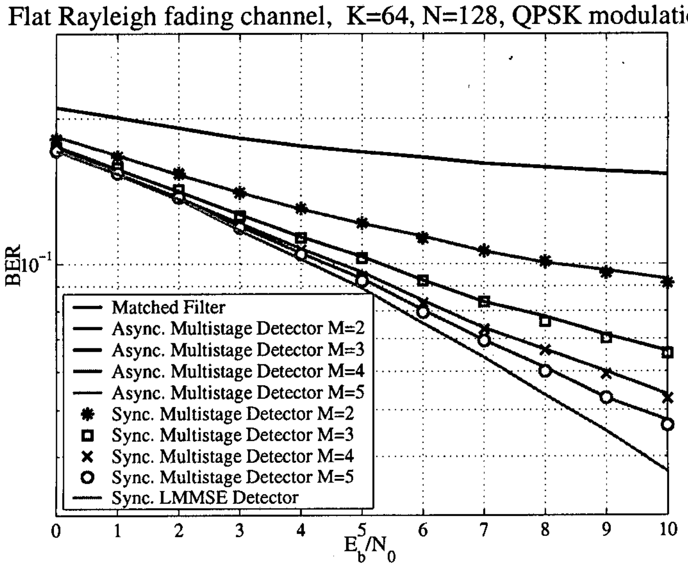


Figure 4.8: Comparison between the BER of multistage detectors for synchronous systems (solid lines) and the BER of multistage detectors for chip synchronous and symbol asynchronous CDMA systems (markers). BER versus $\frac{E_b}{N_0}$ for $\beta = 0.5$ and a varying number of stages.

$k = 1, \dots, K$. Additionally, we focus on the transmission of a single symbol b_k per user as in [82]. Then, the general system model (4.1) can be rewritten as

$$y[p] = \sum_{k=1}^{K} a_{kk} b_k \sum_{u=0}^{N-1} s_k[u] \widetilde{\psi}\left(\left(\frac{p}{r} - u\right) T_c - \tau_k\right) + n[p] \qquad p \in \mathbb{Z} \tag{4.29}$$

where a_{kk}, b_k, $s_k[u]$, $n[p]$, r are defined in Section 4.2.

We also assume that the chip pulse $\widetilde{\psi}(t)$ is much shorter than the symbol waveform, i.e. $\widetilde{\psi}(t)$ becomes negligible for $|t| > t_0$ with some $t_0 \ll T_s$. This is usually valid in systems with large spreading factor, which are considered in this work. In fact, in the systems in use the chip pulse decays rapidly and becomes negligible after 8-10 chip periods[26]. We make use of the previous assumption to neglect the useful signal outside the symbol interval $[0, T_s]$. Thus, the system model (4.29) with $p = 0, 1, \dots, Nr - 1$ reduces to

$$\widetilde{\boldsymbol{y}} = \sum_{k=1}^{K} a_{kk} b_k \boldsymbol{v}_k + \widetilde{\boldsymbol{n}}$$

where $\widetilde{\boldsymbol{y}}$ and $\widetilde{\boldsymbol{n}}$ are the Nr dimensional vectors of received signal and zero mean, complex-valued, circular symmetric, white Gaussian noise with variance $\sigma^2 = \frac{rN_0}{E_\psi T_c}$, respectively. Furthermore, $\boldsymbol{v}_k$ is the Nr dimensional virtual spreading sequence of user k given by

$$\boldsymbol{v}_k = \widetilde{\boldsymbol{\Psi}}_k \boldsymbol{s}_k,$$

where $\widetilde{\boldsymbol{\Psi}}_k$ is an $Nr \times N$ matrix taking into account the effects of the pulse shape and the time delay of user k. It is defined as

$$\widetilde{\boldsymbol{\Psi}}_k = \left(\begin{array}{c|c|c|c}
\widetilde{\psi}(-\tau_k) & \widetilde{\psi}(-T_c-\tau_k) & \dots & \widetilde{\psi}((-N+1)T_c-\tau_k) \\
\vdots & \vdots & \vdots & \vdots \\
\widetilde{\psi}(\frac{(r-1)T_c}{r}-\tau_k) & \widetilde{\psi}((\frac{(r-1)}{r}-1)T_c-\tau_k) & \dots & \widetilde{\psi}((\frac{(r-1)}{r}-N+1)T_c-\tau_k) \\
\hline
\widetilde{\psi}(T_c-\tau_k) & \widetilde{\psi}(-\tau_k) & \dots & \widetilde{\psi}((-N+2)T_c-\tau_k) \\
\vdots & \vdots & \vdots & \vdots \\
\widetilde{\psi}(\frac{(r-1)T_c}{r}+T_c-\tau_k) & \widetilde{\psi}((\frac{(r-1)}{r})T_c-\tau_k) & \dots & \widetilde{\psi}((\frac{(r-1)}{r}-N+2)T_c-\tau_k) \\
\hline
\vdots & \vdots & \vdots & \vdots \\
\hline
\widetilde{\psi}(N-1)T_c-\tau_k) & \widetilde{\psi}((N-2)T_c-\tau_k) & \dots & \widetilde{\psi}(-\tau_k) \\
\vdots & \vdots & \vdots & \vdots \\
\widetilde{\psi}(\frac{r-1}{r}T_c+(N-1)T_c-\tau_k) & \widetilde{\psi}((\frac{r-1}{r}+N-2)T_c-\tau_k) & \dots & \widetilde{\psi}((\frac{r-1}{r})T_c-\tau_k)
\end{array} \right). \tag{4.30}$$

[26] To have an idea of the decaying rate the reader can consider the sinc function that is the slowest decaying function among the raised-cosine functions.

Structuring the matrix $\widetilde{\boldsymbol{\Psi}}_k$ in blocks of dimension $r \times 1$, $\widetilde{\boldsymbol{\Psi}}_k$ is a block-wise Toeplitz matrix.

Let $\widetilde{\boldsymbol{S}}$ be the $rN \times K$ matrix of virtual[27] spreading, i.e. $\widetilde{\boldsymbol{S}} = (\widetilde{\boldsymbol{\Psi}}_1 \boldsymbol{s}_1, \widetilde{\boldsymbol{\Psi}}_2 \boldsymbol{s}_2, \ldots, \widetilde{\boldsymbol{\Psi}}_K \boldsymbol{s}_K)$. Furthermore, let $\boldsymbol{A}$ be the $K \times K$ diagonal matrix of received amplitudes, and $\boldsymbol{b}$ the vector of transmitted symbols. Then, the system model in matrix notation is given by

$$\begin{aligned} \widetilde{\boldsymbol{y}} &= \widetilde{\boldsymbol{S}}\boldsymbol{A}\boldsymbol{b} + \widetilde{\boldsymbol{n}} \qquad (4.31)\\ &= \widetilde{\boldsymbol{H}}\boldsymbol{b} + \widetilde{\boldsymbol{n}} \end{aligned}$$

with $\widetilde{\boldsymbol{H}} = \widetilde{\boldsymbol{S}}\boldsymbol{A}$. Additionally, $\widetilde{\boldsymbol{h}}_k$ denotes the k^{th} column of the matrix $\widetilde{\boldsymbol{H}}$. $\widetilde{\boldsymbol{T}}$ and $\widetilde{\boldsymbol{R}}$ are the correlation matrices defined as $\widetilde{\boldsymbol{T}} = \widetilde{\boldsymbol{H}}\widetilde{\boldsymbol{H}}^H$ and $\widetilde{\boldsymbol{R}} = \widetilde{\boldsymbol{H}}^H\widetilde{\boldsymbol{H}}$, respectively.

4.4.2 Linear Detection

Following the same approach as in Chapter 3, the linear MMSE detector and the multistage detector Type J-I for symbol quasi synchronous but chip asynchronous CDMA systems are given by

$$\begin{aligned} \widehat{\boldsymbol{b}}_{\text{MMSE}} &= \widetilde{\boldsymbol{H}}^H(\widetilde{\boldsymbol{T}} + \sigma^2\boldsymbol{I})^{-1}\widetilde{\boldsymbol{y}} \\ &= (\widetilde{\boldsymbol{R}} + \sigma^2\boldsymbol{I})^{-1}\widetilde{\boldsymbol{H}}^H\widetilde{\boldsymbol{y}} \end{aligned}$$

and

$$\widehat{b}_{J-I,k} = \sum_{m=0}^{M-1} (\widetilde{\boldsymbol{w}}_k)_m \widetilde{\boldsymbol{h}}^H \widetilde{\boldsymbol{T}}^m \widetilde{\boldsymbol{y}} \qquad (4.32)$$

respectively. The weights for the detection of the k^{th} symbol are obtained as

$$\widetilde{\boldsymbol{w}}_k = \widetilde{\boldsymbol{\Phi}}_k^{-1} \widetilde{\boldsymbol{\varphi}}_k$$

with

$$\widetilde{\boldsymbol{\Phi}}_k = \begin{pmatrix} (\widetilde{\boldsymbol{R}}^2)_{kk} + \sigma^2(\widetilde{\boldsymbol{R}})_{kk} & \ldots & (\widetilde{\boldsymbol{R}}^{M+1})_{kk} + \sigma^2(\widetilde{\boldsymbol{R}}^M)_{kk} \\ (\widetilde{\boldsymbol{R}}^3)_{kk} + \sigma^2(\widetilde{\boldsymbol{R}}^2)_{kk} & \ldots & (\widetilde{\boldsymbol{R}}^{M+2})_{kk} + \sigma^2(\widetilde{\boldsymbol{R}}^{M+1})_{kk} \\ \ldots & \ldots & \ldots \\ (\widetilde{\boldsymbol{R}}^{M+1})_{kk} + \sigma^2(\widetilde{\boldsymbol{R}}^M)_{kk} & \ldots & (\widetilde{\boldsymbol{R}}^{2M})_{kk} + \sigma^2(\widetilde{\boldsymbol{R}}^{2M-1})_{kk} \end{pmatrix}$$

$$\widetilde{\boldsymbol{\varphi}}_k = \left((\widetilde{\boldsymbol{R}})_{kk}, (\widetilde{\boldsymbol{R}}^2)_{kk}, \ldots, (\widetilde{\boldsymbol{R}}^M)_{kk}\right)^T. \qquad (4.33)$$

[27]Here, the adjective 'virtual' refers to the fact that $\widetilde{\boldsymbol{S}}$ takes into account the asynchronism and the chip-pulse waveform.

The structure of the Type J-I detector coincides with the structure proposed in Figure 3.2 for synchronous systems by substituting $\boldsymbol{H}$ with $\widetilde{\boldsymbol{H}}$. Once again the design of the asymptotic weights reduces to the computation of the asymptotic values $\widetilde{R}^s_{kk,\infty} = \lim_{K=\beta N \to +\infty}(\widetilde{\boldsymbol{R}}^s)_{kk}$, for $s = 1, 2, \ldots, 2M$.

For further studies it is convenient to define the concept of r-block-wise circulant matrices of order N:

Definition 2 *Let r and N be positive integers. An r-block-wise circulant matrix of order N is an $rN \times N$ matrix of the form*

$$\boldsymbol{C}^{(N)} = \left(\begin{array}{cccc} c_{1,0} & c_{1,1} & \cdots & c_{1,N-1} \\ \vdots & \vdots & \cdots & \vdots \\ c_{r,0} & c_{r,1} & \cdots & c_{r,N-1} \\ \hline c_{1,N-1} & c_{1,0} & \cdots & c_{1,N-2} \\ \vdots & \vdots & \cdots & \vdots \\ c_{r,N-1} & c_{r,0} & \cdots & c_{r,N-2} \\ \hline \cdots & \cdots & \cdots & \cdots \\ \hline c_{1,1} & c_{1,2} & \cdots & c_{1,0} \\ \vdots & \vdots & \cdots & \vdots \\ c_{r,1} & c_{r,2} & \cdots & c_{r,0} \end{array} \right).$$

In the matrix $\boldsymbol{C}^{(N)}$ an $r \times N$ block row is obtained by circularly right shifting of the previous block. Since the matrix $\boldsymbol{C}^{(N)}$ is univocally defined by the unitary Fourier transforms of the sequence[28] $\{c_{s,0}, c_{s,1}, \ldots c_{s,N-1}\}$

$$f_s(x) = \frac{1}{\sqrt{2\pi}} \sum_{k=0}^{N-1} c_{sk} \mathrm{e}^{-j2\pi xk} \qquad s = 1, \ldots, r,$$

we will denote an r-block-wise circulant matrix of order N by $\boldsymbol{C}^{(N)}(f_1(x), f_2(x), \ldots f_r(x))$ in the following[29].

Let $\overline{\boldsymbol{H}}$ be a matrix of form similar to the form of matrix $\widetilde{\boldsymbol{H}}$ but with the block-wise Toeplitz matrices $\widetilde{\boldsymbol{\Psi}}_k$ replaced by block-wise circulant matrices $\boldsymbol{C}_k$, i.e. $\overline{\boldsymbol{H}} =$

[28]Throughout this chapter we utilize unitary Fourier transforms both in the continuous time and in the discrete time domain. With this choice the functions of the complete orthogonal system for the direct Fourier transform are simply the complex conjugate of functions of the complete orthogonal system for the inverse Fourier transform. The unitary Fourier transform of a function $f(t)$ in the continuous time domain is given by $F(j\omega) = \frac{1}{\sqrt{2\pi}} \int f(t)\mathrm{e}^{-j\omega t}\mathrm{d}t$. The unitary Fourier transform of a sequence $\{\ldots, c_{-1}, c_0, c_1, \ldots\}$ in the discrete time domain is given by $C(\mathrm{e}^{j\omega}) = \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{+\infty} c_n \mathrm{e}^{-j\omega n}$.

[29]In this section we denote the argument of a Fourier transform of a continuous time function by f and the argument of a Fourier transform of a sequence by x.

$(\boldsymbol{C}_1\boldsymbol{s}_1, \boldsymbol{C}_2\boldsymbol{s}_2, \ldots, \boldsymbol{C}_K\boldsymbol{s}_K)\boldsymbol{A}$. The matrices $\boldsymbol{C}_k$, $k = 1, \ldots, K$ are chosen in a way such that the spectrum of $\widetilde{\boldsymbol{\Psi}}_k$, $k = 1, \ldots, K$ converges asymptotically to the spectrum of $\boldsymbol{C}_k$, $k = 1, \ldots, K$. This choice will be thoroughly discussed in the following. Firstly, we determine the asymptotic values of the diagonal elements of the matrices $\overline{\boldsymbol{R}}^\ell = (\overline{\boldsymbol{H}}^H\overline{\boldsymbol{H}})^\ell$, $\ell \in \mathbb{Z}^+$, as $N, K \to \infty$ with $\frac{K}{N} \to \beta$. Then, we extend the results to the matrices $\widetilde{\boldsymbol{R}}^\ell$, $\ell \in \mathbb{Z}^+$.

The convergence of the diagonal elements of $\overline{\boldsymbol{R}}^\ell$ to deterministic values is established in the following theorem.

Theorem 10 *Let $\boldsymbol{A} \in \mathbb{C}^{K\times K}$ be a diagonal matrix with k^{th} diagonal element a_{kk}. Given a function $\Xi(j2\pi f) : \mathbb{R} \to \mathbb{C}$ with finite support and bounded in absolute value*[30]*, let*

$$\xi_\tau(x) \triangleq \frac{1}{T_c} \sum_{s=-\infty}^{+\infty} e^{j2\pi \frac{\tau}{T_c}(x+s)} \Xi^* \left(\frac{j2\pi}{T_c}(x+s) \right).$$

Furthermore, let $\boldsymbol{C}_k^{(N)}$, $k = 1, \ldots, K$, be K r-block-wise circulant matrices of order N defined by

$$\boldsymbol{C}_k^{(N)} \triangleq \boldsymbol{C}_k^{(N)}(\xi_{\tau_k}(x), \xi_{\tau_k - \frac{T_c}{r}}(x), \ldots, \xi_{\tau_k - \frac{(r-1)T_c}{r}}(x)),$$

where r is a positive integer[31]*. Finally, let $\overline{\boldsymbol{H}} = \overline{\boldsymbol{S}}\boldsymbol{A}$ with $\overline{\boldsymbol{S}} = (\boldsymbol{C}_1^{(N)}\boldsymbol{s}_1, \boldsymbol{C}_2^{(N)}\boldsymbol{s}_2, \ldots, \boldsymbol{C}_K^{(N)}\boldsymbol{s}_K)$.*

Let us assume that

(a) $\boldsymbol{s}_k$, for $k = 1, \ldots, K$, are K independent N-dimensional column vectors with i.i.d. random elements $s_{nk} \in \mathbb{C}$ such that $\mathrm{E}\{s_{nk}\} = 0$, $\mathrm{E}\{|s_{nk}|^2\} = \frac{1}{N}$, and $\lim_{N\to\infty} \mathrm{E}\{N^3 |s_{nk}^{(N)}|^6\} < +\infty$.[32]

(b) $(\tau_1, \tau_2, \ldots, \tau_K)$ is a sequence of K random variables with $\tau_k \in [0, T_c)$ and T_c positive real[33]*.*

(c) The sequence of the empirical joint distributions $F^{(K)}_{|\boldsymbol{A}|^2,T}(\lambda, \tau) = \frac{1}{K}\sum_{k=1}^K 1(\lambda - |a_{kk}|^2)1(\tau - \tau_k)$ converges almost surely, as $K \to \infty$, to a non-random distribution function $F_{|\boldsymbol{A}|^2,T}(\lambda, \tau)$ with bounded support.

---

[30] $\Xi(j2\pi f)$ models the unitary Fourier transform of the chip pulse waveform $\widetilde{\psi}(t)$.

[31] r is a sampling factor that defines the sampling rate as $\frac{r}{T_c}$. $\boldsymbol{C}_k^{(N)}$ takes into account the chip-pulse waveforms.

[32] These random column vectors model the spreading sequences.

[33] τ_k corresponds to the time delay of user k.

(d) *The spectral radius of the matrix* $\overline{\boldsymbol{R}} = \overline{\boldsymbol{H}}^H \overline{\boldsymbol{H}}$ *is almost surely upper bounded as* $K, N \to +\infty$ *with*[34] $\frac{K}{N} \to \beta$.

(e) $K = K(N)$ *with* $\lim_{N\to\infty} \frac{K(N)}{N} = \beta$.

Then, given $(|a_{kk}|^2, \tau_k)$, *the* k^{th} *diagonal element of the matrix* $\overline{\boldsymbol{R}}^\ell$ *converges with probability one to a deterministic value, conditionally on* $(|a_{kk}|^2, \tau_k)$,

$$\lim_{K=\beta N\to\infty} (\overline{\boldsymbol{R}}^\ell)_{kk} \stackrel{a.s.}{=} \overline{R}^\ell(|a_{kk}|^2, \tau_k)$$

with $\overline{R}(|a_{kk}|^2, \tau_k)$ *determined by the following recursion*

$$\overline{R}^\ell(\lambda, \tau) = \sum_{s=0}^{\ell-1} g(\overline{\boldsymbol{T}}^{\ell-s-1}, \lambda, \tau)\overline{R}^s(\lambda, \tau) \tag{4.34}$$

and

$$\overline{\boldsymbol{T}}^\ell(x) = \sum_{s=0}^{\ell-1} \mathrm{f}(\overline{\boldsymbol{R}}^{\ell-s-1}, x)\overline{\boldsymbol{T}}^s(x) \qquad 0 \le x \le 1 \tag{4.35}$$

$$\mathrm{f}(\overline{\boldsymbol{R}}^s, x) = \beta \int \lambda \boldsymbol{\Delta}_\tau(x)\boldsymbol{\Delta}_\tau^H(x)\overline{R}^s(\lambda, \tau)\mathrm{d}\, F_{|\boldsymbol{A}|^2,T}(\lambda, \tau) \qquad 0 \le x \le 1 \tag{4.36}$$

$$g(\overline{\boldsymbol{T}}^s, \lambda, \tau) = \lambda \int_0^1 \boldsymbol{\Delta}_\tau^H(x)\overline{\boldsymbol{T}}^s(x)\boldsymbol{\Delta}_\tau(x)\mathrm{d}\, x \tag{4.37}$$

where

$$\boldsymbol{\Delta}_\tau(x) = \begin{pmatrix} \xi_\tau(x) \\ \xi_{\tau - \frac{T_c}{r}}(x) \\ \vdots \\ \xi_{\tau - \frac{T_c(r-1)}{r}}(x) \end{pmatrix}.$$

The recursion is initialized by setting $\overline{\boldsymbol{T}}^0(x) = \boldsymbol{I}_r$ *and* $\overline{R}^0(\lambda, \tau) = 1$.

Theorem 10 is proven in Appendix C Section C.3.

For $r = 1$ and $F_{|\boldsymbol{A}|^2,T}(\lambda, \tau) = F_{|\boldsymbol{A}|^2}(\lambda)\delta(\tau)$, i.e. for synchronous systems sampled at the chip rate with $\xi_0(x) = \mathrm{sinc}(x)$ it can be verified that equation (4.34) reduces to equation (3.26). Equation (4.37) becomes $g(\overline{\boldsymbol{T}}^s, \lambda, \tau) = \lambda\beta m_{\boldsymbol{R}}^s$ and the recursion in Theorem 10 coincides with the recursion in Theorem 6 for synchronous systems.

[34] This condition can be replaced by the less restrictive condition that the integer positive eigenvalue moments are upper bounded.

Asymptotically, there exists a strong relation between Toeplitz and circulant matrices. In fact, given a sequence u_m, $m \in \mathbb{Z}$ that is square summable[35], i.e. $\sum_{m=-\infty}^{+\infty} |u_m|^2 < +\infty$, and given the sequence $\{\boldsymbol{U}_N\}$ of Toeplitz matrices

$$\boldsymbol{U}_N = \begin{pmatrix} u_0 & u_{-1} & u_{-2} & \dots & u_{-(N-1)} \\ u_1 & u_0 & u_{-1} & \dots & u_{-(N-2)} \\ \dots & \dots & \dots & \dots & \dots \\ u_{N-1} & u_{N-2} & u_{N-3} & \dots & u_0 \end{pmatrix}$$

the sequence of the empirical eigenvalue distributions of the Toeplitz matrices $\boldsymbol{U}_N$ and of the empirical eigenvalue distributions of the circulant matrices $\{\boldsymbol{C}_N(\frac{1}{\sqrt{2\pi}} \sum_{m=-\infty}^{+\infty} x_m \mathrm{e}^{j2\pi m x})\}$ converge to the same limit eigenvalue distribution for $N \to \infty$ [90–92].

Let us consider the matrix $\widetilde{\boldsymbol{\Psi}}_k^{(N)}$, with $r = 1$ in (4.30) and let us determine the asymptotically equivalent circulant matrix. If $\Xi(j2\pi f)$ is the unitary Fourier transform of $\widetilde{\psi}(t)$, then, $\mathrm{e}^{j2\pi f \tau_k}\Xi(j2\pi f)$ is the unitary Fourier transform of $\widetilde{\psi}(t - \tau_k)$ and

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} \sum_{m=-\infty}^{+\infty} \widetilde{\psi}(mT_c - \tau_k)\mathrm{e}^{j2\pi m x} &= \left(\frac{1}{\sqrt{2\pi}} \sum_{m=-\infty}^{+\infty} \widetilde{\psi}(mT_c - \tau_k)\mathrm{e}^{-j2\pi m x} \right)^* \\ &= \left(\frac{1}{T_c} \sum_{k=-\infty}^{+\infty} \Xi\left(j2\pi \frac{x+k}{T_c} \right) \mathrm{e}^{-j2\pi \frac{\tau_k}{T_c}(x+k)} \right)^* \end{aligned} \tag{4.38}$$

where, in the first step, we assume $\widetilde{\psi}(t)$ to be real and then we make use of the property of the complex number z that $z = (z^*)^*$. Let us notice that $\left(\frac{1}{\sqrt{2\pi}} \sum_{m=-\infty}^{+\infty} \widetilde{\psi}(mT_c - \tau_k)\mathrm{e}^{-j2\pi m x}\right)$ is the unitary Fourier transform of the sequence $\{\widetilde{\psi}(mT_c - \tau_k)\}$. Then, it is possible to use the relation between the unitary Fourier transform of a sequence obtained by sampling a given function at rate $1/T_c$ and the unitary Fourier transform of the continuous time function [93]. Thus, the sequence of the eigenvalue distribution of the Toeplitz matrices $\widetilde{\boldsymbol{\Psi}}_k^{(N)}$ converges to the limiting eigenvalue distribution of the sequence of circulant matrices $C_N\left(\frac{1}{T_c} \sum_{m=-\infty}^{\infty} \mathrm{e}^{j2\pi \frac{\tau_k}{T_c}(x+m)} \Xi^*(j2\pi \frac{x+m}{T_c})\right)$. This property suggests the following conjecture.

Proposition 1 *Let assumptions (a)-(e) of Theorem 10 be satisfied and let the definitions in Theorem 10 hold. Additionally, assume:*

[35]This condition is very general and includes almost all chip-pulse waveforms of practical interest. In fact, it coincides with the condition for the existence of the Fourier transform of a sequence.

(i) Given the function $\widetilde{\psi}(t) : \mathbb{R} \to \mathbb{R}$ *with unitary Fourier transform* $\Xi(j2\pi f)$ *as in Theorem 10, the sequence* $\{\widetilde{\psi}(T_c n - \tau)\}$ *is square root summable, i.e.,* $\sum_{n=-\infty}^{+\infty} |\widetilde{\psi}(T_c n - \tau)|^2 < \infty$*, for any* $\tau \in [0, T_c]$.

(ii) Let $\widetilde{\boldsymbol{\Psi}}_k^{(N)}$ *be as in (4.30).*

(iii) Let $\widetilde{\boldsymbol{S}}^{(N)} = (\widetilde{\boldsymbol{\Psi}}_1^{(N)} \boldsymbol{s}_1, \widetilde{\boldsymbol{\Psi}}_2^{(N)} \boldsymbol{s}_2, \ldots \widetilde{\boldsymbol{\Psi}}_K^{(N)} \boldsymbol{s}_K)$ *and* $\widetilde{\boldsymbol{H}}^{(N)} = \widetilde{\boldsymbol{S}}^{(N)} \boldsymbol{A}^K$.

Then, the diagonal elements of the matrix $(\widetilde{\boldsymbol{R}})^\ell = (\widetilde{\boldsymbol{H}}^H \widetilde{\boldsymbol{H}})^\ell$, $\ell \in \mathbb{Z}^+$, *converge asymptotically to the corresponding diagonal elements of the matrix* $\overline{\boldsymbol{R}}^\ell$, *i.e. conditionally on* $(|a_{kk}|^2, \tau_k)$

$$\lim_{K=\beta N \to +\infty} \widetilde{\boldsymbol{R}}_{kk}^\ell = \overline{R}^\ell(|a_{kk}|^2, \tau_k).$$

Proposition 1 can be proven under the conjecture that not only the spectrum but also the eigenvectors of Toeplitz matrices converge asymptotically to the eigenvectors of circulant matrices. Implicitly, this last conjecture has been already used in the proof of Theorem 6 in [13]. Proposition 1 is supported by numerical simulations. In Figure 4.9, the empirical distributions of $\widetilde{\epsilon}_s = \frac{(\widetilde{\boldsymbol{R}}^{(512)})_{kk}^s - \overline{R}^s(|a_{kk}|^2, \tau_k)}{\overline{R}^s(|a_{kk}|^2, \tau_k)}$, $s = 2, 3, 5, 6$, i.e. the empirical distribution of the relative error made by approximating the diagonal elements of powers of the 512×512 matrix $\widetilde{\boldsymbol{R}}^{(512)}$ by the asymptotic values $\overline{R}^s(|a_{kk}|^2, \tau_k)$, are compared to the corresponding empirical distributions of $\overline{\epsilon}_s = \frac{(\overline{\boldsymbol{R}}^{(512)})_{kk}^s - \overline{R}^s(|a_{kk}|^2, \tau_k)}{\overline{R}^s(|a_{kk}|^2, \tau_k)}$, $s = 2, 3, 5, 6$. The empirical distributions of $\widetilde{\epsilon}_s$ and $\overline{\epsilon}_s$ have been obtained assuming that $\widetilde{\psi}(t)$ is a raised cosine waveform with roll-off factor $\gamma = 0.5$, using a system load $\beta = 0.5$ and sampling factor $r = 2$. The empirical distributions of $\widetilde{\epsilon}_s$, $s = 2, 3, 5, 6$ match very well the empirical distributions of the corresponding $\overline{\epsilon}_s$. Later on, the conjecture in Proposition 1 will be further supported by the comparison between the asymptotic performance and the performance of finite communication systems (e.g. Figure 4.10).

The asymptotic SINR of the multistage detector Type J-I for symbol quasi-synchronous but chip asynchronous systems is obtained from (3.37) by replacing $\boldsymbol{\varphi}_k^\infty$ and $\boldsymbol{\Phi}_k^\infty$ by $\widetilde{\boldsymbol{\varphi}}_k^\infty = \lim_{N=\beta K \to \infty} \widetilde{\boldsymbol{\varphi}}_k$ and $\widetilde{\boldsymbol{\Phi}}_k^\infty = \lim_{N=\beta K \to \infty} \widetilde{\boldsymbol{\Phi}}_k$, respectively. Then,

$$\mathrm{SINR}_{J-I,k}^\infty = \frac{1}{\frac{1}{(\widetilde{\boldsymbol{\varphi}}_k^\infty)^T (\widetilde{\boldsymbol{\Phi}}_k^\infty)^{-1} \widetilde{\boldsymbol{\varphi}}_k^\infty} - 1}. \tag{4.39}$$

The performance of the linear MMSE detector can be approximated with arbitrarily high precision by the performance of the multistage detector Type J-I with a sufficiently large number of stages, as already discussed in Chapter 3.

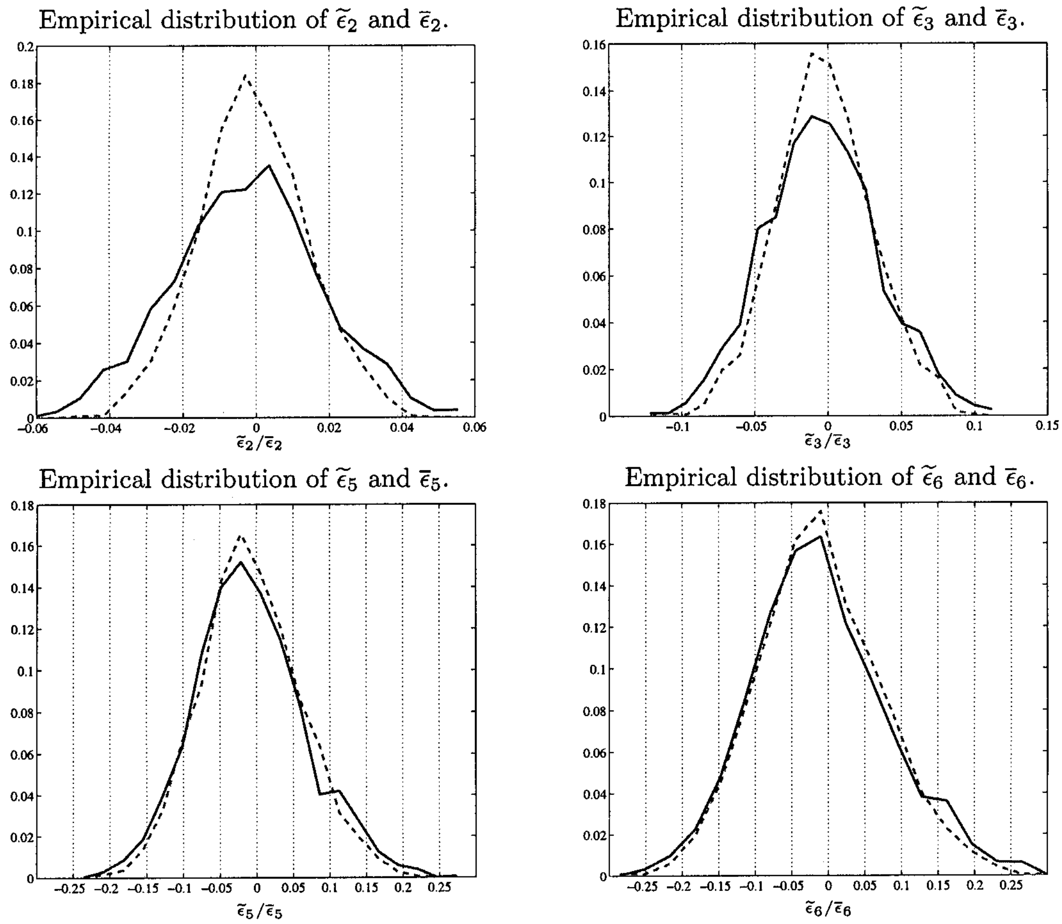


Figure 4.9: Comparison of the empirical distributions of $\widetilde{\epsilon}_s = \dfrac{(\widetilde{\boldsymbol{R}}^{(512)})^s_{kk} - \overline{R}^s(|a|^2_{kk}, \tau_k)}{\overline{R}^s(|a|^2_{kk}, \tau_k)}$ (solid lines) and $\overline{\epsilon}_s = \dfrac{(\overline{\boldsymbol{R}}^{(512)})^s_{kk} - \overline{R}^s(|a|^2_{kk}, \tau_k)}{\overline{R}^s(|a|^2_{kk}, \tau_k)}$ (dashed lines) for $s = 2, 3, 5, 6$.

Let us consider a chip asynchronous but symbol quasi synchronous CDMA system with equal received powers, raised cosine chip pulse waveform of roll-off factor $\gamma = 0.5$, system load $\beta = \frac{1}{2}$, and a Type J-I detector with $M = 4$. In Figure 4.10 the performance of finite systems with increasing dimension ($K = 32$, $K = 128$, and $K = 1024$) is compared to the asymptotic performance. As expected, the variance of the output SINR for the finite systems decreases as the system size increases and the SINRs for the finite systems concentrate more and more around the asymptotic values computed by (4.39).

4.4.3 Effects of Asynchronism and of Chip Pulse Waveforms

Theorem 10 and Proposition 1 are not useful only for the design of multistage detectors with asymptotic weights. In fact, their application to the analysis of CDMA systems with random spreading and linear detectors optimal in a minimum mean square error sense demonstrates general properties of these systems.

In the following we focus on three cases:

- Chip pulse waveforms with bandwidth $B \leq \frac{1}{2T_c}$.
- Chip pulse waveforms with bandwidth $\frac{1}{2T_c} \leq B \leq \frac{1}{T_c}$.
- CDMA systems using square root Nyquist chip pulse waveforms and chip matched filtering at the front-end.

Chip pulse waveforms with $B \leq \frac{1}{2T_c}$

Chip pulse waveforms $\widetilde{\psi}(t)$ with bandwidth B non greater than $\frac{1}{2T_c}$ are analyzed. Thus, the conditions of the sampling theorem are satisfied when sampling at a rate that is equal to or a multiple of the chip rate, the samples are sufficient statistics for the detection of all users in the system.

Denoting by $\Xi(j2\pi x)$ the unitary Fourier transform of the function $\widetilde{\psi}(t)$, the unitary Fourier transform of $\widetilde{\psi}(t-\tau_k)$ is given by $\mathrm{e}^{-j2\pi\tau_k x}\Xi(j2\pi x)$. The r-blockwise circulant matrices of order N asymptotically equivalent to the block-wise Toeplitz matrix (4.30), are

$$\boldsymbol{C}_k = \boldsymbol{C}_k\left(\frac{\mathrm{e}^{j2\pi\tau_k x}}{T_c}\Xi^*\left(j2\pi\frac{x}{T_c}\right), \frac{\mathrm{e}^{j2\pi(\tau_k - \frac{T_c}{r})x}}{T_c}\Xi^*\left(j2\pi\frac{x}{T_c}\right), \dots \frac{\mathrm{e}^{j2\pi(\tau_k - \frac{(r-1)T_c}{r})x}}{T_c}\Xi^*\left(j2\pi\frac{x}{T_c}\right)\right).$$

By specializing the vector $\boldsymbol{\Delta}_\tau(x)$ in Theorem 10 to the case $B \leq \frac{1}{2T_c}$, we obtain, for any time delay τ

$$\boldsymbol{\Delta}_\tau(x) = \frac{\mathrm{e}^{j2\pi\tau_k x}}{T_c}\Xi^*(j2\pi x)\boldsymbol{e}$$

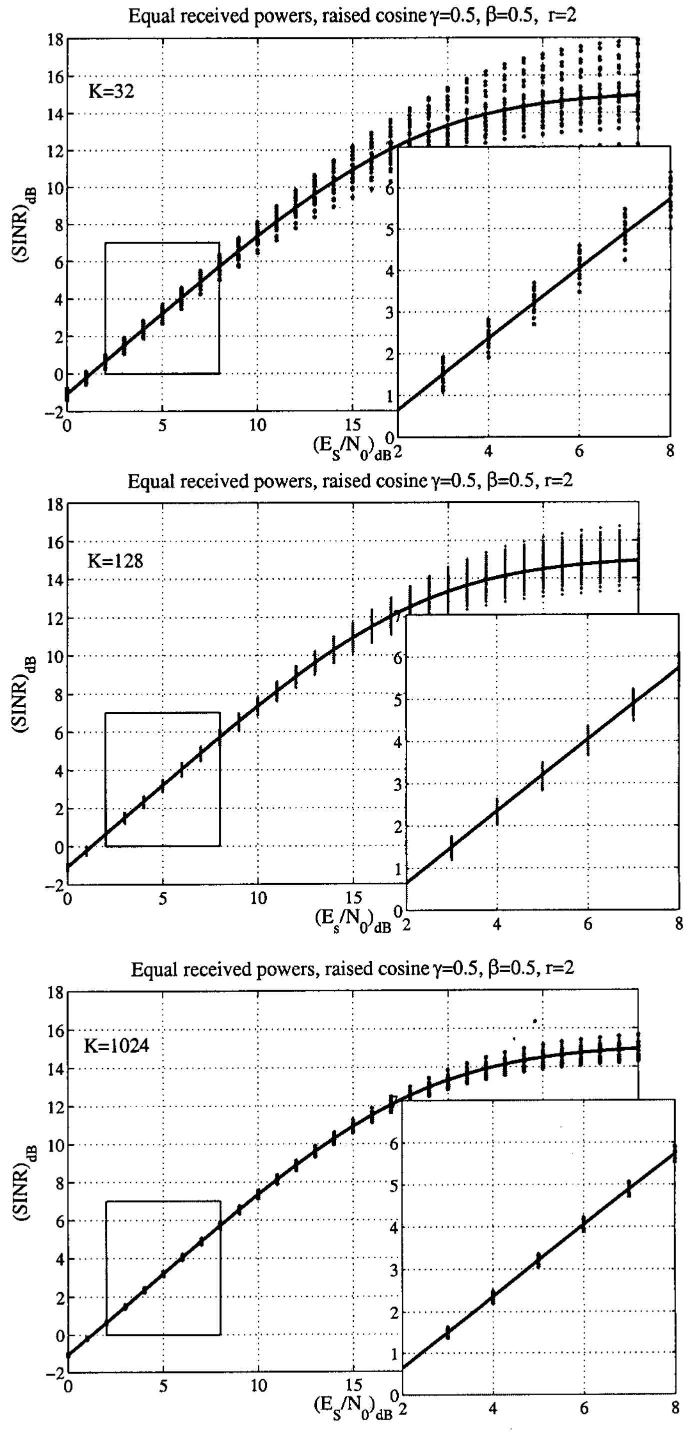


Figure 4.10: SINR_{dB} versus $(\frac{E_s}{N_0})_{\text{dB}}$ as $K = \beta N \to \infty$ (solid line) and with a finite number of users (red dots), namely $K = 32$, $K = 128$, $K = 1024$. A chip asynchronous but symbol quasi synchronous CDMA system with equal received powers and $\beta = \frac{1}{2}$ is considered.

with $\boldsymbol{e} = (1, \mathrm{e}^{-j2\pi\frac{T_c}{r}x}, \dots \mathrm{e}^{-j2\pi\frac{T_c(r-1)}{r}x})^T$.

Theorem 10 applied to the case $B \leq \frac{1}{2T_c}$ yields the following algorithm to derive $\overline{R}^{\ell}(\lambda)$ and the asymptotic eigenvalue moments of the matrix $\overline{\boldsymbol{R}}$.

Algorithm 2

Initialization: *Let $\rho_0(z) = 1$ and $\mu_0(y) = \frac{1}{r}$.*

l^{th} step:

- *Define $u_{\ell-1}(y) = y\mu_{\ell-1}(y)$ and write it as a polynomial in y.*
- *Define $v_{\ell-1}(z) = z\rho_{\ell-1}(z)$ and write it as a polynomial in z.*
- *Define $\mathcal{E}_s = \frac{1}{T_c^{2s}}\int_{-1/2}^{1/2}\left|\Xi(j2\pi\frac{x}{T_c})\right|^{2s}\mathrm{d}\,x$ and replace all monomials $y, y^2, \dots, y^{\ell}$ in the polynomial $u_{\ell-1}(y)$ by $\mathcal{E}_1, \mathcal{E}_2, \dots, \mathcal{E}_{\ell}$, respectively. Denote the result by $U_{\ell-1}$.*
- *Define $m^s_{|\boldsymbol{A}|^2} = \mathrm{E}\{|a_{kk}|^{2s}\}$ and replace all monomials $z, z^2, \dots, z^{\ell}$ in the polynomial $v_{\ell-1}(z)$ by the moments $m^1_{|\boldsymbol{A}|^2}$, $m^2_{|\boldsymbol{A}|^2}, \dots, m^{\ell}_{|\boldsymbol{A}|^2}$, respectively. Denote the result by $V_{\ell-1}$.*
- *Calculate*

$$\rho_{\ell}(z) = \sum_{s=0}^{\ell-1} r^2 z U_{\ell-s-1}\rho_s(z)$$
$$\mu_{\ell}(y) = r\sum_{s=0}^{\ell-1}\beta y V_{\ell-s-1}\mu_s(y).$$

- *Assign $\rho_{\ell}(\lambda)$ to $\overline{R}^{\ell}(\lambda)$.*
  Replace all monomials $z, z^2, \dots, z^{\ell}$ in the polynomial $\rho_{\ell}(z)$ by the moments $m^1_{|\boldsymbol{A}|^2}$, $m^2_{|\boldsymbol{A}|^2}, \dots, m^{\ell}_{|\boldsymbol{A}|^2}$, respectively, and assign the result to $m^{\ell}_{\overline{\boldsymbol{R}}}$.

The derivation of Algorithm 2 is detailed in Appendix C Section C.4.

The limit diagonal element $\widetilde{R}^{\ell}(\lambda)$ of $\widetilde{\boldsymbol{R}}^{\ell}$ and the eigenvalue moments $m^{\ell}_{\widetilde{\boldsymbol{R}}}$ equal $\overline{R}^{\ell}(\lambda)$ and $m^{\ell}_{\overline{\boldsymbol{R}}}$, respectively thanks to Proposition 1.

By applying Algorithm 2 and Proposition 1 we compute the first five asymptotic eigenvalue moments

$$
\begin{aligned}
m^1_{\widetilde{R}} &= r m^1_{|A|^2} \mathcal{E}_1 \\
m^2_{\widetilde{R}} &= r^2[\beta(m^1_{|A|^2})^2 \mathcal{E}_2 + m^2_{|A|^2} \mathcal{E}_1^2] \\
m^3_{\widetilde{R}} &= r^3[\beta^2 \mathcal{E}_3 (m^1_{|A|^2})^3 + 3 m^2_{|A|^2} \mathcal{E}_2 \beta m^1_{|A|^2} \mathcal{E}_1 + m^3_{|A|^2} \mathcal{E}_1^3] \\
m^4_{\widetilde{R}} &= r^4[2\beta^2 \mathcal{E}_2^2 m^2_{|A|^2} (m^1_{|A|^2})^2 + 4\beta \mathcal{E}_1^2 \mathcal{E}_2 m^3_{|A|^2} m^1_{|A|^2} + 4\beta^2 \mathcal{E}_1 \mathcal{E}_3 m^2_{|A|^2} (m^1_{|A|^2})^2 + \beta^3 \mathcal{E}_4 (m^1_{|A|^2})^4 \\
&\quad + 2\beta \mathcal{E}_1^2 \mathcal{E}_2 (m^2_{|A|^2})^2 + \mathcal{E}_1^4 m[4]] \\
m^5_{\widetilde{R}} &= r^5[m^5_{|A|^2} \mathcal{E}_1^5 \beta^4 + \mathcal{E}_5 (m^1_{|A|^2})^5 + 5\beta^3 \mathcal{E}_1 \mathcal{E}_4 m^2_{|A|^2} (m^1_{|A|^2})^3 + 5\beta^3 \mathcal{E}_3 \mathcal{E}_2 m^2_{|A|^2} (m^1_{|A|^2})^3 \\
&\quad + 5\beta^2 \mathcal{E}_3 \mathcal{E}_1^2 m^3_{|A|^2} (m^1_{|A|^2})^2 + 5\beta^2 \mathcal{E}_1^2 \mathcal{E}_3 (m^2_{|A|^2})^2 m^1_{|A|^2} + 5\beta^2 \mathcal{E}_1 \mathcal{E}_2^2 (m^2_{|A|^2})^2 m^1_{|A|^2} \\
&\quad + 5\beta^2 \mathcal{E}_2^2 \mathcal{E}_1 m^3_{|A|^2} (m^1_{|A|^2})^2 + 5\beta \mathcal{E}_2 \mathcal{E}_1^3 m^4_{|A|^2} m^1_{|A|^2} + 5 \mathcal{E}_2 \mathcal{E}_1^3 m^3_{|A|^2} m^2_{|A|^2}].
\end{aligned}
$$

For ideal sinc functions with bandwidth $B = \frac{1}{2T_c}$, we have $\mathcal{E}_s = 1$, $s = 1, 2, \ldots$. It can here be verified that $\widetilde{R}^\ell(|a_{kk}|^2) = r^\ell R^\ell(|a_{kk}|^2)$, where $R^\ell(|a_{kk}|^2)$ are the asymptotic diagonal elements obtained by Algorithm 1 for synchronous systems.

In general, the eigenvalue moments of $\widetilde{\boldsymbol{R}}$ depend only on the system load β, the sampling rate $\frac{\tau}{T_c}$, the eigenvalue distribution of the matrix $\boldsymbol{A}^H \boldsymbol{A}$, and $\mathcal{E}_s$, $s \in \mathbb{Z}^+$. The latter coefficients take into account the effects of the shape of the chip pulse or, equivalently, of the frequency spectrum of the function $\widetilde{\psi}(t)$. The diagonal elements $\widetilde{R}^\ell(|a_{kk}|^2)$ and the eigenvalue moments $m^\ell_{\widetilde{R}}$ are also independent of the delay distribution. In particular, Algorithm 2 can be applied also to synchronous systems with or without oversampling and any kind of chip-pulse waveform provided that $B \leq \frac{1}{T_c}$. Since the performance of the large class of linear detectors that admit a representation as multistage detectors depends only on the diagonal elements $\widetilde{R}^\ell(|a_{kk}|^2)$ and the eigenvalue moments $m^\ell_{\widetilde{R}}$, we can state the following corollary.

Corollary 4 *Let the assumptions of Proposition 1 be satisfied. Additionally, let us assume that B, the bandwidth of $\widetilde{\psi}(t)$, satisfies the constraint $B \leq \frac{1}{2T_c}$. Then, the asymptotic limiting values $\widetilde{R}(|a_k|^2, \tau_k)$ are independent of the distribution of the time delays $f_T(\tau)$ and synchronous and asynchronous CDMA systems have the same performance when a linear detector that admits a representation as multistage detector is used at the receiver.*

Asynchronism does not cause any performance degradation on the system if the time delays and the received amplitudes of the signals are known at the receiver and the sampling rate satisfies the conditions of the sampling theorem. In this way we have generalized the results obtained in [82] for systems using an ideal Nyquist sinc waveform to any kind of chip pulse waveforms with bandwidth $B \leq \frac{1}{2T_c}$.

The output SINR is also independent of the initial sampling time. Therefore, the system does not incur any degradation in SINR if, for all signals of interest, we consider a discrete statistic obtained by sampling the received signal starting at a random instant and with a proper sampling rate, instead of considering K different statistics obtained by sampling the received signal at the exact time delay of each signal to be detected. This property has a very positive impact on the complexity of the system. In fact, without performance degradation we can replace a bank of K different samplers and K different multiuser detectors by a single sampler followed by a single multiuser detector processing all users jointly. Additionally, the prior knowledge of the time delay of each user is not required in order to sample the received signal. The estimation of the time delay can be done afterwards using the discrete time received signal.

As already mentioned, the previous results enable also an asymptotic analysis of the effects of oversampling. We verified numerically that an increase of the sampling rate above the chip rate does not provide any benefit. This is a consequence of the sufficiency of the statistics obtained by sampling at rate $\frac{r^*}{T_c}$, where r^* is the minimum sampling factor such that $B \leq \frac{r^*}{2T_c}$.

Chip pulse waveform with $\frac{1}{2T_c} \leq B \leq \frac{1}{T_c}$

Let $\widetilde{\psi}(t)$ be a chip waveform with unitary Fourier transform $\Xi(j2\pi f)$ and bandwidth $\frac{1}{2T_c} \leq B \leq \frac{1}{T_c}$. Sufficient statistics are obtained sampling at rate $\frac{2}{T_c}$ since the condition of the sampling theorem is satisfied. The asymptotic values of the diagonal elements of $\widetilde{\boldsymbol{R}}^{\ell}$ can be obtained by specializing Theorem 10 to this scenario and then applying Proposition 1. The following Corollary specializes Theorem 10 to a system using chip pulse waveforms with bandwidth $\frac{1}{2T_c} \leq B \leq \frac{1}{T_c}$.

Corollary 5 *Let the definitions of Theorem 10 hold and let us assume that conditions (a)-(d) of Theorem 10 are satisfied. Additionally, assume:*

(1) The random variables λ and τ are statistically independent and $f_T(\tau)$, the probability density function of the random variable T, with support $\tau \in [0, T_c)$ is symmetric around $\tau = \frac{T_c}{2}$, i.e. $f_T(\tau - \frac{T_c}{2})$ is an even function.

(2) $\Xi(j2\pi f) : \mathbb{R} \to \mathbb{R}$ is real[36] *and bandlimited with bandwidth $B \in [\frac{1}{2T_c}, \frac{1}{T_c}]$*

(3) $r = 2$.

Then, given $(|a_{kk}|^2, \tau_k)$, the k^{th} diagonal element of the matrix $(\overline{\boldsymbol{R}}^{(N)})^{\ell} = ((\overline{\boldsymbol{H}}^{(N)})^H \overline{\boldsymbol{H}}^{(N)})^{\ell}$ converges with probability one to a deterministic value, conditionally on $|a_{kk}|^2$,

$$\lim_{K=\beta N \to \infty} (\overline{\boldsymbol{R}}^{(N)})^{\ell}_{kk} \overset{a.s.}{=} \overline{R}^{\ell}(|a_{kk}|^2)$$

[36] This condition corresponds to the assumption that the chip-pulse waveform is an even function. This condition is usually satisfied in practical systems.

with $\overline{R}(\lambda)|_{\lambda=|a_{kk}|^2}$ *determined by the following recursion:*

$$\overline{R}^{\ell}(\lambda) = \sum_{s=0}^{\ell-1} g(\overline{\boldsymbol{T}}^{\ell-s-1}, \lambda)\overline{R}^{s}(\lambda)$$

and

$$\overline{\boldsymbol{T}}^{\ell}(x) = \sum_{s=0}^{\ell-1} f(\overline{\boldsymbol{R}}^{\ell-s-1}, x)\boldsymbol{Q}(x)\overline{\boldsymbol{T}}^{s}(x) \qquad -\frac{1}{2} \le x \le \frac{1}{2}$$

$$f(\overline{\boldsymbol{R}}^{s}) = \beta \int \lambda \overline{R}^{s}(\lambda)\mathrm{d}\, F_{|\boldsymbol{A}|^2}(\lambda)$$

$$g(\overline{\boldsymbol{T}}^{s}, \lambda) = \lambda \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathrm{tr}(\overline{\boldsymbol{T}}^{s}(x)\boldsymbol{Q}(x))\mathrm{d}\, x$$

where

$$\boldsymbol{Q}(x) = \begin{pmatrix} q_1(x) & q_2(x)\mathrm{e}^{-j\pi x} \\ q_2(x)\mathrm{e}^{j\pi x} & q_1(x) \end{pmatrix} \tag{4.40}$$

with

$$q_1(x) = \frac{1}{T_c^2}\begin{cases} \Xi^2(j\frac{2\pi}{T_c}x) + \Xi^2(j\frac{2\pi}{T_c}(x+1)) & -\frac{1}{2} \le x \le 0 \\ \Xi^2(j\frac{2\pi}{T_c}x) + \Xi^2(j\frac{2\pi}{T_c}(x-1)) & 0 \le x \le \frac{1}{2} \end{cases}$$

and

$$q_2(x) = \frac{1}{T_c^2}\begin{cases} \Xi^2(j\frac{2\pi}{T_c}x) - \Xi^2(j\frac{2\pi}{T_c}(x+1)) & -\frac{1}{2} \le x \le 0 \\ \Xi^2(j\frac{2\pi}{T_c}x) - \Xi^2(j\frac{2\pi}{T_c}(x-1)) & 0 \le x \le \frac{1}{2}. \end{cases}$$

The recursion is initialized by setting $\overline{\boldsymbol{T}}^{0}(x) = \boldsymbol{I}_2$ *and* $\overline{R}^{0}(\lambda) = 1$.

Corollary 5 is derived in Appendix C Section C.5.

Applying Corollary 5 we obtain the following algorithm.

Algorithm 3

Initialization: *Let* $\rho_0(z) = 1$ *and* $\mu_0(y) = 1$.

l^{th} step:

- *Define* $u_{\ell-1}(y) = y\mu_{\ell-1}(y)$ *and write it as a polynomial in* y.
- *Define* $v_{\ell-1}(z) = z\rho_{\ell-1}(z)$ *and write it as a polynomial in* z.
- *Define*

$$\mathcal{E}_s = \left(\frac{2}{T_c^2}\right)^s \int_{-1}^{1} \Xi^{2s}\left(j\frac{2\pi}{T_c}x\right)\mathrm{d}\, x \tag{4.41}$$

and replace all monomials $y, y^2, \ldots, y^{\ell}$ *in the polynomial* $u_{\ell-1}(y)$ *by* $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_{\ell}$*, respectively. Denote the result by* $U_{\ell-1}$.

- *Define* $m^s_{|\boldsymbol{A}|^2} = \mathrm{E}\{|a_{kk}|^{2s}\}$ *and replace all monomials* $z, z^2, \ldots, z^\ell$ *in the polynomial* $v_{\ell-1}(z)$ *by the moments* $m^1_{|\boldsymbol{A}|^2}$, $m^2_{|\boldsymbol{A}|^2}, \ldots, m^\ell_{|\boldsymbol{A}|^2}$, *respectively. Denote the result by* $V_{\ell-1}$.
- *Calculate*

$$\rho_\ell(z) = \sum_{s=0}^{\ell-1} zU_{\ell-s-1}\rho_s(z)$$

$$\mu_\ell(y) = \sum_{s=0}^{\ell-1} \beta y V_{\ell-s-1}\mu_s(y).$$

- *Assign* $\rho_\ell(\lambda)$ *to* $\overline{R}^\ell(\lambda)$.

  Replace all monomials $z, z^2, \ldots, z^\ell$ *in the polynomial* $\rho_\ell(z)$ *by the moments* $m^1_{|\boldsymbol{A}|^2}, m^2_{|\boldsymbol{A}|^2}, \ldots, m^\ell_{|\boldsymbol{A}|^2}$, *respectively, and assign the result to* $m^\ell_{\overline{\boldsymbol{R}}}$.

Algorithm 3 is derived in Appendix C Section C.6.

The asymptotic limits $\widetilde{R}^\ell(\lambda)$ and the eigenvalue moments $m^\ell_{\widetilde{\boldsymbol{R}}}$ of the matrix $\widetilde{\boldsymbol{R}}$ can be obtained from Algorithm 3 by applying Proposition 1.

Interestingly, the recursive equations in Corollary 5 do not depend on the time delay τ_k of the signal of user k, i.e. the performance of a CDMA system is independent of the sampling instants and time delays under Condition (1) and (2) on the chip waveforms and on the time delays of Corollary 5.

Additionally, the dependence of $\widetilde{R}^\ell(\lambda)$ on the chip pulse waveforms becomes clear from Algorithm 3: $\widetilde{R}^\ell(\lambda)$ depends on $\Xi(\jmath 2\pi f)$ through the quantities $\mathcal{E}_s$, $s = 1, 2, \ldots$, defined in (4.41).

Proposition 1 is completely general and can be applied to the analysis of synchronous CDMA sampled at a rate that is a multiple of the chip rate. To this aim it is sufficient to consider that the time delays τ_k of all users are deterministic and equal to τ_0. This is modelled mathematically by setting $f_T(\tau)$, the probability density function of the time delay equal to a Dirac, i.e.

$$f_T(\tau) = \delta(\tau - \tau_0).$$

The comparison of synchronous and asynchronous systems with equal chip pulse waveforms enables us to analyze the effects on the system performance of the chip pulse waveforms jointly with the effects of the distribution of the delays . Hereafter, we will elaborate on these aspects focusing on square root raised cosine and on raised cosine chip-pulse waveforms with roll-off $\gamma \in [0, 1]$. To simplify the notation we assume $T_c = 1$. Let

$$\Upsilon(x) = \begin{cases} 1 & 0 \le |x| \le \frac{1-\gamma}{2} \\ \frac{1}{2}\left(1 - \sin\frac{\pi}{\gamma}\left(|x| - \frac{1}{2}\right)\right) & \frac{1-\gamma}{2} \le |x| \le \frac{1+\gamma}{2} \\ 0 & |x| \ge \frac{1+\gamma}{2}. \end{cases}$$

The energy frequency spectrum of a square root raised cosine waveform with unit energy is given by $|\Xi_{\mathrm{sqrc}}(j2\pi x)|^2 = \Upsilon(x)$. The unitary Fourier transform of a raised cosine chip waveform with unit energy is $\Xi_{\mathrm{rc}}(j2\pi x) = \frac{4}{\sqrt{4-\gamma}}\Upsilon(x)$. The corresponding coefficients $\mathcal{E}_{\mathrm{sqrc},s}$ and $\mathcal{E}_{\mathrm{rc},s}$, $s = 1, 2, \ldots$, are given by

$$\mathcal{E}_{\mathrm{sqrt},s}{=}2^s(1-\gamma) + 2\int_{\frac{1-\gamma}{2}}^{\frac{1+\gamma}{2}} \sin^s\left(\frac{\pi}{\gamma}\left(\frac{1}{2}-x\right)\right)\mathrm{d}x$$

and

$$\mathcal{E}_{\mathrm{rc},s}{=}\left(\frac{2}{4-\gamma}\right)^s\left[4^s(1-\gamma)+\int_{\frac{1-\gamma}{2}}^{\frac{1+\gamma}{2}}\left[1-\sin\left(\frac{\pi}{\gamma}\left(x-\frac{1}{2}\right)\right)\right]^{2s}\mathrm{d}x+\int_{-\frac{1+\gamma}{2}}^{-\frac{1-\gamma}{2}}\left[1+\sin\left(\frac{\pi}{\gamma}\left(x+\frac{1}{2}\right)\right)\right]^{2s}\mathrm{d}x\right],$$

respectively.

The performance of synchronous CDMA systems with square root raised cosine chip-pulse waveforms is well known to be independent of the roll-off and given in Chapter 3.

For a synchronous CDMA system with raised cosine chip pulse waveforms, sampled at rate $\frac{2}{T_c}$, and with time delay $\tau_0 = 0$, Proposition 1 can be applied. The recursive equations of Proposition 1 reduce to the recursive equations of Corollary 5 with

$$\boldsymbol{Q}(x) = \begin{cases} \frac{4}{4-\gamma}\begin{pmatrix} 1 & e^{-j\pi x} \\ e^{j\pi x} & 1 \end{pmatrix} & 0 \le |x| \le \frac{1-\gamma}{2} \\ \frac{4}{4-\gamma}\begin{pmatrix} 1 & e^{-j\pi x}\sin\left(\frac{\pi}{\gamma}\left(\frac{1}{2}-|x|\right)\right) \\ e^{j\pi x}\sin\left(\frac{\pi}{\gamma}\left(\frac{1}{2}-|x|\right)\right) & \sin^2\left(\frac{\pi}{\gamma}\left(\frac{1}{2}-|x|\right)\right) \end{pmatrix} & \frac{1-\gamma}{2} \le |x| \le \frac{1}{2}. \end{cases}$$

Figure 4.11 and Figure 4.12 show the large system performance, in terms of asymptotic output SINR, of detectors Type J-I with $M = 4$ and increasing roll-off versus the SNR, $\frac{E_s}{N_0}$ being E_s the energy per symbol in case of both synchronous (lines with markers) and asynchronous CDMA systems (lines without markers). The SINR is obtained assuming equal received powers at the receiver and system load $\beta = \frac{1}{2}$.

While the SINR of synchronous systems with square root Nyquist waveforms is independent of the roll-off, it decreases as the roll-off increases, if the modulation is based on raised cosine waveforms. For $\gamma = 0$, i.e. for an ideal sinc chip pulse waveform, the performance of synchronous and asynchronous systems coincides as already observed in the previous section. In contrast to the case of synchronous systems, asynchronous systems with uniform time delay distribution or time delay distribution satisfying condition (1) in Corollary 5 outperform the corresponding synchronous systems with equal roll-off, both in case of square root raised cosine and raised cosine chip-pulse waveforms. The comparison of Figure 4.11 and Figure 4.12 shows that asynchronous systems with square root raised cosine largely outperform asynchronous systems using raised cosine waveforms and linear multiuser detection. The output SINR of asynchronous systems increases as the roll-off increases.

Increasing the roll-off is equivalent to increasing the system bandwidth and to redistributing the available energy in the additional degrees of freedom added in the frequency domain. Since the bandwidth increases there is room for potential improvements of the SINR. These degrees of freedom can be utilized to reduce the multiple access interference. Let us consider identical chip-pulse waveforms of different users. If all waveforms have the same time delay the correlation between two chip-pulse waveforms is maximum and also the average correlation is maximum. However, if the system is asynchronous the average correlation is lower and the multiple access interference decreases. This gives an intuitive explanation of the reasons why asynchronous systems can outperform synchronous systems.

Figure 4.13 and Figure 4.14 illustrate the output SINR of a Type J-I detector with $M = 4$ as a function of the roll-off for synchronous systems (dashed lines) and chip asynchronous but symbol quasi synchronous systems (solid lines) for three different levels of $\frac{E_s}{N_0}$: 15 dB, 20 dB, and 30 dB. The gap between the solid and the dashed lines corresponding to the same level of $\frac{E_s}{N_0}$ is the improvement achievable with a 3-stage detector ($M = 4$) using an asynchronous system instead of a synchronous system. The gap increases as the roll-off and/or the SNR increase, both for raised cosine (Figure 4.14) and square root raised cosine waveforms (Figure 4.13).

In Figure 4.15 and Figure 4.16 the SINR is plotted as a function of the system load. The improvement achievable by asynchronous systems over synchronous systems increases as the system load increases, both for raised cosine and square root raised cosine waveforms.

Square root Nyquist chip pulse waveforms and chip matched filtering at the front-end.

Let us reconsider the chip asynchronous but symbol quasi synchronous CDMA system slightly modified. We assume that the used chip pulse is a square root Nyquist

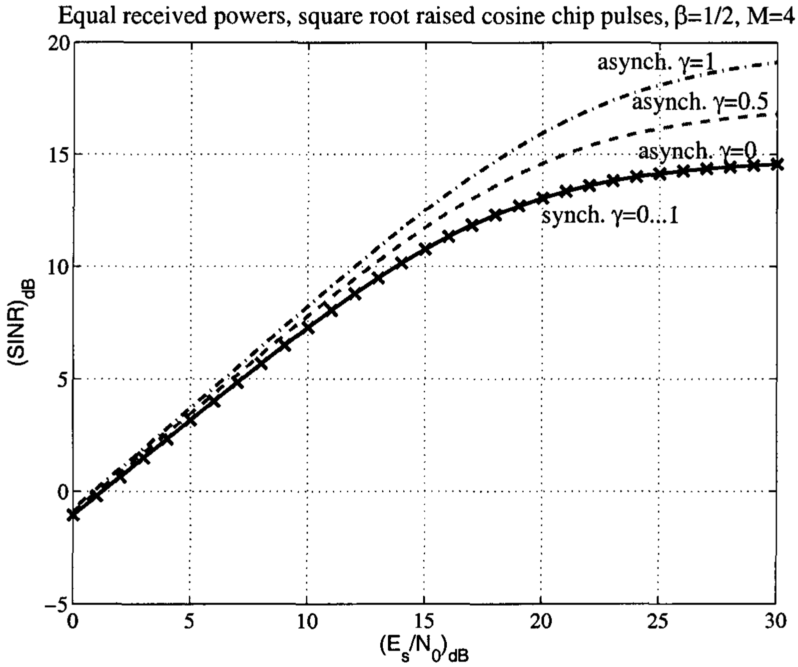


Figure 4.11: Output SINR of a Type J-I detector with $M = 4$ versus $\frac{E_s}{N_0}$ for synchronous systems (lines with markers) and chip asynchronous but symbol quasi synchronous systems (lines without markers). CDMA systems with equal received powers, **square root raised cosine chip pulse waveforms**, sampling rate $\frac{2}{T_c}$, and system load $\beta = \frac{1}{2}$. Three different square root raised cosine waveforms are considered corresponding to different roll-offs: $\gamma = 0$ (dot-dashed line), $\gamma = 0.5$ (dashed line), and $\gamma = 1$ (solid line).

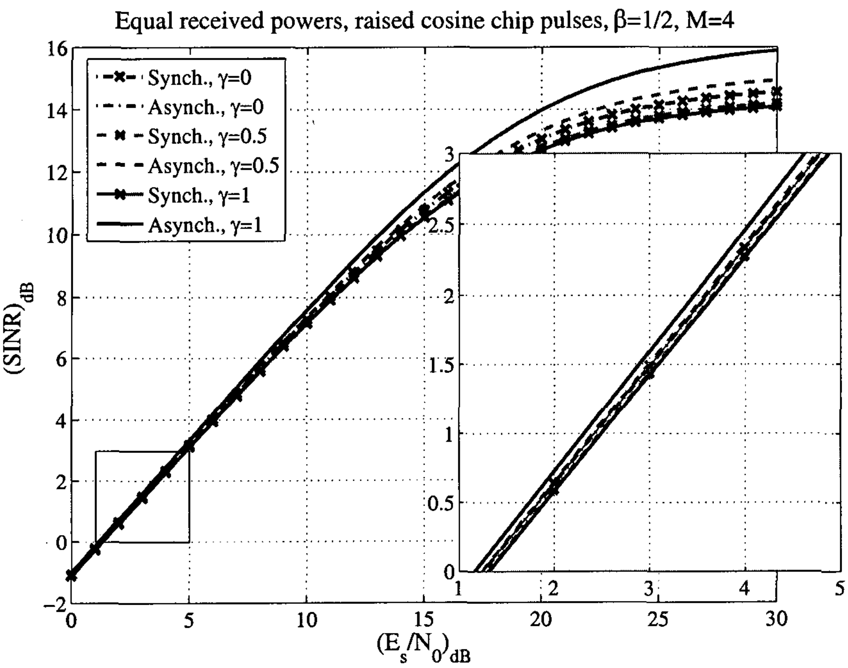


Figure 4.12: As Figure 4.11 for **raised cosine chip pulse waveforms**.

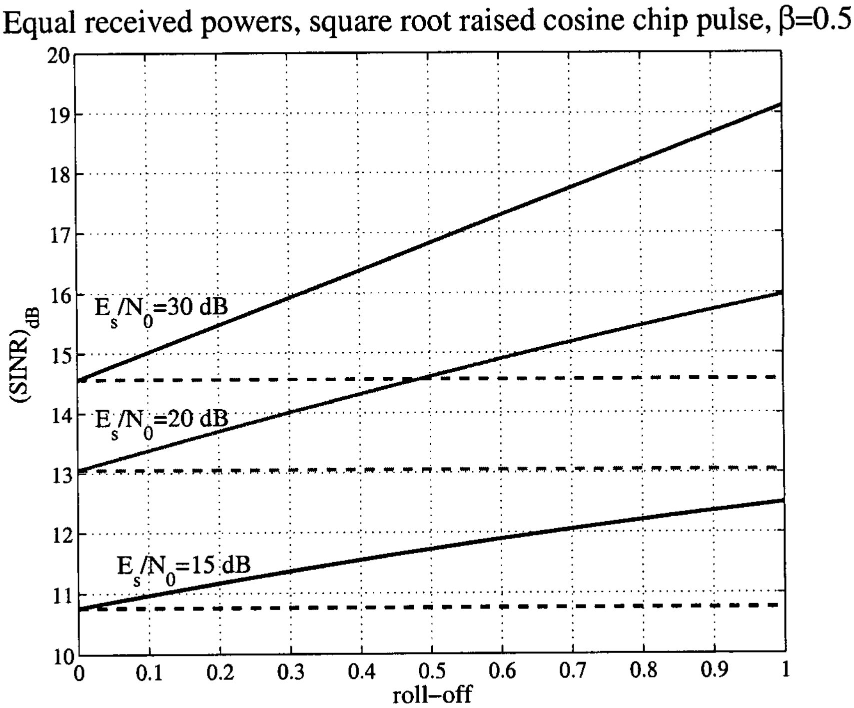


Figure 4.13: Output SINR of a Type J-I detector with $M = 4$ versus the roll-off γ for synchronous systems (dashed lines) and chip asynchronous but symbol quasi synchronous systems (solid lines). CDMA systems with equal received powers, **square root raised cosine chip pulse waveforms**, sampling rate $\frac{2}{T_c}$, and system load $\beta = \frac{1}{2}$. Three different level of SNR are considered: 15 dB, 20 dB, and 30 dB.

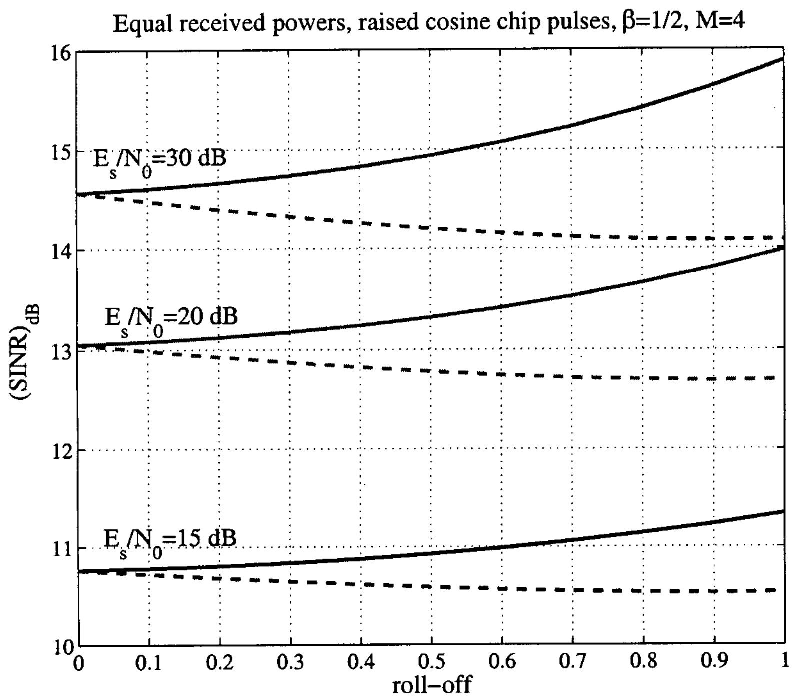


Figure 4.14: Idem as Figure 4.13 for raised cosine chip pulse waveforms.

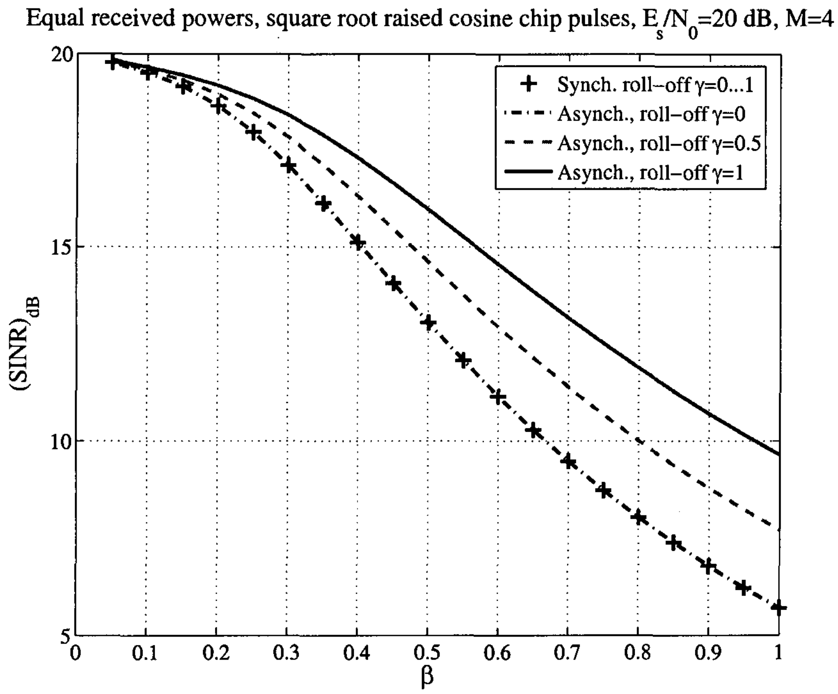


Figure 4.15: Output SINR of a Type J-I detector with $M = 4$ versus the system load β for synchronous systems (lines with markers) and chip asynchronous but symbol quasi synchronous systems (lines without markers). CDMA systems with equal received powers, **square root raised cosine chip waveforms**, sampling rate $\frac{2}{T_c}$, and system load $\beta = \frac{1}{2}$. Three different raised cosine waveforms are considered corresponding to different roll-offs: $\gamma = 0$ (dot-dashed line), $\gamma = 0.5$ (dashed line), and $\gamma = 1$ (solid line).

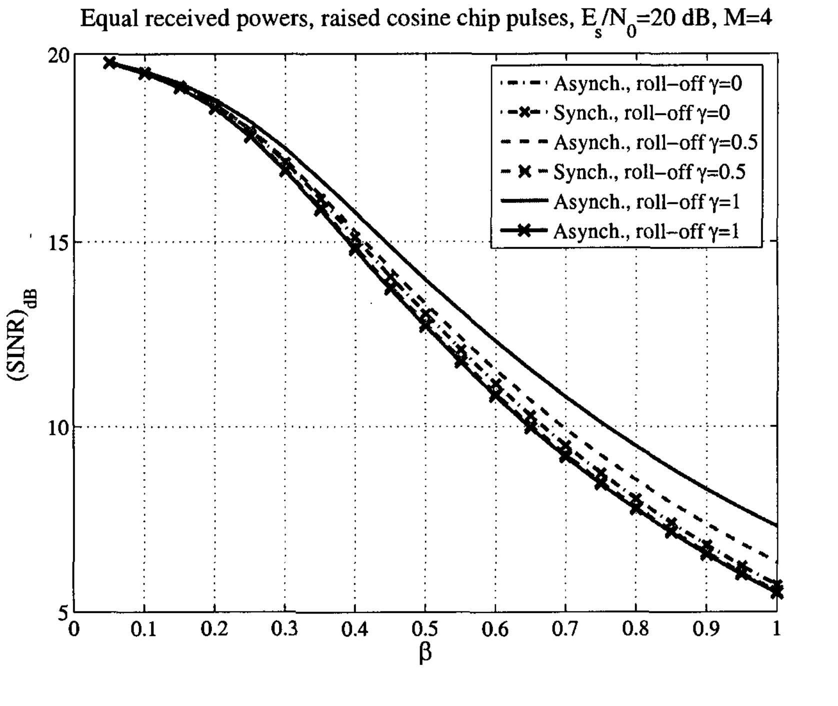


Figure 4.16: As Figure 4.15 for **raised cosine chip waveforms.**

pulse.

The front-end consists of:

- An analog filter $G(f)$ matched to the chip pulse and normalized by the chip pulse energy, i.e. $G(f) = \frac{\psi^*(f)}{\sqrt{E_\psi}}$;
- A subsequent sampler with sampling rate equal to the chip rate.

By sampling the output of the chip matched filter at the chip rate, the discrete-time signal after the sampler is given by

$$y[p] = \sum_{k=1}^{K} a_{kk} \sum_{m=-\infty}^{+\infty} b_k[m] \sum_{u=0}^{N-1} s_{k,m}[u]\widetilde{\phi}\left(\left(\frac{p}{r}-u-mN\right)T_c-\tau_k\right) + n[p]. \tag{4.42}$$

The notation utilized in (4.42) has been introduced in Section 4.4.1 with the exception of $\widetilde{\phi}(t)$, which is the convolution of the chip pulse waveform with the filter at the front-end, i.e. $\widetilde{\phi}(t) = \frac{\psi(t)\psi^*(-t)}{\sqrt{E_\psi}}$. By the definition of a square root Nyquist waveforms, $\phi(t)$ satisfies the Nyquist pulse-shaping criterion. The same criterion is also satisfied by the power spectrum of the noise. Then, after sampling at the chip rate, the discrete time noise process $\{n[p]\}$ is white with variance $\frac{N_0}{E_\psi T_c}$. The system model in matrix notation is given by (4.31) with the matrix $\widetilde{\boldsymbol{\Psi}}$ obtained by substituting $\widetilde{\psi}(t)$ with $\widetilde{\phi}(t)$ and assuming $r = 1$. Then, Theorem 10 can be applied. Hereafter, we refer to this system as System A.

Given the square root raised cosine chip-pulse waveform with roll-off γ [94]

$$\widetilde{\psi}(t) = \frac{4\gamma(\frac{t}{T_c})\cos(\pi(1+\gamma)\frac{t}{T_c}) + \sin(\pi(1-\gamma)\frac{t}{T_c})}{\pi t(1-(4\gamma\frac{t}{T_c})^2)} \qquad \gamma \in [0,1] \tag{4.43}$$

the matrix $\boldsymbol{Q}_\tau(x) = \boldsymbol{\Delta}_\tau(x)\boldsymbol{\Delta}_\tau^H(x)$ occurring in Theorem 10 reduces to the scalar

$$\boldsymbol{Q}_\tau(x) = \begin{cases} \frac{1}{2} + \frac{1}{2}\sin^2\left(\frac{\pi}{\gamma}(x+\frac{1}{2})\right) + \frac{\cos 2\pi\tau}{2}\left(1 - \sin^2\left(\frac{\pi}{\gamma}(x+\frac{1}{2})\right)\right) & -\frac{1}{2} \le x \le -\frac{1-\gamma}{2} \\ 1 & -\frac{1-\gamma}{2} \le x \le \frac{1-\gamma}{2} \\ \frac{1}{2} + \frac{1}{2}\sin^2\left(\frac{\pi}{\gamma}(x-\frac{1}{2})\right) + \frac{\cos 2\pi\tau}{2}\left(1 - \sin^2\left(\frac{\pi}{\gamma}(x-\frac{1}{2})\right)\right) & \frac{1-\gamma}{2} \le x \le \frac{1}{2}. \end{cases}$$

due to the fact that $r = 1$ in this case. In the following the asymptotic analysis of system A is performed. Equal received powers, system load $\beta = \frac{1}{2}$, Type J-I detectors with $M = 3$ define the scenario we consider for the asymptotic analysis.

In Figure 4.17 the asymptotic output SINR (solid lines) as a function of $\frac{E_s}{N_0}$ is compared to the corresponding output SINR for finite systems (dots) with 512 users. The performance depends on the time delay. The two groups of curves illustrate the

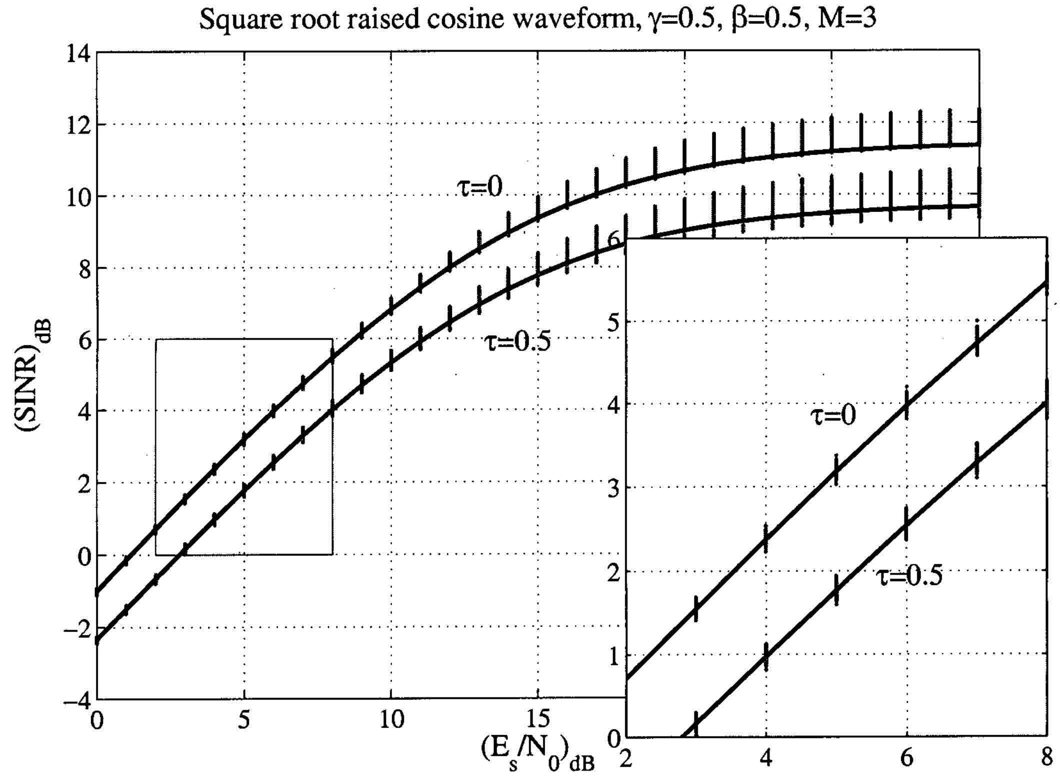


Figure 4.17: Output SINR of a Type J-I detector with $M = 3$ versus $\frac{E_s}{N_0}$ for large systems (solid lines) and finite systems with 512 users (dots). CDMA systems with equal received powers, square root raised cosine chip waveforms ($\gamma = 0.5$), sampling rate $\frac{1}{T_c}$, system load $\beta = \frac{1}{2}$ and chip matched filter at the front-end. Two user signals with time delays $\tau = 0$ and $\tau = 0.5$ are considered.

performance corresponding to the most favorable time delay ($\tau = 0$) and to the least favorable time delay ($\tau = 1/2$). The gap between these two cases is quite significant: it exceeds 1 dB in the analyzed range of $\frac{E_s}{N_0}$.

The dependence of the performance on the time delay is illustrated in Figure 4.18. The solid lines represent the output SINR as a function of the time delay τ parameterized with respect to $\frac{E_s}{N_0}$. The parameter $\frac{E_s}{N_0}$ varies from 0 dB to 20 dB in steps of 2 dB. The performance is maximum at the extreme points of the range of τ, i.e. $\tau = 0$ and $\tau = 1$, and it is minimum in the middle of the interval $[0, 1]$. The output SINR of asynchronous systems is compared with the SINR of synchronous systems using the same chip pulse waveform and sampling the signal at the best time instant ($\tau = 0$). The asynchronous system slightly outperforms the synchronous one around $\tau = 0$ and $\tau = 1$ while the performance is severely worse for signals delayed by $\tau = \frac{1}{2}$.

This observation leads to the following conclusions:

- The use of square root Nyquist chip pulse waveforms jointly with the above described front-end requires a preliminary estimation of the time delay of the user of interest and a good synchronization in order to avoid severe performance degradation.
- In order to obtain good performance, different statistics are needed for the detection of different users and they have to be processed independently. This does not affect the complexity when detection of a single user is required (e.g. downlink). However, it is a significant drawback when several or all users in the system have to be detected (e.g. uplink) since the detection cannot be performed jointly.

In Figure 4.19 system A with roll-off $\gamma = 1$ is compared to a system using raised cosine waveforms with $\gamma = 1$ and sufficient statistics. The latter system is referred to as system B. The SINR versus the time delay for values of the parameter $\frac{E_s}{N_0}$ varying between 0 dB and 20 dB in steps of 2 dB is plotted, both for system A (solid lines) and for system B (dashed lines). For some values of the parameter $\frac{E_s}{N_0}$ system B outperforms system A, although earlier in this section systems using square root raised cosine waveforms were shown to perform better than systems using raised cosine waveforms. The reason for this behavior is that the statistics utilized in system A, sufficient for synchronous systems, are not sufficient for asynchronous systems.

4.5 Asynchronous Systems: General Case

In this section we extend the previous results to a general asynchronous system. Without loss of generality we can assume that the maximum delay among users is the

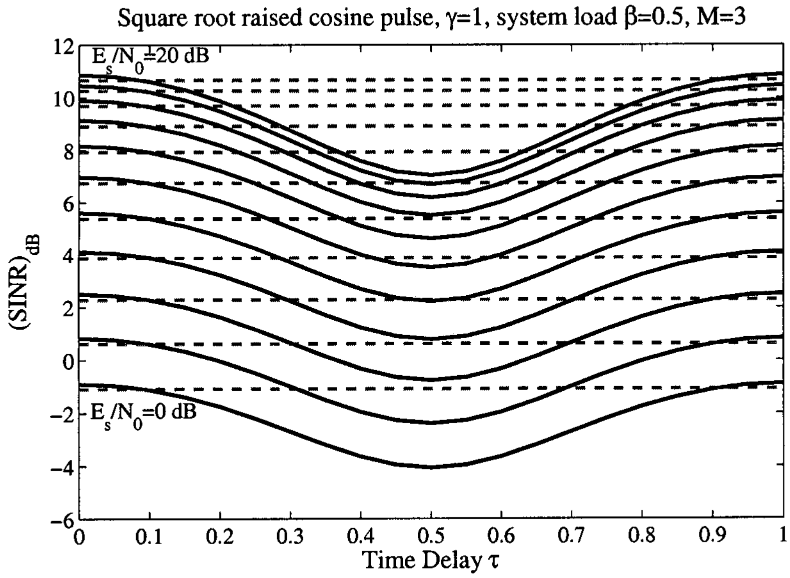


Figure 4.18: Asymptotic output SINR of a Type J-I detector with $M = 3$ versus the time delay τ for asynchronous systems (solid lines) and synchronous systems (dashed lines). CDMA systems with equal received powers, square root raised cosine chip waveforms ($\gamma = 1$), system load $\beta = \frac{1}{2}$, sampling rate $\frac{1}{T_c}$, and chip matched filter at the front-end. The curves are parameterized by $\frac{E_s}{N_0}$ with $\frac{E_s}{N_0}$ varying between 0 dB and 20 dB in steps of 2 dB.

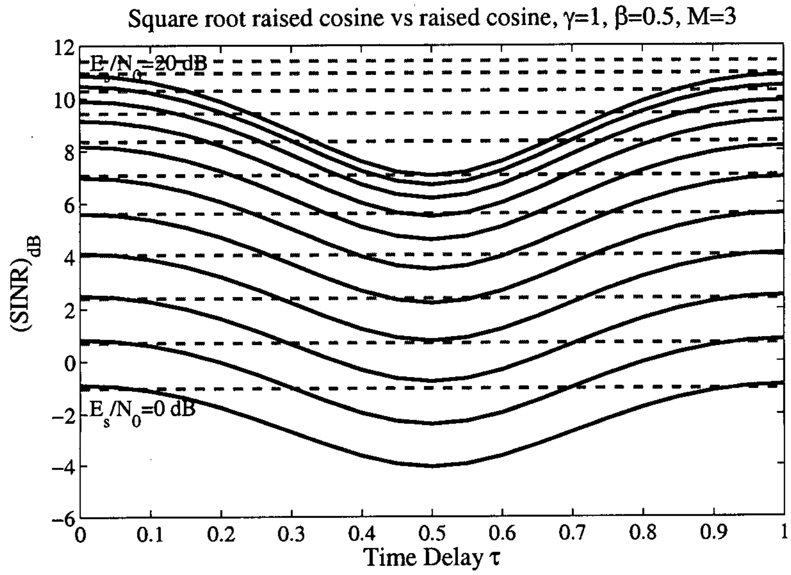


Figure 4.19: Asymptotic output SINR of a Type J-I detector with $M = 3$ versus the time delay τ for asynchronous systems with square root raised cosine chip waveforms ($\gamma = 1$) and sampling rate $\frac{1}{T_c}$ (solid lines), and asynchronous CDMA systems with raised cosine chip waveforms ($\gamma = 1$) and sampling rate $\frac{2}{T_c}$ (dashed lines). The curves are parameterized by $\frac{E_s}{N_0}$ with $\frac{E_s}{N_0}$ varying between 0 dB and 20 dB in steps of 2 dB.

symbol interval T_s [16]. User 1 has time delay $\tau_1 = 0$ and the other users are ranked in ascending order of delay. The chip pulse is much shorter than the symbol waveform, as already assumed for symbol quasi synchronous but chip asynchronous systems. Thus, we can neglect the intersymbol interference between symbols transmitted by the same user in the asymptotic analysis.

The system model in matrix notation is

$$\widetilde{\mathcal{Y}} = \widetilde{\mathcal{H}}\mathcal{B} + \widetilde{\mathcal{N}}$$

where $\widetilde{\mathcal{Y}} = [\dots, \widetilde{\boldsymbol{y}}^T(m-1), \widetilde{\boldsymbol{y}}^T(m), \widetilde{\boldsymbol{y}}^T(m+1), \dots]^T$ is the infinite-length column vector of the received signal sampled at rate $\frac{r}{T_c}$, with $\widetilde{\boldsymbol{y}}(m)$ being the rN-dimensional vector corresponding to the m^{th} transmitted symbol; $\mathcal{B} = [\dots, \boldsymbol{b}^T(m-1), \boldsymbol{b}^T(m), \boldsymbol{b}^T(m+1), \dots]^T$ is the infinite-length vector of transmitted symbols as defined in Section 4.3; and $\widetilde{\mathcal{N}}$ is the infinite-length column vector of white Gaussian noise with variance $\sigma^2 = \frac{rN_0}{E_\psi}$. The matrix $\widetilde{\mathcal{H}} \in \mathbb{C}^{rN \times K}$ models the effects of the spreading sequences, the pulse shape, and the received amplitudes. It is structured as the matrix $\mathcal{H}$ in (4.4):

$$\widetilde{\mathcal{H}} = \begin{bmatrix} \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \dots & \mathbf{0} & \widehat{\boldsymbol{H}}_d(m-2) & \widehat{\boldsymbol{H}}_u(m-1) & \mathbf{0} & \dots & \dots & \dots \\ \dots & \dots & \mathbf{0} & \widehat{\boldsymbol{H}}_d(m-1) & \widehat{\boldsymbol{H}}_u(m) & \mathbf{0} & \dots & \dots \\ \dots & \dots & \dots & \mathbf{0} & \widehat{\boldsymbol{H}}_d(m) & \widehat{\boldsymbol{H}}_u(m+1) & \mathbf{0} & \dots \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}.$$

Here, $\widehat{\boldsymbol{H}}_u(m)$ and $\widehat{\boldsymbol{H}}_d(m) \in \mathbb{C}^{rN \times K}$ are the upper and lower part of the matrix $\widehat{\boldsymbol{H}}(m) = [\widehat{\boldsymbol{H}}_u^T(m), \widehat{\boldsymbol{H}}_d^T(m)]^T$, which is defined as $\widehat{\boldsymbol{H}}(m) = \widehat{\boldsymbol{S}}(m)\boldsymbol{A}$, where $\boldsymbol{A}$ is the $K \times K$ diagonal matrix of received amplitudes defined in Section 4.3. Furthermore, $\widehat{\boldsymbol{S}}(m)$ is a $2rN \times K$ matrix obtained as follows:

- Given the sequence of the time delays $\{\tau_1, \tau_2, \dots, \tau_K\}$, derive from it the two sequences $\{\widehat{\tau}_1, \widehat{\tau}_2, \dots, \widehat{\tau}_K\}$ and $\{\widetilde{\tau}_1, \widetilde{\tau}_2, \dots, \widetilde{\tau}_K\}$ with $\widehat{\tau}_k = \lfloor \frac{\tau_k}{T_c} \rfloor$ and $\widetilde{\tau}_k = \tau_k - \widehat{\tau}_k$, $k = 1, \dots, K$. A system with time delays $\{\widehat{\tau}_1, \widehat{\tau}_2, \dots, \widehat{\tau}_K\}$ reduces to a symbol asynchronous but chip synchronous system, while a system with time delays $\{\widetilde{\tau}_1, \widetilde{\tau}_2, \dots, \widetilde{\tau}_K\}$ is a symbol quasi synchronous but chip asynchronous system.

- Build the $rN \times K$ matrix $\widetilde{\boldsymbol{S}}(m)$ of virtual spreading for symbol quasi synchronous but chip asynchronous systems with time delays $\{\widetilde{\tau}_1, \widetilde{\tau}_2, \dots, \widetilde{\tau}_K\}$ as explained in Section 4.4.1.

- Append $\mathbf{0}_{rN \times K}$, an $rN \times K$ zero matrix, to the matrix $\widetilde{\boldsymbol{S}}(m)$, i.e. build the matrix $[\widetilde{\boldsymbol{S}}^T(m), \mathbf{0}_{rN \times K}^T]^T$.

- The matrix $\widehat{\boldsymbol{S}}(m)$ is obtained from $[\widetilde{\boldsymbol{S}}^T(m), \mathbf{0}_{rN\times K}^T]^T$ by circular downshift of the k^{th} column, for $k = 1, \ldots, K$, by $r\widehat{\tau}_k$ positions.

By the construction of the matrix $\widetilde{\boldsymbol{S}}$ we model the effects of chip asynchronism. The subsequent circulant downshift introduces the effects of delays that are multiples of the chip interval in the model.

Since the system model for asynchronous systems has a structure identical to the structure of the system model for symbol asynchronous but chip synchronous systems, the considerations made in Section 4.3.3 apply. The projection subspace is given by

$$\chi(\widetilde{\mathcal{H}}) = \text{span}\{\widetilde{\hbar}_k^H(m)\widetilde{\mathcal{T}}^m\}_{m=0}^{M-1}$$

where $\widetilde{\mathcal{T}} = \widetilde{\mathcal{H}}\widetilde{\mathcal{H}}^H$ and $\widetilde{\hbar}_k$ is the virtual spreading sequence of the m^{th} symbol transmitted by user k, i.e. the column of $\widetilde{\mathcal{H}}$ containing the k^{th} column vector of the matrix $\widehat{\boldsymbol{H}}(m)$ defined in Section 4.4.1. With this choice of the projection subspace, it is possible to build a multistage detector with sliding observation window that does not suffer from truncation effects. Its structure is identical to the structure of the detector in Figure 4.6 by replacing the $2N \times K$ matrices $\boldsymbol{H}(n), \ldots, \boldsymbol{H}(n-M)$ by the $2rN \times K$ matrices $\widehat{\boldsymbol{H}}(n), \ldots, \widehat{\boldsymbol{H}}(n-M)$, respectively.

The asymptotic analysis and design of multistage detectors follows the same lines as in the case of symbol asynchronous but chip synchronous systems treated in Section 4.3.3. The problem reduces once again to the computation of the diagonal elements of the matrix $\widetilde{\mathcal{R}}^m$, with $\widetilde{\mathcal{R}} = \widetilde{\mathcal{H}}^H\widetilde{\mathcal{H}}$. We conjecture[37] that these diagonal elements coincide with the diagonal elements of $\widetilde{\mathcal{R}}$ for symbol quasi synchronous but chip asynchronous systems. This conjecture is motivated by the observation that the diagonal elements of the matrix $\mathcal{R}^m$ for symbol asynchronous but chip synchronous systems are a periodical repetition of the diagonal elements of the matrix $\boldsymbol{R}^m$ for synchronous systems, i.e. the shift of the column of the matrix $\mathcal{H}$ due to the symbol asynchronism does not influence the asymptotic values of the diagonal elements of the matrix $\mathcal{R}^m$ for the infinite-size matrix $\mathcal{H}$. Numerical simulations support this conjecture.

Numerical simulations were performed for an asynchronous CDMA system with maximum time delay equal to the symbol interval. The 64 users utilized raised cosine chip-pulse waveforms, QPSK modulation, and random spreading sequences with $N = 128$. Perfect power control was applied; all users were received with the same power and sampled at rate $\frac{2}{T_c}$. At the receiver, detection was performed by matched filters, by detectors Type J-I with sliding observation window, and by MSWF with $M = 2$ and 4. The performance of the various detectors is compared in

[37]The author thinks that this conjecture can be proven rigorously following the lines of the proof of Theorem 9 and Theorem 10. This issue is left for further studies.

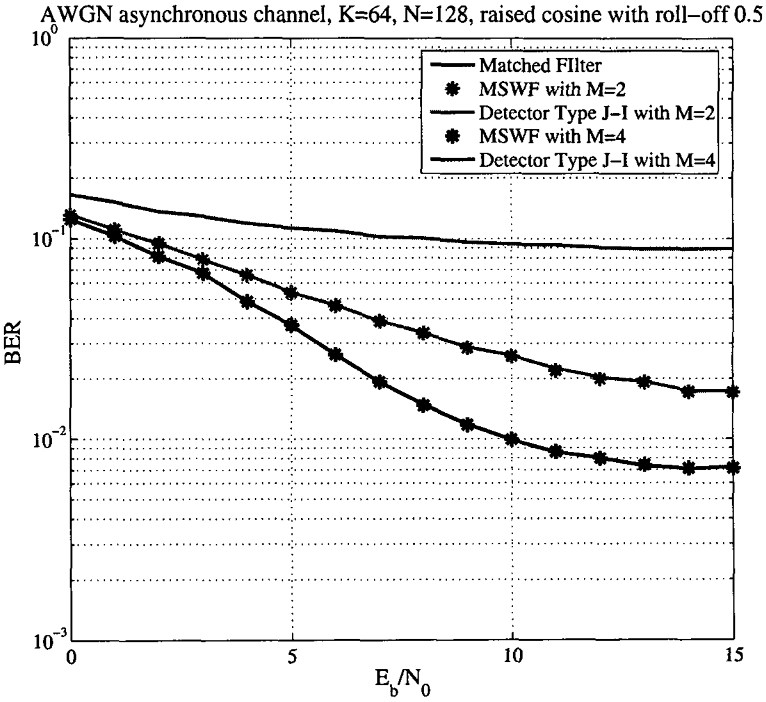


Figure 4.20: BER of matched filters, Type J-I detectors (solid lines), and MSWF (markers) with $M = 2$ and 4 versus $\frac{E_b}{N_0}$. CDMA systems with equal received powers, raised cosine chip waveforms (roll-off $\gamma = 0.5$), sampling rate $\frac{2}{T_c}$, and system load $\beta = \frac{1}{2}$ are considered.

Figure 4.20 by plotting their BER versus $\frac{E_b}{N_0}$. The performance of detectors Type J-I matches completely the BER of the corresponding MSWF. The comparison between the matched filter and the multistage detectors shows the substantial improvements achievable by multiuser detection.

4.6 Conclusions

In this chapter we studied low complexity linear multiuser detection techniques for asynchronous CDMA systems. Since the polynomial expansion detectors are intrinsically suboptimal for asynchronous systems, we focused on multistage Wiener filters and we extended the Type J-I detectors to asynchronous systems. In contrast to the full rank linear MMSE detector, the proposed implementation scheme for Type J-I detectors does not suffer from windowing effects and retains the same complexity, up to some memory elements, as the corresponding Type J-I detector for synchronous CDMA systems. Additionally, the proposed scheme is characterized by a sliding observation window. This yields a performance that is independent of the position of the detected symbol in the observation window.

The performance of linear MMSE detectors with finite observation windows was analyzed assuming a chip synchronous and symbol asynchronous system. Given a finite observation window we proposed an algorithm to determine the SINR at the output of the linear MMSE detector for all transmitted symbols that can be observed within the observation window. Unlike synchronous systems, in this scenario the Type J-I detector can outperform the full rank linear MMSE detector with given finite observation window when a sufficiently large number of stages is utilized. In fact, although the Type J-I detectors do not benefit from the available statistics in an optimal manner, as well as the linear MMSE detector does, they can use a wider set of observables while keeping low complexity.

The effects of different kinds of statistics on the SINR have been investigated. We considered the observables obtained at the output of a front-end performing low-pass filtering and subsequent sampling at a rate equal to twice the filter bandwidth. A second kind of decision statistics was obtained by sampling at the chip rate the output of a filter matched to the chip-pulse waveform. The performance analysis showed that the first group of observables was the most convenient for low complexity linear multiuser detection in the uplink channel. In fact, it enables to perform linear multiuser detection jointly for all users without incurring a degradation of the SINR.

The analysis of asynchronous systems revealed the effects of the chip-pulse waveforms and of the time delay distribution on the system performance.

We generalized the result for the ideal sinc waveform in [82] showing that the performance of synchronous and asynchronous systems is the same when the bandwidth of the chip-pulse waveform is not greater than half the chip rate. The effects of the chip-pulse waveform can be easily taken into account through certain coefficients $\mathcal{E}_s$, $s \in \mathbb{Z}^+$.

The impact of the chip-pulse waveforms on the SINR changes substantially as the bandwidth gets larger. In this case the system performance is significantly affected by the distribution of the time delays and the SINR of linear detectors may depend on the specific time delay of the signal of interest. We identified a large class of chip-pulse waveforms and time delay distributions for which the performance is independent of the time delay of the signal of interest and depends on the chip pulse waveforms through certain coefficients $\mathcal{E}_s$ (see (4.41)).

Specializing the general result on the asymptotic analysis of performance to square root raised cosine and raised cosine waveforms, we showed that the SINR of linear multiuser detectors optimum in a mean square sense increases significantly with the roll-off if the time delay is uniformly distributed. In contrast, it remains constant (square root raised cosine waveform) or decreases (raised cosine waveform) as the roll-off increases if the system is synchronous. Thus, an asynchronous CDMA system with linear multiuser detection and square root raised cosine or raised cosine waveforms outperforms the corresponding synchronous system in terms of SINR.

The analytical tools developed in this chapter for the analysis and design of low complexity multiuser detectors could be utilized for the optimization of chip pulse waveforms in CDMA systems with linear multiuser detectors. This application is beyond the scope of this work.

5 Linear Multiuser Detection and Correlated Spatial Diversity

5.1 Introduction

In this chapter we consider synchronous CDMA systems with spatial diversity both at the transmitting and the receiving sites. The analysis of linear multiuser detectors and the design of low complexity multistage detectors are extended to systems with correlated spatial diversity[1].

Modelling the spreading matrices as random matrices, Hanly and Tse [9] analyzed a CDMA system consisting of users transmitting to a multiuser receiver with spatial diversity. The spatial diversity can be obtained by multiple antenna elements at a single base station, or by combining the signals received at multiple base stations. In [9], these two cases of spatial diversity are referred to as micro-diversity and macro-diversity, respectively. This celebrated work covered many interesting aspects of CDMA systems with spatial diversity:

- There is a simple relation between the degrees of freedom introduced by spatial diversity (L receiving antennas) and the degrees of freedom in frequency given by the spread spectrum techniques (spreading factor N): the multi-antenna system behaves like a system with a single receive antenna but with spreading factor multiplied by the number of receiving antennas, and with the received power of each user being the sum of the received powers at the individual antennas. This behaviour is known as *resource pooling effect.* It shows the possibility to trade bandwidth (spreading factor) for the number of antennas and vice versa according to the peculiarity of the communication system.

- The effect of a single interferer onto the user of interest is captured by the concept of *effective interference.*

- Low complexity power control and admission control algorithms, the power-limited capacity region for a finite number of classes of users, and the interference-limited user capacity region are provided.

[1] Hereafter we refer to spatial diversity due to multiple input multiple output (MIMO) channels with correlation at the transmitting and/or the receiving sites as correlated spatial diversity.

The work in [9] is based on the performance analysis of linear multiuser receivers under the assumption that the spreading sequences are Gaussian and the random channel gains are circular symmetric and independent for all users and antennas, and that for any user the gains to all antennas are identically distributed. The analysis does not address cases of practical interest like multi-antenna systems with correlated channels and/or line-of-sight components.

The pioneering works in [7] and [8] on antenna arrays at the transmitter and receiver promise huge increases in the throughput of wireless communication systems. These works motivated many studies of the capacity of such systems in more realistic situations. In this stream are works that analyze the effects of channel correlation [95–102], line-of-sight components [62,63,103], multiple scattering [104], and keyholes [98] (this list does not claim to be comprehensive). Fading correlation and line-of-sight components were found to affect channel capacity severely. It is natural and of practical interest to consider their effects also in a CDMA system with spatial diversity.

In this work we consider a general framework with one or more antenna arrays at the receive side including combined micro- and macro-diversity scenarios. The transmitting users may use multiple element antennas, but need not do so. The channel gains may be correlated and contain line of sight components, i.e. their mean may be different from zero. The analysis is based on the assumption of independent random spreading. Our results include the results in [9] as special cases. Additionally, we provide a rigorous proof of the results for the macro-diversity case, only conjectured in [9].

In the micro-diversity case with independent channel gains analyzed in [9], the system behaviour is captured by the multiuser efficiency defined in Section 2.3.1. In [9], the multiuser efficiency is shown to converge to a deterministic constant in the large system limit. In the macro-diversity case with L receiving antennas, L constants, $a_1, a_2, \ldots a_L$, characterize the system. With correlated channel gains, we show that the large system behaviour is captured by a deterministic positive definite square Hermitian matrix with size equal to the number of receive antennas. Table 5.1 compares the scenarios investigated in [9] with the general case considered in this work. The results in [9] are revisited in the light of the general results so that all scenarios are represented by a matrix $\boldsymbol{A}$:

- In the micro-diversity scenario with independent channels, $\boldsymbol{A}$ is the identity matrix multiplied by the constant multiuser efficiency a.
- In the macro-diversity case with independent channels, $\boldsymbol{A}$ is a diagonal matrix.

In this chapter we analyze three linear receivers corresponding to different levels of knowledge of the interference structure and noise at the receiver:

Micro-diversity [9] • Channel gains independent for all users and antennas. • For a given user, the channel gains at all receiving antennas are identically distributed.	$\boldsymbol{A} = a\boldsymbol{I}_L$
Macro-diversity [9] • Channel gains independent for all users and antennas.	$\boldsymbol{A} = \begin{pmatrix} a_1 & 0 & \ddots & \ddots \\ 0 & a_2 & 0 & \ddots \\ \ddots & \ddots & \ddots & \ddots \\ \ddots & \ddots & 0 & a_L \end{pmatrix}$
General case • Correlated channel gains.	$\boldsymbol{A} = \boldsymbol{A}^H$

Table 5.1: Asymptotic constants characterizing CDMA systems with correlated and independent spatial diversity.

- The linear MMSE receiver, which requires a complete knowledge of the spreading sequences and the channel gains of the interferers.
- The single user Bayesian receiver, which assumes only a statistical knowledge of the spreading sequences and the channel gains of the interferers.
- The single user matched filter receiver. In this case the receiver has no information about the noise and interference.

Additionally, the design of low complexity multistage detectors is investigated. The universal weights for Type J-J and Type J-I detectors are derived. Type J-J and Type J-I detectors have a complexity order per transmitted bit which is linear in the number of transmitting antennas and in the number of users, as the SUMF receiver. The unified framework for the analysis of multistage linear detectors is extended to CDMA systems with correlated spatial diversity.

Thanks to the assumption of independence among the chips, the analysis shows that the performance of these linear receivers are not affected by channel correlation between transmitting antennas and suffer only from channel correlations among receiving antennas. For large CDMA systems without receive antenna diversity, the multiuser efficiency is identical for all users. Therefore, a single constant fully characterizes the system performance. In contrast, we show that the multiuser efficiency in CDMA systems with spatial diversity changes from user to user, in general. Additionally, we give sufficient conditions under which also a system with spatial diversity and statistically dependent channel gains is characterized by a unique multiuser efficiency. The single user Bayesian receiver and the SUMF receiver are shown to be asymptotically equivalent, in terms of SINR, to an *ideal* finite CDMA system with (i) linear MMSE detector and SUMF, respectively, at the receiver; (ii) spreading factor L; and (iii) spreading sequences equal to the vector of the channel gains.

5.2 System Model

We consider a CDMA system with spreading factor N and K' users. Each user employs a transmit antenna array with N_{T} elements sending independent data streams through each of the elements. Thus, we may speak of a system with $K = K'N_{\mathrm{T}}$ virtual users. The signal is received by L receive antennas. These antennas can be part of an array or can be placed at different locations, but processed jointly.

Assuming the channel to be flat fading, the baseband discrete-time system model is given by[2]

$$\mathfrak{y} = \mathfrak{H}\boldsymbol{b} + \mathfrak{n} \tag{5.1}$$

where $\mathfrak{y}$ is the NL-dimensional vector of received signal-samples, $\boldsymbol{b}$ is the K-dimensional vector of transmitted symbols, and $\mathfrak{n}$ is discrete-time, circularly symmetric complex-valued additive white Gaussian noise with zero mean and variance σ^2. The influence of spreading and fading is described by the $NL \times K$ matrix

$$\mathfrak{H} = \sum_{l=1}^{L} (\boldsymbol{S}\boldsymbol{D}\boldsymbol{\Lambda}_l) \otimes \mathrm{e}_l \tag{5.2}$$

where $\boldsymbol{S}$ is the $N \times K$ spreading matrix whose k^{th} column is the spreading sequence of the k^{th} virtual user. Furthermore, the diagonal square matrix $\boldsymbol{D} \in \mathbb{C}^{K\times K}$ contains the transmitted amplitudes of all virtual users such that its k^{th} diagonal element d_k is the amplitude of the signal transmitted by the virtual user indexed by k. The diagonal matrices $\boldsymbol{\Lambda}_1, \boldsymbol{\Lambda}_2, \ldots, \boldsymbol{\Lambda}_L \in \mathbb{C}^{K\times K}$ take into account the effect of the flat fading channel. The k^{th} diagonal element of $\boldsymbol{\Lambda}_l$ is the channel gain between the transmitting antenna element of the k^{th} virtual user and the l^{th} receive antenna and will be denoted by λ_{lk} in the following. The channel gains can, in general, be

[2] As in the case of synchronous CDMA systems with multipath fading channel, the received signal at the symbol time interval n depends only on the transmitted signal at the same symbol interval. Therefore, the symbol-time index n will be omitted in the system model (5.1).

correlated and contain line of sight components as in Rice channels. $\boldsymbol{e}_l$ is the L-dimensional unit column vector whose elements are zero except the l^{th} element that equals 1, i.e. $\boldsymbol{e}_l = (\delta_{lj})_{j=1}^{L}$. Finally, $\otimes$ denotes the Kronecker product.

In the following, the spreading matrix $\boldsymbol{S}$ is modelled as a random matrix whose elements are independent[3] with zero mean, variance $\frac{1}{N}$, and fourth moment such that there exists a $\gamma > 1$ for which $\mathrm{E}\{|s_{11}|^4\} \leq \frac{1}{N^\gamma}$. This condition is satisfied by all chips of practical use as Gaussian or binary chips. Moreover, we assume the transmitted symbols to be uncorrelated and identically distributed random variables with zero mean and unit variance, i.e. $\mathrm{E}\{\boldsymbol{b}\boldsymbol{b}^H\} = \boldsymbol{I}_K$. In order to simplify the notation, it will be helpful in the following to define the L-dimensional vectors $\boldsymbol{l}_k = d_k[\lambda_{1k}, \lambda_{2k}, \ldots, \lambda_{Lk}]^T$, $k = 1, \ldots, K$, $\boldsymbol{l} = d[\lambda_1, \lambda_2, \ldots, \lambda_L]^T$, $k = 1, \ldots, K$ and the diagonal square matrices $\boldsymbol{L}_\ell = \boldsymbol{D}\boldsymbol{\Lambda}_\ell$, $\ell = 1, \ldots, L$.

5.3 Linear MMSE Receiver

Throughout this chapter we adopt the following notation:

- $\mathfrak{h}_k$ denotes the k^{th} column of $\mathfrak{H}$;
- $\mathfrak{H}_k$ is the $NL \times (K-1)$ matrix obtained from $\mathfrak{H}$ by suppressing the k^{th} column $\mathfrak{h}_k$.

The linear MMSE detector generates a soft decision $\widehat{\boldsymbol{b}}_k = \boldsymbol{c}_k^H \mathfrak{y}$ based on the observation $\mathfrak{y}$. The linear MMSE detector $\boldsymbol{c}_k$ for the detection of b_k, the transmitted symbol of user k, can be derived from the Wiener-Hopf theorem [20] for the estimation of zero-mean random variables. It is given by

$$\boldsymbol{c}_k = \mathrm{E}\{\mathfrak{y}\mathfrak{y}^H\}^{-1}\mathrm{E}\{b_k^*\mathfrak{y}\} \tag{5.3}$$

with the expectation taken over all variables that are unknown to the receiver, i.e. the transmitted symbols $\boldsymbol{b}$ and the noise. Specializing the Wiener-Hopf equation to the system model (5.1) yields

$$\begin{aligned} \boldsymbol{c}_k &= (\mathfrak{H}\mathfrak{H}^H + \sigma^2\boldsymbol{I})^{-1}\mathfrak{h}_k & (5.4)\\ &= c \cdot (\mathfrak{H}_k\mathfrak{H}_k^H + \sigma^2\boldsymbol{I})^{-1}\mathfrak{h}_k & (5.5) \end{aligned}$$

for some $c \in \mathbb{R}$. The second expression follows from the matrix inversion lemma. Its performance is measured by the signal-to-interference-and-noise ratio SINR_k at its output which is well known [105] to be given by

$$\mathrm{SINR}_k = \mathfrak{h}_k^H(\mathfrak{H}_k\mathfrak{H}_k^H + \sigma^2\boldsymbol{I})^{-1}\mathfrak{h}_k. \tag{5.6}$$

[3] Note that the random variables s_{nk} are not required to be identically distributed.

The SINR_k can be conveniently expressed in terms of the multiuser efficiency η_k (see Definition 1 and the subsequent Equation (2.24)):

$$\text{SINR}_k = \frac{||\boldsymbol{l}_k||^2}{\sigma^2}\eta_k. \tag{5.7}$$

5.3.1 General Case

Let us notice that SINR_k depends on the spreading sequences and the channel parameters of all users. To get deeper insights into the behaviour of the linear MMSE detector it is convenient to analyze the performance, as $K, N \to \infty$ with constant ratio $\beta = \frac{K}{N}$. To this aim, we have to define how the matrices $\boldsymbol{D}, \boldsymbol{\Lambda}_1, \boldsymbol{\Lambda}_2, \ldots, \boldsymbol{\Lambda}_L$ behave as the system grows large. Let us consider a system with K virtual users and the K corresponding $(L+1)$-variate random variables $(d_k, \lambda_{1k}, \lambda_{2k}, \ldots, \lambda_{Lk})$ for $k = 1, \ldots, K$. The empirical joint distribution function for the random variables $(d_k, \lambda_{1k}, \lambda_{2k}, \ldots, \lambda_{Lk})$ for $k = 1, \ldots, K$ is the distribution function

$$F^{(K)}_{\boldsymbol{D},\boldsymbol{\Lambda}_1,\boldsymbol{\Lambda}_2,\ldots,\boldsymbol{\Lambda}_L}(d, \lambda_1, \lambda_2, \ldots, \lambda_L) = \frac{1}{K}\sum_{k=1}^{K} 1(d - d_k, \lambda_1 - \lambda_{1k}, \lambda_2 - \lambda_{2k}, \ldots, \lambda_L - \lambda_{Lk})$$

where $1(\cdot)$ is the multidimensional indicator function (see definition in the Glossary). We assume that the joint empirical distribution $F^{(K)}_{\boldsymbol{D},\boldsymbol{\Lambda}_1,\boldsymbol{\Lambda}_2,\ldots,\boldsymbol{\Lambda}_L}(d, \lambda_1, \lambda_2, \ldots, \lambda_L)$ converges weakly with probability 1 to a limit distribution function $F_{\boldsymbol{D},\boldsymbol{\Lambda}_1,\boldsymbol{\Lambda}_2,\ldots,\boldsymbol{\Lambda}_L}(d, \lambda_1, \lambda_2, \ldots, \lambda_L)$ with bounded support. Let us notice that, if the $(d_k, \lambda_{1k}, \lambda_{2k}, \ldots, \lambda_{Lk})$, for all k, are independent realizations of a common distribution function, then the empirical distribution function $F^{(K)}_{\boldsymbol{D},\boldsymbol{\Lambda}_1,\boldsymbol{\Lambda}_2,\ldots,\boldsymbol{\Lambda}_L}$ is the natural estimate of the common c.d.f.. The Glivenko-Cantelli theorem (see e.g. [89]) guarantees that, if $(d_k, \lambda_{1k}, \lambda_{2k}, \ldots, \lambda_{Lk})$ are i.i.d. over k, then the empirical distribution converges weakly to the common distribution function with probability 1. For example, if, for each user k, $(d_k, \lambda_{1k}, \lambda_{2k}, \ldots, \lambda_{Lk})$ is a realization of the same Gaussian distribution $F(d, \lambda_1, \lambda_2, \ldots, \lambda_L)$, then the Glivenko-Cantelli lemma guarantees that the sequence of the empirical distribution functions converges almost surely to the same distribution function $F(d, \lambda_1, \lambda_2, \ldots, \lambda_L)$.

In the following, we simplify, where possible, the notation using the limiting joint distribution $F_{\boldsymbol{l}}(l_1, l_2, \ldots, l_L) = F_{\boldsymbol{l}}(\boldsymbol{l})$ rather than the limit distribution $F_{\boldsymbol{D},\boldsymbol{\Lambda}_1,\boldsymbol{\Lambda}_2,\ldots,\boldsymbol{\Lambda}_L}(d, \lambda_1, \lambda_2, \ldots, \lambda_L)$ (recall that $\boldsymbol{l} = [l_1, l_2, \ldots, l_L]^T = d[\lambda_1, \lambda_2, \ldots, \lambda_L]^T$). Under the above assumptions the asymptotic performance depends on a small set of parameters, as shown by the following theorem.

Theorem 11 *Let $\boldsymbol{S}$ be an $N \times K$ random matrix with independent entries s_{ij} that are zero mean, with variance* $\mathrm{E}\{|s_{ij}|^2\} = \frac{1}{N}$ *and forth moment* $\mathrm{E}\{|s_{ij}|^4\} \leq \frac{1}{N^\gamma}$

where $\gamma > 1$. Let $\boldsymbol{l} = (l_1, l_2, \ldots, l_L)$ and let $\boldsymbol{l}_k$ be the vector of received amplitudes of the virtual user k. Assume that the norm of the channel gain vector $\|\boldsymbol{l}_k\|$ is uniformly bounded for all K. Furthermore, the empirical joint distribution of $\boldsymbol{l}_1, \boldsymbol{l}_2, \ldots, \boldsymbol{l}_{k-1}, \boldsymbol{l}_{k+1}, \ldots \boldsymbol{l}_K$ converges almost surely to some limiting joint distribution $F_l(\boldsymbol{l})$ as $K \to \infty$. Then, as $N, K \to \infty$ with $\frac{K}{N} \to \beta$ and L fixed, the SINR of virtual user k, given the fading amplitude $\boldsymbol{l}_k$, converges in probability to the value

$$\lim_{K,N\to\infty} \mathrm{SINR}_k \stackrel{\mathcal{P}}{=} \frac{\boldsymbol{l}_k^H \boldsymbol{A} \boldsymbol{l}_k}{\sigma^2} \tag{5.8}$$

where $\boldsymbol{A}$ is the unique deterministic $L \times L$ matrix solution to the matrix-valued fixed point equation

$$\boldsymbol{A}^{-1} = \boldsymbol{I}_L + \beta \int \frac{\boldsymbol{l}\boldsymbol{l}^H}{\sigma^2 + \boldsymbol{l}^H \boldsymbol{A} \boldsymbol{l}} \, \mathrm{d}F_l(\boldsymbol{l}) \tag{5.9}$$

such that $\boldsymbol{A}$ is positive definite for any positive value of the noise variance.

Proof: See Appendix D Section D.1.

Theorem 11 provides the asymptotic output SINR of a linear MMSE detector for a synchronous CDMA system with correlated spatial diversity. This result holds under very general conditions on the channel gains and demonstrates interesting and useful properties of synchronous CDMA systems with correlated spatial diversity and linear MMSE detector at the receiver. The remainder of this section is devoted to the discussion of these properties. More specifically, in Subsection 5.3.2 Theorem 11 is specialized to the relevant situation of practical interest where the received amplitudes are correlated Gaussian distributed. In Subsection 5.3.3 Theorem 11 is utilized to derive sufficient conditions under which the resource pooling effect arises. General properties of CDMA systems with correlated spatial diversity evinced from Theorem 11 are presented in Subsection 5.3.4.

5.3.2 Correlated Gaussian Received Amplitudes

In practice, fading amplitudes are often complex Gaussian distributed and correlated. Rayleigh fading also violates the assumption of uniformly bounded channel gains. However, it can be approximated arbitrary closely by a distribution with bounded support. Thus, from an engineering perspective, we need not worry about that fact. Assume that the limiting joint distribution is given as

$$f_L(\boldsymbol{l}) = \frac{1}{\pi^L \det \boldsymbol{C}_l} \exp\left(-\boldsymbol{l}^H \boldsymbol{C}_l^{-1} \boldsymbol{l}\right). \tag{5.10}$$

In the absence of power control, i.e. $\boldsymbol{D} = \boldsymbol{I}_K$, this implies that $\boldsymbol{C}_l$ is the correlation matrix of the fading at the receive side, with entries

$$c_{ij} = \mathrm{E}\left\{\lambda_i \lambda_j^*\right\}. \tag{5.11}$$

Consider the eigenvalue decomposition

$$C_l = M\Psi M^H, \tag{5.12}$$

with $\Psi = \mathrm{diag}(\psi_1, \ldots, \psi_L)$, and the change of variables

$$g = M^H l \tag{5.13}$$

$$g_k = [g_{1k}, \ldots, g_{Lk}]^T = M^H l_k. \tag{5.14}$$

The components of the random vector g are statistically independent. Plugging (5.13) into (5.9), the matrix M also diagonalizes the deterministic limit matrix A, i.e. the eigenvectors of the matrix A coincide with the eigenvectors of the correlation matrix C_l . Thus, we obtain for correlated Rayleigh fading

$$\lim_{K,N\to\infty} \mathrm{SINR}_k \overset{\mathcal{P}}{=} \frac{1}{\sigma^2} \sum_{\ell=1}^{L} a_\ell |g_{\ell k}|^2 \tag{5.15}$$

where the a_i, $i = 1, \ldots, L$ are the solution to the fixed point equations

$$a_\ell = \frac{1}{1 + \frac{\beta}{\pi^L} \int \frac{\psi_l |x_l|^2}{\sigma^2 + \sum_{n=1}^{L} a_n \psi_n |x_n|^2} \prod_{n=1}^{L} \exp(-|x_n|^2) \mathrm{d}\, x_n} \quad \forall \ell = 1, \ldots, L. \tag{5.16}$$

Thus, recalling the characteristics of the macro-diversity in [9] presented in Table 5.1 and comparing the previous result to the macro-diversity case, it is evident that to any correlated Rayleigh fading scenario, there exists an equivalent macro-diversity scenario as in [9] with independent Rayleigh fading.

5.3.3 Uncorrelated Received Amplitudes

It is clear from (5.7) and (5.8) that, unless the matrix A is a multiple of the identity matrix, the multiuser efficiency is, in general, not identical for all users. In this section we analyze under which conditions on the limiting joint distribution $F_l(l_1, l_2, \ldots, l_L)$ or, equivalently, on the corresponding limiting probability density function $f_l(l_1, l_2, \ldots, l_L)$ the matrix A is diagonal or proportional to the identity matrix. In fact, for diagonal A, the general result in Theorem 11 simplifies to the system of fixed-point equations in [9], Theorem 3. The following corollary summarizes some sufficient conditions that yield a diagonal structure of A.

Corollary 6 *Let S and l_k be as in Theorem 11. If the joint probability density function $f_l(l_1, l_2, \ldots, l_L)$ is an even function of $\mathrm{Re}(l_r)$ and $\mathrm{Im}(l_r)$, for any r and for any value of the parameters $(l_1, \ldots, l_{k-1}, l_{k+1}, \ldots l_L)$, then, as $N, K \to \infty$ with*

$\frac{K}{N} \to \beta$ and L fixed, the SINR of virtual user k, given the fading amplitude $\boldsymbol{l}_k$, converges in probability to the value

$$\lim_{K,N\to\infty} \mathrm{SINR}_k \stackrel{\mathcal{P}}{=} \frac{1}{\sigma^2}\sum_{\ell=1}^{L} a_\ell |l_{\ell k}|^2, \tag{5.17}$$

where a_ℓ, $\ell = 1 \ldots L$, are the unique positive solutions to the system of fixed-point equations

$$a_\ell = \frac{1}{1+\beta\displaystyle\int \frac{|l_\ell|^2}{\sigma^2+\sum_{n=1}^{L} a_n|l_n|^2} f_l(l_1,l_2,\ldots,l_L)\mathrm{d}\,l_1\mathrm{d}\,l_2\ldots\mathrm{d}\,l_L} \quad \forall\,\ell = 1,\ldots,L. \tag{5.18}$$

Proof: In order to verify that system (5.9) is equivalent to system (5.18) under the above mentioned conditions on $f_l(l_1, l_2, \ldots, l_L)$, it is sufficient to verify that, for all $i, j = 1, \ldots, L$ with $i \neq j$, the off-diagonal elements of $\boldsymbol{A}$ are zero. The uniqueness of the solution for system (5.9) guarantees that the constants a_ℓ are the solution we are looking for. In fact, $\forall i, j = 1, \ldots, L$ and $i \neq j$ the off-diagonal elements of $\boldsymbol{A}$ are given by

$$\int \frac{l_i l_j\, f_l(l_1, l_2 \ldots l_L)}{\sigma^2+\sum_{\ell=1}^{L} a_\ell|l_\ell|^2}\mathrm{d}\,l_1\ldots\mathrm{d}\,l_L = \int l_i \mathrm{d}\,l_1\ldots\mathrm{d}\,l_{j-1},\mathrm{d}\,l_{j+1}\ldots\mathrm{d}\,l_L \int \frac{l_j f_l(l_1,l_2\ldots l_L)}{\sigma^2+\sum_{\ell=1}^{L} a_\ell|l_\ell|^2}\mathrm{d}\,l_j.$$

Since the function $l_j/(\sigma^2 + \sum_{\ell=1}^{L} a_\ell|l_\ell|^2)$ is an odd function of $\mathrm{Re}(l_j)$ and $\mathrm{Im}(l_j)$, the integral with respect to l_j will be always zero if $f_l(l_1, l_2, \ldots, l_L)$ is an even function in $\mathrm{Re}(l_j)$ and $\mathrm{Im}(l_j)$ for all possible values of l_j with $j = 1, \ldots, L$ and $j \neq i$. Then, the off-diagonal elements of $\boldsymbol{A}$ are zero and this concludes the proof of Corollary 6.

Following the same approach used for Corollary 4 in [9] and using Corollary 6 we find sufficient conditions under which the matrix $\boldsymbol{A}$ is proportional to the identity matrix, i.e. $\boldsymbol{A} = \eta\boldsymbol{I}$. If $\boldsymbol{A} = \eta\boldsymbol{I}$, then the scalar η coincides with the multiuser efficiency of a linear MMSE detector as it is apparent from (5.7) and (5.8). Let us assume that the conditions of Corollary 6 are satisfied. If we additionally assume that the joint probability density function $f_l(l_1, l_2, \ldots, l_L)$ is exchangeable, i.e. for any permutation π of $\{1, \ldots, L\}$

$$f_l(l_1, l_2, \ldots, l_L) = f_l(l_{\pi(1)}, l_{\pi(2)}, \ldots, l_{\pi(L)})$$

then the system of equations (5.18), which defines the diagonal matrix $\boldsymbol{A}$, satisfies $a_\ell = \eta$, for all $\ell = 1, \ldots, L$, $\boldsymbol{A} = \eta\boldsymbol{I}$, and the system of equations (5.18) reduces a single fixed-point equation. This result is stated in the following corollary.

Corollary 7 *Let $\boldsymbol{S}$ and $f_l(l_1, l_2, \ldots, l_L)$ be as in Corollary 6. If the limiting probability density function $f_l(l_1, l_2, \ldots, l_L)$ is exchangeable, i.e. for any permutation π of $\{1, \ldots, L\}$*

$$f_l(l_1, l_2, \ldots, l_L) = f_l(l_{\pi(1)}, l_{\pi(2)}, \ldots, l_{\pi(L)})$$

then, as $N, K \to \infty$ with $\frac{K}{N} \to \beta$ and L fixed, $\frac{\mathrm{SINR}_k \sigma^2}{P_k}$, with $P_k = ||\boldsymbol{l}_k||^2$ converges in probability to the deterministic constant η, which is the unique scalar multiuser efficiency, solution to the fixed point equation

$$\eta = \frac{1}{1 + \frac{\beta}{L} \int \frac{P}{\sigma^2 + \eta P} \mathrm{d}\, F_P(P)}. \tag{5.19}$$

Here P is the random variable defined by $P = ||\boldsymbol{l}||^2$ and $F_P(P)$ is its distribution.

The conditions of Corollaries 6 and 7 imply that $l_1, l_2, \ldots l_L$ are uncorrelated. However, the converse is not true in general, i.e. the matrix $\boldsymbol{A}$ is not typically diagonal for asymptotically uncorrelated received amplitudes. Corollaries 6 and 7 provide sufficient conditions for the matrix $\boldsymbol{A}$ to be diagonal and proportional to the identity matrix, respectively, when the received amplitudes are asymptotically uncorrelated. Under the conditions of Corollary 7 the resource pooling effect arises. In fact, let us compare (5.19) to (2.27). It is apparent that the multiuser efficiency of a synchronous CDMA system with L receive antennas and spreading factor N is equal to the multiuser efficiency of a synchronous CDMA system with single receive antenna, spreading factor NL, and with received power of each user being the sum of the received powers at the individual antennas.

5.3.4 Remarks

The empirical joint distributions $F_l^{(K)}(l_1, l_2, \ldots, l_L)$ and the limiting joint distribution $F_l(l_1, l_2, \ldots, l_L)$ are not able to capture and describe the effects of the correlation due to antenna coupling at the transmitter side. Since the effects of the channel gains on the system performance are taken into account only by $F_l(l_1, l_2, \ldots, l_L)$, we can conclude that the correlations of the channel gains due to coupling effects at the transmitter side do not affect the asymptotic performance of the linear MMSE receiver. This property is intrinsically related to the assumption of the statistical independence of the spreading sequences of the transmitting antennas. It does not hold true if the condition of independence is not satisfied. In fact, in this case, $F_l(l_1, l_2, \ldots, l_L)$ would not be sufficient to describe the system behaviour.

As a consequence of Theorem 11, the asymptotic behaviour of the general system is completely described by an $L \times L$ matrix $\boldsymbol{A}$. In contrast to the case of a single receive antenna or the cases in which the resource pooling effect arises, the multiuser efficiency of the linear MMSE receiver varies from user to user, in general. In particular, for user k, it depends on the direction of the channel gains, $\boldsymbol{l}_k$, with respect to the eigenvectors of $\boldsymbol{A}$: The SINR is maximum if $\boldsymbol{l}_k$ has the same direction as the eigenvector corresponding to the maximum eigenvalue of $\boldsymbol{A}$.

Typically, in order to determine the eigenvectors of the matrix $\boldsymbol{A}$ the solution of the matrix fixed point equation (5.9) is required. However, in the special case where the limiting p.d.f. $f_l(l_1, l_2, \ldots, l_L)$ is Gaussian, the eigenvectors of $\boldsymbol{A}$ coincide with the eigenvectors of the covariance matrix $\mathrm{E}\{\boldsymbol{l}\boldsymbol{l}^H\}$[4].

In the light of Theorem 11 we can revisit the known results in [9]. Theorem 1 in [9] states that if the elements s_{ij} are i.i.d. Gaussian random variables with zero mean and variance $\frac{1}{N}$, if the received amplitudes $l_{\ell k}$ are independent for all users k and antennas ℓ, if for any given user, the received amplitudes are identically distributed, and if asymptotically the sequence of the empirical distributions converges to a bounded distribution function, then the matrix $\boldsymbol{A}$ in (5.8) is given by

$$\boldsymbol{A} = a\boldsymbol{I}_L$$

with

$$a = \left(1 + \frac{\beta}{L}\mathrm{E}\left(\frac{P}{\sigma^2 + aP}\right)\right)^{-1}$$

where $P = \boldsymbol{l}^H\boldsymbol{l}$.

By comparing this result to the result in Corollary 6 it becomes evident that Corollary 7 implies Theorem 1 in [9].

The following result is conjectured in Theorem 3 in [9]. If the chip elements s_{ij} are independent, zero mean, Gaussian distributed, the received signal amplitudes are independent and if, asymptotically, the sequence of the empirical distribution converges to a bounded distribution function, then Theorem 3 in [9] conjectures that the matrix $\boldsymbol{A}$ in (5.8) is given by

$$\boldsymbol{A} = \mathrm{diag}(a_1, a_2, \ldots, a_L)$$

with

$$a_\ell = \left(1 + \beta\mathrm{E}\left(\frac{|l_\ell|^2}{\sigma^2 + \sum_{n=1}^{L} a_n |l_n|^2}\right)\right)^{-1} \qquad \ell = 1, \ldots, L.$$

By comparing the previous conjecture to Corollary 6, we notice that Corollary 6 includes and proves rigorously Theorem 3 in [9].

In Table 5.2 we recapitulate the results of Corollary 6 and Corollary 7 and summarize the sufficient conditions under which the resource pooling effect arises.

[4] We assume here that the asymptotic channel gains $\lambda_1, \lambda_2, \ldots, \lambda_L$ are zero mean as is typical in a baseband model.

[5] As already noticed in Section 5.3.3, Hypothesis A implies that $l_1, l_2, \ldots l_L$ are uncorrelated.

Assumptions	Hypothesis A	The joint probability density function $f(l_1, l_2, \ldots, l_L)$ is an even function of $\mathrm{Re}(l_i)$ and $\mathrm{Im}(l_i)$ for $i = 1, \ldots, L$. $\Downarrow$ Received amplitudes uncorrelated at the receiver[5].
	Hypothesis B	$f(l_1, l_2, \ldots, l_L)$ is exchangeable, i.e. for any permutation π of $\{1, 2, \ldots, L\}$ $f(l_1, l_2, \ldots, l_L) = f(l_{\pi(1)}, l_{\pi(2)}, \ldots, l_{\pi(L)})$
Implications	Hypothesis B	$\boldsymbol{A}$ is diagonal
	Hypothesis A and B	$\boldsymbol{A} = a\boldsymbol{I}$
Conclusions	The resource pooling effect arises if both hypotheses A and B are satisfied.	

Table 5.2: Summary of Corollary 6 and Corollary 7.

5.4 Single User Bayesian Receiver

The single user Bayesian receiver is a linear detector that is suitable when the receiver is synchronized and has complete information about the user of interest, i.e. spreading sequence, and received power, but it does not know the spreading sequences of the interferers and has only statistical knowledge of the interference. More specifically, we assume that the following information is known at the receiver:

- Knowledge of the signature sequence, channel gains, and transmit power of user k;
- Knowledge of the statistics of the signature sequences, channel gains, and transmit powers of all interferers.

This detector has been analyzed for the case of independent channel gains in [9] under the denomination of *matched filter*.

The Bayesian single user detector $\boldsymbol{c}_k$ for the user of interest k minimizes the mean square error between its output, $\widehat{b}_{\mathrm{Bf},k} = \boldsymbol{c}_k^H \boldsymbol{y}$, and the transmitted symbol b_k. It is given by the Wiener-Hopf equation

$$\boldsymbol{c}_k = \mathrm{E}\{\mathfrak{y}\mathfrak{y}^H\}^{-1}\mathrm{E}\{b_k^* \mathfrak{y}\} \tag{5.20}$$

as for the linear MMSE receiver (5.3). However, in this case the expectation is taken not only over the transmitted signals and the noise, as for the linear MMSE receiver, but also with respect to the signature sequences, the channel gains, and the transmit powers of all interferers. Equation (5.20) yields the following explicit expression for the Bayesian filter

$$\boldsymbol{c}_k = \frac{[\boldsymbol{I}_N \otimes [(\beta - \frac{1}{N})\boldsymbol{C}_l + \sigma^2 \boldsymbol{I}_L]^{-1}]\mathfrak{h}_k}{1 + [\boldsymbol{l}_k^H [(\beta - \frac{1}{N})\boldsymbol{C}_l + \sigma^2 \boldsymbol{I}_L]^{-1} \boldsymbol{l}_k] \boldsymbol{s}_k^H \boldsymbol{s}_k}. \tag{5.21}$$

A better insight into the Bayesian filter receiver can be obtained from (5.21) by performing a permutation Π of the elements of $\boldsymbol{c}_k$ and $\mathfrak{y}_k$ and such that the elements corresponding to the same antenna are relocated next to each other ($\Pi : i \rightarrow [(i - 1)\mathrm{mod}L]N + \lfloor \frac{i}{L} \rfloor + 1$). Let us denote with $\boldsymbol{c}_k^{\Pi}$ and $\mathfrak{y}^{\Pi}$ the Bayesian filter receiver and the received signal vector respectively obtained by such a permutation. Let ξ_l be the l^{th} element of the vector $\boldsymbol{\xi}_k = [(\beta - \frac{1}{N})\boldsymbol{C}_l + \sigma^2 \boldsymbol{I}_L]^{-1} \boldsymbol{l}_k$. Then,

$$\boldsymbol{c}_k^{\Pi} = \left[\frac{\xi_1 \boldsymbol{s}_k^T}{1 + (\boldsymbol{l}_k^H \boldsymbol{\xi}_k)(\boldsymbol{s}_k^H \boldsymbol{s}_k)}, \ldots, \frac{\xi_L \boldsymbol{s}_k^T}{1 + (\boldsymbol{l}_k^H \boldsymbol{\xi}_k)(\boldsymbol{s}_k^H \boldsymbol{s}_k)} \right]^T. \tag{5.22}$$

Equation (5.22) shows that, similarly to the case of completely independent channel gains in [9], the Bayesian filter despreads the received signal at each antenna using the spreading sequence $\boldsymbol{s}_k$ and, then, it performs a maximal ratio combining of the despread signals using as weights the coefficients

$$\frac{\xi_l}{1+(\boldsymbol{l}_k^H\boldsymbol{\xi}_k)(\boldsymbol{s}_k^H\boldsymbol{s}_k)} \qquad l=1,\ldots,L. \tag{5.23}$$

The coefficients for maximal ratio combining depend on the correlation matrix of the channel gains of the interferers.

The following theorem provides the performance of the Bayesian filter in terms of its limiting SINR as the system dimensions grow large with constant ratio.

Theorem 12 *Let $\boldsymbol{l}_k$ be the vector of received amplitudes of user k. Let us assume that, almost surely, the empirical joint distribution of $\boldsymbol{l}_1, \boldsymbol{l}_2, \ldots, \boldsymbol{l}_{k-1}, \boldsymbol{l}_{k+1}, \ldots \boldsymbol{l}_K$ converges to some limiting joint distribution $F_{\boldsymbol{l}}$ as $K \to \infty$. Additionally, the elements of the spreading vector $\boldsymbol{s}_k$ are assumed to be independent and identically distributed. Then, if $N, K \to \infty$ with $\frac{K}{N} \to \beta$ and L fixed, SINR_k of the Bayesian filter for the transmitted signal k, conditioned on the vector of received amplitudes $\boldsymbol{l}_k$, converges almost surely to a constant value*

$$\lim_{\substack{K,N\to\infty \\ \frac{K}{N}\to\beta}} \mathrm{SINR}_k \overset{a.s.}{=} \boldsymbol{l}_k^H(\beta \mathrm{E}\{\boldsymbol{l}\boldsymbol{l}^H\}+\sigma^2\boldsymbol{I}_L)^{-1}\boldsymbol{l}_k \tag{5.24}$$

where $\boldsymbol{l}$ is the L-variate random variable with joint distribution F.

Proof: See Appendix D Section D.2.

The asymptotic analysis provides a result of simple interpretation: the SINR of user k is equivalent to the SINR at the output of a linear MMSE detector for a CDMA system with:

- Spreading factor equal to the number of receiving antennas;
- Spreading sequence of the user of interest equal to the vector $\boldsymbol{l}_k$ of channel gains;
- Spreading sequences of the interferers equal to the vectors of the channel gains attenuated by a factor $\sqrt{\beta}$. This takes into account the beneficial effects of the spreading in the original CDMA system.

In contrast to the case of independent channel gains in [9], the performance depends on the direction of the vector of the channel gains. For a given received power the SINR is maximized as $\boldsymbol{l}_k$ has the direction of the eigenvector corresponding to the minimum eigenvalue of the matrix $(\beta\mathrm{E}\{\boldsymbol{l}\boldsymbol{l}^H\}+\sigma^2\boldsymbol{I})^{-1}$.

5.5 Matched Filter

The single user matched filter requires only the knowledge of the spreading sequence of the user of interest. Its output is given by

$$\widehat{b}_{\mathrm{mf},k} = \mathfrak{h}_k^H \mathfrak{y}.$$

As in the case of the single user Bayesian receiver, the matched filter despreads the received signals at each antenna and then it combines the despread signals using as weight coefficients the received energy at each antenna.

The asymptotic performance of the single user matched filter is given by the following theorem:

Theorem 13 *Let $\boldsymbol{l}_k$, $\boldsymbol{s}_k$, and $F_{\boldsymbol{l}}(l_1, l_2, \dots l_L)$ be as in Theorem 12.*

Then, if $N, K \to \infty$ with $\frac{K}{N} \to \beta$ and L fixed, SINR$_k$ *of the matched filter for the transmitted signal k, conditioned on the vector of received amplitudes $\boldsymbol{l}_k$, converges almost surely to a constant value*

$$\lim_{\substack{K,N\to\infty \\ \frac{K}{N}\to\beta}} \mathrm{SINR}_k \stackrel{a.s.}{=} \frac{(\boldsymbol{l}_k^H \boldsymbol{l}_k)^2}{\boldsymbol{l}_k^H(\beta \mathrm{E}\{\boldsymbol{l}\boldsymbol{l}^H\} + \sigma^2 \boldsymbol{I}_L)\boldsymbol{l}_k} \tag{5.25}$$

where $\boldsymbol{l}$ is the L-variate random variable with joint distribution $F_{\boldsymbol{l}}(l_1, l_2, \dots l_L)$.

Theorem 13 is proven in Appendix D Section D.3.

Like for the single user Bayesian filter, the SINR of the matched filter is equivalent to the SINR of a matched filter for a CDMA system with spreading factor L, spreading sequence of the user of interest equal to the vector of the channel gains, and spreading sequence of the interferers equal to their channel gains attenuated by a factor $\sqrt{\beta}$.

The spectral efficiency depends on the direction of $\boldsymbol{l}_k$. It is maximized when $\boldsymbol{l}_k$ has the direction of the eigenvector corresponding to the maximum eigenvalue of the correlation matrix of the interferers, $\mathrm{E}\{\boldsymbol{l}\boldsymbol{l}^H\}$.

5.6 Multistage Detection

The design of detectors Type J-J and Type J-I and the general result for the analysis of linear multistage detectors for CDMA systems with spatial diversity follows along the same lines as for the synchronous systems presented in Chapter 3.

The projection subspace that enables joint projection is given by

$$\chi_{M,k}(\mathfrak{H}) = \mathrm{span}\{\mathfrak{T}^m \mathfrak{h}_k\}_{m=0}^{M-1}$$

where $\mathfrak{T} = \mathfrak{H}\mathfrak{H}^H$. The design of the asymptotic weights reduces to the determination of the asymptotic values of the diagonal elements of $\mathfrak{R}^m = (\mathfrak{H}^H\mathfrak{H})^m$, $m = 1, \dots, 2M$. With obvious substitutions, equations (3.11) and (3.10) define detector Type J-J and Type J-I, respectively, for systems with spatial diversity. The same holds for equations (3.32) and (3.31).

The following theorem states that $(\mathfrak{R}^m)_{kk}$ converges almost surely to a deterministic value conditionally on $\boldsymbol{l}_k$. A recursive algorithm to compute such a limiting value is also provided.

Theorem 14 *Let the matrix $\boldsymbol{S}$ be as in Theorem 6. Let the vectors $\boldsymbol{l}_k$, $k = 1, \dots K$ and the c.d.f. $F_l(l_1, l_2, \dots, l_L)$ be defined as in Theorem 11. $\boldsymbol{L}_\ell$, $\ell = 1, \dots, L$, is a $K \times K$ diagonal matrix whose k^{th} element coincides with the ℓ^{th} component of $\boldsymbol{l}_k$, i.e. $(\boldsymbol{L})_{kk} = (\boldsymbol{l}_\ell)_k$. Define[6] $\mathfrak{H} = \sum_{\ell=1}^{L} \boldsymbol{S}\boldsymbol{L}_\ell \otimes \boldsymbol{e}_\ell$ and assume that the spectral radius of the matrix $\mathfrak{R} = \mathfrak{H}^H\mathfrak{H}$ is upper bounded. Then, as $N, K \to \infty$ with $\frac{K}{N} \to \beta$ and L fixed, the diagonal elements of the matrix $\mathfrak{R}^m$ corresponding to the virtual user k, with given fading amplitude $\boldsymbol{l}_k$, converges with probability 1 to the deterministic value*

$$\mathfrak{R}^m(\boldsymbol{l}_k) \stackrel{a.s.}{=} \lim_{K=\beta N \to \infty} (\mathfrak{R}^m)_{kk}$$

with $\mathfrak{R}^m(\boldsymbol{l})$ determined by the following recursion

$$\mathfrak{R}^m(\boldsymbol{l}) = \sum_{s=0}^{m-1} g(\mathfrak{Z}^{m-s-1}, \boldsymbol{l})\mathfrak{R}^s(\boldsymbol{l}) \tag{5.26}$$

$$\mathfrak{Z}^m = \sum_{s=0}^{m-1} \beta \mathrm{E}\{\mathfrak{R}^{m-s-1}(\boldsymbol{l})\boldsymbol{l}\boldsymbol{l}^H\}\mathfrak{Z}^s \tag{5.27}$$

$$g(\mathfrak{Z}^{m-1}, \boldsymbol{l}) = \boldsymbol{l}^H\mathfrak{Z}^s\boldsymbol{l}.$$

The recursion is initialized by $\mathfrak{R}^0(\boldsymbol{l}) = 1$ and $\mathfrak{Z}^0 = \boldsymbol{I}_L$.

The proof is in Appendix D Section D.4 . This theorem yields the following algorithm to compute $\mathfrak{R}^m(\boldsymbol{l})$ and $m_{\mathfrak{R}}^m$, $m \in \mathbb{Z}^+$, the asymptotic eigenvalue moments of the matrix $\mathfrak{R}^m$.

Algorithm 4

Initialization: *Let $\rho_0(\boldsymbol{l}) = 1$ and $\boldsymbol{\mu}_0 = \boldsymbol{I}$.*

ℓ^{th} step:
- *Define $u_{\ell-1}(\boldsymbol{l}) = \boldsymbol{l}^H\boldsymbol{\mu}_{\ell-1}\boldsymbol{l}$.*

[6] Note that $\mathfrak{H}$ models the transfer matrix in the system model (5.1) and its definition coincides with (5.2).

- *Define* $Z_{\ell-1}(l) = \rho_{\ell-1}(l)ll^H$ *and write it as a polynomial in the monomials* $l_1^{r_1}\dots l_L^{r_L}(l_1^*)^{s_1}\dots(l_1^*)^{s_L}$.
- *Define* $m_l^{(r_1,\dots,r_L,s_1,\dots,s_L)} = \mathrm{E}\{\prod_{\ell=1}^L l_\ell^{r_\ell}(l_\ell^*)^{s_\ell}\}$ *and replace all monomials* $\prod_{\ell=1}^L l_\ell^{r_\ell}(l_\ell^*)^{s_\ell}$ *in* $Z_{\ell-1}(l)$ *by the corresponding* $m_l^{(r_1,\dots,r_L,s_1,\dots,s_L)}$. *Assign the result to* $\mathbf{V}_{\ell-1}$.
- *Calculate*

$$\rho_\ell(l) = \sum_{s=0}^{\ell-1} u_{\ell-s-1}(l)\rho_s(l)$$

$$\boldsymbol{\mu}_\ell = \sum_{s=0}^{\ell-1} \beta \mathbf{V}_{\ell-s-1}\boldsymbol{\mu}_s.$$

- *Assign* $\rho_\ell(l)$ *to* $\mathfrak{R}^\ell(l)$.
- *Write* $\rho_l(l)$ *as a polynomial in* $\prod_{\ell=1}^L l_\ell^{r_\ell}(l_\ell^*)^{s_\ell}$ *and replace all monomials* $\prod_{\ell=1}^L l_\ell^{r_\ell}(l_\ell^*)^{s_\ell}$ *in* $\rho_\ell(l)$ *by the correspondent moments* $m_l^{(r_1,r_2,\dots,r_L,s_1,s_2,\cdots,s_L)}$. *Assign the result to* $m_{\mathfrak{R}}^\ell$.

If the received amplitudes l_k are independent, the previous algorithm simplifies since the matrix $\mathfrak{Z}^s$, $s \in \mathbb{Z}^+$ is a diagonal matrix. If the received amplitudes l_k are asymptotically independent and identically distributed as in the micro-diversity scenario analyzed in [9], the limiting diagonal elements of the matrix $\mathfrak{R}^\ell$ and the eigenvalue moments $m_{\mathfrak{R}}^\ell$ can be derived from Algorithm 1 for synchronous single receiving antenna systems by replacing (i) β with $\beta' = \frac{K}{LN}$ and (ii) the received power of user k at a single antenna, $|a_{kk}|^2$, by the total received power of user k at all antennas, $l^H l$. This result can be obtained directly from Theorem 1 in [9] as proposed in [106] or, alternatively, from Algorithm 4 (noting that $\mathbf{V}_s$ is proportional to the identity matrix and $\mathfrak{R}^\ell(l)$ is a function of $l^H l$).

As for synchronous CDMA systems with single receiving antennas, the limiting values of $\mathfrak{R}_{ii}^\ell$ enable an asymptotic analysis of any multistage detector in the Krylov subspace $\chi_{M,k}(\mathfrak{H})$ for CDMA systems with correlated spatial diversity.

5.7 Numerical Results

In this section, we assess the performance of the detectors Type J-J and Type J-I proposed in 5.6 and compare them to the performance of (i) the exact polynomial expansion detector in Section 2.2.5, (ii) the multistage Wiener receiver, and (iii) the full rank linear MMSE receiver. The assessment is performed assuming independent and identically distributed Gaussian received amplitudes.

Furthermore, the effects of the correlated spatial diversity on the output SINR of multistage detectors are analyzed by using the theoretical results in Section 5.6.

Figure 5.1 and 5.2 show the assessment of the performance of multistage detectors assuming independent and identically distributed Gaussian received amplitudes. The simulation results presented there were obtained for uncoded transmission using $\frac{\pi}{4}$-offset QPSK modulation, and assuming perfect knowledge of the channel. The receivers are compared in terms of their BER evaluated as a function of the normalized signal-to-noise ratio E_b/N_0 where E_b is the mean energy per bit and N_0 is the one sided noise spectral density. Figure 5.1 shows the BER versus E_b/N_0 for 5-stage detectors, $\beta = 2$ ($K' = 64$, $N = 64$, $N_{\mathrm{T}} = 2$), and $L = 4$. More specifically, the dashed lines show the performance of an exact polynomial expansion detector and an exact multistage Wiener filter with 5-stages. The solid lines plot the BER of a Type J-J and a Type J-I detector, the corresponding approximation of the polynomial expansion detector and the MSWF with asymptotic weights, respectively. The performance degradation due to the large system approximation for the weights is completely negligible. Figure 5.2 shows the performance improvements of a Type J-J detector with large-system weights for an increasing number of stages.

The effects of the correlated spatial diversity were analyzed assuming the received amplitudes to be Gaussian with limiting joint distribution (5.10) and correlation matrix

$$\boldsymbol{C}_l = \begin{bmatrix} 1 & 0.5 & 0.3 \\ 0.5 & 1 & 0.5 \\ 0.3 & 0.5 & 1 \end{bmatrix}.$$

The numerical results presented in the following were obtained using $L = 3$ receiving antennas at the base station and assuming a system load $\beta = \frac{3}{2}$.

In case of correlated received amplitudes, the multiuser efficiency of a linear MMSE detector depends on the direction of the received amplitude vector of the user of interest as discussed in Section 5.3.4. For correlated Gaussian received amplitudes the performance of a linear MMSE detector is maximum or minimum when the received amplitude vector is parallel to some of the eigenvectors of the correlation matrix $\boldsymbol{C}_l$ (see Section 5.3.2). The same property holds also for Type J-I and Type J-J detectors, as verified numerically. Let us denote by $\boldsymbol{l}_{\mathrm{MAX}}$ and $\boldsymbol{l}_{\mathrm{MIN}}$ the eigenvectors corresponding to the maximum and minimum eigenvalues of the matrix $\boldsymbol{C}_l$, respectively. Figure 5.3 shows the asymptotic multiuser efficiency of a Type J-I detector with $M = 4$ when the received amplitude vector $\widetilde{\boldsymbol{l}}$ span the subspace $\{\boldsymbol{l}_{\mathrm{MIN}}, \boldsymbol{l}_{\mathrm{MAX}}\}$, i.e. it is a linear combination $\widetilde{\boldsymbol{l}} = u\boldsymbol{l}_{\mathrm{MAX}} + (1-u)\boldsymbol{l}_{\mathrm{MIN}}$. The solid lines plot the output SINR as a function of u, the coefficient of the linear combination $\widetilde{\boldsymbol{l}}$, for different values of the input SNR. The performance is maximum when the channel gain vector is parallel to $\boldsymbol{l}_{\mathrm{MIN}}$, i.e., $u = 0$ and minimum when the channel gain vector is parallel to $\boldsymbol{l}_{\mathrm{MAX}}$, i.e., $u = 1$. The dashed lines illustrate the asymptotic multiuser efficiency of a Type J-I detector for a multiuser MIMO system with independent received amplitudes for the sake of comparison. For independent received amplitudes, the multiuser efficiency does not depend on the direction of the channel gain vectors and it has an intermediate value between the maximum and the minimum multiuser efficiency obtained in the case of correlated spatial diversity.

In Figure 5.4 the asymptotic output SINR of a polynomial expansion detector or Type J-J detector (dashed lines) and of a MSWF or Type J-I detector (solid lines) is plotted as a function of the input SNR for three different received amplitude vectors $\widetilde{\boldsymbol{l}}$, with $u = 0, 0.5, 1$ and perfect power control, i.e. the sum of received powers for each user k, $\boldsymbol{l}_k^H \boldsymbol{l}_k$ is identical for all users. In the case of a single receive antenna or multiple antennas with independent and identically distributed received amplitudes, the MSWF and the polynomial expansion detectors are equivalent if perfect power control is performed (see Chapter 3). On the contrary, for correlated received amplitudes, even in case of perfect power control, the detector Type J-I outperforms the detector Type J-J with equal number of stages. The difference between the SINRs of the two detectors increases as the input SNR increases and/or u decreases.

5.8 Conclusions

In this chapter we determined the asymptotic performance of the linear MMSE receiver, the single user Bayesian filter receiver, and the single user matched filter receiver in CDMA systems with random spreading and spatial diversity. We considered the general case where the channel gains are correlated and there are line of sight components. Our results include as special cases the results in [9] that were derived under the constraints of independence of the channel gains and uniformly distributed phases. Deriving the results in [9] from the general equations (5.8) and (5.9), we could prove the results for the macro-diversity case, which was only conjectured in [9].

Our Theorem 11 shows that the system is asymptotically described by an $L \times L$ matrix $\boldsymbol{A}$ that characterizes completely the effects of channel correlation and line of sight components. The efficiency of the system in recovering the symbol transmitted by the physical user k strongly depends on the direction of the channel gain vector $\boldsymbol{l}_k$ with respect to the eigenvectors of $\boldsymbol{A}$.

Conditions under which the resource pooling effect occurs have been given for the general case.

The single user Bayesian filter and the single user matched filter in a large CDMA scenario with correlated spatial diversity were shown to be equivalent, in terms of performance, to a linear MMSE detector and a matched filter, respectively, in a

CDMA system with spreading factor L and spreading sequences equal to the channel gains.

The design of low complexity multiuser detectors was extended to CDMA systems with correlated spatial diversity. A general framework for the asymptotic analysis of any multiuser detector in a "natural" Krylov projection subspace was also provided.

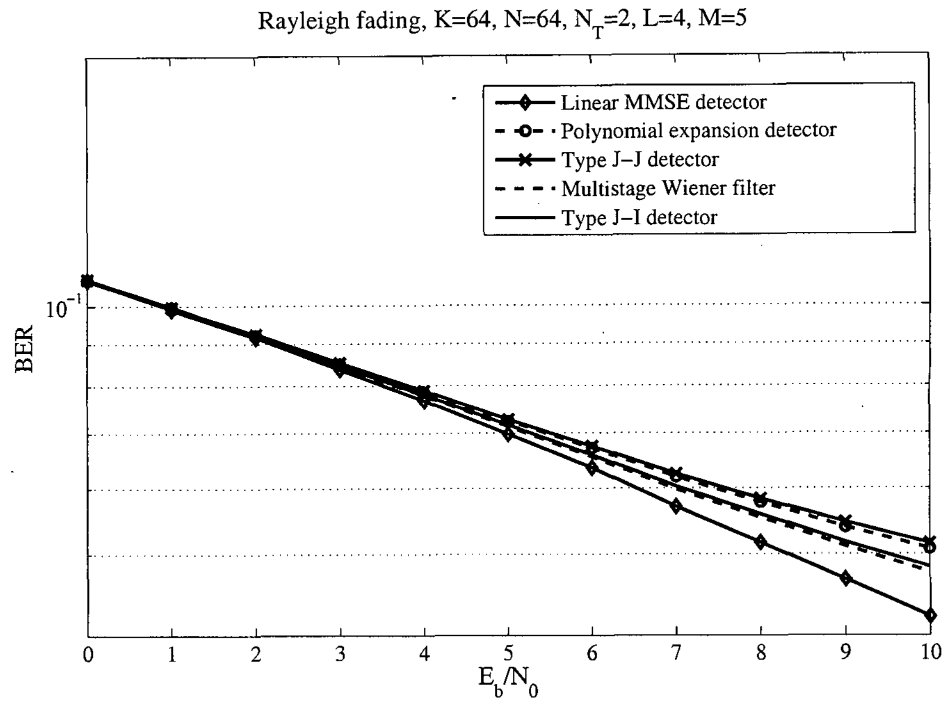


Figure 5.1: BER versus $\frac{E_b}{N_0}$ for $\beta = 2$ and $L = 4$.

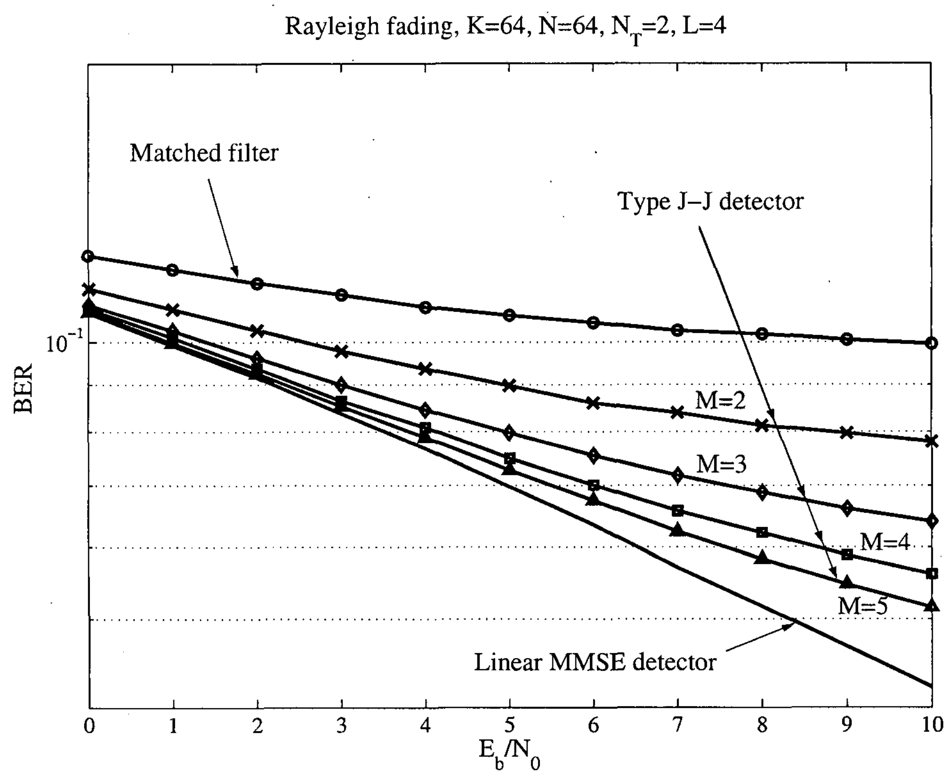


Figure 5.2: BER of a Type J-J detector versus $\frac{E_b}{N_0}$ for $\beta = 2$, $L = 4$, and different numbers of stages $M = 2, \ldots, 5$.

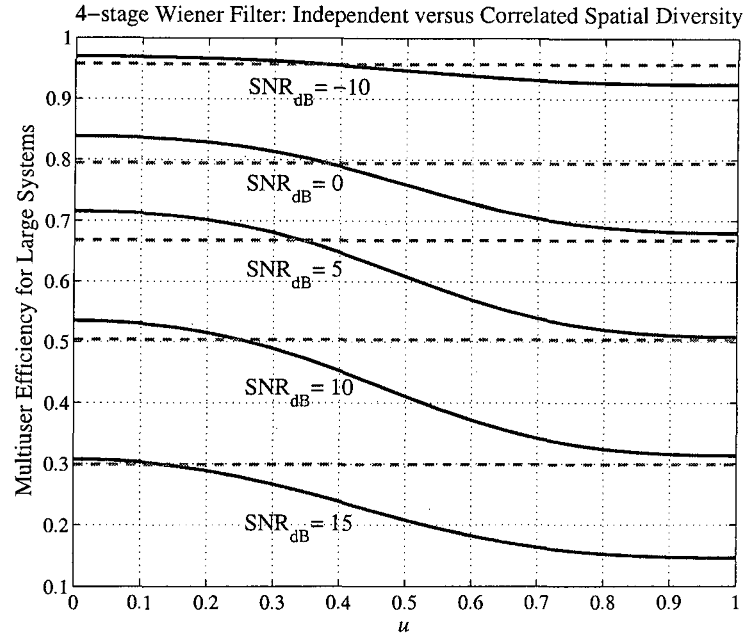


Figure 5.3: Asymptotic multiuser efficiency of a Type J-I detector with $M = 4$, $L = 3$, and $\beta = \frac{3}{2}$ versus the coefficient u of the linear combination $\widetilde{\boldsymbol{l}} = u\boldsymbol{l}_{\mathrm{MAX}}+(1-u)\boldsymbol{l}_{\mathrm{MIN}}$ for different values of the input SNR and correlated received amplitudes (solid lines) or independent and identically distributed received amplitudes (dashed lines).

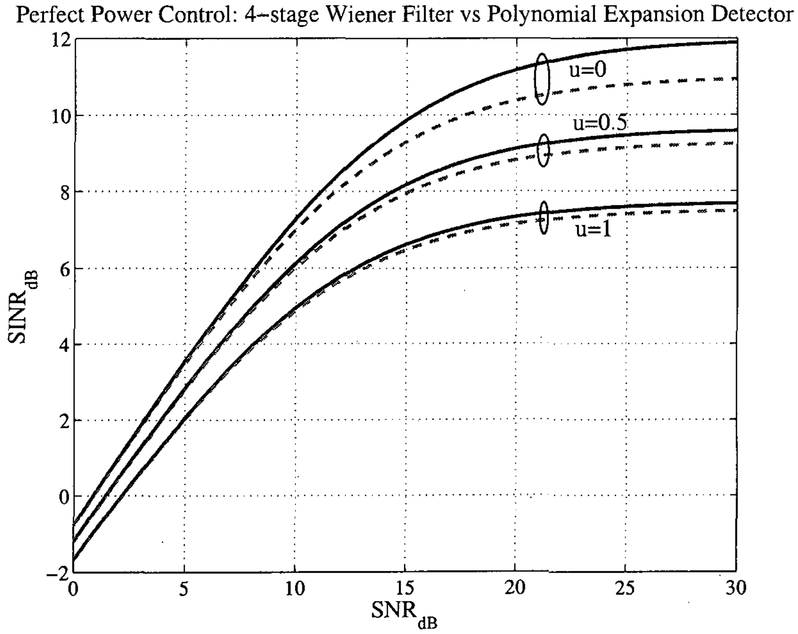


Figure 5.4: Asymptotic output SINR in decibels of polynomial expansion detectors/Type J-J detectors (dashed lines) and MSWF/Type J-I detectors (solid lines) with $M = 4$ versus SNR for different coefficients u of the linear combination $\widetilde{\boldsymbol{l}} = u\boldsymbol{l}_{\mathrm{MAX}} + (1-u)\boldsymbol{l}_{\mathrm{MIN}}$ ($u = 0$, 0.5, 1).

6 Conclusions and Perspectives

6.1 Summary and Conclusions

Multiuser detection with its appealing benefits is a viable approach also for large CDMA systems thanks to the low complexity of the multistage detectors proposed in this work. These detectors achieve performance close to the performance of linear MMSE detectors with the same complexity order per bit as the single user matched filter.

A large number of users in a CDMA system makes the use of multiuser detection challenging because of the high required computation power. The multistage detectors Type J-I and Type J-J proposed in this work benefit of the large system size. They take advantage of the self averaging properties of the transfer matrix of the system as its size becomes large, and the possibility to jointly project all users onto the projection subspaces.

Detectors Type J-I and Type J-J have been designed for synchronous CDMA systems with flat fading and frequency selective fading and for synchronous CDMA systems with correlated spatial diversity.

The multistage detectors with universal weights have been efficiently extended to asynchronous scenarios. An implementation with a sliding observation window and observation window length expanding with the number of stages enables to keep the same complexity per bit as for synchronous systems. This is in contrast to the behaviour of classical MSWF, polynomial expansion detectors, and linear MMSE detectors.

The design of low complexity detectors has been based on a property of some Gram random matrices established first in this work: the diagonal elements of integer powers of these Gram random matrices converge to deterministic values when the matrix size grows large. Utilizing this property, we could develop simple algorithms to determine both the diagonal elements and the eigenvalue moments. Additionally, this approach enabled also the limiting spectral analysis of some Gram matrices not available yet in the literature (e.g. the Gram matrix that models asynchronous CDMA systems with random spreading).

A unified framework for the large system performance analysis of a wide class of linear multiuser detectors has been provided benefitting from this new property of random matrices. This framework includes well known detectors as the multistage Wiener filters, the polynomial expansion detectors, and parallel interference cancelling detectors. The analysis disproves the widespread belief of the equivalence of the polynomial expansion detectors and the multistage Wiener filters. The MSWF detectors outperform the former and they are equivalent only in the case of synchronous CDMA systems with perfect power control, i.e. when all users are received with the same power level.

The proposed multistage detectors for asynchronous systems with observation window expanding with the number of stages can outperform linear MMSE detectors with fixed observation window. Thus, for asynchronous systems the multistage detectors are beneficial not only from a complexity point of view but also from a performance perspective.

The effects of chip pulse waveforms have been investigated jointly with the asynchronism. As long as the chip pulse bandwidth is not greater than half the chip rate, synchronous and asynchronous CDMA systems are equivalent. Above that threshold, the output SINRs of multistage detectors and linear MMSE detectors increase with the bandwidth if the system is asynchronous while it remains constant or decreases as the bandwidth decreases if the users are synchronized. First, for chip pulse waveforms of practical interest asynchronism better exploits the available bandwidth. Second, it does not require synchronization procedures and the proposed multistage detectors for asynchronous systems do not imply an increase in complexity compared to the equivalent detectors for synchronous systems. Therefore, asynchronous CDMA systems are the best solution both in terms of performance and in terms of total complexity for large CDMA systems when linear multiuser detection is performed at the receiver. Additionally, asynchronism and multiuser detection enable to compensate to some extent for the loss in spectral efficiency due to the roll-off.

We investigated the large system performance of linear MMSE detectors, of single user Bayesian detectors, and matched filters in scenarios with correlated spatial diversity. This analysis extends the results in [9] for scenarios with multiple receiving antennas and independent channel gains to the practically more relevant scenarios with correlation at the receivers. Our general analysis includes the micro-diversity and macro-diversity cases discussed in [9] and proves rigorously the results for the macro-diversity case only conjectured in [9].

The large system performance analysis of such systems if fully characterized by a square matrix with size equal to the number of receiving antennas. The multiuser efficiency is not identical for all users and strongly depends on the direction of the channel gain vector.

The correlation of the channels at the transmitting sites does not affect the system performance; only the correlation at the receiver plays a major role.

We generalized the conditions under which the resource pooling effect arises. Additionally, we provided some equivalence results for systems of practical interest. For any scenario with correlated Rayleigh fading, there exists a macro-diversity scenario with independent Rayleigh fading which obtains the same SINR in the case of linear MMSE detection. A CDMA system with correlated spatial diversity and single user Bayesian receiver is equivalent to a CDMA system with linear MMSE detection at the receiver, spreading factor equal to the number of receiving antennas, and spreading sequences equal to the channel gain vectors. Similar equivalence results hold also for the single user matched filter.

In conclusions, the design of low complexity multistage detectors based on universal weights resulted fruitful in three different ways. From a signal processing point of view, we designed detectors that are an excellent compromise between complexity and performance. From the point of view of communication theory, the system analysis revealed important characteristics of CDMA communication systems. From a mathematical perspective, useful properties of some random Gram matrices have been discovered. These properties have been extended also to wide classes of random matrices including classes whose spectral analysis was not available yet in the literature (e.g. Gram matrices for chip asynchronous and symbol quasi-synchronous systems, see Section 4.4).

6.2 Perspectives

The possible further developments of this work reflect its threefold nature.

- *Design of low complexity detectors for scenarios with multipath fading channels and intersymbol interference.* In this kind of scenario, the use of an observation window of a single symbol interval is significantly suboptimal because the multiple access interference is correlated from symbol to symbol as in the case of asynchronous systems investigated in Chapter 4. The multistage detectors with an observation window length expanding with the number of stages, which we introduced for asynchronous systems, can be efficiently adapted to this scenario. In this kind of scenarios characterized by a band transfer matrix, the approach with universal weights provides the greatest advantages.

- *Gram matrices obtained from isometric random matrices: convergence of diagonal elements.* The asymptotic convergence of the diagonal elements of powers of Gram matrices $\boldsymbol{G} = \boldsymbol{H}^H \boldsymbol{H}$ has been shown for special classes of random matrices. A common characteristic of these classes was to be built around random vectors, that are statistically independent with i.i.d. entries. It is of theoretical and practical interest to investigate if this property extends to Gram matrices built around random orthogonal vectors.

- *Convergence rate of the diagonal elements of Gram matrices* $\boldsymbol{G}^s = (\boldsymbol{H}^H\boldsymbol{H})^s$, $s \in \mathbb{Z}^+$. The analysis of the convergence rate of $(\boldsymbol{G}^s)_{kk}$ was beyond of the scope of this work. Its investigation is of theoretical and practical interest.

- *Analysis of asynchronous CDMA systems with orthogonal spreading.* The use of orthogonal spreading has several advantages. Furthermore, other orthogonal access schemes (e.g. MC-CDMA) can be regarded as special cases of orthogonal CDMA systems. Because these schemes are attractive for the fourth generation of wireless systems, they are being studied intensively at the moment. However, the orthogonality is easily destroyed in the uplink due to non-idealities of the channel like asynchronism or frequency selective fading. The impact of asynchronism on these access schemes is not clear. An investigation of the level of asynchronism that can be tolerated without loosing the benefits of orthogonal spreading is of great interest to evaluate the cost of the system.

- *Chip-pulse waveform optimization.* With conventional detection, the roll-off of the chip-pulse waveforms has a detrimental effect on the spectral efficiency. In this work we have shown that multiuser detection can compensate this loss to some extent if the system is asynchronous. Additionally, we provided a fundamental tool, i.e. Theorem 10, for the analysis of the system performance with any kind of chip-pulse waveform of practical interest. This opens the way to an optimization of the chip pulse waveform with linear multiuser detectors.

- *Low complexity power control and admission control algorithms for CDMA systems with correlated spatial diversity.* In [9], the performance analysis of large systems with statistically independent spatial diversity turned out to be a useful tool for the development of low complexity algorithms for power control and call admission control. Similarly, the results presented in Chapter 5 could support low complexity algorithms for power control and call admission control in more realistic scenarios that take into account spatial correlation of the channels.

A Proofs of Chapter 2

In order to prove Lemma 2 in Section 2.4, let us notice that the Lyapunov inequality (Appendix E Lemma 12) and the bound $\lim_{N\to\infty} N^3 \mathrm{E}\{|x_j|^6\} < \infty$ imply $\lim_{N\to\infty} N^2 \mathrm{E}\{|x_j|^4\} < \infty$.

By applying Lemma 1 for $p = 3$, we obtain

$$\begin{aligned}
\mathrm{E}\left\{\left|\boldsymbol{x}_N^H \boldsymbol{C}\boldsymbol{x}_N - \frac{\mathrm{tr}\boldsymbol{C}_N}{N}\right|^3\right\} &\leq K_3\left[\left(\mathrm{E}\{|x_j|^4\}\mathrm{tr}(\boldsymbol{C}\boldsymbol{C}^H)\right)^{\frac{3}{2}} + \mathrm{E}\{|x_j|^6\}\mathrm{tr}(\boldsymbol{C}\boldsymbol{C}^H)^{\frac{3}{2}}\right] \\
&\leq K_3\left[\left(N^2\mathrm{E}\{|x_j|^4\}\frac{\mathrm{tr}(\boldsymbol{C}\boldsymbol{C}^H)}{N}\right)^{\frac{3}{2}} N^{-\frac{3}{2}}\right. \\
&\left. + \left(N^3\mathrm{E}\{|x_j|^6\}\frac{\mathrm{tr}(\boldsymbol{C}\boldsymbol{C}^H)^{\frac{3}{2}}}{N}\right) N^{-2}\right].
\end{aligned}$$

Thanks to the bounds on $\mathrm{E}\{|x_j|^4\}$ and $\mathrm{E}\{|x_j|^6\}$ and on the moments of $F_{\boldsymbol{C}}(\lambda)$ the quantities $\left(N^2\mathrm{E}\{|x_j|^4\}\frac{\mathrm{tr}(\boldsymbol{C}\boldsymbol{C}^H)}{N}\right)^{\frac{3}{2}}$ and $\left(N^3\mathrm{E}\{|x_j|^6\}\frac{\mathrm{tr}(\boldsymbol{C}\boldsymbol{C}^H)^{\frac{3}{2}}}{N}\right)$ are upper bounded. This observation yields

$$\mathrm{E}\left\{\left|\boldsymbol{x}_N^H \boldsymbol{C}\boldsymbol{x}_N - \frac{\mathrm{tr}\boldsymbol{C}_N}{N}\right|^3\right\} \leq K_3' N^{\frac{3}{2}} \tag{A.1}$$

Fixing $\epsilon > 0$ and using Markov's inequality (Appendix E Lemma 10), it follows

$$\Pr\left\{\left|\boldsymbol{x}^H\boldsymbol{C}\boldsymbol{x} - \frac{\mathrm{tr}(\boldsymbol{C})}{N}\right| \geq \epsilon\right\} < \frac{\mathrm{E}\left\{\left|\boldsymbol{x}^H\boldsymbol{C}\boldsymbol{x} - \frac{\mathrm{tr(C)}}{N}\right|^3\right\}}{\epsilon^3}. \tag{A.2}$$

Therefore, the inequalities (A.1) and (A.2) yield

$$\sum_{N=1}^{\infty}\Pr\left\{\left|\boldsymbol{x}^H\boldsymbol{C}\boldsymbol{x} - \frac{\mathrm{tr}(\boldsymbol{C})}{N}\right| \geq \epsilon\right\} < \sum_{N=1}^{\infty}\frac{K_3}{\epsilon^6 N^{\frac{3}{2}}} < \infty. \tag{A.3}$$

Using the Borel-Cantelli lemma (See Appendix E Lemma 13), we conclude that with probability one only finitely many of the events $\left\{\left|\boldsymbol{x}^H\boldsymbol{C}\boldsymbol{x} - \frac{\mathrm{tr}\boldsymbol{C}}{N}\right| \leq \epsilon\right\}$ occur, i.e., as $N \to \infty$,

$$\boldsymbol{x}^H\boldsymbol{C}\boldsymbol{x} - \frac{\mathrm{tr}(\boldsymbol{C})}{N} \xrightarrow{a.s.} 0.$$

This concludes the proof of Lemma 2.

B Proofs of Chapter 3

B.1 Proof of Theorem 6

Let us consider any realization of the random matrix $\boldsymbol{T}^n_{\sim k}$ of size $N \times N$. Thanks to the almost sure convergence of the empirical eigenvalue distribution [107], $\forall \varepsilon$ and $\delta > 0$ there exists an N' such that $\forall N > N'$

$$\Pr\left\{\left|\frac{1}{N}\mathrm{tr}\boldsymbol{T}^n_{\sim k} - m^n_{\boldsymbol{T}}\right| < \delta\right\} > 1 - \epsilon \tag{B.1}$$

where $m^n_{\boldsymbol{T}}$ denotes the limiting eigenvalue moment of order n of the matrix $\boldsymbol{T}$. Since the support of the limiting eigenvalue distribution $F_{|\boldsymbol{A}|^2}$ is upper bounded, all eigenvalue moments $m^n_{|\boldsymbol{A}|^2}$, $n \in \mathbb{Z}^+$ are finite. Then, the same property holds for $m^n_{\boldsymbol{T}}$ (see [77]).

By appealing to Lemma 1 we obtain the inequality

$$\mathrm{E}\left\{\left|\boldsymbol{s}_k^H \boldsymbol{T}^n_{\sim k}\boldsymbol{s}_k - \frac{\mathrm{tr}\boldsymbol{T}^n_{\sim k}}{N}\right|^3\right\} \leq CN\left[(\mathrm{E}\{|s_{11}|^4\})^{\frac{3}{2}} + \mathrm{E}\{|s_{11}|^6\}\right] \tag{B.2}$$

$$\leq \frac{C'}{N^2} \tag{B.3}$$

where C and C' are constants depending on $\max((m^{2n}_{\boldsymbol{T}})^{\frac{3}{2}}, m^{6n}_{\boldsymbol{T}})$ but not on N, and $\boldsymbol{s}_k$ is the k^{th} column of $\boldsymbol{S}$. We use the Lyapunov inequality (see Lemma 12) to bound $\mathrm{E}\{|s_{11}|^4\}$.

The almost sure convergence as $N \to \infty$

$$\boldsymbol{s}_k^H \boldsymbol{T}^n_{\sim k}\boldsymbol{s}_k \xrightarrow{a.s.} m^n_{\boldsymbol{T}} \tag{B.4}$$

follows along the same lines as the proof of Lemma 2.

The strong law of large numbers (see e.g. [108]) yields the almost sure convergence $\boldsymbol{s}_k^H \boldsymbol{s} \xrightarrow{a.s.} 1$ as $N \to \infty$. Then,

$$\boldsymbol{R}_{kk} = |a_{kk}|^2 \boldsymbol{s}_k^H \boldsymbol{s}_k \xrightarrow{a.s.} |a_{kk}|^2. \tag{B.5}$$

For $\ell \geq 2$

$$
\begin{aligned}
(\boldsymbol{R}^{\ell})_{kk} &= |a_{kk}|^2 \boldsymbol{s}_k^H \boldsymbol{T}^{\ell-1} \boldsymbol{s}_k \\
&= |a_{kk}|^4 \boldsymbol{s}_k^H \boldsymbol{s}_k \boldsymbol{s}_k^H \left(\boldsymbol{T}_{\sim k} + |a_{kk}|^2 \boldsymbol{s}_k \boldsymbol{s}_k^H\right)^{\ell-2} \boldsymbol{s}_k \\
&+ |a_{kk}|^2 \boldsymbol{s}_k^H \left(|a_{kk}|^2 \boldsymbol{s}_k \boldsymbol{s}_k^H + \boldsymbol{T}_{\sim k}\right)^{\ell-2} \boldsymbol{s}_k
\end{aligned}
\tag{B.6}
$$

Expanding the product we can rewrite the first term on the right hand side of (B.6) as

$$
|a_{kk}|^4 \boldsymbol{s}_k^H \boldsymbol{s}_k \boldsymbol{s}_k^H \left(\boldsymbol{T}_{\sim k} + |a_{kk}|^2 \boldsymbol{s}_k \boldsymbol{s}_k^H\right)^{\ell-2} \boldsymbol{s}_k = |a_{kk}|^2 \boldsymbol{s}_k^H \boldsymbol{s}_k \left(\boldsymbol{R}^{\ell-1}\right)_{kk}. \tag{B.7}
$$

The second term on the right hand side of (B.6) can be further decomposed as

$$
\begin{aligned}
|a_{kk}|^2 \boldsymbol{s}_k^H \boldsymbol{T}_{\sim k} \left(\boldsymbol{T}_{\sim k} + |a_{kk}|^2 \boldsymbol{s}_k \boldsymbol{s}_k^H\right)^{\ell-2} \boldsymbol{s}_k &= |a_{kk}|^2 \boldsymbol{s}_k^H \boldsymbol{T}_{\sim k} \boldsymbol{s}_k \left(\boldsymbol{R}^{\ell-2}\right)_{kk} \\
&+ |a_{kk}|^2 \boldsymbol{s}_k^H \boldsymbol{T}_{\sim k}^2 \left(\boldsymbol{T}_{\sim k} + |a_{kk}|^2 \boldsymbol{s}_k \boldsymbol{s}_k^H\right)^{\ell-3} \boldsymbol{s}_k.
\end{aligned}
\tag{B.8}
$$

By further expansion of the term $|a_{kk}|^2 \boldsymbol{s}_k^H \boldsymbol{T}_{\sim k}^2 (\boldsymbol{T}_{\sim k} + |a_{kk}|^2 \boldsymbol{s}_k \boldsymbol{s}_k^H)^{\ell-3} \boldsymbol{s}_k$ we obtain

$$
\begin{aligned}
\left(\boldsymbol{R}^{\ell}\right)_{kk} &= |a_{kk}|^2 \boldsymbol{s}_k^H \boldsymbol{s}_k \left(\boldsymbol{R}^{\ell-1}\right)_{kk} + |a_{kk}|^2 \boldsymbol{s}_k^H \boldsymbol{T}_{\sim k} \boldsymbol{s}_k \left(\boldsymbol{R}^{\ell-2}\right)_{kk} \\
&+ |a_{kk}|^2 \boldsymbol{s}_k^H \boldsymbol{T}_{\sim k}^2 \boldsymbol{s}_k \left(\boldsymbol{R}^{\ell-3}\right)_{kk} + |a_{kk}|^2 \boldsymbol{s}_k^H \boldsymbol{T}_{\sim k}^3 \left(\boldsymbol{T}_{\sim k} + |a_{kk}|^2 \boldsymbol{s}_k \boldsymbol{s}_k^H\right)^{\ell-4} \boldsymbol{s}_k.
\end{aligned}
$$

Iterating the expansion (B.8) we get

$$
\left(\boldsymbol{R}^{\ell}\right)_{kk} = \sum_{s=0}^{\ell-1} |a_{kk}|^2 \boldsymbol{s}_k^H \boldsymbol{T}_{\sim k}^{\ell-1-s} \boldsymbol{s}_k \left(\boldsymbol{R}^{s}\right)_{kk} \tag{B.9}
$$

Therefore, (B.5), (B.4), and the recursion yield

$$
R_{kk,\infty}^{\ell} = \lim_{\substack{K,N\to\infty \\ \frac{K}{N}\to\beta}} (\boldsymbol{R}^{s})_{kk} \overset{a.s.}{\to} \sum_{s=0}^{\ell-1} |a_{kk}|^2 m_{\boldsymbol{T}}^{\ell-1-s} R_{kk,\infty}^{s}. \tag{B.10}
$$

Making use of the relation $m_{\boldsymbol{T}}^{n} = \beta m_{\boldsymbol{R}}^{n}$ we obtain (3.26).

B.2 Asymptotic Diagonal Elements: A Closed Form Expression

Theorem 15 *Let $\boldsymbol{A}$, $\boldsymbol{S}$, and $\boldsymbol{R}$ be as in Theorem 6. Conditioned on a_{kk}, the k^{th} diagonal element of $\boldsymbol{A}$, $(\boldsymbol{R}^{\ell})_{kk}$, converges almost surely, as $N, K \to \infty$ with $\frac{K}{N} \to \beta$,* *to the following deterministic quantity* $R^{\ell}_{kk,\infty}$ *depending on* $|a_{kk}|^2$*:*

$$R^{\ell}_{kk,\infty} = \sum_{\substack{(i_0,i_1,\dots i_{\ell-1}):\\ i_0+\sum_{j=1}^{\ell-1} j i_j=\ell \\ i_0-\sum_{j=1}^{\ell-1} i_j\geq 0}} \left(i_0-\sum_{j=1}^{\ell-1} i_j, i_1, \dots i_{\ell-1}\right)! |a_{kk}|^{2i_0} \prod_{s=1}^{\ell-1} (\beta m^s_{\boldsymbol{R}})^{i_s} \tag{B.11}$$

for any $k, \ell \in \mathbb{Z}^+$*. Here,* $(i_0, i_1, \dots i_{\ell-1})$ *is an* ℓ*-tuple of nonnegative integers and* $(\cdot, \cdot, \dots, \cdot)!$ *denotes the multinomial coefficient.*

Proof: By expanding $(\boldsymbol{R}^{\ell})_{kk} = |a_{kk}|^2 \boldsymbol{s}_k^H \left(\boldsymbol{T}_{\sim k} + |a_{kk}|^2 \boldsymbol{s}_k \boldsymbol{s}_k^H\right)^{\ell-1} \boldsymbol{s}_k$ and using the asymptotic convergence in (B.4) and (B.5) we obtain the asymptotic convergence

$$(\boldsymbol{R}^{\ell})_{kk} \xrightarrow{a.s.} \sum_{\substack{(i_0,i_1,\dots i_{\ell-1}):\\ i_0+\sum_{j=1}^{\ell-1} j i_j=\ell}} \varphi(i_0, i_1, \dots, i_{\ell-1}) |a_{kk}|^{2i_0} \prod_{s=1}^{\ell-1} (\beta m^s_{\boldsymbol{R}})^{i_s}. \tag{B.12}$$

where the coefficients $\varphi(i_0, i_1, \dots, i_{\ell-1})$ are obtained expanding the binomial $\left(\boldsymbol{T}_{\sim k} + |a_{kk}|^2 \boldsymbol{s}_k \boldsymbol{s}_k^H\right)^{\ell-1}$.

Finding a closed-form expression for $R^{\ell}_{kk,\infty}$ is equivalent to the combinatorial problem of determining the coefficients $\varphi(i_0, i_1, \dots, i_{\ell-1})$ since $m^s_{\boldsymbol{R}}$ are given in closed-form in [77].

Let us consider the set $\mathcal{S}$ of all binary strings of length $\ell-1$. We define two elements in $\mathcal{S}$ equivalent if both of them contain the same number of runs of ones with the same length, i.e. both of them contain i_1 runs of length 1, i_2 runs of length 2, i_s runs of length s for $1 \leq s \leq \ell-1$. This equivalence relation induces a partition of $\mathcal{S}$ into classes of equivalence. The subset of the equivalent strings with i_s runs of ones with length s and, by convention, $i_0 = \ell - \sum_{s=1}^{\ell-1} s i_s$ with $i_0 \geq 1$ is denoted by $\mathcal{S}_{i_0,i_1\dots i_{\ell-1}}$. It is straightforward to recognize that the number of terms $c^{i_0} \prod_{s=1}^{\ell-1} \left(\boldsymbol{x}^H \boldsymbol{Y} \boldsymbol{x}\right)^{i_s}$ obtained from the expansion of $c^{i_0} \boldsymbol{x}^H \left(\boldsymbol{Y} + c\boldsymbol{x}\boldsymbol{x}^H\right)^{\ell-1} \boldsymbol{x}$ is equal to the cardinality of $\mathcal{S}_{i_0,i_1\dots i_{\ell-1}}$. The latter equals the number of distinct permutations of a multiset with $i'_0 = i_0 - \sum_{k=1}^{\ell-1} i_k \geq 0$ elements equal to zero, i_s elements equal to s, for $1 \leq s \leq \ell-1$, i.e. the multinomial coefficient $(i'_0, i_1, \dots i_{\ell-1})!$ [109].

B.3 Proof of Theorem 7

For $k = 1 \dots K$ we define:

- The $L(K-1) \times (K-1)$ matrix $\boldsymbol{A}_{\sim k} = \operatorname{diag}(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots \boldsymbol{a}_{k-1}, \boldsymbol{a}_{k+1} \ldots \boldsymbol{a}_K)$;
- The $N \times N$ permutation matrix corresponding to a cyclic down-shift by ℓ positions, $\boldsymbol{\Pi}_\ell$;
- The $N \times L$ spreading block of user k $\boldsymbol{S}_k = (\boldsymbol{s}_{(k-1)L+1}, \boldsymbol{s}_{(k-1)L+2}, \ldots, \boldsymbol{s}_{kL}) = (\boldsymbol{\Pi}_0 \boldsymbol{s}_{(k-1)L+1}, \boldsymbol{\Pi}_1 \boldsymbol{s}_{(k-1)L+1}, \ldots, \boldsymbol{\Pi}_{L+1} \boldsymbol{s}_{(k-1)L+1})$.

For further studies it is useful to define $r_v^{(u)}(\boldsymbol{a}_k) \triangleq \boldsymbol{a}_k^H \boldsymbol{S}_k^H \boldsymbol{T}^{u-1} \boldsymbol{\Pi}_v \boldsymbol{S}_k \boldsymbol{a}_k$ for $u \in \mathbb{Z}^+$ and to notice that $r_0^{(u)}(\boldsymbol{a}_k) = (\boldsymbol{R}^u)_{kk}$. By substituting $\boldsymbol{T} = \boldsymbol{S}_k \boldsymbol{a}_k \boldsymbol{a}_k^H \boldsymbol{S}_k + \boldsymbol{T}_{\sim k}$ and proceeding as in Theorem 6 we obtain

$$r_v^{(u)}(\boldsymbol{a}_k) = \sum_{\ell=0}^{u-2} \boldsymbol{a}_k^H \boldsymbol{S}_k^H \boldsymbol{T}_{\sim k}^\ell \boldsymbol{S}_k \boldsymbol{a}_k r_v^{(u-\ell-1)}(\boldsymbol{a}_k) + \boldsymbol{a}_k^H \boldsymbol{S}_k^H \boldsymbol{T}_{\sim k}^{u-1} \boldsymbol{\Pi}_v \boldsymbol{S}_k \boldsymbol{a}_k \qquad \text{for} \quad u \in \mathbb{Z}^+ \tag{B.13}$$

with the convention $\sum_{\ell=0}^{-1}(\cdot) = 0$.

Let us define

$$\boldsymbol{T}_s^\ell = \lim_{\substack{K,N\to\infty \\ \frac{K}{N}\to\beta}} \begin{bmatrix} \frac{1}{N}\operatorname{tr}(\boldsymbol{T}^\ell \boldsymbol{\Pi}_s) & \frac{1}{N}\operatorname{tr}(\boldsymbol{T}^\ell \boldsymbol{\Pi}_{s+1}) & \ddots & \frac{1}{N}\operatorname{tr}(\boldsymbol{T}^\ell \boldsymbol{\Pi}_{s+L-1}) \\ \frac{1}{N}\operatorname{tr}(\boldsymbol{T}^\ell \boldsymbol{\Pi}_{s-1}) & \frac{1}{N}\operatorname{tr}(\boldsymbol{T}^\ell \boldsymbol{\Pi}_s) & \ddots & \ddots \\ \ddots & \ddots & \ddots & \ddots \\ \frac{1}{N}\operatorname{tr}(\boldsymbol{T}^\ell \boldsymbol{\Pi}_{s-L+1}) & \ddots & \ddots & \frac{1}{N}\operatorname{tr}(\boldsymbol{T}^\ell \boldsymbol{\Pi}_s) \end{bmatrix}. \tag{B.14}$$

Using the same arguments as in Theorem 6 it is straightforward to show the following limits for $N \to \infty$

$$\begin{aligned} \boldsymbol{s}_k^H \boldsymbol{\Pi}_s \boldsymbol{s}_k &\xrightarrow{a.s.} \delta_{0,s} \\ \boldsymbol{S}_k^H \boldsymbol{\Pi}_s \boldsymbol{S}_k &\xrightarrow{a.s.} \boldsymbol{I}_L(s) \\ \boldsymbol{S}_k^H \boldsymbol{T}_{\sim k}^\ell \boldsymbol{\Pi}_s \boldsymbol{S}_k &\xrightarrow{a.s.} \boldsymbol{T}_s^\ell \end{aligned} \tag{B.15}$$

where $\boldsymbol{I}_L(s)$ denotes an $L \times L$ matrix with $i_{u,u+s} = 1$, for $u = 1, \ldots, L-s$, and zero elsewhere.

By using the limits (B.15) in (B.13) it follows that $r_v^{(u)}(\boldsymbol{a}_k)$ converges almost surely to the deterministic limit $\rho_v^{(u)}(\boldsymbol{a}_k)$ where $\rho_v^{(u)}(\boldsymbol{a})$ satisfies the recursive expression

$$\rho_v^{(u)}(\boldsymbol{a}) = \sum_{\ell=0}^{u-2} \boldsymbol{a}^H \boldsymbol{T}_0^\ell \boldsymbol{a} \rho_v^{(u-\ell-1)}(\boldsymbol{a}) + \boldsymbol{a}^H \boldsymbol{T}_v^{u-1} \boldsymbol{a}. \tag{B.16}$$

In order to determine $\boldsymbol{T}_s^0$ note that

$$\frac{1}{N}\operatorname{tr}\boldsymbol{\Pi}_s = \begin{cases} 1 & \text{for } s = 0, \\ 0 & \text{otherwise.} \end{cases} \tag{B.17}$$

Then, by using (B.17) in the definition of $\mathcal{T}_s^\ell$ for $\ell = 0$ we obtain $\mathcal{T}_s^0 = \boldsymbol{I}_L(s)$. Let us focus now on computing $\mathcal{T}_s^\ell$ for $\ell \neq 0$. The property of the trace[1] and the definitions of $\boldsymbol{T}$ and $\boldsymbol{R}$ yield

$$\frac{1}{N}\mathrm{tr}\boldsymbol{T}^\ell\boldsymbol{\Pi}_s = \frac{1}{N}\mathrm{tr}\left(\boldsymbol{A}^H\boldsymbol{S}^H\boldsymbol{T}^{\ell-1}\boldsymbol{\Pi}_s\boldsymbol{S}\boldsymbol{A}\right) \tag{B.18}$$

$$= \frac{1}{N}\sum_{k=1}^{K} r_s^{(\ell)}(\boldsymbol{a}_k). \tag{B.19}$$

Thus, $\lim_{\substack{K,N\to\infty \\ \frac{K}{N}\to\beta}} \frac{1}{N}\mathrm{tr}\boldsymbol{T}^\ell\boldsymbol{\Pi}_s = \beta\mathrm{E}\{\rho_s^\ell\}$ and the matrix $\mathcal{T}_s^\ell$ is given by

$$\mathcal{T}_s^\ell = \beta\begin{bmatrix} \overline{\rho}_s^{(\ell)} & \overline{\rho}_{s+1}^{(\ell)} & \ddots & \overline{\rho}_{s+L-1}^{(\ell)} \\ \overline{\rho}_{s-1}^{(\ell)} & \overline{\rho}_s^{(\ell)} & \ddots & \overline{\rho}_{s+L-2}^{(\ell)} \\ \ddots & \ddots & \ddots & \ddots \\ \overline{\rho}_{s-L+1}^{(\ell)} & \ddots & \ddots & \overline{\rho}_s^{(\ell)} \end{bmatrix} \tag{B.20}$$

where $\overline{\rho}_j^{(\ell)} = \mathrm{E}\{\rho_j^\ell\}$. Note that $\overline{\rho}_{-v}^{(\ell)} = (\overline{\rho}_v^{(\ell)})^*$.

In order to prove Equivalence 1 let us observe that, if

(A) $\mathcal{T}_0^u$ is proportional to the identity matrix with $\mathcal{T}_0^u = \beta\overline{\rho}_0^{(u)}\boldsymbol{I}_L$, $\forall\, u \in \mathbb{Z}^+$ and $\overline{\rho}_0^{(0)} = \frac{1}{\beta}$;

(B) $\mathcal{T}_v^u = \boldsymbol{0}$, for all $v \neq 0$ and $u \in \mathbb{Z}^+$;

then Equivalence 1 holds. In fact, by using assumptions (A) and (B), the recursive expression (B.16) reduces to

$$\rho_0^{(u)}(\boldsymbol{a}) = \sum_{\ell=0}^{u-1} \boldsymbol{a}^H\boldsymbol{a}\beta\overline{\rho}_0^{(\ell)}\rho_0^{(u-\ell-1)}(\boldsymbol{a}). \tag{B.21}$$

Since $\rho_0^{(u)}$ depends on $\boldsymbol{a}$ only through the scalar $\mathcal{P} = \boldsymbol{a}^H\boldsymbol{a}$ the previous recursion is rewritten as

$$\widetilde{\rho}_0^{(u)}(\mathcal{P}) = \sum_{\ell=0}^{u-1} \mathcal{P}\beta\overline{\rho}_0^{(\ell)}\widetilde{\rho}_0^{(u-\ell-1)}(\mathcal{P}). \tag{B.22}$$

[1] We refer here to the well know property of the trace $\mathrm{tr}\boldsymbol{A}\boldsymbol{B} = \mathrm{tr}\boldsymbol{B}\boldsymbol{A}$, where $\boldsymbol{A}$ is an $m \times n$ matrix and $\boldsymbol{B}$ is an $n \times m$ matrix.

If Assumptions (A) and (B) are satisfied, then $R^u_{kk,\infty} \xrightarrow{a.s.} \widetilde{\rho}^u_0(\mathcal{P}_k)$, where $\mathcal{P}_k = \boldsymbol{a}_k^H \boldsymbol{a}_k$. By using the following substitutions

$$\begin{aligned} \mathcal{P}_k &\to |a_{kk}|^2 \\ \widetilde{\rho}_0^{(\ell)} &\to m^\ell_{\boldsymbol{R}} \end{aligned} \tag{B.23}$$

it becomes apparent that (B.22) coincides with (3.26), i.e., $R^u_{kk,\infty}$ for the matrices $\boldsymbol{R}$ defined in Theorems 6 and 7 in Section 3.4 are equal. The identity of the moments follows from the identity of $R^u_{kk,\infty}$, $u \in \mathbb{Z}^+$. Therefore, Equivalence 1 holds under Assumptions (A) and (B).

The proof of Equivalence 1 reduces to show that Assumptions (A) and (B) are satisfied if the probability density function $f_{\boldsymbol{a}}(a_1, a_2, \dots, a_L)$ is an even function. We prove first that $\overline{\rho}_v^{(u)} = 0$ for $v \neq 0$ and $u \in \mathbb{Z}^+$. This property is shown by induction. We have

$$\overline{\rho}_v^{(1)} = \begin{cases} \mathrm{E}\{\sum_{i=1}^{L-v} a_i^* a_{i+v}\} & 1 \le v \le L-1 \\ 0 & v \ge L \end{cases}. \tag{B.24}$$

Since the argument of the expectation $\mathrm{E}\{a_i^* a_{i+v}\}$ in (B.24) is an odd function in a_i while the probability density function $f_{\boldsymbol{a}}(a_1, a_2, \dots, a_L)$ is an even function, then the function $a_i^* a_{i+v} f_{\boldsymbol{a}}(a_1, a_2, \dots, a_L)$ is an odd function in a_i and $\overline{\rho}_v^{(1)} = \int a_i^* a_{i+v} f_{\boldsymbol{a}}(a_1, a_2, \dots, a_L) = 0 \;\forall v \ge 1$. Note that $\boldsymbol{T}_0^{(0)}$ is proportional to the identity matrix. For the induction in s, we assume that $\overline{\rho}_v^{(s)} = 0$ for $v \neq 0$ and $s < u$. Then, all diagonal elements of $\boldsymbol{T}_v^{(s)}$, for $v \neq 0$ and and $s < u$, are zero and $\boldsymbol{T}_0^u$ is proportional to the identity matrix. This implies that $\rho_v^{(u)} = \boldsymbol{a}^H \boldsymbol{T}_v^{u-1} \boldsymbol{a}$, for $v \in \mathbb{Z}^+$, is an odd function in all variables that appear in it so that $\overline{\rho}_v^{(u)} = 0 \;\forall v \neq 0$. This completes the induction. Moreover, $\boldsymbol{T}_0^u$ is proportional to the identity matrix. Therefore, Assumptions (A) and (B) are satisfied and Equivalence 1 is proven.

To prove Equivalence 2 we introduce the $N \times LK$ matrix $\overline{\boldsymbol{S}}$ whose elements are i.i.d, zero mean with variance $\mathrm{E}\{|\overline{s}_{11}|^2\} = \frac{1}{N}$ and sixth moment such that $\lim_{N\to\infty} \mathrm{E}\{N^3 |\overline{s}_{ij}|^6\} < \infty$. Let $\overline{\boldsymbol{s}}_k$ denote the k^{th} column of $\overline{\boldsymbol{S}}$. $\overline{\boldsymbol{S}}_k = (\overline{\boldsymbol{s}}_{(k-1)L+1}, \dots, \overline{\boldsymbol{s}}_{kL})$ is the $N \times L$ "spreading" block of user k and $\overline{\boldsymbol{S}}_{\sim k} = (\overline{\boldsymbol{S}}_1, \dots, \overline{\boldsymbol{S}}_{k-1}, \overline{\boldsymbol{S}}_{k+1}, \dots \overline{\boldsymbol{S}}_K)$. $\overline{\boldsymbol{T}}$, $\overline{\boldsymbol{T}}_{\sim k}$, and $\overline{\boldsymbol{R}}$ are defined similarly to $\boldsymbol{T}$, $\boldsymbol{T}_{\sim k}$, and $\boldsymbol{R}$, respectively, substituting $\boldsymbol{S}$ with $\overline{\boldsymbol{S}}$. For $K, N \to \infty$ with $\frac{K}{N} \to \beta$, Lemma 3 yields

$$\overline{\boldsymbol{s}}_\ell^H \overline{\boldsymbol{T}}^u_{\sim k} \overline{\boldsymbol{s}}_m \xrightarrow{a.s.} 0 \tag{B.25}$$

An expansion of $(\overline{\boldsymbol{R}}^u)_{kk}$ along the lines of (B.13) yields

$$(\overline{\boldsymbol{R}}^u)_{kk} = \sum_{\ell=0}^{u-2} \boldsymbol{a}_k^H \overline{\boldsymbol{S}}_k^H \overline{\boldsymbol{T}}^\ell_{\sim k} \overline{\boldsymbol{S}}_k \boldsymbol{a}_k (\overline{\boldsymbol{R}}^{u-\ell-1})_{kk}. \tag{B.26}$$

Since $\overline{\boldsymbol{S}}_k^H \overline{\boldsymbol{T}}_k^{\ell} \overline{\boldsymbol{S}}_k \xrightarrow{a.s.} m_{\overline{\boldsymbol{T}}}^{\ell} \boldsymbol{I}_L(0)$ it is straightforward to recognize that (B.26) and (B.22) coincide asymptotically. Since $(\boldsymbol{R}^u)_{kk}$ and $(\overline{\boldsymbol{R}}^u)_{kk}$, for $u \in \mathbb{Z}^+$ converge to the same limit, then also the eigenvalue moments of $\boldsymbol{R}$ and $\overline{\boldsymbol{R}}$ are equal. This implies that the eigenvalue distribution of $\boldsymbol{R}$ converges to the same eigenvalue distribution of $\overline{\boldsymbol{R}}$ and Equivalence 2 is proven.

When the probability density function $f_{\boldsymbol{a}}(a_1, a_2, \dots, a_L)$ is not even the recursion provided by (B.16) and (B.20) does not simplify to the recursion in (B.21) and a more complex algorithm is required. The following algorithm obtained from the general recursion (B.16) and (B.20) determines $R_{kk,\infty}^l$ and $m_{\mathbf{R}}^l$ for channels where the limiting probability density function is not even.

Algorithm 5

INITIALIZATION: *Let* $\boldsymbol{x} = (x_1, x_2, \dots x_L)^T$, $\boldsymbol{T}_0^s = \boldsymbol{I}_L(s)$, *and* $\rho_v^{(1)}(\boldsymbol{x}) = \sum_{r=1}^{L-v} x_{v+r}^* x_r$, $v = 1, \dots L-1$.

RECURSION:

- *Assume* $\rho_v^{(s)} = 0$ *for* $s = 1, \dots, \ell-1$ *and* $(\ell-1)(L-1) \leq v \leq \ell(L-1)$.
- *Define*

$$\rho_v^{(\ell)}(x_1, x_2, \dots, x_L) = \sum_{s=0}^{\ell-2} \boldsymbol{x}^H \boldsymbol{T}_0^s \boldsymbol{x} \rho_v^{\ell-s-1} + \boldsymbol{x}^H \boldsymbol{T}_v^{\ell-1} \boldsymbol{x} \qquad 0 \leq v \leq \ell(L-1) \tag{B.27}$$

  and write them as polynomials in $x_1, x_2, \dots, x_L$.

- *Replace all monomials* $\prod_{\ell=1}^{L} x_\ell^{i_\ell}$ *by the mixed moments* $m_{\boldsymbol{A}}^{(i_1 \dots i_L)} = \mathrm{E}\{a_1^{i_1} a_2^{i_2} \dots a_L^{i_L}\}$ *in* $\rho_v^{(\ell)}(x_1, x_2, \dots, x_L)$, $v = 1, \dots, \ell(L-1)$ *and assign the result to* $\overline{\rho}_v^{(\ell)}$.
- *Build the matrices*

$$\boldsymbol{T}_v^{\ell} = \beta \begin{bmatrix} \overline{\rho}_v^{\ell} & \overline{\rho}_{v+1}^{\ell} & \ddots & \ddots & \overline{\rho}_{v+L-1}^{\ell} \\ \overline{\rho}_{v-1}^{\ell} & \overline{\rho}_v^{\ell} & \ddots & \ddots & \overline{\rho}_{v+L-2}^{\ell} \\ \ddots & \ddots & \ddots & \ddots & \ddots \\ \overline{\rho}_{v-L+2}^{\ell} & \ddots & \ddots & \ddots & \overline{\rho}_{v+1}^{\ell} \\ \overline{\rho}_{v-L+1}^{\ell} & \overline{\rho}_{v-L+2}^{\ell} & \ddots & \overline{\rho}_{v-1}^{\ell} & \overline{\rho}_v^{\ell} \end{bmatrix} \tag{B.28}$$

  by using the relation $\overline{\rho}_{-v}^{(\ell)} = (\overline{\rho}_v^{(\ell)})^*$.

- *Assign* $\rho_0^{(\ell)}(\mathbf{a}_k)$ *to* $R_{kk,\infty}^{\ell}$ *and* $\overline{\rho}_0^{(\ell)}$ *to* $m_{\boldsymbol{R}}^{\ell}$.

C Proofs of Chapter 4

C.1 Proof of Theorem 8

In this section we use the following definition.

Definition 3 *A $p \times p$ principal submatrix of an $n \times n$ matrix $\boldsymbol{A}$, with $p < n$ is a $p \times p$ submatrix of $\boldsymbol{A}$ obtained by selecting rows and columns of $\boldsymbol{A}$ with the same indices.*

The following lemmas are useful in the proof of Theorem 8.

Lemma 4 *Let $\mathcal{H}_N$ be an $N \times K$ random matrix and $\mathcal{R}_N = \mathcal{H}_N^H \mathcal{H}_N$ and $\mathcal{T}_N = \mathcal{H}_N \mathcal{H}_N^H$. If the spectral radius of $\mathcal{H}$ is upper bounded with probability 1 as $N, K \to \infty$ with $\frac{K}{N} \to \beta$, then also the spectral radius of any principal submatrix extracted from $\mathcal{R}_N^m$ and $\mathcal{T}_N^m$, for any finite m, is upper bounded with probability 1.*

Proof: Let us consider any realization of the random matrix $\mathcal{H}_N$ and any $q \times q$ principal submatrix of $\mathcal{R}_N^m$ or $\mathcal{T}_N^m$ with $q = q(N)$ and $\frac{q}{N} \to \nu > 0$ for $N \to \infty$. Since for $N \to \infty$ the spectral radius of the random matrix $\mathcal{H}_N$, is finite with probability one, for N sufficiently large any realization of $\mathcal{H}_N$, and thus of $\mathcal{R}_N^m$ and $\mathcal{T}_N^m$ (with m finite) has upper bounded spectral radius, except possibly in a set of matrices of probability zero. Applying Theorem 16 for interlacing eigenvalues of bordered matrices (see Section E.1) the spectral radius of any $q \times q$ principal submatrix of $\mathcal{R}_N^m$ or $\mathcal{T}_N^m$ is not greater than the spectral radius of $\mathcal{R}_N^m$ or $\mathcal{T}_N^m$, respectively. Therefore, the spectral radius of any $q \times q$ principal submatrix of $\mathcal{R}_N^m$ or $\mathcal{T}_N^m$ is upper bounded as $N \to \infty$, except possibly in a set of probability zero.

Lemma 5 *Let $\mathcal{X}^N$ be an $N \times N$ semi-definite positive random matrix and let $q = q(N)$ be a positive integer function of the positive integer N, i.e., $q : \mathbb{Z}^+ \to \mathbb{Z}^+$ with $\frac{q(N)}{N} \to c_2 > 0$ as $N \to \infty$. Assume*

(a) Any diagonal element $\chi_{nn}^{(N)}$, with $n = n(N)$ and $\lim_{N\to\infty} \frac{n(N)}{N} = x$, of the matrix $\mathcal{X}^{(N)}$ converges almost surely to the deterministic limiting value χ_x^{∞}.

(b) $\forall n \in \mathbb{Z}^+$ and $\forall \epsilon > 0$, $\Pr\{|\chi_{nn}^{(N)} - \chi_x^{(\infty)}| \geq \epsilon\} \leq o\left(\frac{1}{N^c}\right)$ with $c > 1$.

(c) *The normalized trace of the matrix* $\boldsymbol{\mathcal{X}}^{(N)}$ *converges to a finite value* $\lim_{N\to\infty} \frac{1}{N}\mathrm{tr}\boldsymbol{\mathcal{X}}^{(N)} = c_1 < +\infty$.

(d) $q = q(N)$ *with* $\frac{q(N)}{N} \to c_2 > 0$ *as* $N \to \infty$.

Then, the trace of any $q \times q$ *principal submatrix of* $\boldsymbol{\mathcal{X}}^{(N)}$ *normalized by* q *converges to a finite deterministic limiting value with probability 1.*

Proof: The convergence of the trace any $q \times q$ principal submatrix of $\boldsymbol{\mathcal{X}}^{(N)}$ to a finite deterministic value is a straightforward consequence of assumption (c) and the properties of absolutely convergent series. In fact, the normalized trace of $\boldsymbol{\mathcal{X}}^{(N)}$, for $N \to \infty$ is a series of non negative elements absolutely convergent thanks to the fact that $\boldsymbol{\mathcal{X}}^{(N)}$ is semi-definite positive and to assumption (c). The normalized trace of a principal submatrix of $\boldsymbol{\mathcal{X}}^{(N)}$ is an absolutely convergent series obtained from the normalized trace of $\boldsymbol{\mathcal{X}}^{(N)}$ considering only a subset of terms. The property of absolutely convergent series guarantees the convergence of the normalized trace of any principal submatrix of $\boldsymbol{\mathcal{X}}^{(N)}$.

The following considerations demonstrate the convergence with probability 1. Without loss of generality let us consider the principal submatrix including rows and columns with indices between $n(N)$ and $n'(N) = n(N) + q(N) - 1$. Then, $\forall \epsilon > 0$

$$
\begin{aligned}
\Pr\left\{\left|\sum_{k=n(N)}^{n'(N)} (\chi_{kk}^{(N)} - \chi_x^{\infty})\right| < q(N)\epsilon\right\} &\geq \Pr\left\{\sum_{k=n(N)}^{n'(N)} |\chi_{kk}^{(N)} - \chi_x^{\infty}| < q(N)\epsilon\right\} \\
&\geq \Pr\left\{\bigcap_{k=n(N)}^{n'(N)} \left\{|\chi_{kk}^{(N)} - \chi_x^{\infty}| < \epsilon\right\}\right\} \\
&= 1 - \Pr\left\{\bigcup_{k=n(N)}^{n'(N)} \left\{|\chi_{kk}^{(N)} - \chi_x^{\infty}| > \epsilon\right\}\right\} \\
&\geq 1 - \sum_{k=n(N)}^{n'(N)} \Pr\left\{|\chi_{kk}^{(N)} - \chi_x^{\infty}| > \epsilon\right\} \\
&\geq 1 - q(n) \max_{k\in[n(N),n'(N)]} \Pr\{|\chi_{kk}^{(N)} - \chi_x^{\infty}| \geq \epsilon\}
\end{aligned} \tag{C.1}
$$

Bound (C.1) and assumption (b) yield the convergence of the normalized trace of any principal submatrix of $\boldsymbol{\mathcal{X}}^{(N)}$ with probability 1.

Lemma 6 *Let the assumptions of Theorem 8 or Theorem 9 be satisfied. Then,*

$$\sup_N \max_{\substack{i=1,\ldots,WN \\ j=1,\ldots,(W+1)N}} \left[\sum_{j=1}^{(W+1)K} \mathrm{E}\{|h_{ij}^{(N)}|^2\} + \sum_{i=1}^{WN} \mathrm{E}\{|h_{ij}^{(N)}|^2\} \right] < +\infty \tag{C.2}$$

and the Lindeberg condition is satisfied, i.e. for every $\tau > 0$

$$\lim_{N\to\infty} \max_{\substack{i=1,\ldots,WN \\ j=1,\ldots,(W+1)N}} \left[\sum_{j=1}^{(W+1)K} \mathrm{E}\{|h_{ij}^{(N)}|^2 1_{\{|h_{ij}^{(N)}|>\tau\}}(h_{ij}^{(N)})\} + \sum_{i=1}^{WN} \mathrm{E}\{|h_{ij}^{(N)}|^2 1_{\{|h_{ij}^{(N)}|>\tau\}}(h_{ij}^{(N)})\} \right] = 0 \tag{C.3}$$

where $1_{\mathcal{A}}(\cdot)$ *is the indicator function on the set* $\mathcal{A}$ *(see Glossary).*

Proof: We consider first the assumptions of Theorem 9. The inequality (C.2) follows from assumptions (a) and (c) of Theorem 9. In fact,

$$\begin{aligned}
&\sup_N \max_{\substack{i=1,\ldots,WN \\ j=1,\ldots,(W+1)N}} \left[\sum_{j=1}^{(W+1)K} \mathrm{E}\{|h_{ij}^{(N)}|^2\} + \sum_{i=1}^{WN} \mathrm{E}\{|h_{ij}^{(N)}|^2\} \right] \\
&< \sup_N \max_{\substack{i=1,\ldots,WN \\ j=1,\ldots,(W+1)N}} \left[\sum_{j=1}^{(W+1)K} \lambda_j \mathrm{E}\{|s_{ij}^{(N)}|^2\} + \sum_{i=1}^{WN} \lambda_j \mathrm{E}\{|s_{ij}^{(N)}|^2\} \right] \\
&< \sup_N \left[\frac{(W+1)K}{N} + W \right] < +\infty.
\end{aligned} \tag{C.4}$$

In order to verify Lindeberg condition (C.3), we first show that

$$\lim_{N\to\infty} \max_{i=1,\ldots,WN} \sum_{j=1}^{(W+1)K} \mathrm{E}\{|h_{ij}^{(N)}|^2 1_{\{|h_{ij}^{(N)}|>\tau\}}(h_{ij}^{(N)})\} = 0.$$

Let us observe that $\forall i, j$

$$\begin{aligned}
\mathrm{E}\{|h_{ij}^{(N)}|^2 1_{\{|h_{ij}^{(N)}|>\tau\}}(h_{ij}^{(N)})\} &= |a_{jj}|^2 \int_{\{|s_{ij}^{(N)}|^2 > \frac{\tau^2}{|a_{kk}|^2}\}} |s_{ij}^{(N)}|^2 \mathrm{d}F(s_{ij}^{(N)}) \\
&\le \frac{|a_{jj}|^{2+\delta}}{\tau^\delta} \int_{\{|s_{ij}^{(N)}|\ge 0\}} |s_{ij}^{(N)}|^{2+\delta} \mathrm{d}F(s_{ij}^{(N)})
\end{aligned} \tag{C.5}$$

where $F(s_{ij}^{(N)})$ is the distribution function of $s_{ij}^{(N)}$ and $\delta \in \mathbb{R}^+$. From the conditions of Theorem 8 and Theorem 9 $\mathrm{E}\{|s_{ij}^{(N)}|^6\} \le \frac{c}{N^3}$ with c finite constant. Then, for $\delta = 4$ $\mathrm{E}\{|h_{ij}^{(N)}|^2 1_{\{|h_{ij}^{(N)}|>\tau\}}(h_{ij}^{(N)})\} \le \frac{|a_{jj}|^6 c}{\tau^4 N^3}$. Let $m = \max_{j=1\ldots K} |a_{jj}|^6$. Since $F_{|\boldsymbol{A}|^2}(\lambda)$ has upper bounded support m is also bounded and $\max_{ij} \mathrm{E}\{|h_{ij}^{(N)}|^2 1_{\{|h_{ij}^{(N)}|>\tau\}}(h_{ij}^{(N)})\} \leq \frac{mc}{\tau^4 N^3}$. Therefore,

$$\lim_{N\to\infty} \max_{i=1\ldots WN} \sum_{j=1}^{(W+1)K} \mathrm{E}\{|h_{ij}^{(N)}|^2 1_{\{|h_{ij}^{(N)}|>\tau\}}(h_{ij}^{(N)})\} \leq \lim_{N\to\infty} \frac{mc}{\tau^4 N^2} = 0. \tag{C.6}$$

The proof that

$$\lim_{N\to\infty} \max_{i=1\ldots(W+1)K} \sum_{j=1}^{WN} \mathrm{E}\{|h_{ji}^{(N)}|^2 1_{\{|h_{ji}^{(N)}|>\tau\}}(h_{ji}^{(N)})\} = 0 \quad .$$

follows the same lines as the proof of (C.6). Thus, we conclude that Lindeberg condition (C.3) is satisfied.

The hypotheses of Theorem 9 reduce to the hypotheses of Theorem 8 for $\lambda_i \equiv 1$. This concludes the proof of Lemma 6.

Let us consider now a matrix $\boldsymbol{\mathcal{S}}_{N,W}$ or a matrix $\boldsymbol{\mathcal{H}}_{N,W}$ with $\boldsymbol{A} = \boldsymbol{I}$ with empirical distribution $F_P^{(N)}(\rho)$ of the time delays. In Figure C.1 we illustrate the structure of the matrix $\boldsymbol{\mathcal{H}}_{N,W}$. The elements of the matrix $\boldsymbol{\mathcal{H}}_{N,W}$ in the shaded region are i.i.d. with zero mean and variance $\frac{1}{N}$. Outside the shaded region the elements of the matrix $\boldsymbol{\mathcal{H}}_{N,W}$ are zero. Since the time delays are random, the shaded region is also random for finite N and can be described by the empirical distribution of the time delays $F_P^{(N)}(\rho)$. In the following lemma we show that the random shaded region converges to a deterministic region as $N \to \infty$. This deterministic region is illustrated in Figure 4.2. The convergence of the shaded region to a deterministic region is a consequence of the convergence of the empirical distribution $F_P^{(N)}(\rho)$ to a limit distribution function $F_P(\rho)$. The following lemma describes the limit region by the functions $r(x)$, $c(y)$, and $l(y)$ shown in Figure 4.2. The shaded region is described in the following lemma by the function $v(x,y)$ and it is the region where $\mathrm{E}\{|h_{ij}|^2\} = \frac{1}{N}$.

Lemma 7 *Let the definitions of Theorem 8 hold. Furthermore, let the conditions of Theorem 8 be satisfied.*

(a) For each N, let $v_N : [0,W] \times [0,(W+1)\beta] \to \mathbb{R}$ be the variance of the elements of the matrix $\boldsymbol{\mathcal{S}}_{N,W}$ normalized by $\frac{1}{N}$, i.e.

$$v_N(x,y) = N\mathrm{E}\{|h_{ij}|^2\}$$

where $x \in [0,W]$, $y \in [0,(W+1)\beta]$, $i = 1,\ldots,WN$ and $j = 1,\ldots(W+1)N$ satisfy

$$\frac{i}{N} \leq x < \frac{i+1}{N} \qquad and \qquad \frac{j}{N} \leq y < \frac{j+1}{N}.$$

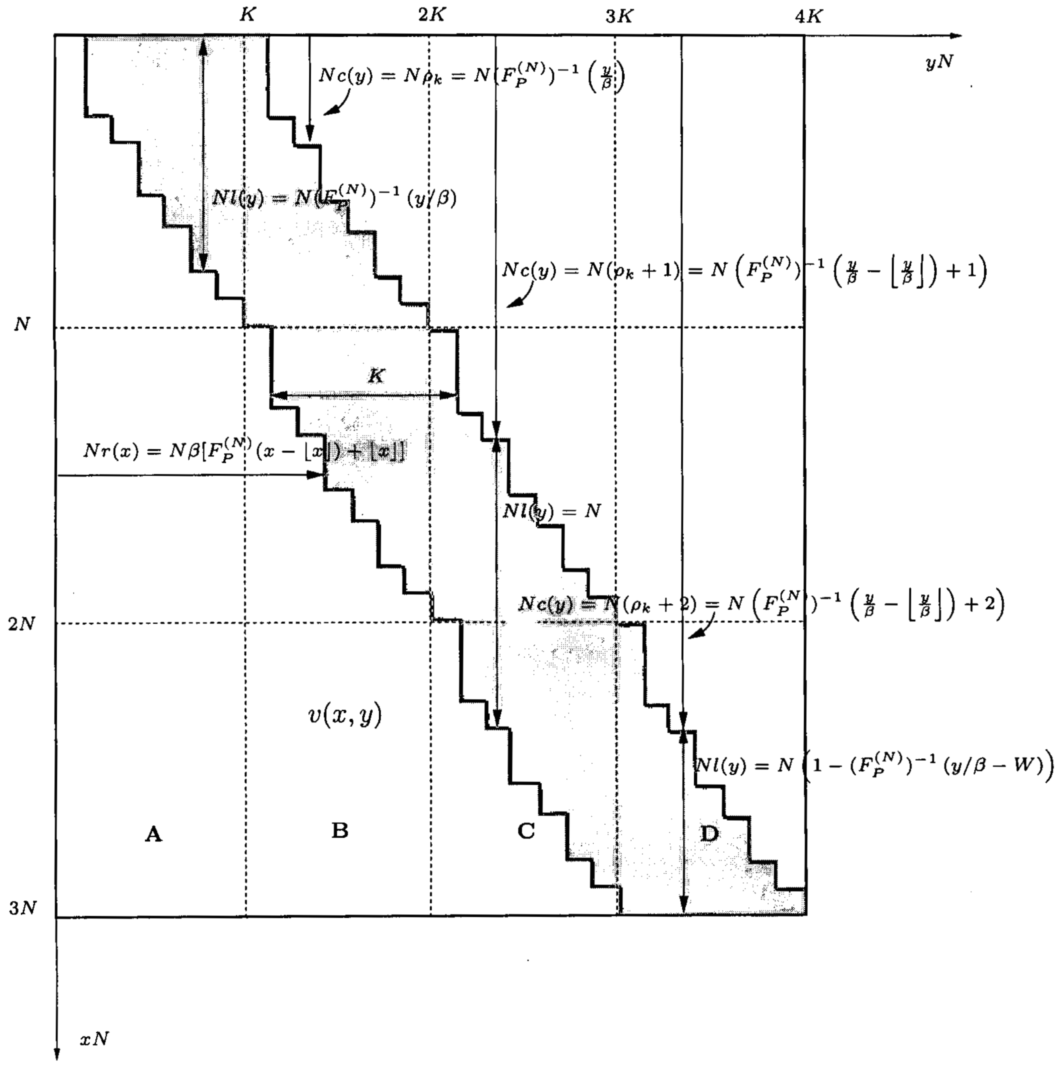


Figure C.1: Graphical representation of the matrix $\mathcal{H}_{N,W}$ or, equivalently, of the function $v(x, y)$. The elements of the matrix $\mathcal{H}_{N,W}$ in the shaded region are i.i.d. with zero mean and variance $\frac{1}{N}$. Outside the shaded region the elements of the matrix $\mathcal{H}_{N,W}$ are zero. The function $v(x, y)$ is equal to 1 in the shaded region and zero elsewhere. The width of the shaded region is constant and equal to K. The height of the shaded region is constant and equal to N in B and C, whereas it varies on A and D.

Then, $v_N(x,y)$ converges to a limited bounded function

$$v(x,y) = \begin{cases} 1, & [0,W] \times [r(x), r(x)+\beta], \\ 0, & \textit{elsewhere} \end{cases}$$

with

$$r(x) = \beta[F_P(x - \lfloor x \rfloor) + \lfloor x \rfloor] \qquad \textit{and} \qquad 0 \le x \le W.$$

Equivalently,

$$v(x,y) = \begin{cases} 1, & [\max(0, c(y)-1), \min(W, c(y))] \times [0, \beta(W+1)], \\ 0, & \textit{elsewhere} \end{cases}$$

with

$$c(y) = \beta\left[\left\lfloor \frac{y}{\beta} \right\rfloor + F_P^{-1}\left(\frac{y}{\beta} - \left\lfloor \frac{y}{\beta} \right\rfloor\right)\right] \qquad \textit{and} \qquad 0 \le y \le \beta(W+1).$$

(b) *Denote by $\boldsymbol{\hbar}_{N,W}(k)$ and $\boldsymbol{\eth}_{N,W}(n)$ the k^{th} column and the n^{th} row of $\boldsymbol{\mathcal{H}}_{N,W}$, respectively. If $\mathcal{L}(\boldsymbol{x})$ is the number of nonzero elements of the vector $\boldsymbol{x}$, then*

$$\lim_{N\to\infty} \frac{\mathcal{L}(\boldsymbol{\eth}_{N,W}(n))}{N} = \beta \tag{C.7}$$

$$\lim_{N\to\infty} \frac{\mathcal{L}(\boldsymbol{\hbar}_{N,W}(n))}{N} = l(y) \tag{C.8}$$

where y is defined by $k = k(N)$ and $\lim_{N\to\infty} \frac{k}{N} = y$ and

$$l(y) = \begin{cases} F_P^{-1}\left(\frac{y}{\beta}\right) & 0 \le y < \beta \\ 1 & \beta \le y < \beta W \\ 1 - F_P^{-1}\left(\frac{y}{\beta} - W\right) & \beta W \le y \le \beta(W+1) \end{cases}$$

Proof: From the definition of the matrix $\boldsymbol{H}^{(N)}(\ell)$ and assumption (a) of Theorem 8 the variance of the element $h_{ij}^{(N)}(\ell)$ of the matrix $\boldsymbol{H}^{(N)}(\ell)$ is equal to $\frac{1}{N}$ for $\rho_j^{(N)}N + 1 \le i \le \rho_j^{(N)}N + 1$ and $1 \le j \le K$ and it is zero elsewhere. Let us consider the matrix $\boldsymbol{\mathcal{H}}_{N,W}$ described in (4.8). The matrices $\boldsymbol{H}^{(N)}(2), \boldsymbol{H}^{(N)}(3), \ldots, \boldsymbol{H}^{(N)}(W)$ are submatrices of $\boldsymbol{\mathcal{H}}_{N,W}$ completely contained in $\boldsymbol{\mathcal{H}}_{N,W}$. Thus the normalized variance $v(x,y)$ can be completely derived from the variance of the elements of $\boldsymbol{H}^{(N)}(\ell)$, $\ell = 2, \ldots, W$ by normalizing the variance by $\frac{1}{N}$, and by appropriate shifting and normalization of the indices describing the interval

$$\rho_j^{(N)}N + 1 \le i \le \rho_j^{(N)}N + 1, \qquad 1 \le j \le K.$$

Then, $v(x,y)=1$ for

$$\rho^{(N)}_{(yN-1)\bmod K+1}+\left\lfloor\frac{Ny}{K}-1\right\rfloor-1\leq x\leq\rho^{(N)}_{(yN-1)\bmod K+1}+\left\lfloor\frac{Ny}{K}-1\right\rfloor$$

and $\beta\leq y\leq\beta W$. Note that the term $\lfloor\frac{Ny}{K}-1\rfloor$ is due to the down shift of the matrix $\boldsymbol{H}(\ell)$ by $\ell-2$ blocks in $\mathcal{H}_{NW}$. Let us focus now on the matrix $\boldsymbol{H}^{(N)}(1)$. Only the block $\boldsymbol{H}^{(N)}_d(1)$ appears in $\mathcal{H}_{N,W}$. Therefore, the variance of the matrix $\mathcal{H}_{N,W}$ for $0\leq j\leq K$ is equal to $\frac{1}{N}$ for $1\leq i\leq\rho^{(N)}_j N+N$. Then, $v(x,y)=1$ for $0\leq x\leq\rho^{(N)}_{\lfloor yN\rfloor}$ and $0\leq y\leq\beta$. Finally, let us consider the matrix $\boldsymbol{H}^{(N)}(W+1)$. Only the block $\boldsymbol{H}^{(N)}_u(W+1)$ of the matrix $\boldsymbol{H}^{(N)}(W+1)$ appears in $\mathcal{H}_{N,W}$. Using the same arguments as for the matrices $\boldsymbol{H}^{(N)}(\ell)$, $\ell=1,\dots,W$ we obtain that $v(x,y)=1$ for

$$\rho^{(N)}_{(yN-1)\bmod K+1}+\left\lfloor\frac{Ny}{K}-1\right\rfloor-1\leq x\leq\frac{K(W+1)}{N}\quad\text{and}\quad\frac{KW}{N}\leq y\leq\frac{K(W+1)}{N}.$$

By using the definition $\aleph(yN,K)=(yN-1)\bmod K+1$ the previous results can be rewritten in the following way. The function $v(x,y)=1$ for

$$\left(\max\left(\frac{1}{N},\rho^{(N)}_{\aleph(yNK)}+\lfloor\frac{Ny}{K}-1\rfloor-1+\frac{1}{N}\right)\leq x\leq\min\left(W,\rho^{(N)}_{\aleph(yNK)}+\lfloor\frac{Ny}{K}-1\rfloor\right)\right)\bigcap\left(\frac{1}{N}\leq y\leq\frac{K(W+1)}{N}\right).\tag{C.9}$$

The maximum and the minimum in (C.9) takes into account the upper and down "truncation" of the matrix $\mathcal{H}_{N,W}$, respectively. $v(x,y)=0$ elsewhere. Therefore, for $N\to\infty$ $v(x,y)=1$ in the region

$$\max\left(0,F_P^{-1}\left(\frac{y}{\beta}-\lfloor\frac{y}{\beta}\rfloor\right)+\left\lfloor\frac{y}{\beta}\right\rfloor-1\right)\leq x\leq\min\left(\beta(W+1),F_P^{-1}\left(\frac{y}{\beta}-\lfloor\frac{y}{\beta}\rfloor\right)+\left\lfloor\frac{y}{\beta}\right\rfloor-1\right)$$

and $0\leq y\leq\beta(W+1)$. By using the definition $c(y)=F_P^{-1}\left(\frac{y}{\beta}-\lfloor\frac{y}{\beta}\rfloor\right)+\lfloor\frac{y}{\beta}\rfloor$, $v(x,y)=1$ in the region

$$[\max(0,c(y)-1)\leq x\leq\min(c(y),(W+1)\beta)]\times[0,\beta(W+1)].\tag{C.10}$$

Note that $v(x,y)$ corresponds to the shaded region in Figure 4.2. The set (C.10) describes the shaded region in Figure 4.2 by defining the subinterval in $[0,W]$ where $v(x,y)=1$ for each $y\in[0,\beta(W+1)]$. We can find an equivalent representation of the same set by defining the subinterval of $y\in[0,(W+1)\beta]$ where $v(x,y)=1$ for each $x\in[0,W]$. This equivalent representation can be easily derived considering Figure C.1. The shaded region is bounded on the left by the function $N\beta[F_P^{(N)}(xN-\lfloor xN\rfloor)+\lfloor xN\rfloor]$ and on the right by the function $N\beta[F_P^{(N)}(xN-\lfloor xN\rfloor)+\lfloor xN\rfloor+1]$. Thus, the region where $v(x,y)=1$ can also be written as

$$\left(\frac{1}{N}\leq x\leq W\right)\bigcap\left(\frac{K}{N}\left(F_P^{(N)}(x+\lfloor x\rfloor)+\lfloor x\rfloor\right)\leq y\leq\frac{K}{N}\left(F_P^{(N)}(x+\lfloor x\rfloor)+\lfloor x\rfloor+1\right)\right).$$

Therefore, for $N \to \infty$, $v(x,y) = 1$ in the region

$$(0 \le x \le W) \bigcap (r(x) \le y \le r(x) + \beta)$$

with $r(x) = \beta\left(F_P(x + \lfloor x \rfloor) + \lfloor x \rfloor\right)$.

Let us consider the row vector $\boldsymbol{\eth}_{N,W}(n)$. Because of the structure of $\boldsymbol{\mathcal{H}}_{N,W}$ there are K nonzero elements in $\boldsymbol{\eth}_{N,W}(n)$ for any index n[1]. This can easily be seen in Figure C.1. The "width" of the shaded region is equal to K in each row. Therefore,

$$\lim_{N\to\infty} \frac{\mathcal{L}(\boldsymbol{\eth}_{N,W}(n))}{N} = \beta$$

Given the vector $\boldsymbol{\hbar}_{N,W}(k)$, with $k = k(N)$, the number of nonzero elements in $\boldsymbol{\hbar}_{N,W}(k)$ is $N\rho_k^{(N)} = NF_P^{(N)\,-1}\left(\frac{k}{K}\right) = NF_P^{(N)\,-1}\left(\frac{k}{\beta N}\right)$ when $1 \le k \le K$. If $K+1 \le k \le WK$, $\boldsymbol{\hbar}_{N,W}(k)$ has N nonzero elements. Finally, if $WK + 1 \le k \le (W+1)K$,

$$\begin{aligned}
\mathcal{L}(\boldsymbol{\hbar}_{N,W}(k)) &= N - N\rho_{k-KW}^{(N)} \\
&= N - NF_P^{(N)\,-1}\left(\frac{k - KW}{K}\right) \\
&= N\left(1 - F_P^{(N)\,-1}\left(\frac{k}{\beta N} - W\right)\right).
\end{aligned}$$

Then,

$$\lim_{N\to\infty} \frac{\mathcal{L}(\boldsymbol{\hbar}_{N,W}(k))}{N} = \begin{cases} F_P^{-1}\left(\frac{y}{\beta}\right) & 0 \le y < \beta \\ 1 & \beta \le y < \beta W \\ 1 - F_P^{-1}\left(\frac{y}{\beta} - W\right) & \beta W \le y \le \beta(W+1) \end{cases}$$

with $y = \lim_{N\to\infty} \frac{k(N)}{N}$.

Proof of Theorem 8: In this proof we adopt the following notation. For $k = 1, \ldots, (W+1)K$ and $n = 1, \ldots, WN$ we define:

- $\boldsymbol{\hbar}_{N,W}(k)$, the k^{th} column of $\boldsymbol{\mathcal{H}}_{N,W}$;
- $\boldsymbol{\eth}_{N,W}(n)$, the n^{th} row of $\boldsymbol{\mathcal{H}}_{N,W}$;
- $\boldsymbol{\mathcal{H}}_{N,W,\sim k}$, the $WN \times (W+1)K - 1$ matrix obtained from $\boldsymbol{\mathcal{H}}_{N,W}$ by suppressing the k^{th} column;

[1] From a physical point of view this reflects the assumption that K users are active at the same time and transmit infinite streams of data.

- $\mathcal{H}_{N,W,\models n}$, the $WN-1 \times (W+1)K$ matrix obtained from $\mathcal{H}_{N,W}$ by suppressing the n^{th} row;
- $\mathcal{R}_{N,W,\models n} = \mathcal{H}^H_{N,W,\models n}\mathcal{H}_{N,W,\models n}$;
- $\mathcal{T}_{N,W,\sim k} = \mathcal{H}_{N,W,\sim k}\mathcal{H}^H_{N,W,\sim k}$.

Theorem 8 is proven by induction. The first step proves that $(\mathcal{T}_{N,W})_{nn} \xrightarrow{a.s.} \beta$ and $(\mathcal{R}_{N,W})_{kk} \xrightarrow{a.s.} l(y)$, as $N \to \infty$, with $y = \frac{k}{N}$ and $l(y)$ defined in (8). In the ℓ^{th} step we assume the almost sure convergence of $(\mathcal{T}^m_{N,W})_{nn}$ and $(\mathcal{R}^m_{N,W})_{kk}$ for $1 \le m \le \ell - 1$ and we prove that $(\mathcal{T}^\ell_{N,W})_{nn}$ and $(\mathcal{R}^\ell_{N,W})_{kk}$ converge almost surely to the deterministic value in (4.11) and (4.12) respectively.

First step: Let us consider $(\mathcal{R}_{N,W})^1_{kk} = \boldsymbol{\hbar}^H_{N,W}(k)\boldsymbol{\hbar}_{N,W}(k)$ and $(\mathcal{T}_{N,W})^1_{nn} = \boldsymbol{\eth}_{N,W}(n)\boldsymbol{\eth}^H_{N,W}(n)$, the diagonal elements of $\mathcal{R}^1_{N,W}$ and $\mathcal{T}^1_{N,W}$. We distinguish the following two cases: (i) The number of nonzero elements in $\boldsymbol{\hbar}^H_{N,W}(k)$ or $\boldsymbol{\eth}_{N,W}(n)$ goes to infinity as $N \to \infty$; (ii) The number of nonzero elements in $\boldsymbol{\hbar}^H_{N,W}(k)$ or $\boldsymbol{\eth}_{N,W}(n)$ keeps finite as $N \to \infty$.

In the first case we can apply the strong law of large numbers (see e.g. [110]) to prove the almost sure convergence of $(\mathcal{R}^1_{N,W})_{kk}$ and $(\mathcal{T}^1_{N,W})_{nn}$. Since the variance of the nonzero i.i.d elements is given by $\mathrm{E}\{|\widehat{h}^{(N)}_{ij}|^2\} = \frac{1}{N}$, the strong law of large numbers guarantees that $(\mathcal{R}^1_{N,W})_{kk}$ and $(\mathcal{T}^1_{N,W})_{nn}$ converge almost surely to the limiting value $\mathcal{R}^1_W(y) = \lim_{N\to\infty} \frac{\mathcal{L}(\boldsymbol{\hbar}_{N,W}(k))}{N}$ and $\mathcal{T}^1_W(x) = \lim_{N\to\infty} \frac{\mathcal{L}(\boldsymbol{\eth}_{N,W}(n))}{N}$, where $\mathcal{L}(\boldsymbol{x})$ is the number of nonzero elements in the vector $\boldsymbol{x}$. From Lemma 7, $\mathcal{T}^1_W(x) = \beta$ for $x \in [0, W]$ and $\mathcal{R}^1_W(y) = l(y)$ for $y \in [0, \beta(W+1)]$.

Let us consider case (ii). From Lemma 7, it is apparent that case (ii) is never verified for any $\boldsymbol{\eth}_{N,W}(n)$ (note that $\lim_{N\to\infty} \frac{\mathcal{L}(\boldsymbol{\eth}_{N,W}(k))}{N} = \beta > 0\ \forall n$). Then, we focus on $\boldsymbol{\hbar}_{N,W}(k)$. Case (ii) corresponds to values of k such that $l(y) = 0$ with $y = \lim_{N\to\infty} \frac{k}{N}$. Let κ be the finite number of elements in $\boldsymbol{\hbar}_{N,W}(k)$ as $N \to \infty$, then the following inequality holds[2]

$$
\begin{aligned}
\mathrm{E}\left\{(\boldsymbol{\hbar}^H_{N,W}(n)\boldsymbol{\hbar}_{N,W}(n))^2\right\} &= \mathrm{E}\left\{\left(\sum_{i=1}^{NW} |h_{ik}|^2\right)^2\right\} \\
&= \mathrm{E}\left\{\sum_{i=1}^{NW} |h_{ik}|^4\right\} + \mathrm{E}\left\{\sum_{\substack{i,j=1 \\ i\neq j}}^{NW} |h_{ik}|^2 |h_{jk}|^2\right\} \\
&= \kappa \mathrm{E}\{|\widehat{h}_{ik}|^4\} + 2\kappa(\kappa-1)(\mathrm{E}\{|\widehat{h}_{ik}|^2\})^2 < \frac{\kappa'}{N^2},
\end{aligned}
\tag{C.11}
$$

[2] We recall here that $h^{(N)}_{ik}$ denotes an element of the matrix $\mathcal{H}$. The element $h^{(N)}_{ik}$ can be zero or $h^{(N)}_{ik} = \widehat{h}^{(N)}_{ik}$, where $\widehat{h}^{(N)}_{ik}$ are random i.i.d. with zero mean and variance $\frac{1}{N}$.

where κ' is a constant independent of N. The last inequality derives from the property of the sixth moment of $\widehat{h}_{ik}$ and the Lyapunov inequality that yield $\mathrm{E}\{|\widehat{h}_{ik}|^4\} < \frac{\kappa''}{N^2}$. Applying the Bienaymé inequality (see Lemma 11 in Appendix E) and the inequality (C.11), for any $\epsilon > 0$

$$\Pr\{|\boldsymbol{\hbar}_{N,W}^H(n)\boldsymbol{\hbar}_{N,W}(n)| > \epsilon\} < \frac{\mathrm{E}\{(\boldsymbol{\hbar}_{N,W}^H(n)\boldsymbol{\hbar}_{N,W}(n))^2\}}{\epsilon^2} < \frac{\kappa'}{\epsilon^2 N^2}.$$

The fist inequality is due to Bienaymé inequality, the second inequality follows from (C.11). Therefore, $\sum_{N=1}^{\infty} \Pr\{|\boldsymbol{\hbar}_{N,W}^H(n)\boldsymbol{\hbar}_{N,W}(n)| > \epsilon\} < \sum_{N=1}^{+\infty} \frac{\kappa'}{\epsilon^2 N^2} < +\infty$. This bound is the condition of the Borel-Cantelli lemma (see Lemma 13 in Appendix E) to prove that $\boldsymbol{\hbar}_{N,W}^H(n)\boldsymbol{\hbar}_{N,W}(n)$ converges almost surely to 0. Therefore, by appealing to the Borel-Cantelli lemma we obtain

$$(\boldsymbol{\mathcal{R}}_{N,W})_{nn} = \boldsymbol{\hbar}_{N,W}^H(n)\boldsymbol{\hbar}_{N,W}(n) \xrightarrow{a.s.} 0.$$

<u>l^{th} step:</u> We assume the almost sure convergence of $(\boldsymbol{\mathcal{T}}_{N,W}^m)_{nn}$ and $(\boldsymbol{\mathcal{R}}_{N,W}^m)_{kk}$ for $1 \leq m \leq \ell - 1$ to the deterministic values $\mathcal{R}_W^m(y)$ and $\mathcal{T}_W^m(x)$, respectively. Let us notice that the conditions of Theorem 5 for the matrix $\boldsymbol{\mathcal{H}}_{N,W}$ are satisfied. In fact, Lemma 6 holds for matrix $\boldsymbol{\mathcal{H}}_{N,W}$. It follows from Lemma 6 that the conditions of Theorem 5 are also satisfied. Theorem 5 guarantees the almost sure convergence of the eigenvalue distribution of $\boldsymbol{\mathcal{T}}_{N,W}$ and $\boldsymbol{\mathcal{R}}_{N,W}$ to a unique deterministic probability density function. This result, along with assumption (c), implies also the almost sure convergence of all eigenvalue moments of $\boldsymbol{\mathcal{R}}_W^m$ and $\boldsymbol{\mathcal{T}}_W^m$, in particular, of the m^{th} moments[3] $m_{\boldsymbol{\mathcal{T}}_{N,W}}^m = \lim_{N\to\infty} \frac{\mathrm{tr}\boldsymbol{\mathcal{T}}_{N,W}^m}{NW}$ and $m_{\boldsymbol{\mathcal{R}}_{N,W}}^m = \lim_{N\to\infty} \frac{\mathrm{tr}\boldsymbol{\mathcal{R}}_{N,W}^m}{K(W+1)}$. Then, the assumptions of Lemma 5 are satisfied. Let $(\boldsymbol{\mathcal{R}}_{N,W}^m)_{k_1:k_2}$ and $(\boldsymbol{\mathcal{T}}_{N,W}^m)_{n_1:n_2}$ be the principal submatrices extracted from $\boldsymbol{\mathcal{R}}_{N,W}^m$ and $\boldsymbol{\mathcal{T}}_{N,W}^m$ and including rows and columns from k_1 to k_2 and from n_1 to n_2, respectively. By appealing to Lemma 5 $\frac{1}{N}\mathrm{tr}\{(\boldsymbol{\mathcal{R}}_{N,W}^m)_{k_1:k_2}\}$ and $\mathrm{tr}\{(\boldsymbol{\mathcal{T}}_{N,W}^m)_{n_1:n_2}\}$ converge almost surely to the deterministic limits

$$\lim_{N\to\infty} \frac{1}{N}\mathrm{tr}\{(\boldsymbol{\mathcal{T}}_{N,W}^m)_{n_1:n_2}\} = \int_{x_1}^{x_2} \mathcal{T}_W^m(x)\mathrm{d}x \qquad \text{with } \lim_{N\to\infty} \frac{n_i}{N} = x_i \tag{C.12}$$

$$\lim_{N\to\infty} \frac{1}{N}\mathrm{tr}\{(\boldsymbol{\mathcal{R}}_{N,W}^m)_{k_1:k_2}\} = \int_{y_1}^{y_2} \mathcal{R}_W^m(y)\mathrm{d}y \qquad \text{with } \lim_{N\to\infty} \frac{k_i}{N} = y_i. \tag{C.13}$$

Additionally, from Lemma 4 in this section the trace of any $q \times q$ principal submatrix of $\boldsymbol{\mathcal{R}}_{N,W}^s$ and $\boldsymbol{\mathcal{T}}_{N,W}^s$ is upper bounded with probability 1. Given the almost sure

[3]Let us consider an $N \times N$ random matrix $\boldsymbol{A}$ with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_N$ and let assume that the sequence of the empirical eigenvalue distribution of $\boldsymbol{A}$ converges almost surely to a limit eigenvalue distribution function $F(\lambda)$. The m^{th} eigenvalue moment of the matrix $\boldsymbol{A}$ is given by $m_{\boldsymbol{A}}^m = \int \lambda^m \mathrm{d}F(\lambda)$. The almost sure convergence guarantees that $m_{\boldsymbol{A}}^m = \lim_{N\to\infty} \frac{1}{N}\mathrm{tr}\boldsymbol{A}^m$.

convergence to a deterministic limit of the trace of any principal submatrix $q \times q$ of $\mathcal{R}^m_{N,W}, \mathcal{T}^m_{N,W}$ for $1 \leq m \leq \ell - 1$ and the finite bound on its spectral radius, we can consider the convergence of $(\mathcal{R}^\ell_{N,W})_{kk}$ and $(\mathcal{T}^\ell_{N,W})_{nn}$ for $\ell \geq 2$. Following the same lines as in the proof of Theorem 6, Appendix B we can expand $(\mathcal{R}^\ell_{N,W})_{nn}$ and $(\mathcal{T}^\ell_{N,W})_{kk}$ as follows:

$$
\begin{aligned}
(\mathcal{R}^\ell_{N,W})_{kk} =& \frac{\mathcal{L}(\boldsymbol{\hbar}_{N,W}(k))}{N} \widetilde{\boldsymbol{\hbar}}^H_{N,W}(k) \left(\mathcal{T}_{N,W,\sim k} + \frac{\mathcal{L}(\boldsymbol{\hbar}_{N,W}(k))}{N} \widetilde{\boldsymbol{\hbar}}_{N,W}(k) \widetilde{\boldsymbol{\hbar}}^H_{N,W}(k) \right)^{\ell-1} \widetilde{\boldsymbol{\hbar}}_{N,W}(k) \\
=& \frac{\mathcal{L}(\boldsymbol{\hbar}_{N,W}(k))}{N} \sum_{s=0}^{\ell-1} \widetilde{\boldsymbol{\hbar}}^H_{N,W}(k) \mathcal{T}^{\ell-s-1}_{N,W,\sim k} \widetilde{\boldsymbol{\hbar}}_{N,W}(k) (\mathcal{R}^s_{N,W})_{kk} \qquad \text{(C.14)} \\
(\mathcal{T}^\ell_{N,W})_{nn} =& \frac{\mathcal{L}(\boldsymbol{\eth}_{N,W}(k))}{N} \widetilde{\boldsymbol{\eth}}_{N,W}(n) \left(\mathcal{R}_{N,W,\models n} + \frac{\mathcal{L}(\boldsymbol{\eth}_{N,W}(k))}{N} \widetilde{\boldsymbol{\eth}}^H_{N,W}(n) \widetilde{\boldsymbol{\eth}}_{N,W}(n) \right)^{\ell-1} \widetilde{\boldsymbol{\eth}}^H_{N,W}(n) \\
=& \frac{\mathcal{L}(\boldsymbol{\eth}_{N,W}(k))}{N} \sum_{s=0}^{\ell-1} \widetilde{\boldsymbol{\eth}}_{N,W}(n) \mathcal{R}^{\ell-s-1}_{N,W,\models n} \widetilde{\boldsymbol{\eth}}^H_{N,W}(n) (\mathcal{T}^s_{N,W})_{nn} \qquad \text{(C.15)}
\end{aligned}
$$

where $\widetilde{\boldsymbol{\hbar}} = \sqrt{\frac{N}{\mathcal{L}(\boldsymbol{\hbar}_{N,W}(k))}} \boldsymbol{\hbar}$ and $\widetilde{\boldsymbol{\eth}} = \sqrt{\frac{N}{\mathcal{L}(\boldsymbol{\eth}_{N,W}(k))}} \boldsymbol{\eth}$. Thanks to the recursive expressions (C.14) and (C.15), the almost sure convergence of $(\mathcal{R}^\ell_{N,W})_{kk}$ and $(\mathcal{T}^\ell_{N,W})_{nn}$ reduces to the almost sure convergence of $\widetilde{\boldsymbol{\hbar}}^H_{N,W}(k) \mathcal{T}^s_{N,W,\sim k} \widetilde{\boldsymbol{\hbar}}_{N,W}(k)$ and $\widetilde{\boldsymbol{\eth}}_{N,W}(k) \mathcal{R}^s_{N,W,\models k} \widetilde{\boldsymbol{\eth}}^H_{N,W}(k)$, respectively, to a deterministic value for $s = 1, \dots \ell - 1$.

Again we distinguish between case (i) and case (ii). We consider first case (i). $\widetilde{\boldsymbol{\hbar}}_{N,W}(k)$ and $\widetilde{\boldsymbol{\eth}}_{N,W}(k)$ have nonzero i.i.d. elements in the interval $[k_1, k_2]$ and $[n_1, n_1 + K]$, with

$$k_1 = \max \left(1, \left\lfloor \frac{k-1}{K} \right\rfloor N + \rho^{(N)}_{(k-1) \mod K+1} N - N + 1 \right) \qquad \text{(C.16)}$$

$$k_2 = \min \left(WN, \left\lfloor \frac{k-1}{K} \right\rfloor N + \rho^{(N)}_{(k-1) \mod K+1} N \right) \qquad \text{(C.17)}$$

$$n_1 = KF^{(N)}_P \left(\frac{n}{N} - \left\lfloor \frac{n}{N} \right\rfloor - \frac{1}{N} \right) + \left\lfloor \frac{n-1}{N} \right\rfloor K. \qquad \text{(C.18)}$$

Since Lemma 4 and Lemma 5 guarantee that for N sufficiently large the empirical eigenvalue distribution of the principal submatrices $(\mathcal{T}^s_{W,N})_{k_1:k_2}$ and $(\mathcal{R}^s_{N,W})_{n_1:n_1+K-1}$ have upper bounded support and their normalized traces converge to a deterministic limiting value with probability one, we can apply Lemma 2 in

Chapter 2 to obtain

$$
\begin{aligned}
g(\boldsymbol{\mathcal{T}}^{s}_{W}, y) &\triangleq \lim_{K=\beta N\to\infty} \widetilde{\boldsymbol{\hbar}}^{H}_{N,W}(k)\boldsymbol{\mathcal{T}}^{s}_{N,W,\sim k}\widetilde{\boldsymbol{\hbar}}_{N,W}(k) \\
&= \lim_{K=\beta N\to\infty} \frac{1}{\mathcal{L}(\boldsymbol{\hbar}(\boldsymbol{k})_{N,W})} \sum_{r=k_1}^{k_2} \boldsymbol{\mathcal{T}}^{s}_{N,W}(r) \\
&\overset{a.s.}{\to} \frac{1}{l(y)} \int_{\max\left(0,\lfloor\frac{y}{\beta}\rfloor+F_P^{-1}\left(y-\lfloor\frac{y}{\beta}\rfloor\beta\right)-1\right)}^{\min\left(W,\lfloor\frac{y}{\beta}\rfloor+F_P^{-1}\left(y-\lfloor\frac{y}{\beta}\rfloor\beta\right)\right)} \mathcal{T}^{s}_{W}(x)dx
\end{aligned}
$$

and

$$
\begin{aligned}
f(\boldsymbol{\mathcal{R}}^{s}_{W}, x) &\triangleq \lim_{K=\beta N\to\infty} \widetilde{\boldsymbol{\eth}}_{N,W}(n)\boldsymbol{\mathcal{R}}^{s}_{N,W,\vDash n}\widetilde{\boldsymbol{\eth}}^{H}_{N,W}(n) \\
&= \lim_{K=\beta N\to\infty} \frac{1}{K} \sum_{r=n_1}^{n_1+K} \boldsymbol{\mathcal{R}}^{s}_{N,W}(r) \\
&\overset{a.s.}{\to} \frac{1}{\beta} \int_{\beta(F_P(x-\lfloor x\rfloor)+\lfloor x\rfloor)}^{\beta(F_P(x-\lfloor x\rfloor)+\lfloor x\rfloor+1)} \mathcal{R}^{s}_{W}(y)dy \\
&\overset{a.s.}{\to} \frac{1}{\beta} \int_{r(x)}^{r(x)+\beta} \mathcal{R}^{s}_{W}(y)dy.
\end{aligned}
$$

Therefore, defining by convention $\mathcal{R}^{0}_{W}(y) = 1$ for $y \in [0, (W+1)\beta]$, and $\mathcal{T}^{0}_{W}(x) = 1$ for $x \in [0, W]$,

$$
\begin{aligned}
\mathcal{R}^{\ell}_{W}(y) &= l(y) \sum_{s=0}^{\ell-1} g(\boldsymbol{\mathcal{T}}^{\ell-s-1}, y)\mathcal{R}^{s}_{W}(y) \\
\mathcal{T}^{\ell}_{W}(x) &= \sum_{s=0}^{\ell-1} f(\boldsymbol{\mathcal{R}}^{\ell-s-1}, x)\mathcal{T}^{s}_{W}(x).
\end{aligned}
$$

Case (ii) can be verified for the vectors $\boldsymbol{\hbar}_{N,W}(k)$ but not for the vectors $\boldsymbol{\eth}_{N,W}(n)$ In fact, as already mentioned, $\lim_{N\to\infty} \frac{\mathcal{L}(\boldsymbol{\eth}_{N,W}(n))}{N} = \beta > 0$ for any n. Thus, we can focus on the case as $\lim_{N\to\infty} \frac{\mathcal{L}(\boldsymbol{\hbar}_{N,W}(k))}{N} = 0$. This corresponds to values of k such that $l(y) = 0$. Assumption (c) implies that the spectral radius of $\boldsymbol{\mathcal{R}}^{s}_{N,W}$ for N sufficiently large is finite and bounded with probability one. Let $\lambda_{\max}$ denote this upper bound. Lemma 8 in Appendix E guarantees that for any vector $\boldsymbol{\hbar}_{N,W}(k)$ the scalar $\boldsymbol{\hbar}^{H}_{N,W}(k)\boldsymbol{\mathcal{R}}^{s}_{N,W}\boldsymbol{\hbar}_{N,W}(k)$ is upper bounded by $\lambda_{\max}\boldsymbol{\hbar}^{H}_{N,W}(k)\boldsymbol{\hbar}_{N,W}(k)$. Then, for any realization of $\boldsymbol{\mathcal{R}}^{s}_{N,W,\sim k}$ and $\boldsymbol{\hbar}_{N,W}(k)$ there results

$$
|\boldsymbol{\hbar}^{H}_{N,W}(k)\boldsymbol{\mathcal{R}}^{s}_{N,W,\sim k}\boldsymbol{\hbar}_{N,W}(k)| \leq \lambda_{\max}\boldsymbol{\hbar}^{H}_{N,W}(k)\boldsymbol{\hbar}_{N,W}(k). \tag{C.19}
$$

Then, the convergence in probability 1 of (C.19) reduces to the almost sure convergence of $\boldsymbol{\hbar}_{N,W}^H(k)\boldsymbol{\hbar}_{N,W}(k)$ to zero as $N \to \infty$ for case (ii). This convergence has been already shown in the proof of the first step of the strong induction for this theorem. This concludes the proof of Theorem 8.

C.2 Proof of Theorem 9

Beside the notation introduced in Section C.1 for the proof of Theorem 8, in this section we adopt the following notation.

For $k = 1, \ldots, (W-1)K$ and $n = 1, \ldots, WN$ we define:

- $\varsigma_{N,W}(k)$, the k^{th} column of $\boldsymbol{\mathcal{S}}_{N,W}$.
- $\boldsymbol{\sigma}_{N,W}(n)$, the n^{th} row of $\boldsymbol{\mathcal{S}}_{N,W}$.
- $\alpha(k)$, the k^{th} diagonal element of the matrix $\boldsymbol{\mathcal{A}}$.

The proof of Theorem 9 follows along the lines of the proof of Theorem 8 by strong induction. The first step proves that $(\boldsymbol{\mathcal{T}}_{N,W})_{nn} \xrightarrow{a.s.} \beta \mathrm{E}_P\{\lambda(\rho)\}$ and $(\boldsymbol{\mathcal{R}}_{N,W})_{kk} \xrightarrow{a.s.} \lambda(y - \lfloor \frac{y}{\beta} \rfloor \beta) l(y)$, as $N \to \infty$, with $y = \lim_{K=\beta N \to \infty} \frac{k}{N}$, $l(y)$ defined in (4.25), and $\lambda(\rho)$ defined in the statement of Theorem 9. In the ℓ^{th} step we assume the almost sure convergence of $(\boldsymbol{\mathcal{T}}_{N,W}^m)_{nn}$ and $(\boldsymbol{\mathcal{R}}_{N,W}^m)_{kk}$ for $1 \le m \le \ell - 1$ and we prove that $(\boldsymbol{\mathcal{T}}_{N,W}^\ell)_{nn}$ and $(\boldsymbol{\mathcal{R}}_{N,W}^\ell)_{kk}$ converge almost surely to the deterministic values in (4.19) and (4.20) respectively.

First step: Let us consider $(\boldsymbol{\mathcal{R}}_{N,W}^1)_{kk} = \boldsymbol{\hbar}_{N,W}^H(k)\boldsymbol{\hbar}_{N,W}(k) = |\alpha(k)|^2 \varsigma_{N,W}^H(k)\varsigma_{N,W}(k)$. Given k,

$$\lim_{K=\beta N \to \infty} (\boldsymbol{\mathcal{R}}_{N,W}^1)_{kk} \overset{a.s.}{=} \lim_{K=\beta N \to \infty} |a_{k-\lfloor \frac{k-1}{\beta} \beta \rfloor, k-\lfloor \frac{k-1}{\beta} \beta \rfloor}|^2 \frac{\mathcal{L}(\varsigma_{N,W}(k))}{N}$$
$$= \lambda\left(y - \left\lfloor \frac{y}{\beta} \right\rfloor \beta\right) l(y)$$

where a_{kk} is the k^{th} element of the matrix $\boldsymbol{A}$ and $\mathcal{L}(\boldsymbol{x})$ denotes the number of nonzero elements in the vector $\boldsymbol{x}$ as in Theorem 8.

Let us focus on $(\boldsymbol{\mathcal{T}}_{N,W}^1)_{nn} = \boldsymbol{\eth}_{N,W}(n)\boldsymbol{\eth}_{N,W}^H(n) = \boldsymbol{\sigma}_{N,W}(n)\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{A}}^H\boldsymbol{\sigma}_{N,W}^H(n)$. $\boldsymbol{\sigma}_{N,W}(n)$ has nonzero i.i.d. elements in the interval $[n, n+K]$. Since the diagonal elements of $\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{A}}^H$ are a periodical repetition of the elements of the matrix $\boldsymbol{A}\boldsymbol{A}^H$, any principal submatrix $(\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{A}}^H)_{n:n+K}$ has the same trace as $\boldsymbol{A}\boldsymbol{A}^H$. Thanks to assumption (c) in Theorem 9 the spectral radius of the submatrix $(\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{A}}^H)_{n_1:n_1+K}$ is upper bounded. Furthermore, using the assumption that the sequence of the empirical distribution of $|a_{kk}|^2$, where a_{kk} are the diagonal elements of the matrix $\boldsymbol{A}$, converges almost surely to the deterministic distribution $F_\Lambda(\lambda)$, the normalized trace of the matrix $\boldsymbol{A}\boldsymbol{A}^H$ converges to a deterministic limit value

$$\lim_{K\to\infty}\frac{1}{K}\mathrm{tr}\boldsymbol{A}\boldsymbol{A}^H = \lim_{K\to\infty}\frac{1}{K}\sum_{k=1}^{K}|a_{kk}|^2 \overset{a.s.}{=} \mathrm{E}_P(\lambda(\rho)). \tag{C.20}$$

Appealing to Lemma 1 in Chapter 2 and following the same line as in the proof of Lemma 2, we obtain

$$\lim_{K\to\beta N\to\infty}(\boldsymbol{\mathcal{T}}^1_{N,W})_{nn} = \beta\mathrm{E}_\rho(\lambda(\rho)).$$

<u>*Step ℓ:*</u> Following along the line of Theorem 8 and making use of the same lemmas we can expand $(\boldsymbol{\mathcal{R}}^\ell_{N,W})_{nn}$ and $(\boldsymbol{\mathcal{T}}^\ell_{N,W})_{kk}$:

$$\begin{aligned}(\boldsymbol{\mathcal{R}}^\ell_{N,W})_{kk} &= |\alpha(k)|^2\varsigma^H_{N,W}(k)\left(\boldsymbol{\mathcal{T}}_{N,W,\sim k} + |\alpha(k)|^2\varsigma_{N,W}(k)\varsigma_{N,W}(k)\right)^{\ell-1}\varsigma_{N,W}(k)\\ &= \sum_{s=0}^{\ell-1}|\alpha(k)|^2\varsigma^H_{N,W}(k)\boldsymbol{\mathcal{T}}^{\ell-s-1}_{N,W,\sim k}\varsigma_{N,W}(k)(\boldsymbol{\mathcal{R}}^s_{N,W})_{kk} \end{aligned}\tag{C.21}$$

$$\begin{aligned}(\boldsymbol{\mathcal{T}}^\ell_{N,W})_{nn} &= \boldsymbol{\sigma}_{N,W}(n)\boldsymbol{\mathcal{A}}\left(\boldsymbol{\mathcal{R}}_{N,W,\vDash n} + \boldsymbol{\mathcal{A}}^H\boldsymbol{\sigma}^H_{N,W}(n)\boldsymbol{\sigma}_{N,W}(n)\boldsymbol{\mathcal{A}}\right)^{\ell-1}\boldsymbol{\mathcal{A}}^H\boldsymbol{\sigma}^H_{N,W}(n)\\ &= \sum_{s=0}^{\ell-1}\boldsymbol{\sigma}_{N,W}(n)\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{R}}^{\ell-s-1}_{N,W,\vDash n}\boldsymbol{\mathcal{A}}^H\boldsymbol{\sigma}^H_{N,W}(n)(\boldsymbol{\mathcal{T}}^s_{N,W})_{nn}. \end{aligned}\tag{C.22}$$

Then, the proof of almost sure convergence of $(\boldsymbol{\mathcal{R}}^\ell_{N,W})_{kk}$ and $(\boldsymbol{\mathcal{T}}^\ell_{N,W})_{nn}$ reduces to the proof of the almost sure convergence of $|\alpha(k)|^2\varsigma_{N,W}(k)\boldsymbol{\mathcal{T}}^s_{N,W,\sim k}\varsigma^H_{N,W}(k)$ and $\boldsymbol{\sigma}_{N,W}(n)\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{R}}^s_{N,W,\vDash k}\boldsymbol{\mathcal{A}}^H\boldsymbol{\sigma}^H_{N,W}(n)$, respectively, to a deterministic value.

First let us focus on the case that the number of nonzero elements in $\varsigma_{NW}(k)$ and $\boldsymbol{\sigma}_{NW}(n)$ is infinite as $N\to\infty$ and use the same argument as in Theorem 8 to obtain

$$\begin{aligned} g(\mathcal{T}^s_W, y) &\triangleq \lim_{K=\beta N\to\infty}|\alpha(k)|^2\varsigma_{N,W}(k)\boldsymbol{\mathcal{T}}^s_{N,W,\sim k}\varsigma^H_{N,W}(k)\\ &= \lim_{N=\beta K\to\infty}|a_{k-\lfloor\frac{k-1}{K}\rfloor K,k-\lfloor\frac{k-1}{K}\rfloor}K|^2\sum_{r=k_1}^{k_2}(\boldsymbol{\mathcal{T}}^s_{N,W,\sim k})_{rr}\\ &\overset{a.s.}{=} \lambda\left(y-\left\lfloor\frac{y}{\beta}\right\rfloor\beta\right)\int_{\max\left(0,\lfloor\frac{y}{\beta}\rfloor+F_P^{-1}\left(y-\lfloor\frac{y}{\beta}\rfloor\beta\right)-1\right)}^{\min\left(W,\lfloor\frac{y}{\beta}\rfloor+F_P^{-1}\left(y-\lfloor\frac{y}{\beta}\rfloor\beta\right)\right)}\mathcal{T}^s_W(x)\mathrm{d}\,x \end{aligned}$$

where k_1 and k_2 are defined in (C.16) and (C.17), respectively.

In order to prove that $\boldsymbol{\sigma}_{N,W}(n)\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{R}}^s_{N,W,\vDash k}\boldsymbol{\mathcal{A}}^H\boldsymbol{\sigma}^H_{N,W}(n)$ converges almost surely to a deterministic value and to determine that value, we introduce the functions $\widetilde{F}_\rho(x)$

and $\widetilde{\lambda}(\rho)$ corresponding to the limit distribution $F_P(x)$ and the limit function $\lambda(\rho)$:

$$\widetilde{F}_\rho(x) \triangleq \begin{Bmatrix} F_P(x), & 0 \le x < 1 \\ F_P(x-1)+1, & 1 \le x < 2 \\ \cdots & \\ F_P(x-W+1)+W-1, & W-1 \le x < W \end{Bmatrix} = F_P(x - \lfloor x \rfloor) + \lfloor x \rfloor$$

$$\widetilde{\lambda}(\rho) \triangleq \lambda(\rho - \lfloor \rho \rfloor), \qquad 0 \le \rho \le W.$$

In a similar way we also define $\widetilde{F}_\rho^{(N)}(n)$ and $\widetilde{\lambda}^{(N)}(n)$ corresponding to the empirical distribution $F_\rho^{(N)}(n)$ and $\lambda^{(N)}(n)$ for a finite N. Then, using the definition of n_1 in (C.18) and taking into account the relation $\lim_{N=\beta K\to\infty} \frac{n_1}{N} = \beta\widetilde{F}_\rho(x)$, there results

$$\begin{aligned}
f(\mathcal{R}_W^s, x) &\triangleq \lim_{N=\beta K\to\infty} \boldsymbol{\sigma}_{N,W}(n)\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{R}}_{N,W,\vDash k}^s\boldsymbol{\mathcal{A}}^H\boldsymbol{\sigma}_{N,W}^H(n) \\
&= \lim_{N=\beta K\to\infty} \frac{1}{N} \sum_{k=N\beta\widetilde{F}_\rho^{(N)}(n)}^{N\beta\widetilde{F}_\rho^{(N)}(n)+K} \widetilde{\lambda}^{(N)}(\rho^{(N)}(k))(\boldsymbol{\mathcal{R}}_{N,W,\vDash k}^s)_{kk} \\
&= \int_{\beta\widetilde{F}_\rho(x)}^{\beta\widetilde{F}_\rho(x)+\beta} \widetilde{\lambda}(\widetilde{\rho}(z))\mathcal{R}_W^s(z)\mathrm{d}\, z \\
&= \int_x^{x+1} \widetilde{\lambda}(\rho)\mathcal{R}_W^s(\beta\widetilde{F}(\rho))\beta\mathrm{d}\widetilde{F}_\rho(\rho) \\
&= \int_x^{\lfloor x+1\rfloor} \widetilde{\lambda}(\rho-\lfloor\rho\rfloor)\mathcal{R}_W^s(\beta F_P(\rho-\lfloor\rho\rfloor)+\beta\lfloor\rho\rfloor)\beta\mathrm{d}F_P(\rho-\lfloor\rho\rfloor) \\
&+ \int_{\lfloor x+1\rfloor}^{x+1} \widetilde{\lambda}(\rho-\lfloor\rho\rfloor)\mathcal{R}_W^s(\beta F_P(\rho-\lfloor\rho\rfloor)+\beta\lfloor\rho\rfloor)\beta\mathrm{d}F_P(\rho-\lfloor\rho\rfloor). \qquad (C.23)
\end{aligned}$$

Equation (C.23) holds since $F_P(\rho-\lfloor\rho\rfloor)$ is differentiable in the intervals $(x, \lfloor x+1\rfloor)$ and $(\lfloor x+1\rfloor, x+1)$ thanks to assumption (b). Therefore, stating by convention $\mathcal{R}_W^0(y) = 1$ for $y \in [0, (W+1)\beta]$, and $\mathcal{T}_W^0(x) = 1$ for $x \in [0, W]$,

$$\mathcal{R}_W^\ell(y) = \lim_{K=\beta N\to\infty} (\mathcal{R}_{N,W}^\ell)_{kk} = \sum_{s=0}^{\ell-1} \lambda\left(y - \left\lfloor\frac{y}{\beta}\right\rfloor\beta\right) g(\mathcal{T}^{\ell-s-1}, y)\mathcal{R}_W^s(y)$$

$$\mathcal{T}_W^\ell(x) = \lim_{K=\beta N\to\infty} (\boldsymbol{\mathcal{T}}_{N,W}^\ell)_{nn} = \sum_{s=0}^{\ell-1} f(\mathcal{R}^{\ell-s-1}, x)\mathcal{T}_W^s(x).$$

In the case that the number of the nonzero elements in $\varsigma_{NW}(k)$ keeps finite as $N\to\infty$, we use the same argument as in the proof of Theorem 8 to prove that

$$\lim_{N=\beta K\to\infty} |\alpha(k)|^2\varsigma_{NW}^H(k)\boldsymbol{\mathcal{T}}_{NW,\sim k}^s\varsigma_{NW}(k) \overset{a.s.}{=} 0.$$

This concludes the proof of Theorem 9.

C.3 Proof of Theorem 10

Let us consider an r-block-wise circulant matrix of order N, $\boldsymbol{C}_k^{(N)}$ defined in Theorem 10, and let us denote with $\boldsymbol{F}_N^H$ the unitary Fourier transform matrix of dimensions $N \times N$

$$\boldsymbol{F}_N = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega^1 & \omega^2 & \cdots & \omega^{N-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{(N-1)} & \omega^{2(N-1)} & \cdots & \omega^{(N-1)(N-1)} \end{pmatrix}$$

with $\omega = \mathrm{e}^{\frac{j2\pi}{N}}$. We can extend the well known result on the diagonalization of circulant matrices[4] to decompose the r-block-wise circulant matrix $\boldsymbol{C}_k^{(N)}$ as

$$\boldsymbol{C}_k^{(N)} = (\boldsymbol{F}_N \otimes \boldsymbol{I}_r)\boldsymbol{\Delta}_k \boldsymbol{F}_N$$

where $\boldsymbol{\Delta}_k$ is an $rN \times N$ block diagonal matrix with ℓ^{th} block

$$(\boldsymbol{\Delta}_k)_{\ell\ell} = \begin{pmatrix} \xi_{\tau_k}(\frac{\ell-1}{N}) \\ \xi_{\tau_k - \frac{T_c}{r}}(\frac{\ell-1}{N}) \\ \vdots \\ \xi_{\tau_k - \frac{(r-1)T_c}{r}}(\frac{\ell-1}{N}) \end{pmatrix},$$

$$\xi_\tau(x) \triangleq \frac{1}{T_c} \sum_{s=-\infty}^{+\infty} \mathrm{e}^{j2\pi\frac{\tau}{T_c}(x+s)} \Xi^* \left(\frac{j2\pi}{T_c}(x+s)\right),$$

and $(\boldsymbol{F}_N \otimes \boldsymbol{I}_r)$ is a unitary matrix.

The matrix $\overline{\boldsymbol{S}}$ can then be rewritten as

$$\overline{\boldsymbol{S}} = (\boldsymbol{F}_N \otimes \boldsymbol{I}_r)(\boldsymbol{\Delta}_1\overline{\boldsymbol{s}}_1, \boldsymbol{\Delta}_2\overline{\boldsymbol{s}}_2, \ldots, \boldsymbol{\Delta}_K\overline{\boldsymbol{s}}_K),$$

with $\overline{\boldsymbol{s}}_k = \boldsymbol{F}_N^H \boldsymbol{s}_k$. Assuming the elements of the spreading sequence $\boldsymbol{s}_k$ i.i.d. Gaussian distributed, $\overline{\boldsymbol{s}}_k$ is also a vector with i.i.d. Gaussian distributed elements having the same distribution as the elements of $\boldsymbol{s}_k$. This assumption will be removed later on in the proof.

In the following we focus on the asymptotic spectral analysis of the matrix $\overline{\boldsymbol{R}} = \boldsymbol{A}^H \overline{\boldsymbol{S}}^H \overline{\boldsymbol{S}} \boldsymbol{A} = \overline{\boldsymbol{H}}^H \overline{\boldsymbol{H}}$ with $\overline{\boldsymbol{H}} = (\boldsymbol{\Delta}_1\overline{\boldsymbol{s}}_1, \boldsymbol{\Delta}_2\overline{\boldsymbol{s}}_2, \ldots, \boldsymbol{\Delta}_K\overline{\boldsymbol{s}}_K)\boldsymbol{A}$. Considering that the absolute value of $\xi_\tau(x)$ is upper bounded thanks to the assumption of Theorem 10 that $\Xi(j2\pi f)$ has finite support and is bounded in absolute value and applying the same arguments as in Lemma 6, the matrix $\overline{\boldsymbol{H}}$ satisfies the conditions of Theorem 5.

[4] A circulant matrix $\boldsymbol{C}^{(N)}(f(x))$ can be decomposed as $\boldsymbol{C}^{(N)}(f(x)) = \boldsymbol{F}_N \boldsymbol{D} \boldsymbol{F}_N^H$, with $\boldsymbol{D} = \mathrm{diag}(f(0), f(\frac{1}{N}), \ldots, f(\frac{N-1}{N}))$.

Then, with probability one the eigenvalue distribution of the matrix $\overline{\boldsymbol{R}}$ converges to a deterministic distribution. Additionally, its finite moments are also finite thanks to assumption (d) of Theorem 10.

As the proofs of Theorem 8 and Theorem 9, the proof of Theorem 10 is based on strong induction. In the first step we prove the following facts:

1. The diagonal elements of the matrix $\overline{\boldsymbol{R}}$ converge almost surely, as $N \to \infty$, to deterministic values $\overline{R}(|a_{kk}|^2, \tau_k)$, conditionally on $(|a_{kk}|^2, \tau_k)$.

2. $(\overline{\boldsymbol{T}})_{nn}$, the $r \times r$ block diagonal elements of the matrix $\overline{\boldsymbol{T}} = \overline{\boldsymbol{H}}\overline{\boldsymbol{H}}^H$, converge almost surely to deterministic blocks $\overline{\boldsymbol{T}}(x)$, with $x = \lim_{N\to\infty} \frac{n(N)}{N}$.

Then, in the recursion step, we use the following induction assumptions:

1. For $s = 1, \dots, \ell - 1$, the diagonal elements of the matrix $\overline{\boldsymbol{R}}^s$, converge, as $N \to \infty$, to deterministic values $\overline{R}^s(|a_{kk}|^2, \tau_k)$, conditionally on $(|a_{kk}|^2, \tau_k)$,

2. For $s = 1, \dots, \ell-1$, $(\overline{\boldsymbol{T}}^s)_{nn}$, the $r \times r$ block diagonal elements of the matrix $\overline{\boldsymbol{T}}^s$ converge almost surely to deterministic blocks $\boldsymbol{T}^s(x)$, with $x = \lim_{N\to\infty} \frac{n(N)}{N}$.

We prove:

1. The diagonal elements of the matrix $\overline{\boldsymbol{R}}^\ell$, converge, as $N \to \infty$, to deterministic values $R^\ell(|a_{kk}|^2, \tau_k)$, conditionally on $(|a_{kk}|^2, \tau_k)$.

2. The blocks $(\overline{\boldsymbol{T}}^\ell)_{nn}$, converge almost surely to deterministic blocks $\overline{\boldsymbol{T}}^\ell(x)$.

Throughout this proof we adopt the following notation. For $k = 1, \dots, K$ and $n = 1, \dots, N$

- $\overline{\boldsymbol{h}}_k$ is the k^{th} column of the matrix $\overline{\boldsymbol{H}}$;
- $\overline{\boldsymbol{\delta}}_n$ is the n^{th} block row of $\overline{\boldsymbol{H}}$ of dimension $r \times K$;
- $\overline{\boldsymbol{H}}_{\vDash n}$ is the matrix obtained from $\overline{\boldsymbol{H}}$ by suppressing $\overline{\boldsymbol{\delta}}_n$;
- $\overline{\boldsymbol{H}}_{\sim k}$ is the matrix obtained from $\overline{\boldsymbol{H}}$ by suppressing $\overline{\boldsymbol{h}}_k$.
- $\overline{\boldsymbol{T}}_{\sim k} = \overline{\boldsymbol{H}}_{\sim k}\overline{\boldsymbol{H}}_{\sim k}^H$;
- $\overline{\boldsymbol{R}}_{\vDash n} = \overline{\boldsymbol{H}}_{\vDash n}^H \overline{\boldsymbol{H}}_{\vDash n}$;
- $\overline{\boldsymbol{\sigma}}_n = (\overline{s}_{n1}, \overline{s}_{n2}, \dots, \overline{s}_{nK})$.

- $\boldsymbol{\nabla}_{n,t}$, for $t=1,\dots,r$ and $n=1,\dots,N$, is the $K\times K$ diagonal matrix with the k^{th} element equal to $\widetilde{\xi}_{\tau_k-\frac{(t-1)T_c}{r}}(\frac{n-1}{N})$. Note that $\overline{\boldsymbol{\sigma}}_n\boldsymbol{\nabla}_{n,t}\boldsymbol{A}$ coincides with the $(t+(n-1)r)^{\text{th}}$ row of the matrix $\overline{\boldsymbol{H}}$.

- $(\overline{\boldsymbol{T}}^s)_{nn}$ is the n^{th} diagonal block of $\overline{\boldsymbol{T}}^s$ of dimension $r\times r$.

First step: Consider $\overline{\boldsymbol{R}}_{kk}=\overline{\boldsymbol{h}}_k^H\overline{\boldsymbol{h}}_k=|a_{kk}|^2\overline{\boldsymbol{s}}_k^H\boldsymbol{\Delta}_k^H\boldsymbol{\Delta}_k\overline{\boldsymbol{s}}_k$. Thanks to the assumption that $\Xi(j2\pi f)$ is bounded in absolute value with finite support also $\xi_\tau(x)$ is upper bounded. Because of the property of circulant matrices the limit eigenvalues of the matrix $\boldsymbol{\Delta}_k^H\boldsymbol{\Delta}_k$ are given by $\sum_{t=1}^r|\widetilde{\xi}_{\tau_k-\frac{(t-1)T_c}{r}}(x)|^2$. Therefore, the limit eigenvalue distribution of the matrix $\boldsymbol{\Delta}_k^H\boldsymbol{\Delta}_k$ has upper bounded support. Then, we can apply Lemma 1 in Chapter 2, and following the same lines as in the proof of Lemma 2 we prove that $\overline{\boldsymbol{R}}_{kk}$ converges almost surely to the deterministic value

$$
\begin{aligned}
\overline{R}(\lambda,\tau)|_{(\lambda,\tau)=(|a_{kk}|^2,\tau_k)} &= \lim_{K=\beta N\to\infty}\overline{\boldsymbol{R}}_{kk}\\
&= \lim_{K=\beta N\to\infty}\frac{|a_{kk}|^2}{N}\mathrm{tr}(\overline{\boldsymbol{\Delta}}_k^H\overline{\boldsymbol{\Delta}}_k)\\
&= \lim_{K=\beta N\to\infty}\frac{|a_{kk}|^2}{N}\sum_{\ell=1}^{N}(\boldsymbol{\Delta}_k^H)_{\ell,\ell}(\boldsymbol{\Delta}_k)_{\ell,\ell}\\
&= \lambda\int_0^1\boldsymbol{\Delta}_\tau^H(x)\boldsymbol{\Delta}_\tau(x)\mathrm{d}\,x. \qquad \text{(C.24)}
\end{aligned}
$$

Let us now consider the block matrix $\overline{\boldsymbol{T}}_{nn}$ whose (u,v) element $(\overline{\boldsymbol{T}}_{nn})_{uv}$ is given by

$$
(\overline{\boldsymbol{T}}_{nn})_{uv}=\overline{\boldsymbol{\sigma}}_n\boldsymbol{A}\boldsymbol{\nabla}_{n,u}\boldsymbol{\nabla}_{n,v}^H\boldsymbol{A}^H\overline{\boldsymbol{\sigma}}_n^H.
$$

Thanks to assumption (c) of Theorem 10 on the support of $F_{|\boldsymbol{A}|^2,T}(\lambda,\tau)$ and the fact that $\xi_\tau(x)$ is bounded in absolute value, the spectral radius of the matrix $\boldsymbol{A}\boldsymbol{\nabla}_{n,u}\boldsymbol{\nabla}_{n,v}^H\boldsymbol{A}^H$ is upper bounded. Thus, we can apply Lemma 1 in Chapter 2 and proceed as in the proof of Lemma 2 to obtain

$$
\begin{aligned}
\lim_{K=\beta N\to\infty}(\overline{\boldsymbol{T}}_{nn})_{uv} &= \beta\lim_{K=\beta N\to\infty}\frac{1}{K}\mathrm{tr}\boldsymbol{A}\boldsymbol{\nabla}_{n,u}\boldsymbol{\nabla}_{n,v}^H\boldsymbol{A}^H\\
&= \lim_{K=\beta N\to\infty}\frac{\beta}{K}\sum_{k=1}^{K}|a_{kk}|^2\xi_{\tau_k-\frac{u-1}{r}T_c}\left(\frac{n-1}{N}\right)\xi^*_{\tau_k-\frac{v-1}{r}T_c}\left(\frac{n-1}{N}\right)\\
&= \beta\int\lambda\xi_{\tau-\frac{u-1}{r}T_c}(x)\xi_{\tau-\frac{v-1}{r}T_c}(x)\mathrm{d}\,F(\lambda,\tau), \qquad \text{(C.25)}
\end{aligned}
$$

with

$$\boldsymbol{\Delta}_\tau(x) = \begin{pmatrix} \xi_\tau(x) \\ \xi_{\tau-\frac{T_c}{r}}(x) \\ \vdots \\ \xi_{\tau-\frac{(r-1)T_c}{r}}(x) \end{pmatrix}.$$

Therefore,

$$\begin{aligned}\overline{\boldsymbol{T}}(x) &= \lim_{K=\beta N\to\infty} \overline{\boldsymbol{T}}_{nn} \\ &= \beta \int \lambda \boldsymbol{\Delta}_\tau(x)\boldsymbol{\Delta}_\tau^H(x) \mathrm{d}\, F(\lambda,\tau)\end{aligned}$$

with $0 \le x \le 1$. This concludes the proof of the first step.

Step ℓ: Following the same approach as in the proof of Theorem 6, Appendix B, we can expand $(\overline{\boldsymbol{R}}^\ell)_{kk}$ and $(\overline{\boldsymbol{T}}^\ell)_{nn}$ as follows:

$$(\overline{\boldsymbol{R}}^\ell)_{kk} = \sum_{s=0}^{\ell-1} \overline{\boldsymbol{h}}_k^H \overline{\boldsymbol{T}}_{\sim k}^{\ell-s-1} \overline{\boldsymbol{h}}_k (\overline{\boldsymbol{R}}^s)_{kk} \tag{C.26}$$

$$(\overline{\boldsymbol{T}}^\ell)_{nn} = \sum_{s=0}^{\ell-1} \overline{\boldsymbol{\delta}}_n \overline{\boldsymbol{R}}_{\vDash n}^{\ell-s-1} \overline{\boldsymbol{\delta}}_n^H (\overline{\boldsymbol{T}}^s)_{nn}. \tag{C.27}$$

The almost sure convergence of $(\overline{\boldsymbol{R}}^\ell)_{kk}$ and $(\overline{\boldsymbol{T}}^\ell)_{nn}$ to a deterministic limit reduces to the almost sure convergence of $\overline{\boldsymbol{h}}_k^H \overline{\boldsymbol{T}}_{\sim k}^{\ell-s-1} \overline{\boldsymbol{h}}_k$ and $\overline{\boldsymbol{\delta}}_n \overline{\boldsymbol{R}}_{\vDash n}^{\ell-s-1} \overline{\boldsymbol{\delta}}_n^H$, for $s = 0, \ldots, \ell-1$, to a deterministic value. Using again the same approach as in the proof of Theorem 8 we obtain

$$\begin{aligned}\lim_{N=\beta K\to\infty} (\overline{\boldsymbol{\delta}}_n)_u \overline{\boldsymbol{R}}_{\vDash n}^s (\overline{\boldsymbol{\delta}}_n^H)_v &= \lim_{K=\beta N\to\infty} \frac{1}{N} \mathrm{tr} \boldsymbol{A} \boldsymbol{\nabla}_{nu} \overline{\boldsymbol{R}}_{\vDash n}^s \boldsymbol{\nabla}_{nv}^H \boldsymbol{A}^H \\ &= \lim_{K=\beta N\to\infty} \frac{1}{N} \sum_{k=1}^K |a_{kk}|^2 \xi_{\tau-\frac{u-1}{r}T_c}\left(\frac{n-1}{N}\right) \xi^*_{\tau-\frac{v-1}{r}T_c}\left(\frac{n-1}{N}\right) (\overline{\boldsymbol{R}}_{\vDash n}^s)_{kk} \\ &= \beta \int \lambda \xi^*_{\tau-\frac{u-1}{r}T_c}(x) \xi_{\tau-\frac{v-1}{r}T_c}(x) \overline{R}^s(\lambda,\tau) \mathrm{d}\, F(\lambda,\tau). \end{aligned} \tag{C.28}$$

We denote with $\mathrm{f}(\overline{\boldsymbol{R}}^s, x)$ the $r \times r$ matrix

$$\begin{aligned}\mathrm{f}(\overline{\boldsymbol{R}}^s, x) &= \lim_{K=\beta N\to\infty} \overline{\boldsymbol{\delta}}_n \overline{\boldsymbol{R}}_{\vDash n}^s \overline{\boldsymbol{\delta}}_n^H \\ &= \beta \int \lambda \boldsymbol{\Delta}_\tau \boldsymbol{\Delta}_\tau^H \overline{R}^s(\lambda,\tau) \mathrm{d}\, F(\lambda,\tau). \end{aligned} \tag{C.29}$$

Furthermore, we define

$$\begin{aligned}
g(\overline{\boldsymbol{T}}^s, \lambda, \tau) &= \lim_{N=\beta K\to\infty} \overline{\boldsymbol{h}}_k^H \overline{\boldsymbol{T}}^s_{\sim k} \overline{\boldsymbol{h}}_k \\
&= \lim_{N=\beta K\to\infty} \frac{1}{N} \mathrm{tr}\left(|a_{kk}|^2 \boldsymbol{\Delta}_k^H \overline{\boldsymbol{T}}^s_{\sim k} \boldsymbol{\Delta}_k\right) \\
&= \lim_{N=\beta K\to\infty} \frac{|a_{kk}|^2}{N} \sum_{n=1}^{N} (\boldsymbol{\Delta})_{nn}^H \overline{\boldsymbol{T}}_{nn} (\boldsymbol{\Delta})_{nn} \\
&= \lambda \int_0^1 \boldsymbol{\Delta}_\tau^H(x) \overline{\boldsymbol{T}}^s(x) \boldsymbol{\Delta}_\tau(x) \mathrm{d}\, x
\end{aligned}$$

where $\sqrt{\lambda}$ is the absolute value of the received amplitude of user k and τ its time delay. From (C.26)

$$\begin{aligned}
\overline{R}^\ell(\lambda, \tau) &= \lim_{N=\beta K\to\infty} (\overline{\boldsymbol{R}}^\ell)_{kk} \\
&= \sum_{s=0}^{\ell-1} g(\overline{\boldsymbol{T}}^{\ell-s-1}, \lambda, \tau) \overline{R}^s(\lambda, \tau)
\end{aligned}$$

and from (C.27)

$$\begin{aligned}
\overline{\boldsymbol{T}}^\ell(x) &= \lim_{N=\beta K\to\infty} (\overline{\boldsymbol{T}}^\ell)_{nn} \\
&= \sum_{s=0}^{\ell-1} \mathsf{f}(\overline{\boldsymbol{R}}^{\ell-s-1}, x) \overline{\boldsymbol{T}}^s(x).
\end{aligned}$$

Let us remove the assumption that the elements of the vectors $\boldsymbol{s}_k$ are Gaussian. For this wider class of spreading sequences the elements of the vector $\overline{\boldsymbol{s}}_k = \boldsymbol{F}_N^H \boldsymbol{s}_k$ are not i.i.d. and Lemma 1 in Section 2.4 cannot be applied. In the proof of Theorem 6 in [13] it is shown that Lemma 1 can be extended to any vector $\boldsymbol{x}_N = \boldsymbol{U}\boldsymbol{v}$ where $\boldsymbol{U}$ is a unitary matrix and $\boldsymbol{v}$ is a vector with elements satisfying assumption (a) of Theorem 10. By appealing to this result the extension of Theorem 10 to any distribution of the spreading elements is straightforward.

C.4 Derivation of Algorithm 2

For a signal with bandwidth $B \le \frac{1}{2T_c}$,

$$\boldsymbol{\Delta}_\tau(x) = \frac{1}{T_c} \Xi(j2\pi x/T_c) \mathrm{e}^{-\frac{j2\pi\tau x}{T_c}} \boldsymbol{e}, \qquad -\frac{1}{2} \le x \le \frac{1}{2}$$

with $\boldsymbol{e} = (1, \mathrm{e}^{j2\pi\frac{x}{r}}, \ldots \mathrm{e}^{j2\pi\frac{(r-1)}{r}x})$.

Specializing the recursive equation of Theorem 10 to this case the recursion can be drastically simplified as

$$\overline{R}^{\ell}(\lambda,\tau) = \sum_{s=0}^{\ell-1} g(\overline{\boldsymbol{T}}^{\ell-s-1},\lambda,\tau) \qquad \overline{R}^{s}(\lambda,\tau) \tag{C.30}$$

$$\overline{\boldsymbol{T}}^{\ell}(x) = \sum_{s=0}^{\ell-1} \boldsymbol{f}(\overline{\boldsymbol{R}}^{\ell-s-1},\lambda,\tau)\overline{\boldsymbol{T}}^{s}(x) \qquad -\frac{1}{2} \le x \le \frac{1}{2}$$

$$\boldsymbol{f}(\overline{\boldsymbol{R}}^{s},x) = \frac{\beta}{T_c^2}|\Xi(j2\pi x/T_c)|^2 \int \lambda \overline{R}^{s}(\lambda,\tau)\boldsymbol{e}\boldsymbol{e}^H \mathrm{d}\, F_{|\boldsymbol{A}|^2,T}(\lambda,\tau) \qquad -\frac{1}{2} \le x \le \frac{1}{2}$$

$$g(\boldsymbol{T}^s,\lambda,\tau) = \frac{\lambda}{T_c^2}\int_{-1/2}^{1/2} |\Xi(j2\pi x/T_c)|^2 \boldsymbol{e}^H \overline{\boldsymbol{T}}^{s}(x)\boldsymbol{e}\,\mathrm{d}\,x$$

with $\overline{\boldsymbol{T}}^{0}(x) = \boldsymbol{I}$ and $\overline{R}^{0}(\lambda,\tau) = 1$.

Let us observe that $g(\overline{\boldsymbol{T}}^{s},\lambda,\tau)$ is independent of τ for any ℓ. Considering the recursion on $\overline{R}^{\ell}(\lambda,\tau)$ and the initializing value $\overline{R}^{0}(\lambda,\tau) = 1$, it is apparent that also $\overline{R}^{\ell}(\lambda,\tau)$ is independent of τ.

Additionally, $(\boldsymbol{e}\boldsymbol{e}^H)^m = r^{m-1}\boldsymbol{e}\boldsymbol{e}^H$ where $r \in \mathbb{Z}^+$ is the dimension of the vector[5] $\boldsymbol{e}$. Then, it is straightforward to verify by recursion that the matrix $\overline{\boldsymbol{T}}^{s}(x)$, $s = 1,2,\ldots,\ell-1$, is proportional to the matrix $\boldsymbol{e}\boldsymbol{e}^H$ and we can express it as $\overline{\boldsymbol{T}}^{s}(x) = \overline{T}^{s}(x)\boldsymbol{e}\boldsymbol{e}^H$, $s = 1,2,\ldots$. The previous considerations yield

$$\overline{R}^{\ell}(\lambda) = \sum_{s=0}^{\ell-1} g(\overline{\boldsymbol{T}}^{\ell-s-1},\lambda)\overline{R}^{s}(\lambda)$$

$$\overline{T}^{\ell}(x)\boldsymbol{e}\boldsymbol{e}^H = \sum_{s=1}^{\ell-1} \boldsymbol{f}(\overline{\boldsymbol{R}}^{\ell-s-1},x)\overline{T}^{s}(x)\boldsymbol{e}\boldsymbol{e}^H + \boldsymbol{f}(\overline{\boldsymbol{R}}^{\ell-1},x)\overline{\boldsymbol{T}}^{0}(x) \tag{C.31}$$

$$\boldsymbol{f}(\overline{\boldsymbol{R}}^{s},x) = f(\overline{\boldsymbol{R}}^{s},x)\boldsymbol{e}\boldsymbol{e}^H \tag{C.32}$$

$$f(\overline{\boldsymbol{R}}^{s},x) = \frac{\beta}{T_c^2}|\Xi(j2\pi x/T_c)|^2 \int \lambda \overline{R}^{s}(\lambda)\mathrm{d}\, F_{|\boldsymbol{A}|^2}(\lambda) \qquad -\frac{1}{2} \le x \le \frac{1}{2}$$

$$g(\boldsymbol{T}^s,\lambda) = \frac{r^2\lambda}{T_c^2}\int_{-1/2}^{1/2} |\Xi(j2\pi x/T_c)|^2 \overline{T}^{s}(x)\mathrm{d}\,x$$

with $\overline{\boldsymbol{T}}^{0}(x) = \boldsymbol{I}$ and $\overline{R}^{0}(\lambda) = 1$.

[5] In system model (4.31) r is the sampling rate normalized by $\frac{1}{T_c}$.

Substituting (C.32) for $\boldsymbol{f}(\overline{\boldsymbol{R}}^s, x)$ in (C.31) we obtain

$$\begin{aligned}\overline{T}^{\ell}(x)\boldsymbol{e}\boldsymbol{e}^H &= \sum_{s=1}^{\ell-1} f(\overline{\boldsymbol{R}}^{\ell-s-1}, x)\overline{T}^s(x)\boldsymbol{e}\boldsymbol{e}^H\boldsymbol{e}\boldsymbol{e}^H + f(\overline{\boldsymbol{R}}^{\ell-1}, x)\overline{\boldsymbol{T}}^0(x)\boldsymbol{e}\boldsymbol{e}^H \\ &= r\sum_{s=1}^{\ell-1} f(\overline{\boldsymbol{R}}^{\ell-s-1}, x)\overline{T}^s(x)\boldsymbol{e}\boldsymbol{e}^H + f(\overline{\boldsymbol{R}}^{\ell-1}, x)\overline{\boldsymbol{T}}^0(x)\boldsymbol{e}\boldsymbol{e}^H \qquad \text{(C.33)}\end{aligned}$$

Recalling that $\overline{\boldsymbol{T}}^0(x) = \boldsymbol{I}$ and stating by convention $\overline{T}^0(x) = \frac{1}{r}$, we obtain from (C.33) the scalar $\overline{T}^{\ell}(x)$:

$$\begin{aligned}\overline{T}^{\ell}(x) &= r\left(\sum_{s=1}^{\ell-1} f(\overline{\boldsymbol{R}}^{\ell-s-1}, x)\overline{T}^s(x) + \frac{1}{r}f(\overline{\boldsymbol{R}}^{\ell-1}, x)\overline{\boldsymbol{T}}^0(x)\right) \\ &= r\sum_{s=0}^{\ell-1} f(\overline{\boldsymbol{R}}^{\ell-s-1}, x)\overline{T}^s(x) \qquad \text{(C.34)}\end{aligned}$$

where the sum in (C.34) includes the term $f(\overline{\boldsymbol{R}}^{\ell-1}, x))\overline{\boldsymbol{T}}^0(x)$ for $s = 0$.

The following equations summarize the final recursion.

$$\begin{aligned}\overline{R}^{\ell}(\lambda) &= \sum_{s=0}^{\ell-1} g(\overline{\boldsymbol{T}}^{\ell-s-1}, \lambda)\overline{R}^s(\lambda) \\ \overline{T}^{\ell}(x) &= r\sum_{s=0}^{\ell-1} f(\overline{\boldsymbol{R}}^{\ell-s-1}, x)\overline{T}^s(x) \\ f(\overline{\boldsymbol{R}}^s, x) &= \frac{\beta}{T_c^2}|\Xi(j2\pi x/T_c)|^2 \int \lambda\overline{R}^s(\lambda)\mathrm{d}\,F_{|\boldsymbol{A}|^2}(\lambda) \qquad -\frac{1}{2} \leq x \leq \frac{1}{2} \\ g(\boldsymbol{T}^s, \lambda) &= \left(\frac{r}{T_c}\right)^2 \lambda \int_{-1/2}^{1/2} |\Xi(j2\pi x/T_c)|^2\overline{T}^s(x)\mathrm{d}\,x\end{aligned}$$

with $\overline{T}^0(x) = \frac{1}{r}$ and $\overline{R}^0(\lambda) = 1$.

Algorithm 2 is derived from the previous set of equations by using the following substitutions[6]:

$$
\begin{array}{lcl}
\lambda & \longrightarrow & z \\
\frac{1}{T_c^2}\left|\Xi(j2\pi\frac{x}{T_c})\right|^2 & \longrightarrow & y \\
\overline{R}^s(\lambda) & \longrightarrow & \rho_s(z) \\
\overline{T}^s(x) & \longrightarrow & \mu_s(y) \\
\frac{1}{T_c^2}\left|\Xi(j2\pi\frac{x}{T_c})\right|^2 \overline{T}^s(x) & \longrightarrow & u_s(y) \\
\lambda\overline{R}^s(\lambda) & \longrightarrow & v_s(z) \\
\int_{-1/2}^{1/2}\frac{1}{T_c^2}\left|\Xi(j2\pi\frac{x}{T_c})\right|^2 \overline{T}^s(x)\mathrm{d}\,x & \longrightarrow & U_s \\
\int \lambda\overline{R}^s(\lambda)\mathrm{d}\,F_{|A|^2}(\lambda) & \longrightarrow & V_s.
\end{array}
$$

C.5 Proof of Corollary 5

Corollary 5 is derived by specializing Theorem 10 to a unitary Fourier transform $\Xi(j2\pi f)$ that is *real* and has bandwidth B with $\frac{1}{2T_c} \le B \le \frac{1}{T_c}$. The unitary Fourier transform in the discrete time domain is given by

$$\xi_\tau(x) = \frac{\mathrm{e}^{-j2\pi\frac{\tau x}{T_c}}}{T_c}\begin{cases}\Xi(j\frac{2\pi x}{T_c}) + \mathrm{e}^{-j2\pi\tau}\Xi(j\frac{2\pi(x+1)}{T_c}) & -\frac{1}{2} \le x \le 0 \\ \Xi(j\frac{2\pi x}{T_c}) + \mathrm{e}^{j2\pi\tau}\Xi(j\frac{2\pi(x-1)}{T_c}) & 0 \le x \le \frac{1}{2}.\end{cases}$$

Let $\widetilde{\boldsymbol{Q}}(x,\tau) \triangleq \boldsymbol{\Delta}_\tau(x)\boldsymbol{\Delta}_\tau^H(x)$ with $r=2$. $\widetilde{\boldsymbol{Q}}(x,\tau)$ can be decomposed as

$$\widetilde{\boldsymbol{Q}}(x,\tau) = \boldsymbol{Q}(x) + \overline{\boldsymbol{Q}}(x,\tau)$$

with $\boldsymbol{Q}(x)$ defined in (4.40) and

$$\overline{\boldsymbol{Q}}(x,\tau) = 2\overline{q}(x)\begin{pmatrix}\cos(2\pi\tau) & -j\mathrm{e}^{-j\pi x}\sin(2\pi\tau) \\ j\mathrm{e}^{j\pi x}\sin(2\pi\tau) & -\cos(2\pi\tau)\end{pmatrix} \tag{C.35}$$

where

$$\overline{q}(x) = \frac{1}{T_c}\begin{cases}\Xi(j\frac{2\pi x}{T_c})\Xi(j\frac{2\pi(x+1)}{T_c}) & -\frac{1}{2} \le x \le 0 \\ \Xi(j\frac{2\pi x}{T_c})\Xi(j\frac{2\pi(x-1)}{T_c}) & 0 \le x \le \frac{1}{2}.\end{cases} \tag{C.36}$$

[6]Note that the substitution of λ with z is redundant. It is used to obtain polynomials in the commonly used variable z.

Equations (4.36) and (4.37) can be rewritten as

$$\begin{aligned}\boldsymbol{f}(\overline{\boldsymbol{R}}^s, x) &= \beta\boldsymbol{Q}(x)\int \lambda\overline{R}^s(\lambda,\tau)\mathrm{d}F_{|\boldsymbol{A}|^2,T}(\lambda,\tau)\\ &\quad+\beta\int \lambda\overline{R}^s(\lambda,\tau)\overline{\boldsymbol{Q}}(x,\tau)\mathrm{d}F_{|\boldsymbol{A}|^2,T}(\lambda,\tau), \qquad -\frac{1}{2}\leq x\leq\frac{1}{2} \qquad \text{(C.37)}\\ g(\overline{\boldsymbol{T}}^s,\lambda,\tau) &= \lambda\int_{-\frac{1}{2}}^{\frac{1}{2}}\mathrm{tr}(\overline{\boldsymbol{T}}^s\boldsymbol{Q}(x))\mathrm{d}x+\lambda\int_{-\frac{1}{2}}^{\frac{1}{2}}\mathrm{tr}(\overline{\boldsymbol{T}}^s\overline{\boldsymbol{Q}}(x,\tau))\mathrm{d}x. \qquad \text{(C.38)}\end{aligned}$$

If condition (1) of Corollary 2 is verified, it can be shown that $g(\overline{\boldsymbol{T}}^s,\lambda,\tau)$ and $\overline{\boldsymbol{R}}^s(\lambda,\tau)$, $s\in\mathbb{Z}$, are independent of τ and $\overline{\boldsymbol{T}}^\ell(x)$ is a matrix of the form

$$\overline{\boldsymbol{T}}^\ell(x)=\begin{pmatrix} t_{\ell,1}(x) & t_{\ell,2}(x)\mathrm{e}^{-j\pi x}\\ t_{\ell,2}(x)\mathrm{e}^{j\pi x} & t_{\ell,1}(x)\end{pmatrix}. \qquad \text{(C.39)}$$

These properties can be proven by strong induction. It is straightforward to verify that they are satisfied for $s=0$. In fact, $\overline{R}^0(\lambda,\tau)=1$ is independent of τ and $\overline{\boldsymbol{T}}^0(x)=\boldsymbol{I}$ is of the form (C.39) with $t_{0,1}(x)=1$ and $t_{0,2}(x)=0$. Since $\mathrm{tr}(\overline{\boldsymbol{Q}}(x,\tau))=0$, $g(\widetilde{\boldsymbol{T}}^0,\lambda,\tau)=\lambda\int_{-\frac{1}{2}}^{\frac{1}{2}}\mathrm{tr}(\boldsymbol{Q}(x))\mathrm{d}x$ and hence $g(\widetilde{\boldsymbol{T}}^0,\lambda,\tau)$ is independent of τ.

The induction step is proven using the following induction assumptions:

- $\overline{R}^s(\lambda,\tau)$ and $g(\overline{\boldsymbol{T}}^s,\lambda,\tau)$ are independent of τ;
- For $s=0,1,\dots\ell-1$, $\overline{\boldsymbol{T}}^s(x)$ is of the form (C.39).

We have $\mathrm{tr}(\overline{\boldsymbol{T}}^s\overline{\boldsymbol{Q}}(x,\tau))=0$, for $s=0,1,\dots,\ell-1$, thanks to the form (C.39) of $\overline{\boldsymbol{T}}^s$. Furthermore, $g(\overline{\boldsymbol{T}}^s,\lambda,\tau)$ is independent of τ. Therefore, all quantities that appear in the right hand side of (4.34) are independent of τ and $\overline{R}^\ell(\lambda,\tau)$ is also independent of τ. In the following we will shortly write $\overline{R}^\ell(\lambda)$. Observing that $\int\overline{\boldsymbol{Q}}(x,\tau)\mathrm{d}F_T(\tau)=0$ thanks to the assumption (1) of Corollary 5 on the probability density function $f_T(\tau)$, (C.37) can be rewritten as

$$\begin{aligned}\boldsymbol{f}(\overline{\boldsymbol{R}},x) &= \beta\boldsymbol{Q}(x)\int\lambda\overline{R}^s(\lambda)\mathrm{d}F_{|\boldsymbol{A}|^2}(\lambda)\\ &= f(\overline{\boldsymbol{R}}^s)\boldsymbol{Q}(x) \qquad \text{(C.40)}\end{aligned}$$

with $f(\overline{\boldsymbol{R}}^s)=\beta\int\lambda\overline{R}^s(\lambda)\mathrm{d}F_{|\boldsymbol{A}|^2}(\lambda)$. Substituting (C.40) in (4.35) yields

$$\overline{\boldsymbol{T}}^\ell(x)=\sum_{s=0}^{\ell-1}f(\overline{\boldsymbol{R}}^s)\boldsymbol{Q}(x)\overline{\boldsymbol{T}}^s(x), \qquad -\frac{1}{2}\leq x\leq\frac{1}{2}. \qquad \text{(C.41)}$$

For $s = 1, 2, \ldots, \ell - 1$,

$$Q(x)\overline{T}^{s}(x) = \begin{pmatrix} q_1(x)t_{s1}(x) + q_2(x)t_{s2}(x) & [q_1(x)t_{s2}(x) + q_2(x)t_{s1}(x)]e^{-j\pi x} \\ [q_1(x)t_{s2}(x) + q_2(x)t_{s1}(x)]e^{j\pi x} & q_1(x)t_{s1}(x) + q_2(x)t_{s2}(x) \end{pmatrix}, \quad -\frac{1}{2} \le x \le \frac{1}{2}. \tag{C.42}$$

From (C.42) it is apparent that $\boldsymbol{Q}(x)\overline{\boldsymbol{T}}^{s}(x)$ is of the form (C.39). Since $\overline{\boldsymbol{T}}^{\ell}(x)$ is a linear combination of matrices of the form (C.39), $\overline{\boldsymbol{T}}^{\ell}(x)$ is also a matrix of the form (C.39).

Let us summarize the results in the following set of recursive equations:

$$\overline{R}^{\ell}(\lambda) = \sum_{s=0}^{\ell-1} g(\overline{\boldsymbol{T}}^{\ell-s-1}, \lambda)\overline{R}^{s}(\lambda) \tag{C.43}$$

$$\overline{\boldsymbol{T}}^{\ell}(x) = \sum_{s=0}^{\ell-1} f(\overline{\boldsymbol{R}}^{\ell-s-1})\boldsymbol{Q}(x)\overline{\boldsymbol{T}}^{s}(x) \tag{C.44}$$

$$f(\overline{\boldsymbol{R}}^{s}) = \beta \int \lambda \overline{R}^{s}(\lambda) \mathrm{d}\, F_{|\boldsymbol{A}|^2}(\lambda), \qquad -\frac{1}{2} \le x \le \frac{1}{2} \tag{C.45}$$

$$g(\overline{\boldsymbol{T}}^{s}, \lambda) = \lambda \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathrm{tr}(\overline{\boldsymbol{T}}^{s}(x)\boldsymbol{Q}(x))\mathrm{d}\, x \tag{C.46}$$

with $\overline{\boldsymbol{T}}^{0}(x) = \boldsymbol{I}_r$ and $\overline{R}^{0}(\lambda) = 1$.

Then, applying again Theorem 10 we obtain that

$$\lim_{K=\beta N \to \infty} (\overline{\boldsymbol{R}}^{\ell})_{kk} \overset{a.s.}{=} \overline{R}^{\ell}(\lambda)|_{\lambda=|a_{kk}|^2}.$$

This concludes the derivation of Corollary 5 from Theorem 10.

C.6 Derivation of Algorithm 3

The eigenvalues of the matrix $\boldsymbol{Q}(x)$ in (4.40) are

$$\begin{aligned} d_1(x) &= \frac{2}{T_c^2}\Xi^2\left(j\frac{2\pi}{T_c}x\right), \qquad -\frac{1}{2} \le x \le \frac{1}{2} \\ d_2(x) &= \frac{2}{T_c^2}\begin{cases} \Xi^2(j\frac{2\pi}{T_c}(x+1)), & -\frac{1}{2} \le x \le 0 \\ \Xi^2(j\frac{2\pi}{T_c}(x-1)), & 0 \le x \le \frac{1}{2}. \end{cases} \end{aligned} \tag{C.47}$$

Let us express the eigenvalue decomposition of the Hermitian matrix $\boldsymbol{Q}(x)$ as $\boldsymbol{Q}(x) = \boldsymbol{U}\boldsymbol{D}(x)\boldsymbol{U}^H$ where $\boldsymbol{U}$ is a unitary matrix and $\boldsymbol{D}(x)$ a diagonal matrix with elements $d_1(x)$ and $d_2(x)$. Considering (C.41) and the fact that $\overline{\boldsymbol{T}}^{0}(x) = \boldsymbol{I}$, it is apparent that

$\overline{\boldsymbol{T}}^{\ell}(x)$ is a polynomial in $\boldsymbol{Q}(x)$. Therefore, it has the same eigenvectors as $\boldsymbol{Q}(x)$ and can be decomposed as

$$\overline{\boldsymbol{T}}^{s}(x) = \boldsymbol{U}\boldsymbol{F}_{s}(x)\boldsymbol{U}^{H} \tag{C.48}$$

where $\boldsymbol{F}_s(x)$ is a diagonal matrix. Then, substituting $\boldsymbol{Q}(x) = \boldsymbol{U}\boldsymbol{D}(x)\boldsymbol{U}^H$ in (C.41) yields

$$\overline{\boldsymbol{T}}^{\ell}(x) = \sum_{s=0}^{\ell-1} f(\overline{\boldsymbol{R}}^{s})\boldsymbol{U}\boldsymbol{D}(x)\boldsymbol{U}^{H}\overline{\boldsymbol{T}}^{s}(x). \tag{C.49}$$

By substituting (C.48) in (C.49) we obtain

$$\begin{aligned}\boldsymbol{U}\boldsymbol{F}_{\ell}(x)\boldsymbol{U}^{H} &= \sum_{s=0}^{\ell-1} f(\overline{\boldsymbol{R}}^{s})\boldsymbol{U}\boldsymbol{D}(x)\boldsymbol{F}_{s}(x)\boldsymbol{U}^{H}\\ &= \boldsymbol{U}\left(\sum_{s=0}^{\ell-1} f(\overline{\boldsymbol{R}}^{s})\boldsymbol{D}(x)\boldsymbol{F}_{s}(x)\right)\boldsymbol{U}^{H}.\end{aligned} \tag{C.50}$$

Equation (C.50) yields

$$\boldsymbol{F}_{\ell}(x) = \sum_{s=0}^{\ell-1} f(\overline{\boldsymbol{R}}^{s})\boldsymbol{D}(x)\boldsymbol{F}_{s}(x). \tag{C.51}$$

From (C.51) it is apparent that $\boldsymbol{F}_{\ell}(x)$ is a polynomial of degree ℓ in $\boldsymbol{D}(x)$. Using the identity of the traces of similar matrices, we obtain

$$\mathrm{tr}(\overline{\boldsymbol{T}}^{s}(x)\boldsymbol{Q}(x)) = \mathrm{tr}(\boldsymbol{F}_{s}(x)\boldsymbol{D}(x)).$$

Therefore, (C.46) can be rewritten as

$$g(\overline{\boldsymbol{T}}^{\ell}, \lambda) = \lambda \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathrm{tr}(\boldsymbol{F}_{s}(x)\boldsymbol{D}(x))\mathrm{d}x \tag{C.52}$$

$$= \lambda \int_{-\frac{1}{2}}^{\frac{1}{2}} [(\boldsymbol{F}_{s})_{11}d_1 + (\boldsymbol{F}_{s})_{22}d_2]\mathrm{d}x \tag{C.53}$$

where $(\boldsymbol{F}^{s})_{11}d_1$ and $(\boldsymbol{F}^{s})_{22}d_2$ are polynomials in d_1 and d_2 with identical coefficients. Denoting by $u_s(y)$ the polynomial in a generic variable y with these coefficients, we can rewrite (C.53) as

$$g(\overline{\boldsymbol{T}}^{s}, \lambda) = \lambda \int_{-\frac{1}{2}}^{\frac{1}{2}} [u_s(d_1(x)) + u_s(d_2(x))]\,\mathrm{d}x.$$

Let $u_s(y) = \sum_{j=0}^{s} \alpha_r y^r$. Then,

$$\begin{aligned} g(\overline{\boldsymbol{T}}^s, \lambda) &= \lambda \sum_{r=0}^{s} \alpha_r \int_{-\frac{1}{2}}^{\frac{1}{2}} (d_1(x))^r + (d_2(x))^r \mathrm{d}x. \\ &= \lambda \sum_{r=0}^{s} \alpha_r \mathcal{E}_r \\ &= \lambda U_s \end{aligned} \tag{C.54}$$

with

$$\begin{aligned} \mathcal{E}_r &= \int_{-\frac{1}{2}}^{\frac{1}{2}} (d_1^r + d_2^r)\, \mathrm{d}x \\ &= \left(\frac{2}{T_c^2}\right)^r \int_{-1}^{1} \left| \Xi\left(j\frac{2\pi}{T_c}x\right)\right|^{2r} \mathrm{d}x. \end{aligned} \tag{C.55}$$

and

$$U_s = \sum_{r=0}^{s} \alpha_r \mathcal{E}_r$$

i.e. $g(\overline{\boldsymbol{T}}^s, \lambda)$ can be computed by substituting the integral $\int u_s(d_1) + u_s(d_2)\mathrm{d}x$ with the quantity obtained by replacing the monomials $y, y^2, \dots y^s$ in $u_s(y)$ by $\mathcal{E}_1, \mathcal{E}_2, \dots \mathcal{E}_s$, respectively.

Using (C.54) in (C.43) and replacing (C.44) by (C.51) the recursive equations of Corollary 5 can be rewritten as

$$\overline{R}^{\ell}(\lambda) = \sum_{s=0}^{\ell-1} \lambda \overline{R}^s(\lambda) U_{\ell-s-1} \tag{C.56}$$

$$\boldsymbol{F}_\ell(x) = \sum_{s=0}^{\ell-1} f(\overline{\boldsymbol{R}}^{\ell-s-1}) \boldsymbol{D}(x) \boldsymbol{F}_s(x) \tag{C.57}$$

$$f(\overline{\boldsymbol{R}}^s) = \beta \int \lambda \overline{R}^s(\lambda) \mathrm{d}F_{|\boldsymbol{A}|^2}(\lambda) \tag{C.58}$$

$$g(\overline{\boldsymbol{T}}^s, \lambda) = \lambda U_s, \tag{C.59}$$

with U_s defined in (C.54), $\boldsymbol{F}_0(x) = \boldsymbol{I}$, and $\overline{R}^0(\lambda) = 1$.

Let us observe that the computation of U_s requires to determine only $u_s(y)$ and not the diagonal matrix $\boldsymbol{D}(x)\boldsymbol{F}_s(x)$. We can easily recognize that $u_\ell(y)$ can be derived by replacing $\boldsymbol{F}_s(x)$ by $\mu_s(y)$ and $\boldsymbol{D}(x)$ by y in (C.57) to obtain

$$\mu_\ell(y) = \sum_{s=0}^{\ell-1} f(\overline{\boldsymbol{R}}^{\ell-s-1}) y \mu_s(y)$$

and by computing $u_\ell(y) = y\mu_\ell(y)$. The scalar U_ℓ is obtained by writing $u_\ell(y)$ as a polynomials in y and by replacing all monomials $y, y^2, \ldots, y^\ell$ with $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_\ell$, respectively.

For the computation of $\overline{R}^\ell(\lambda)$, we use the following substitutions[7]:

$$\begin{array}{lcl} \lambda & \longrightarrow & z \\ \overline{R}^s(\lambda) & \longrightarrow & \rho_s(z) \\ \lambda\overline{R}^s(\lambda) & \longrightarrow & v_s(z). \end{array}$$

Then, (C.56) is rewritten as

$$\rho_\ell(z) = \sum_{s=0}^{\ell-1} zU_{\ell-s-1}\rho_s(z)$$

and $V_\ell = \frac{f(\overline{R}^\ell)}{\beta}$ can be obtained from $v_\ell(z) = z\rho_\ell(z)$ by writing $v_\ell(z)$ as a polynomial in z and by replacing the monomials $z^1, z^2, \ldots, z^\ell$ by the moments $m^1_{|\boldsymbol{A}|^2}, m^2_{|\boldsymbol{A}|^2}, \ldots, m^\ell_{|\boldsymbol{A}|^2}$, respectively.

We conclude the derivation of Algorithm 3 by summarizing the previous considerations and substitutions:

- $$\begin{aligned} \rho_\ell(z) &= \sum_{s=0}^{\ell-1} zU_{\ell-s-1}\rho_s(z) \\ \mu_\ell(y) &= \sum_{s=0}^{\ell-1} \beta y V_{\ell-s-1}\mu_s(y). \end{aligned}$$
- U_s and V_s are obtained from $u_s(y) = y\mu_s(y)$ and $v_s(z) = z\rho_s(z)$, respectively by
  - expanding $u_s(y)$ and $v_s(z)$ as polynomials in y and z, respectively,
  - replacing the monomials y^r and z^r, $r = 1, \ldots, s$ with $\mathcal{E}_s$ and $m^s_{|\boldsymbol{A}|^2}$, respectively.

Then, $\overline{R}^\ell(\lambda) = \rho_\ell(\lambda)$ and the eigenvalue moment $m^\ell_{\overline{\boldsymbol{R}}} = \mathrm{E}\{\overline{R}^\ell(\lambda)\}$ is obtained by replacing all monomials $z, z^2, \ldots, z^\ell$ in the polynomial $\rho_\ell(z)$ by the moments $m^1_{|\boldsymbol{A}|^2}, m^2_{|\boldsymbol{A}|^2}, \ldots, m^\ell_{|\boldsymbol{A}|^2}$, respectively.

[7]Note that the substitution of λ with z is redundant. It is used to obtain polynomials in the commonly used variable z.

D Proofs of Chapter 5

D.1 Proof of Theorem 11

The proof of Theorem 11 is based on Theorem 5 in Section 2.4.

First of all, let us verify that the matrix $\mathfrak{H}$ satisfies conditions (2.34), (2.35), and (2.36). $\mathfrak{H}_{ij}$ is the $(i,j)^{\text{th}}$ $L \times 1$ block of the matrix $\mathfrak{H}$. All blocks are independent since the channel gains are deterministic and the spreading sequence elements are independent. Condition (2.34) of Theorem 5 specializes for Theorem 11 to the following inequality

$$\sup_N \left[\max_{i=1,\dots N} \sum_{j=1}^{K} \mathrm{E}\{\|\mathfrak{H}_{ij}\|^2\} + \max_{i=1,\dots K} \sum_{j=1}^{N} \mathrm{E}\{\|\mathfrak{H}_{ji}\|^2\} \right] \leq \sup_N \left[(\beta+1) \max_{i=1\dots K} \boldsymbol{l}_i^H \boldsymbol{l}_i \right] < +\infty. \tag{D.1}$$

The second inequality in (D.1) holds thanks to the assumption that $\|\boldsymbol{l}_i\|$ is uniformly bounded for all N. Thus, condition (2.34) is satisfied.

In order to verify Lindeberg condition (2.35), let us observe that $\forall i, j$

$$\begin{aligned} \mathrm{E}\{\|\mathfrak{H}_{ij}\|^2 1_{\{\|\mathfrak{H}_{ij}\|>\tau\}}(\mathfrak{H}_{ij})\} &= \int_{\{\|\mathfrak{H}_{ij}\|^2>\tau^2\}} \|\mathfrak{H}_{ij}\|^2 \mathrm{d}F(s_{ij}) \\ &= \boldsymbol{l}_j^H \boldsymbol{l}_j \int_{\{|s_{ij}|^2 > \frac{\tau^2}{\boldsymbol{l}_j^H \boldsymbol{l}_j}\}} |s_{ij}|^2 \mathrm{d}F(s_{ij}) \\ &= \frac{(\boldsymbol{l}_j^H \boldsymbol{l}_j)^{1+\delta}}{\tau^{2\delta}} \int_{\{|s_{ij}|^2 \geq 0\}} |s_{ij}|^2 \mathrm{d}F(s_{ij}) \end{aligned} \tag{D.2}$$

where $1_{\mathcal{A}}(\cdot)$ is the indicator function on the set $\mathcal{A}$ (see Glossary), $F(s_{ij})$ is the distribution function of s_{ij}, and $\delta \in \mathbb{R}^+$. From the condition of Theorem 11 $\mathrm{E}\{|s_{ij}|^2\} \leq \frac{1}{N^\gamma}$ with $\gamma > 1$. Since $\|\boldsymbol{l}_k\|$ are uniformly bounded for all N, there exists an $m < +\infty$ such that $\max\limits_{i=1\dots K} \boldsymbol{l}_i^H \boldsymbol{l}_i \leq m$ and

$$\max_{ij} \mathrm{E}\{\|\mathfrak{H}_{ij}\|^2 1_{\{\|\mathfrak{H}_{ij}\|>\tau\}}(\mathfrak{H}_{ij})\} \leq \frac{m^3}{\tau^2 N^\gamma} \qquad \text{with} \qquad \gamma > 1. \tag{D.3}$$

Then, we conclude that

$$\lim_{\substack{K,N\to\infty \\ \frac{K}{N}\to\beta}} \sum_{i=1}^{K} \mathrm{E}\{\|\mathfrak{H}_{ij}\|^2\} 1_{\{\|\mathfrak{H}_{ij}\|>\tau\}}(\mathfrak{H}_{ij}) \leq \lim_{N\to\infty} \frac{m\beta}{\tau^2 N^{\gamma-1}} = 0. \tag{D.4}$$

The proof that

$$\lim_{N\to\infty} \max_{i=1\ldots K} \sum_{j=1}^{N} \mathrm{E}\{\|\mathfrak{H}_{ji}\|^2 1_{\{\|\mathfrak{H}_{ji}\|>\tau\}}(\mathfrak{H}_{ji})\} = 0$$

follows the same lines as the proof of (D.4). Thus, we conclude that Lindeberg condition (2.35) is satisfied.

Condition (2.36) is trivially verified. In fact, the entries of matrix $\mathfrak{H}$ are all zero mean, i.e., $\boldsymbol{A}_{ij} = \mathbf{0}$, $\forall i, j \in \mathbb{Z}^+$ and also their spectral norms $|\boldsymbol{A}_{ij}|$ are zero. Then, the sums in (2.36) are also zero and condition (2.36) is verified.

Equation (2.37) can be rewritten as

$$\boldsymbol{C}_{kk}^{(1)}(\alpha) = \alpha \boldsymbol{I}_L + \sum_{j=1}^{K} \mathrm{E}\left\{\mathfrak{H}_{kj}([\widetilde{\boldsymbol{C}}_{K\times K}^{(2)}(\alpha)]^{-1})_{jj}\mathfrak{H}_{kj}^H\right\} \tag{D.5}$$

$$= \alpha \boldsymbol{I}_L + \sum_{j=1}^{K}([\widetilde{\boldsymbol{C}}_{K\times K}^{(2)}(\alpha)]^{-1})_{jj}\mathrm{E}\left\{\mathfrak{H}_{kj}\mathfrak{H}_{kj}^H\right\} \tag{D.6}$$

$$= \alpha \boldsymbol{I}_L + \boldsymbol{K}^{(1)}(\alpha) = \boldsymbol{C}^{(1)}(\alpha), \tag{D.7}$$

with

$$\boldsymbol{K}^{(1)}(\alpha) \triangleq \sum_{j=1}^{K}([\widetilde{\boldsymbol{C}}_{K\times K}^{(2)}(\alpha)]^{-1})_{jj}\mathrm{E}\left\{\mathfrak{H}_{kj}\mathfrak{H}_{kj}^H\right\}. \tag{D.8}$$

The step from (D.5) to (D.6) is justified by the fact that $([\widetilde{\boldsymbol{C}}_{K\times K}^{(2)}(\alpha)]^{-1})_{jj}$ is a scalar (1×1 matrix). Equation (D.7) emphasizes that the matrix $\boldsymbol{K}^{(1)}(\alpha)$ is independent of k. This is because all rows of the matrix $\mathfrak{H}$ have the same statistics.

Equation (2.38) can be specialized to system (5.1) as follows.

$$C_{ll}^{(2)}(\alpha) = 1 + \sum_{j=1}^{N} \mathrm{E}\left\{\mathfrak{H}_{jl}^{H}([\widetilde{C}_{N\times N}^{(1)}(\alpha)]^{-1})_{jj}\mathfrak{H}_{jl}\right\} \tag{D.9}$$

$$= 1 + \sum_{j=1}^{N} \mathrm{E}\left\{\mathrm{tr}[\mathfrak{H}_{jl}^{H}([C^{(1)}(\alpha)]^{-1})\mathfrak{H}_{jl}]\right\} \tag{D.10}$$

$$= 1 + \sum_{j=1}^{N} \mathrm{E}\left\{\mathrm{tr}[[C^{(1)}(\alpha)]^{-1}\mathfrak{H}_{jl}\mathfrak{H}_{jl}^{H}]\right\} \tag{D.11}$$

$$= 1 + \sum_{j=1}^{N} \mathrm{tr}\left([C^{(1)}(\alpha)]^{-1}\mathrm{E}\left\{\mathfrak{H}_{jl}\mathfrak{H}_{jl}^{H}\right\}\right)$$

$$= 1 + \mathrm{tr}\left([C^{(1)}(\alpha)]^{-1}\sum_{j=1}^{N}\mathrm{E}\left\{\mathfrak{H}_{jl}\mathfrak{H}_{jl}^{H}\right\}\right) \tag{D.12}$$

$$= 1 + N\mathrm{tr}\left([C^{(1)}(\alpha)]^{-1}\mathrm{E}\left\{\mathfrak{H}_{1l}\mathfrak{H}_{1l}^{H}\right\}\right) \tag{D.13}$$

$$= 1 + \mathrm{tr}\left([C^{(1)}(\alpha)]^{-1}l_l l_l^{H}\right) \tag{D.14}$$

$$= 1 + l_l^{H}[C^{(1)}(\alpha)]^{-1}l_l. \tag{D.15}$$

The step from (D.9) to (D.10) is justified by the fact that the argument in the expectation operator is a scalar. From (D.10) to (D.11) the trace property $\mathrm{tr}(\boldsymbol{AB}) = \mathrm{tr}(\boldsymbol{BA})$ is applied. From (D.11) to (D.12) we use the distributive property of trace, expectation, and sum. Equation (D.12) yields (D.13) thanks to the fact that the expectation of $\mathfrak{H}_{jl}\mathfrak{H}_{jl}^{H}$ is independent of j, i.e. $\mathrm{E}\left\{\mathfrak{H}_{jl}\mathfrak{H}_{jl}^{H}\right\} = \mathrm{E}\left\{\mathfrak{H}_{j'l}\mathfrak{H}_{j'l}^{H}\right\}$ $\forall j, j' \in [1,\ldots,K]$. Let us notice that by the definition of $\widetilde{C}_{K\times K}^{(2)}(\alpha)$ in the statement of Theorem (5) in Section (2.4) $([\widetilde{C}_{K\times K}^{(2)}(\alpha)]^{-1})_{jj} = C_{jj}^{(2)}(\alpha)$. Using (D.15) in (D.8), we obtain $K^{(1)}(\alpha) = \sum_{j=1}^{K} \frac{\mathrm{E}\left\{\mathfrak{H}_{kj}\mathfrak{H}_{kj}^{H}\right\}}{1+l_j^{H}[C^{(1)}(\alpha)]^{-1}l_j}$. Thus, (D.7) can be rewritten as

$$C^{(1)}(\alpha) = \alpha I_L + \sum_{j=1}^{K}\frac{\mathrm{E}\left\{\mathfrak{H}_{kj}\mathfrak{H}_{kj}^{H}\right\}}{1+l_j^{H}[C^{(1)}(\alpha)]^{-1}l_j}$$

$$= \alpha I_L + \frac{1}{N}\sum_{j=1}^{K}\frac{l_j l_j^{H}}{1+l_j^{H}[C^{(1)}(\alpha)]^{-1}l_j}.$$

Then, considering the limit for $K, N \to \infty$

$$C^{(1)}(\alpha) = \alpha I_L + \beta\int\frac{ll^{H}}{1+l^{H}[C^{(1)}(\alpha)]^{-1}l}\mathrm{d}\,F_l(l_1,l_2,\ldots,l_L) \tag{D.16}$$

and using the definition[1] $(\boldsymbol{C}^{(1)}(\alpha))^{-1} \triangleq \frac{\boldsymbol{A}}{\alpha}$ we obtain (5.9) with $\alpha = \sigma^2$.

As intermediate results of the proof of Theorem 5, Girko proved (see [61] proof[2] of Theorem 30.2) that the $L \times L$ matrix block element $\boldsymbol{U}_{ij}$, $i, j = 1, \ldots, N$, of the matrix $\boldsymbol{U} = \left(\alpha \boldsymbol{I}_{n_1} + \Xi_{n_1 \times n_2} \Xi^H_{n_1 \times n_2}\right)^{-1}$ in Theorem 5 converges in probability to the $L \times L$ matrix block element $\boldsymbol{V}_{ij}$ of the matrix $\boldsymbol{V} = (\widetilde{\boldsymbol{\Psi}}^{(1)}(\alpha) + \widetilde{\boldsymbol{A}}_{p_1 \times p_2}(\widetilde{\boldsymbol{\Psi}}^{(2)}(\alpha))^{-1} \widetilde{\boldsymbol{A}}^H_{p_1 \times p_2})^{-1}$, $\widetilde{\boldsymbol{\Psi}}^{(i)}(\alpha) = \mathrm{diag}\{\boldsymbol{\Psi}^{(i)}_{kk}(\alpha)\}_{k=1,\ldots,p_i}$ and $i = 1, 2$. Here, $\widetilde{\boldsymbol{\Psi}}^{(1)}(\alpha)$ and $\widetilde{\boldsymbol{\Psi}}^{(2)}(\alpha)$ are deterministic matrices such that

$$\lim_{\substack{K,N \to \infty \\ \frac{K}{N} \to \beta}} \widetilde{\boldsymbol{\Psi}}^{(i)}(\alpha) = \widetilde{\boldsymbol{C}}^{(i)}(\alpha) \qquad i = 1, 2 \tag{D.17}$$

and $\widetilde{\boldsymbol{C}}^{(i)}(\alpha)$, $i = 1, 2$ solutions of the canonical system of equations in Theorem 5. Furthermore,

$$\lim_{\substack{K,N \to \infty \\ \frac{K}{N} \to \beta}} \mathrm{E}\{|\boldsymbol{U}_{ij} - \boldsymbol{V}_{ij}|\} = 0. \tag{D.18}$$

and the spectral norm of $\boldsymbol{V}_{ij}$ is bounded by $|\boldsymbol{V}_{ij}| < \sigma^2$. By applying these intermediate results to the matrix $\boldsymbol{\mathfrak{U}} = (\boldsymbol{\mathfrak{H}}_k \boldsymbol{\mathfrak{H}}_k^H + \sigma^2 \boldsymbol{I})^{-1}$ we obtain

$$\lim_{\substack{K,N \to \infty \\ \frac{K}{N} \to \beta}} \mathrm{E}\{|\boldsymbol{\mathfrak{U}}_{ij} - \boldsymbol{V}_{ij}|\} = 0 \tag{D.19}$$

with $\boldsymbol{V}_{ij} = [\boldsymbol{\Psi}^{(1)}_{ii}(\sigma^2)]^{-1} \delta_{ij}$ from (D.18) and

$$\lim_{\substack{K,N \to \infty \\ \frac{K}{N} \to \beta}} \boldsymbol{\Psi}^{(1)}_{ii}(\sigma^2) = \boldsymbol{C}^{(1)}(\sigma^2). \tag{D.20}$$

from (D.20).

Let us denote by $\boldsymbol{\mathfrak{L}}_k$ the $LN \times N$ block diagonal matrix whose blocks are identically equal to $\boldsymbol{l}_k$. Its maximum singular value is equal to $\sqrt{\boldsymbol{l}_k^H \boldsymbol{l}_k} < +\infty$ since $\|\boldsymbol{l}_k\|$ is uniformly bounded for all K. Then, $\boldsymbol{\mathfrak{h}}_k = \boldsymbol{\mathfrak{L}}_k \boldsymbol{s}_k$ where $\boldsymbol{s}_k$ is the k^{th} column of the matrix $\boldsymbol{S}$.

[1] This definition is motivated by the fact that the expression of the SINR (5.9) in Theorem 11 in Section 5.3.1 becomes more intuitive using $\frac{\boldsymbol{A}}{\alpha}$ instead of $(\boldsymbol{C}^{(1)}(\alpha))^{-1}$.

[2] We note that there are several typos in the statement of Theorem 30.2 and in its proof. So, $\boldsymbol{Q}$ should be defined as $\boldsymbol{Q} = [\boldsymbol{I}\alpha + (\boldsymbol{A} + \Xi)(\boldsymbol{A} + \Xi)^T]$ and not as $\boldsymbol{Q} = [\boldsymbol{I}\alpha + (\boldsymbol{A} + \Xi)^T(\boldsymbol{A} + \Xi)]$, consistently also the definition of $\boldsymbol{G}$ changes. In the theorem statement $\boldsymbol{X} = \widetilde{\boldsymbol{G}}_{p2 \times p_2}$ and $\boldsymbol{Y} = \widetilde{\boldsymbol{Q}}_{p_1 \times p_1}$. These typos have been discussed with the theorem's author in personal correspondence.

The convergence in probability of $\mathrm{SINR}_k = \mathfrak{h}_k^H \mathfrak{U} \mathfrak{h}_k$ to the quantity $\boldsymbol{l}_k^H[\boldsymbol{C}^{(1)}(\sigma^2)]^{-1}\boldsymbol{l}_k$ is proven if $\eta_1 = \mathrm{E}|\mathfrak{h}_k^H \mathfrak{U} \mathfrak{h}_k - \boldsymbol{l}_k^H[\boldsymbol{C}^{(1)}(\sigma^2)]^{-1}\boldsymbol{l}_k|$ vanishes asymptotically, i.e.

$$\lim_{\substack{K,N\to\infty \\ \frac{K}{N}\to\beta}} \eta_1 = 0. \tag{D.21}$$

The rest of the proof is focused on showing (D.21). Let us observe

$$\eta_1 \leq \mathrm{E}|\mathfrak{h}_k^H \mathfrak{U} \mathfrak{h}_k - \mathfrak{h}_k^H \boldsymbol{V} \mathfrak{h}_k| + \mathrm{E}|\mathfrak{h}_k^H \boldsymbol{V} \mathfrak{h}_k - \boldsymbol{l}_k^H[\boldsymbol{C}^{(1)}(\sigma^2)]^{-1}\boldsymbol{l}_k|$$

where the triangular inequality of the spectral norm is applied and $\boldsymbol{V} = diag([\boldsymbol{\Phi}_{kk}^{(1)}(\sigma^2)]^{-1})_{k=1,\dots,N}$.

By applying the submultiplicative inequality for spectral norms (9) and the triangular inequality to the first term we obtain

$$\begin{aligned}
\mathrm{E}|\mathfrak{h}_k^H(\mathfrak{U}-\boldsymbol{V})\mathfrak{h}_k| &= \mathrm{E}|\sum_{i,\ell} s_{ik}^* \boldsymbol{l}_k^H(\mathfrak{U}-\boldsymbol{V})_{i\ell}\boldsymbol{l}_k s_{jk}| \\
&\leq \sum_{i,\ell} \mathrm{E}|\mathfrak{U}_{i\ell} - \boldsymbol{V}_{i\ell}|\boldsymbol{l}_k^H\boldsymbol{l}_k \mathrm{E}|s_{ik}^* s_{ik}| \\
&= \sum_i \mathrm{E}|\mathfrak{U}_{ii} - \boldsymbol{V}_{ii}|\frac{\boldsymbol{l}_k^H\boldsymbol{l}_k}{N} \\
&\leq \boldsymbol{l}_k^H\boldsymbol{l}_k \max_i \mathrm{E}|\mathfrak{U}_{ii} - \boldsymbol{V}_{ii}|.
\end{aligned} \tag{D.22}$$

Thanks to (D.19) and the fact that $\boldsymbol{l}_k^H\boldsymbol{l}_k < +\infty$

$$\lim_{\substack{K,N\to\infty \\ \frac{K}{N}\to\beta}} \mathrm{E}|\mathfrak{h}_k^H(\mathfrak{U}-\boldsymbol{V})\mathfrak{h}_k| = 0.$$

In order to prove the convergence to zero of $\eta_2 = \mathrm{E}|\mathfrak{h}_k^H \boldsymbol{V} \mathfrak{h}_k - \boldsymbol{l}_k^H[\boldsymbol{C}^{(1)}(\sigma^2)]^{-1}\boldsymbol{l}_k|$ we consider

$$\eta_2^2 \leq \mathrm{E}|\boldsymbol{h}_k^H\boldsymbol{V}\boldsymbol{h}_k - \boldsymbol{l}_k^H[\boldsymbol{C}^{(1)}(\sigma^2)]^{-1}\boldsymbol{l}_k|^2 \tag{D.23}$$

$$= \mathrm{E}((\mathfrak{h}_k^H\boldsymbol{V}\mathfrak{h}_k)^2 - 2\mathfrak{h}_k^H\boldsymbol{V}\mathfrak{h}_k\boldsymbol{l}_k^H[\boldsymbol{C}^{(1)}(\sigma^2)]^{-1}\boldsymbol{l}_k + \boldsymbol{l}_k^H[\boldsymbol{C}^{(1)}(\sigma^2)]^{-1}\boldsymbol{l}_k) \tag{D.24}$$

$$= \mathrm{E}\left(\sum_{ij}\boldsymbol{l}_k^H\boldsymbol{V}_{ii}\boldsymbol{l}_k\boldsymbol{l}_k^H\boldsymbol{V}_{jj}\boldsymbol{l}_k|s_{ik}|^2|s_{jk}|^2 - 2\boldsymbol{l}_k^H[\boldsymbol{C}^{(1)}(\sigma^2)]^{-1}\boldsymbol{l}_k\sum_i\boldsymbol{l}_k^H\boldsymbol{V}_{ii}\boldsymbol{l}_k|s_{ik}|^2 \right.$$

$$\left. +(\boldsymbol{l}_k^H[\boldsymbol{C}^{(1)}]^{-1}\boldsymbol{l}_k)^2\right) \tag{D.25}$$

$$= \sum_i(\boldsymbol{l}_k^H\boldsymbol{V}_{ii}\boldsymbol{l}_k)^2\frac{1}{N\gamma} + \sum_{\substack{i,j \\ i\neq j}}(\boldsymbol{l}_k^H\boldsymbol{V}_{ii}\boldsymbol{l}_k)(\boldsymbol{l}_k^H\boldsymbol{V}_{jj}\boldsymbol{l}_k)\frac{1}{N}$$

$$-\frac{2}{N}\boldsymbol{l}_k^H[\boldsymbol{C}^{(1)}(\sigma^2)]^{-1}\boldsymbol{l}_k\sum_i\boldsymbol{l}_k^H\boldsymbol{V}_{ii}\boldsymbol{l}_k + (\boldsymbol{l}_k^H[\boldsymbol{C}^{(1)}(\sigma^2)]^{-1}\boldsymbol{l}_k)^2. \tag{D.26}$$

From (D.25) to (D.26) we make use of the assumptions on the second and fourth moments of s_{ij}. Let us observe that the spectral norm of $[\boldsymbol{C}^{(1)}(\sigma^2)]^{-1}$ and $\boldsymbol{V}_{ii}$, for any i, are bounded by $|[\boldsymbol{C}^{(1)}(\sigma^2)]^{-1}| < \sigma^2$ and $|\boldsymbol{V}_{ii}| < \sigma^2$. Then, the first term in (D.26) vanishes as $N \to \infty$ since $\gamma > 1$. By applying (D.20), for any i, $\boldsymbol{V}_{ii} \to [\boldsymbol{C}^{(1)}(\sigma^2)]^{-1}$ as $K, N \to \infty$ with $\frac{K}{N} \to \beta$. Then, the second and third terms in (D.26) converge to $(\boldsymbol{l}_k^H[\boldsymbol{C}^{(1)}(\sigma^2)]^{-1}\boldsymbol{l}_k)^2$ and $-2(\boldsymbol{l}_k^H[\boldsymbol{C}^{(1)}(\sigma^2)]^{-1}\boldsymbol{l}_k)^2$, respectively. We can conclude that

$$\lim_{\substack{K,N\to\infty \\ \frac{K}{N}\to\beta}} \eta_2^2 = 0$$

and $\eta_2 \to 0$ as $K, N \to \infty$ as $\frac{K}{N} \to \beta$. Therefore, (D.21) is proven. The Markov inequality implies that, $\forall \varepsilon > 0$

$$\lim_{\substack{K,N\to\infty \\ \frac{K}{N}\to\beta}} \Pr\{|\mathfrak{h}_k^H \mathfrak{U} \mathfrak{h}_k - \boldsymbol{l}_k^H[\boldsymbol{C}^{(1)}(\sigma^2)]^{-1}\boldsymbol{l}_k|\} \le \frac{1}{\varepsilon} \lim_{\substack{K,N\to\infty \\ \frac{K}{N}\to\beta}} \mathrm{E}|\mathfrak{h}_k^H \mathfrak{U} \mathfrak{h}_k - \boldsymbol{l}_k^H[\boldsymbol{C}^{(1)}(\sigma^2)]^{-1}\boldsymbol{l}_k| = 0$$

and the convergence in probability stated in Theorem 11 is proven.

This concludes the proof of Theorem 11.

D.2 Proof of Theorem 12

Let us derive first the single user Bayesian filter. To this aim we calculate $\mathrm{E}\{\mathfrak{y}\mathfrak{y}^H\}^{-1}$ and $\mathrm{E}\{b_k^*\mathfrak{y}\}$ with the expectation taken over the noise, over all transmitted signals, and over the transmitted powers, channel gains, and spreading sequences of all interferers. Then, $\mathrm{E}\{\mathfrak{y}\mathfrak{y}^H\} = \mathrm{E}\{\mathfrak{H}_k\mathfrak{H}_k^H\} + \mathfrak{h}_k\mathfrak{h}_k^H + \sigma^2\boldsymbol{I}_{NL}$. Because of the independence and zero mean of the elements of the spreading sequences, $\mathrm{E}\{\mathfrak{H}_k\mathfrak{H}_k^H\}$ is a block diagonal matrix with N blocks of size $L \times L$. Each block is given by $(\beta - \frac{1}{N})\boldsymbol{C}_l$ with $\boldsymbol{C}_l = \mathrm{E}\{\boldsymbol{l}\boldsymbol{l}^H\}$. It follows

$$\mathfrak{C}_k = \mathrm{E}\{\mathfrak{H}_k\mathfrak{H}_k^H\} = \boldsymbol{I}_N \otimes \left(\beta - \frac{1}{N}\right)\boldsymbol{C}_l. \tag{D.27}$$

By applying the Sherman-Morrison formula (see Appendix E.1) to $(\mathrm{E}\{\mathfrak{y}\mathfrak{y}^H\})^{-1}$ we obtain

$$\begin{aligned}(\mathrm{E}\{\mathfrak{y}\mathfrak{y}^H\})^{-1} = {} & (\mathfrak{C}_k + \sigma^2\boldsymbol{I}_{NL})^{-1} - (\mathfrak{C}_k + \sigma^2\boldsymbol{I}_{NL})^{-1}\mathfrak{h}_k^H[1 + \mathfrak{h}_k^H(\mathfrak{C}_k + \sigma^2\boldsymbol{I}_{NL})^{-1}\mathfrak{h}_k] \\ & \times \mathfrak{h}_k^H(\mathfrak{C}_k + \sigma^2\boldsymbol{I}_{NL})^{-1}.\end{aligned}$$

Let us observe that $\mathrm{E}\{b_k^*\mathfrak{y}\} = \mathfrak{h}_k$. The single user Bayesian receiver is given by

$$\boldsymbol{c}_k = \frac{(\mathfrak{C}_k + \sigma^2\boldsymbol{I}_{NL})^{-1}\mathfrak{h}_k}{1 + \mathfrak{h}_k^H(\mathfrak{C}_k + \sigma^2\boldsymbol{I}_{NL})^{-1}\mathfrak{h}_k}. \tag{D.28}$$

The energy of the useful signal k at the output of the single user Bayesian filter is given by

$$\mathrm{E}\{|\boldsymbol{c}_k^H \mathfrak{h}_k b_k|^2\} = \left(\frac{\mathfrak{h}_k^H (\mathfrak{C}_k + \sigma^2 \boldsymbol{I}_{NL})^{-1} \mathfrak{h}_k}{1 + \mathfrak{h}_k^H (\mathfrak{C}_k + \sigma^2 \boldsymbol{I}_{NL})^{-1} \mathfrak{h}_k} \right)^2. \tag{D.29}$$

The energy of the noise at the output of the single user Bayesian filter is

$$\mathrm{E}\{|\boldsymbol{c}_k^H \mathfrak{n}|^2\} = \sigma^2 \frac{\mathfrak{h}_k^H (\mathfrak{C}_k + \sigma^2 \boldsymbol{I}_{NL})^{-2} \mathfrak{h}_k}{[1 + \mathfrak{h}_k^H (\mathfrak{C}_k + \sigma^2 \boldsymbol{I}_{NL})^{-1} \mathfrak{h}_k]^2}. \tag{D.30}$$

Finally, the energy of the interferers is

$$\begin{aligned}
\mathrm{E}\left\{ \sum_{\substack{j=1 \\ j \neq k}}^{K} |\boldsymbol{c}_k^H \mathfrak{h}_j b_j|^2 \right\} &= \mathrm{E}\left\{ \sum_{\substack{j=1 \\ j \neq k}}^{K} \boldsymbol{c}_k^H \mathfrak{h}_j \mathfrak{h}_j^H \boldsymbol{c}_k \right\} \\
&= \boldsymbol{c}_k^H \mathrm{E}\left\{ \sum_{\substack{j=1 \\ j \neq k}}^{K} \mathfrak{h}_j \mathfrak{h}_j^H \right\} \boldsymbol{c}_k \\
&= \boldsymbol{c}_k^H \mathrm{E}\{\mathfrak{H}_k \mathfrak{H}_k^H\} \boldsymbol{c}_k \\
&= \boldsymbol{c}_k^H \mathfrak{C}_k \boldsymbol{c}_k.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathrm{SINR}_k &= \frac{\mathrm{E}\{|\boldsymbol{c}_k^H \mathfrak{h}_k b_k|^2\}}{\mathrm{E}\{|\boldsymbol{c}_k^H \mathfrak{n}|^2\} + \mathrm{E}\left\{ \sum\limits_{\substack{j=1 \\ j \neq k}}^{K} |\boldsymbol{c}_k^H \mathfrak{h}_j b_j|^2 \right\}} \\
&= \frac{[\mathfrak{h}_k^H (\mathfrak{C}_k + \sigma^2 \boldsymbol{I}_{NL})^{-1} \mathfrak{h}_k]^2}{\sigma^2 \mathfrak{h}_k^H (\mathfrak{C}_k + \sigma^2 \boldsymbol{I}_{NL})^{-2} \mathfrak{h}_k + \mathfrak{h}_k^H (\mathfrak{C}_k + \sigma^2 \boldsymbol{I}_{NL})^{-1} \mathfrak{C}_k (\mathfrak{C}_k + \sigma^2 \boldsymbol{I}_{NL})^{-1} \mathfrak{h}_k} \\
&= \mathfrak{h}_k^H (\mathfrak{C}_k + \sigma^2 \boldsymbol{I}_{NL})^{-1} \mathfrak{h}_k \\
&= \boldsymbol{l}_k^H \left[\left(\beta - \frac{1}{N} \right) \boldsymbol{C}_l + \sigma^2 \boldsymbol{I}_L \right]^{-1} \boldsymbol{l}_k \sum_{n=1}^{N} s_{nk} s_{nk}^*.
\end{aligned} \tag{D.31}$$

Applying the strong law of large numbers to $\sum_{n=1}^{N} s_{nk} s_{nk}^*$, we obtain $\lim_{N \to \infty} \sum_{n=1}^{N} s_{nk} s_{nk}^* \overset{a.s.}{=} 1$. This limit and (D.31) yield the convergence of SNIR_k to $\boldsymbol{l}_k^H (\beta \boldsymbol{C}_l + \sigma^2 \boldsymbol{I}_L)^{-1} \boldsymbol{l}_k$ with probability 1 as $N \to \infty$.

D.3 Proof of Theorem 13

The proof of Theorem 13 follows the same lines as the proof of Theorem 12, taking into account that $\boldsymbol{c}_k = \mathfrak{h}_k$. Then,

$$\begin{aligned} \mathrm{SINR}_k &= \frac{\mathrm{E}\{|\mathfrak{h}_k^H \mathfrak{h}_k b_k|^2\}}{\mathrm{E}\{|\mathfrak{h}_k^H \mathfrak{n}|^2\} + \mathrm{E}\{\sum_{\substack{j=1\\j\neq k}}^{K} |\mathfrak{h}_k^H \mathfrak{h}_j b_j|^2\}} \\ &= \frac{(\mathfrak{h}_k^H \mathfrak{h}_k)^2}{\sigma^2 \mathfrak{h}_k^H \mathfrak{h}_k + \mathfrak{h}_k^H \mathfrak{C}_k \mathfrak{h}_k} \\ &= \frac{(\boldsymbol{l}_k^H \boldsymbol{l}_k \sum_{n=1}^N s_{nk} s_{nk}^*)^2}{\boldsymbol{l}_k^H \left[\left(\beta - \frac{1}{N}\right) \boldsymbol{C}_l + \sigma^2 \boldsymbol{I}_L\right] \boldsymbol{l}_k \sum_{n=1}^N s_{nk} s_{nk}^*} \end{aligned} \tag{D.32}$$

Applying the strong law of large numbers, we obtain the almost sure convergence of SINR_k as $K, N \to \infty$ with $\frac{K}{N} \to \beta$. More specifically, we obtain

$$\lim_{K=\beta N \to \infty} \mathrm{SNIR}_k \overset{a.s.}{=} \frac{(\boldsymbol{l}_k^H \boldsymbol{l}_k)^2}{\boldsymbol{l}_k^H (\beta \boldsymbol{C}_l + \sigma^2 \boldsymbol{I}_L) \boldsymbol{l}_k}.$$

D.4 Proof of Theorem 14

The proof of Theorem 14 follows the lines of the proof of Theorem 10. It is useful to rewrite $\mathfrak{H}$ with a structure similar to the structure of $\widetilde{\boldsymbol{H}}$ in (4.31).

Let us define for each user k a block diagonal matrix $\mathfrak{L}_k = \boldsymbol{I}_N \otimes \boldsymbol{l}_k$ of dimensions $LN \times N$, with identical diagonal blocks equal to $\boldsymbol{l}_k$. The matrix $\mathfrak{H}$ can be rewritten as

$$\begin{aligned} \mathfrak{H} &= (\mathfrak{L}_1 \boldsymbol{s}_1, \mathfrak{L}_2 \boldsymbol{s}_2, \ldots, \mathfrak{L}_K \boldsymbol{s}_K) \\ &= (\mathfrak{h}_1, \mathfrak{h}_2, \ldots, \mathfrak{h}_K). \end{aligned}$$

Additionally, we introduce the following notation. For $n = 1, \ldots N$ and $k = 1, \ldots, K$

- $\boldsymbol{\delta}_n$ is the n^{th} row block of $\mathfrak{H}$ of dimensions $L \times K$.
- $\mathfrak{H}_{\models n}$ is the matrix obtained from $\mathfrak{H}$ by suppressing the n^{th} row block $\boldsymbol{\delta}_n$.
- $\mathfrak{R}_{\models n} = \mathfrak{H}_{\models n}^H \mathfrak{H}_{\models n}$.
- $\boldsymbol{\sigma}_n$ is the n^{th} row of the matrix $\boldsymbol{S}$, i.e. $\boldsymbol{\sigma}_n = (s_{n1}, s_{n2}, \ldots s_{nN})$.
- $\boldsymbol{L}_\ell$, $\ell = 1, \ldots, L$ is a $K \times K$ diagonal matrix with k^{th} element equal to $l_{\ell k}$. Note that $\boldsymbol{\sigma}_n \boldsymbol{L}_\ell$ coincides with the $((n-1)L + \ell)^{\text{th}}$ row of the matrix $\mathfrak{H}$.

- $\mathfrak{T}^s_{nn}$ is the n^{th} diagonal block of $\mathfrak{T}^s$ of dimensions $L \times L$.
- $\mathfrak{T}_{\sim k} = \mathfrak{H}_k^H \mathfrak{H}_k$.

The proof of Theorem 14 is based on strong induction. First, we prove the following:

- For a given user k, conditionally on $\boldsymbol{l}_k$, $\mathfrak{R}_{kk}$ converges almost surely to a deterministic limit $\mathfrak{R}^1(\boldsymbol{l}_k)$ as $K, N \to \infty$ with $\frac{K}{N} \to \beta$.
- The block matrix $\mathfrak{T}^1_{nn}$ converges almost surely to a deterministic limiting matrix $\mathfrak{Z}$ independent of n as $K, N \to \infty$ with $\frac{K}{N} \to \beta$.

Then, in the recursion step, we use the following induction assumptions:

- For $s = 0, \ldots, \ell - 1$, $(\mathfrak{R}^s)_{kk}$ converges to a deterministic limit $\mathfrak{R}^s(\boldsymbol{l}_k)$ and the diagonal blocks $(\mathfrak{T}^s)_{nn}$ converge to a deterministic limiting matrix $\mathfrak{Z}^s$ independent of n as $K, N \to \infty$ with $\frac{K}{N} \to \beta$.
- For $s = 0, \ldots, \ell - 1$, the limit values $\mathfrak{R}^s(\boldsymbol{l}_k)$ and $\mathfrak{Z}^s$ are given by the recursive equations (5.26) and (5.27), respectively.

We prove that $(\mathfrak{R}^\ell)_{kk}$ converges to a deterministic limit $\mathfrak{R}^\ell(\boldsymbol{l})|_{\boldsymbol{l}=\boldsymbol{l}_k}$ and $(\mathfrak{T}^\ell)_{nn}$ converges to a deterministic limiting matrix independent of n as $K, N \to \infty$ with $\frac{K}{N} \to \beta$. Furthermore, we prove that these limit values satisfy also the recursive equations (5.26) and (5.27).

First step: $\mathfrak{R}_{kk} = \boldsymbol{s}_k^H \mathfrak{L}_k^H \mathfrak{L}_k \boldsymbol{s}_k$ and the diagonal elements of the matrix $\mathfrak{L}_k^H \mathfrak{L}_k$ are bounded thanks to the condition on the distribution function F_l. Thus, appealing to Lemma 2 we can show that, given $\boldsymbol{l}_k$, $\mathfrak{R}_{kk}$ converges almost surely to the deterministic value

$$\begin{aligned}
\mathfrak{R}^1(\boldsymbol{l}_k) &= \lim_{K=\beta N \to \infty} \mathfrak{R}_{kk} \\
&= \lim_{K=\beta N \to \infty} \boldsymbol{s}_k^H \mathfrak{L}_k^H \mathfrak{L}_k \boldsymbol{s}_k \\
&= \lim_{K=\beta N \to \infty} \frac{1}{N} \mathrm{tr}(\mathfrak{L}_k^H \mathfrak{L}_k) = \boldsymbol{l}_k^H \boldsymbol{l}_k.
\end{aligned} \tag{D.33}$$

We apply a similar argument to the element (u, v) of the bock matrix $\mathfrak{T}_{nn}$ to show its almost sure convergence to a deterministic value

$$\begin{aligned}
\mathfrak{Z}_{uv} &= \lim_{K=\beta N \to \infty} (\mathfrak{T}_{nn})_{uv} \\
&= \lim_{K=\beta N \to \infty} \boldsymbol{\sigma}_n \boldsymbol{L}_u \boldsymbol{L}_v^H \boldsymbol{\sigma}_n^H \\
&= \lim_{K=\beta N \to \infty} \frac{1}{N} \mathrm{tr}(\boldsymbol{L}_u \boldsymbol{L}_v^H) \\
&= \beta \mathrm{E}\{l_u l_v^*\}.
\end{aligned}$$

Equivalently,

$$\mathbf{\mathfrak{Z}} = \lim_{K=\beta N\to\infty} \mathbf{\mathfrak{T}}_{nn} = \beta \mathrm{E}\{\boldsymbol{l}\boldsymbol{l}^H\}.$$

ℓ^{th} *step:* Following the same approach as in the proof of Theorem 6, Appendix B, we can expand $(\mathbf{\mathfrak{R}}^\ell)_{kk}$ and the $L \times L$ diagonal block of the matrix $\mathbf{\mathfrak{T}}^\ell$, $\mathbf{\mathfrak{T}}^\ell_{nn}$ as follows:

$$(\mathbf{\mathfrak{R}}^\ell)_{kk} = \sum_{s=0}^{\ell-1} \mathbf{\mathfrak{h}}_k^H \mathbf{\mathfrak{T}}_{\sim k}^{\ell-s-1} \mathbf{\mathfrak{h}}_k (\mathbf{\mathfrak{R}}^s)_{kk} \tag{D.34}$$

$$\mathbf{\mathfrak{T}}^\ell_{nn} = \sum_{s=0}^{\ell-1} \boldsymbol{\delta}_n \mathbf{\mathfrak{R}}_{\vDash n}^{\ell-s-1} \boldsymbol{\delta}_n^H \mathbf{\mathfrak{T}}^s_{nn}. \tag{D.35}$$

Thanks to the assumptions of the strong induction, for $s = 1, \ldots, \ell-1$, $(\mathbf{\mathfrak{R}}^s)_{kk}$ and $\mathbf{\mathfrak{T}}^s_{nn}$ converge almost surely to the deterministic limits $\mathfrak{R}^s(\boldsymbol{l}_k)$ and $\mathbf{\mathfrak{Z}}^s$, respectively. Therefore, the almost sure convergence of $(\mathbf{\mathfrak{R}}^\ell)_{kk}$ and $\mathbf{\mathfrak{T}}^\ell_{nn}$ reduces to the almost sure convergence of $\mathbf{\mathfrak{h}}_k^H \mathbf{\mathfrak{T}}_{\sim k}^{\ell-s-1} \mathbf{\mathfrak{h}}_k$ and $\boldsymbol{\delta}_n \mathbf{\mathfrak{R}}_{\vDash n}^{\ell-s-1} \boldsymbol{\delta}_n^H$, respectively.

Let us define

$$\begin{aligned} g(\mathbf{\mathfrak{Z}}^s, \boldsymbol{l}_k) &\triangleq \lim_{K=\beta N\to\infty} \mathbf{\mathfrak{h}}_k^H \mathbf{\mathfrak{T}}^s_{\sim k} \mathbf{\mathfrak{h}}_k \\ &= \lim_{K=\beta N\to\infty} \boldsymbol{s}_k^H \mathbf{\mathfrak{L}}_k^H \mathbf{\mathfrak{T}}^s_{\sim k} \mathbf{\mathfrak{L}}_k \boldsymbol{s}_k. \end{aligned} \tag{D.36}$$

In a similar way, we denote

$$\boldsymbol{f}(\mathfrak{R}^s) \triangleq \lim_{K=\beta N\to\infty} \boldsymbol{\delta} \mathbf{\mathfrak{R}}^s_{\vDash n} \boldsymbol{\delta}^H.$$

The $(u,v)^{th}$ element of $\boldsymbol{f}(\mathfrak{R}^s)$ is given by

$$\begin{aligned} (\boldsymbol{f}(\mathfrak{R}^s))_{uv} &\triangleq \lim_{K=\beta N\to\infty} (\boldsymbol{\delta} \mathbf{\mathfrak{R}}^s_{\vDash n} \boldsymbol{\delta}^H)_{uv} \\ &= \lim_{K=\beta N\to\infty} \boldsymbol{\sigma}_u \boldsymbol{\Lambda}_u \mathbf{\mathfrak{R}}^s_{\vDash n} \boldsymbol{\Lambda}_v^H \boldsymbol{\sigma}_v^H. \end{aligned} \tag{D.37}$$

Let us compute the limits (D.38) and (D.39). The condition on the spectral radius of $\mathbf{\mathfrak{R}}$ guarantees that the spectral radius of $\mathbf{\mathfrak{R}}^u$ and $\mathbf{\mathfrak{T}}^u$ is upper bounded for any finite integer u. Applying the submultiplicative inequality of spectral norms (see Lemma 9 in Appendix E.1) the same property holds for the matrix $\mathbf{\mathfrak{L}}_k^H \mathbf{\mathfrak{T}}^s_{\sim k} \mathbf{\mathfrak{L}}_k$ and its powers and the matrix $\boldsymbol{\Lambda}_u \mathbf{\mathfrak{R}}^s_{\vDash n} \boldsymbol{\Lambda}_u^H$ and its powers. Then, we can apply Lemma 2 to the right hand side of (D.36) and we obtain the following almost sure convergence

$$\begin{aligned} g(\mathbf{\mathfrak{Z}}^s, \boldsymbol{l}_k) &\overset{a.s.}{=} \lim_{K=\beta N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \boldsymbol{l}_k^H \mathbf{\mathfrak{T}}^s_{nn} \boldsymbol{l}_k \\ &\overset{a.s.}{=} \boldsymbol{l}_k^H \mathbf{\mathfrak{Z}}^s \boldsymbol{l}_k. \end{aligned} \tag{D.38}$$

Applying Lemma 2 in Section 2.4 to the right hand side of (D.37) we obtain the almost sure convergence

$$\begin{aligned}
(\boldsymbol{f}(\mathfrak{R}^s))_{uv} &\overset{a.s.}{=} \lim_{K=\beta N\to\infty} \frac{\beta}{K}\mathrm{tr}(\boldsymbol{\Lambda}_u \boldsymbol{\mathfrak{R}}^s_{\models n}\boldsymbol{\Lambda}_v^H) \\
&\overset{a.s.}{=} \lim_{K=\beta N\to\infty} \frac{\beta}{K}\sum_{k=1}^{K} l_{uk}l_{vk}^*(\boldsymbol{\mathfrak{R}}^s_{\models n})_{kk} \\
&\overset{a.s.}{=} \beta\mathrm{E}\{l_u l_v^* \mathfrak{R}^s(\boldsymbol{l})\}.
\end{aligned}$$

Thus, the matrix $\boldsymbol{f}(\mathfrak{R}^s)$ is given by

$$\boldsymbol{f}(\mathfrak{R}^s) = \beta\mathrm{E}\{\boldsymbol{l}\boldsymbol{l}^H \mathfrak{R}^s(\boldsymbol{l})\}. \tag{D.39}$$

Finally, the limit (D.38) and the recursive equation (D.34) yield

$$\begin{aligned}
\mathfrak{R}^\ell(\boldsymbol{l}) &= \lim_{K=\beta N\to\infty} (\boldsymbol{\mathfrak{R}}^\ell)_{kk} \\
&\overset{a.s.}{=} \sum_{s=0}^{\ell-1} g(\boldsymbol{\mathfrak{Z}}^{\ell-s-1}, \boldsymbol{l})\mathfrak{R}^s(\boldsymbol{l}).
\end{aligned}$$

In a similar way, using the limit (D.39) and the recursive equation (D.35) we obtain

$$\begin{aligned}
\boldsymbol{\mathfrak{Z}}^\ell &= \lim_{K=\beta N\to\infty} \boldsymbol{\mathfrak{T}}^\ell_{nn} \\
&\overset{a.s.}{=} \sum_{s=0}^{\ell-1} \beta\mathrm{E}\{\mathfrak{R}^{\ell-s-1}(\boldsymbol{l})\boldsymbol{l}\boldsymbol{l}^H\}\boldsymbol{\mathfrak{Z}}^s.
\end{aligned}$$

Thus, the induction step is proven and this concludes the proof of Theorem 14.

E Mathematical Tools

E.1 Linear Algebra

Lemma 8 *[111] If $\boldsymbol{B}$ is an Hermitian matrix, then for any vector $\boldsymbol{x}$*

$$|\boldsymbol{x}^H \boldsymbol{B}\boldsymbol{x}| \leq \max\{|\lambda| : \lambda \textit{ is an eigenvalue of } \boldsymbol{B}\}\|\boldsymbol{x}\|^2. \tag{E.1}$$

Theorem 16 ***(Interlacing Eigenvalue Theorem for Bordered Matrices)*** *[111] Let $\boldsymbol{A}$ be a given $n \times n$ Hermitian matrix, let $\boldsymbol{y} \in \mathbb{C}^n$ be a given column vector, and let $a \in \mathbb{R}$ be a given real number. Let $\widehat{\boldsymbol{A}}$ be the $(n+1) \times (n+1)$ Hermitian matrix obtained by bordering $\boldsymbol{A}$ with $\boldsymbol{y}$ and a as follows:*

$$\widehat{\boldsymbol{A}} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{y} \\ \boldsymbol{y}^H & a \end{bmatrix}.$$

Let the eigenvalues of $\boldsymbol{A}$ and $\widehat{\boldsymbol{A}}$ be denoted by $\{\lambda_i\}$ and $\{\widehat{\lambda}_i\}$, respectively, and assume that they have been arranged in increasing order, i.e., $\lambda_1 \leq \lambda_2 \ldots \leq \lambda_n$ and $\widehat{\lambda}_1 \leq \widehat{\lambda}_2 \ldots \leq \widehat{\lambda}_{n+1}$. Then

$$\widehat{\lambda}_1 \leq \lambda_1 \leq \widehat{\lambda}_2 \leq \lambda_2 \ldots \leq \lambda_n \leq \widehat{\lambda}_{n+1}.$$

Theorem 17 ***(Sherman-Morrison Formula)*** *[111] Let $\boldsymbol{u}$ and $\boldsymbol{v}$ be n-dimensional column vectors and let $\boldsymbol{A}$ be an $n \times n$ matrix. Then,*

$$(\boldsymbol{A} + \boldsymbol{u}\boldsymbol{v}^H)^{-1} = \boldsymbol{A}^{-1} - \frac{\boldsymbol{A}^{-1}\boldsymbol{u}\boldsymbol{v}^H\boldsymbol{A}^{-1}}{1 + \boldsymbol{v}\boldsymbol{A}^{-1}\boldsymbol{u}}.$$

Lemma 9 ***(Submultiplicative Inequality of Spectral Norm)*** *[111] Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be two $n \times n$ matrices, then*

$$|\boldsymbol{A}\boldsymbol{B}| \leq |\boldsymbol{A}||\boldsymbol{B}|$$

E.2 Probability Theory

Lemma 10 ***(Markov Inequality)*** *[112] Let x be a nonnegative random variable with $\mathrm{E}\{x\} = \eta$, then, for any $\epsilon > 0$*

$$Pr\{x > \epsilon\} \leq \frac{\eta}{\epsilon}. \tag{E.2}$$

Lemma 11 *(**Bienaymé Inequality**) [112] Let x be an arbitrary random and let a and n be two arbitrary numbers. Then, for any $\epsilon > 0$*

$$Pr\{|x-a| \geq \epsilon\} \leq \frac{\mathrm{E}\{|x-a|^n\}}{\epsilon^n}. \tag{E.3}$$

Lemma 12 *(**Lyapunov Inequality**) [112] Let $\theta_k = \mathrm{E}\{|x|^k\} < +\infty$ represent the absolute moments of the random variable x. Then,*

$$\theta_{k-1}^{\frac{1}{k-1}} \leq \theta_k^{\frac{1}{k}} \qquad k \geq 1. \tag{E.4}$$

Lemma 13 *(**Borel-Cantelli Lemma**) [108] [112] Let $A_1, A_2, \ldots$ be an infinite sequence of events, each of which depends only on a finite number of trials. In other words, there exists an integer n_k such that A_k is an event in the sample space of the first n_k Bernoulli trials. Put $p_k = Pr\{A_k\}$, $k = 1, 2, \ldots$.*

(i) Suppose

$$\sum_{k=1}^{+\infty} p_k < +\infty$$

that is, the series on the left converges, Then, with probability one only finitely many of the events $A_1, A_2, \ldots,$ occur. More precisely, to every $\epsilon > 0$ there is an integer r such that the probability that n trials produce one or more among the events $A_{r+1}, A_{r+2}, \ldots$ is less than ϵ for all n.

(ii) If the events A_k are mutually independent, and if $\sum_{k=1}^{+\infty} p_k$ diverges, then, with probability one infinitely many A_k occur. In other words, for every r the probability that n trials produce more than r among the events A_k tends to one as $n \to \infty$.

Bibliography

[1] N. P. Negroponte, *Being Digital.* New York: Random House, Inc., January 1996, vol. 1.

[2] S. Verdú, "Optimum sequence detection of asynchronous multiple-access communications," in *Proc. of IEEE International Symposium on Information Theory (ISIT)*, Sept. 1983, p. 80.

[3] ——, "Minimum probability of error for asynchronous multiple access communications systems," in *Proc. of Military Communications Conference*, Nov. 1983, pp. 1:213–219.

[4] ——, "Optimal multiuser signal detection," Ph.D. dissertation, University of Illinois at Urbana-Champaign, Urbana-Champaign, Illinois, aug 1994.

[5] ——, "Demodulation in presence of multiaccess interference: Progress and misconceptions," in *Intelligent Methods in Signal Processing and Communications*, D. Docampo, A. Figueiras, and F.Perez, Eds. Boston: Birkhauser, 1997, ch. 2, pp. 15–46.

[6] G. Caire, R. R. Müller, and T. Tanaka, "Iterative multiuser joint decoding: Optimal power allocation and low-complexity implementation," *IEEE Transactions on Information Theory*, vol. 50, no. 9, pp. 1950–1973, Sept. 2004.

[7] G. Foschini and M. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311–335, 1998.

[8] I. E. Telatar, "Capacity of multi–antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, Nov./Dec. 1999.

[9] S. V. Hanly and D. N. Tse, "Resource pooling and effective bandwidth in CDMA networks with multiuser receivers and spatial diversity," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1328–1351, May 2001.

[10] D. Tse and S. Hanly, "Linear multiuser receivers: Effective interference, effective bandwidth and user capacity," *IEEE Transactions on Information Theory*, vol. 45, no. 2, pp. 641–657, Mar. 1999.

[11] S. Verdú and S. Shamai (Shitz), "Spectral efficiency of CDMA with random spreading," *IEEE Transactions on Information Theory*, vol. 45, no. 2, pp. 622–640, Mar. 1999.

[12] M. Debbah, W. Hachem, P. Loubaton, and M. d. Courville, "MMSE analysis of certain large isometric random precoded systems," *IEEE Transactions on Information Theory*, vol. 49, no. 5, May 2003.

[13] L. Li, A. Tulino, and S. Verdú, "Design of reduced-rank MMSE multiuser detectors using random matrix methods," *IEEE Transactions on Information Theory*, vol. 50, no. 6, pp. 986 – 1008, June 2004.

[14] W. Hachem, "Simple polynomial detectors for CDMA downlink transmissions on frequency-selective channels," *IEEE Transactions on Information Theory*, vol. 50, no. 1, pp. 164–172, Jan. 2004.

[15] S. Verdú, "Minimum probability of error for asynchronous Gaussian multiple–access channels," *IEEE Transactions on Information Theory*, vol. IT–32, no. 1, pp. 85–96, Jan. 1986.

[16] ——, "Multiuser detection," in *Advances in Statistical Signal Processing: Signal Detection*, H. Poor and J. Thomas, Eds. Greenwich, CT: JAI Press, 1993, pp. 369–410.

[17] R. R. Müller, *Power and Bandwidth Efficiency of Multiuser Systems with Random Spreading.* Aachen, Germany: Shaker–Verlag, 1999.

[18] R. R. Müller and S. Verdú, "Design and analysis of low–complexity interference mitigation on vector channels," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 8, pp. 1429–1441, Aug. 2001.

[19] Z. Xie, R. T. Short, and C. K. Rushforth, "A family of suboptimum detectors for coherent multiuser communications," *IEEE Journal on Selected Areas in Communications*, vol. 8, no. 4, pp. 683–690, May 1990.

[20] S. Kay, *Fundamentals of Statistical Signal Processing, Estimation Theory*, ser. Prentice Hall Signal Processing Series. Prentice Hall, 1993, vol. 1.

[21] L. W. Brooks and I. S. Reed, "Equivalence of the likelihood ratio processor, the maximum signal–to–noise ratio filter, and the Wiener filter," *IEEE Transactions on Aerospace and Electronic Systems*, pp. 690–692, Sept. 1972.

[22] S. Buzzi, M. Lops, and A. M. Tulino, "A new family of MMSE multiuser receivers for interference suppression in DS/CDMA systems employing BPSK modulation," *IEEE Transactions on Communications*, vol. 49, no. 1, pp. 154–167, Jan. 2001.

[23] D. Divsalar and M. K. Simon, "Improved CDMA performance using parallel interference cancellation," in *Proc. of Military Communications Conference*, Oct. 1994, pp. 911–917.

[24] D. Divsalar, M. K. Simon, and D. Raphaeli, "Improved parallel interference cancellation for CDMA," *IEEE Transactions on Communications*, vol. 46, no. 2, pp. 258–268, Feb. 1998.

[25] M. K. Varanasi and B. Aazhang, "Multistage detection in asynchronous code–division multiple–access communications," *IEEE Transactions on Communications*, vol. COM–38, no. 4, pp. 509–519, Apr. 1990.

[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, Jan. 1996.

[27] J. S. Goldstein, I. S. Reed, and L. L. Scharf, "A multistage representation of the Wiener filter based on orthogonal projections," *IEEE Transactions on Information Theory*, vol. 44, no. 7, Nov. 1998.

[28] M. Honig and W. Xiao, "Performance of reduced–rank linear interference suppression," *IEEE Transactions on Information Theory*, vol. 47, no. 5, pp. 1928–1946, July 2001.

[29] W. Xiao, "Large system performance of linear interference suppression for DS-CDMA," Ph.D. dissertation, Northwestern University, Evanston, Illinois, dec 2001.

[30] J. S. Goldstein and I. S. Reed, "Reduced rank adaptive filtering," *IEEE Transactions on Signal Processing*, vol. 42, no. 2, pp. 492–496, Feb. 1997.

[31] H. V. Poor and S. Verdú, "Single-user detectors for multiuser channels," *IEEE Transactions on Communications*, vol. 36, no. 1, pp. 50–60, Jan. 1988.

[32] D. Tse and O. Zeitouni, "Linear multiuser receivers in random environments," *IEEE Transactions on Information Theory*, vol. 46, no. 1, pp. 171–188, Jan. 2000.

[33] S. Shamai (Shitz) and S. Verdú, "The impact of frequency–flat fading on the spectral efficiency of CDMA," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1302–1327, May 2001.

[34] L. G. F. Trichard, J. S. Evans, and I. B. Collings, "Large system analysis of linear multistage parallel interference cancellation," *IEEE Transactions on Communications*, vol. 50, no. 11, pp. 1778–1786, Nov. 2002.

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

[36] W. Chen, U. Mitra, and P. Schniter, "On the equivalence of three reduced rank linear estimators with applications to DS-CDMA," *IEEE Transactions on Information Theory*, vol. 48, no. 9, pp. 2609–2614, Sept. 2002.

[37] L. Li, A. Tulino, and S. Verdú, "Asymptotic eigenvalue moments for linear multiuser detection," *Communications in Information and Systems*, vol. 1, no. 3, pp. 273–304, Fall 2001.

[38] L. G. F. Trichard, J. S. Evans, and I. B. Collings, "Optimal multistage linear multiuser receiver," *IEEE Transactions on Wireless Communications*, vol. 4, no. 3, pp. 1092–1101, May 2005.

[39] P. Loubaton and W. Hachem, "Asymptotic analysis of reduced rank Wiener filters," in *Proc. of IEEE Information Theory Workshop (ITW)*, Paris, France, Apr. 2003, pp. 328–331.

[40] J. W. Silverstein and P. L. Combettes, "Signal detection via spectral theory of large dimensional random matrices," *IEEE Transactions on Signal Processing*, vol. 40, no. 8, pp. 2100–2105, Aug. 1992.

[41] S. Verdú and S. Shamai (Shitz), "Multiuser detection with random spreading and error–correction codes: Fundamental limits," in *Proc. of 35th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Sept./Oct. 1997, pp. 470–482.

[42] S. Verdú, "Multiple–access channels with memory with and without frame synchronism," *IEEE Transactions on Information Theory*, vol. 35, no. 3, pp. 605–619, May 1989.

[43] D. Tse and S. Hanly, "Multiuser demodulation: Effective interference, effective bandwidth and capacity," in *Proc. of 35th Annual Allerton Conference on Communication, Control and Computing*, Monticello, IL, Sept./Oct. 1997, pp. 281–290.

[44] Z. D. Bai, "Methodologies in spectral analysis of large dimensional random matrices, a review," *Statistica Sinica*, vol. 9, no. 3, pp. 611–677, July 1999.

[45] P. Biane, "Free probability for probabilists," Lecture Notes, Paris, France, 2000.

[46] R. R. Müller, "Applications of large random matrices in communications engineering," in *Invited for International Conference on Advances in the Internet, Processing, Systems, and Interdisciplinary Research (IPSI)*, Sveti Stefan, Montenegro, Oct. 2003.

[47] M. L. Mehta, *Random Matrices*, 2nd ed. Boston, MA: Academic Press, 1991.

[48] V. L. Girko, *Theory of Random Determinants.* Dordrecht, The Netherlands: Kluwer Academic Publishers, 1990.

[49] D. V. Voiculescu, K. J. Dykema, and A. Nica, *Free Random Variables.* Providence, RI: American Mathematical Society, 1992.

[50] F. Hiai and D. Petz, *The Semicircle Law, Free Random Variables and Entropy.* Providence, RI: American Mathematical Society, 2000.

[51] A. Tulino and S. Verdú, *Random Matrix Theory and Wireless Communications*, ser. Foundations and Trends in Communications and Information Theory. now, June 2004, vol. 1.

[52] R. R. Müller and S. Verdú, "Spectral efficiency of low–complexity multiuser detection," in *Proc. of IEEE International Symposium on Information Theory (ISIT)*, Sorrento, Italy, June 2000, p. 439.

[53] T. J. Stieltjes, "Recherches sur les fractions continues," *Annales de la Faculté des Sciences de Toulouse*, vol. 8, pp. 1–122, 1894.

[54] V. A. Marčenko and L. A. Pastur, "Distribution of eigenvalues for some sets of random matrices," *Mathematics of the U.S.S.R — Sbornik*, vol. 1, no. 4, pp. 457–483, 1967.

[55] J. W. Silverstein and Z. Bai, "On the empirical distribution of eigenvalues of a class of large dimensional random matrices," *Journal of Multivariate Analysis*, vol. 54, pp. 175–192, 1995.

[56] M. L. Honig and W. Xiao, "Large system performance of reduced–rank linear filters," in *Proc. of 37th Annual Allerton Conference on Communication, Control and Computing*, Monticello, IL, Sept. 1999.

[57] D. Guo, S. Verdú, and L. K. Rasmussen, "Asymptotic normality of linear multiuser receiver outputs," *IEEE Transactions on Information Theory*, vol. 48, no. 12, pp. 3080–3095, Dec. 2002.

[58] J. Zhang, E. K. Chong, and D. N. Tse, "Output MAI distributions of linear MMSE multiuser receivers in DS–CDMA systems," *IEEE Transactions on Information Theory*, vol. 47, no. 3, pp. 1128–1144, Mar. 2001.

[59] J. Zhang and E. K. P. Chong, "Linear MMSE multiuser receivers: MAI conditional weak convergence and network capacity," *IEEE Transactions on Information Theory*, vol. 48, no. 7, pp. 2114–2122, July 2002.

[60] J. Zhang and X. Wang, "Large-system performance analysis of blind and group-blind multiuser receivers," *IEEE Transactions on Information Theory*, vol. 48, no. 9, pp. 2507–2523, Sept. 2002.

[61] V. L. Girko, *Theory of Stochastic Canonical Equations.* Dordrecht, The Netherlands: Kluwer Academic Publishers, 2001, vol. 1.

[62] L. Cottatellucci and M. Debbah, "On the capacity of MIMO Rice channels," in *Proc. 42nd Allerton Conf. on Communication, Control and Computing*, Monticello, Illinois, Sept./Oct. 2004.

[63] ——, "The effects of line of sight components on the capacity of MIMO systems," *Submitted to IEEE Transactions on Information Theory*, July 2004, to be revised.

[64] Z. D. Bai and J. W. Silverstein, "No eigenvalues outside the support of the limiting spectral distribution of large dimensional sample covariance matrices," *The Annals of Probability*, vol. 26, no. 1, pp. 316–345, 1998.

[65] J. Evans and D. N. Tse, "Large system performance of linear multiuser receivers in multipath fading channels," *IEEE Transactions on Information Theory*, vol. 46, no. 6, pp. 2059–2078, Sept. 2000.

[66] L. Cottatellucci and R. R. Müller, "Performance assessment of polynomial detector in MIMO channels with measured data," in *Proc. of IEE Technical Seminar on MIMO Communication Systems*, London, U.K., Dec. 2001, pp. 5/1–5/6.

[67] ——, "Asymptotic design and analysis of multistage detectors with unequal powers," in *Proc. of Information Theory Workshop*, Bangalore, India, Oct. 2002.

[68] ——, "Multiuser interference mitigation with multistage detectors: Design and analysis for unequal powers," in *Proc. of Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2002, pp. 1948–1952.

[69] ——, "Asymptotic analysis of multistage receivers for multipath fading channels," in *Proc. of IEEE International Symposium on Information Theory (ISIT)*, Yokohama, Japan, June/July 2003, p. 509.

[70] W. Hachem, "Low complexity polynomial receivers for downlink CDMA," in *Proc. of Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2002, pp. 1919–1923.

[71] C. Dumard and T. Zemen, "Double Krylov subspace approximation for low complexity iterative multi-user decoding and time-variant channel estimation," in *Proc. of Sixth IEEE Workshop on Signal Processing Advances in Wireless Communications*, New York, USA, June 2005, pp. 346–350.

[72] D. Segovia-Vargas, F. Iñigo, and M. Sierra-Pérez, "Generalized eigenspace beamformer based on CG-lanczos algorithm," *IEEE Transactions on Antennas and Propagation*, vol. 51, no. 8, pp. 2146–2154, Aug. 2003.

[73] X. Wang and H. V. Poor, "Blind multiuser detection: a subspace approach," *IEEE Transactions on Information Theory*, vol. 44, no. 3, pp. 677–690, Mar. 1998.

[74] M. L. Honig, "A comparison of adaptive filtering techniques for DS-CDMAinterference suppression," in *Proc. of Military Communications Conference*, Monterey, CA, Nov. 1997, pp. 836–840.

[75] M. L. Honig and J. S. Goldstein, "Adaptive reduced-rank residual correlation algorithms for DS-CDMA interference suppression," in *Proc. of 32nd Asilomar Conference on Signals, Systems and Computing*, Pacific Grove, CA, Nov. 1998.

[76] U. Madhow and M. L. Honig, "MMSE interference suppression for direct–sequence spread–spectrum CDMA," *IEEE Transactions on Communications*, vol. 42, no. 12, pp. 3178–3188, Dec. 1994.

[77] Y. Q. Yin, "Limiting spectral distribution for a class of random matrices," *Journal of Multivariate Analysis*, vol. 20, pp. 50–68, 1986.

[78] M. Rupf, F. Tarköy, and J. L. Massey, "User–separating demodulation for code–division multiple–access systems," *IEEE Journal on Selected Areas in Communications*, vol. 12, no. 5, pp. 786–795, June 1994.

[79] U. Madhow and M. L. Honig, "On the average near–far resistance for MMSE detection of direct sequence CDMA signals with random spreading," *IEEE Transactions on Information Theory*, vol. 45, no. 6, pp. 2039–2045, Sept. 1999.

[80] P. Schramm and R. R. Müller, "Spectral efficiency of CDMA systems with linear MMSE interference suppression," *IEEE Transactions on Communications*, vol. 47, no. 5, pp. 722–731, May 1999.

[81] Kiran and D. N. Tse, "Effective interference and effective bandwidth of linear multiuser receivers in asynchronous CDMA systems," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1426–1447, Sept. 2000.

[82] A. Mantravadi and V. V. Veeravalli, "MMSE detection in asynchronous CDMA systems: An equivalence result," *IEEE Transactions on Information Theory*, vol. 48, no. 12, pp. 3128–3137, Dec. 2002.

[83] R. Lupas and S. Verdú, "Near–far resistance of multiuser detectors in asynchronous channels," *IEEE Transactions on Communications*, vol. COM–38, no. 4, pp. 496–508, Apr. 1990.

[84] M. L. Honig, U. Madhow, and S. Verdú, "Blind adaptive interference suppression for near–far resistant CDMA," in *Proc. of IEEE Global Telecommunications Conference (Globecom)*, San Francisco, CA, 1994, pp. 379–384.

[85] A. Mantravadi and V. V. Veeravalli, "On chip-matched filtering and discrete sufficient statistics for asynchronous band-limited CDMA systems," *IEEE Transactions on Communications*, vol. 49, no. 8, pp. 1457–1467, Aug. 2001.

[86] M. A. Landolsi and W. E. Stark, "DS-CDMA chip waveform design for optimal power-bandwidth performance," in *Proc. of 6th International Symposium Personal, Indoor and Mobile Radio Communications*, Toronto, ON, Sept. 1995, pp. 706–710.

[87] ——, "DS-CDMA chip waveform design for minimal interference under bandwidth , phase and envelope constraints," *IEEE Transactions on Communications*, vol. 47, no. 11, pp. 1737–1746, Nov. 1999.

[88] G. Zang and C. Ling, "Performance evaluation for band-limited DS-CDMA systems based on simplified improved Gaussian approximation," *IEEE Transactions on Communications*, vol. 51, no. 7, pp. 1204–1213, July 2003.

[89] P. Billingsley, *Probability and Measure*, 3rd ed., ser. Wiley in probability and mathematical statistics. John Wiley & Sons, 1995.

[90] U. Grenander and G. Szegö, Eds., *Toeplitz Forms and Their Applications.* Berkeley, CA: University of California Press, 1958.

[91] R. M. Gray, Ed., *Toeplitz and Circulant Matrices: A Review.* Standford, CA: http://ee.standford.edu/ ~ gray/toeplitz.pdf, 2002.

[92] R. M. Gray, "Toeplitz and circulant matrices: A review," *IEEE Transactions on Information Theory*, vol. 18, no. 6, pp. 725–730, Nov. 1972.

[93] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing.* New Jersey: Prentice–Hall, 1975.

[94] J. Huber, *Trelliscodierung.* Berlin, Germany: Springer–Verlag, 1992.

[95] A. L. Moustakas, S. H. Simon, and A. M. Sengupta, "MIMO capacity through correlated channels in the presence of correlated interferers ond noise: A (not so) large n analysis," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2545–2561, Oct. 2003.

[96] M. Chiani, M. Z. Win, and A. Zanella, "On the capacity of spacially correlated MIMO rayleigh–fading channels," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2363–2371, Oct. 2003.

[97] C. Chuah, D. Tse, J. M. Kahn, and R. A. Valenzuela, "Capacity scaling in dual–antenna–array wireless systems under correlated fading," *IEEE Transactions on Information Theory*, vol. 48, no. 3, pp. 637–650, Mar. 2002.

[98] H. Shin and J. H. Lee, "Capacity of multiple–antenna fading channels: Spatial fading correlation, double scattering, and keyhole," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2636–2647, oct 2003.

[99] R. R. Müller, "A random matrix model for communication via antenna arrays," *IEEE Transactions on Information Theory*, vol. 48, no. 9, pp. 2495–2506, Sept. 2002.

[100] D. Chizhik, G. J. Foschini, and R. A. Valenzuela, "Capacities of multi–element transmit and receive antennas: Correlations and keyholes," *IEE Electronics Letters*, vol. 36, no. 3, pp. 1099–1100, June 2000.

[101] D. Chizhik, F. Rashid-Farrokhi, J. Ling, and A. Lozano, "Effect of antenna separation on the capacity of BLAST in correlated channels," *IEEE Communications Letters*, vol. 4, no. 11, pp. 337–339, Nov. 2000.

[102] M. Debbah and R. Müller, "MIMO channel modelling and the principle of maximum entropy," *IEEE Transactions on Information Theory*, vol. 51, no. 5, pp. 1667–1690, May 2005.

[103] D. Hösli and A. Lapidoth, "The capacity of a MIMO ricean channel is monotonic in the singular values of the mean," in *Proc. of the 5nd International ITG Conference on Source and Channel Coding*, Erlangen, Germany, Jan. 2004.

[104] R. R. Müller, "On the asymptotic eigenvalue distribution of concatenated vector–valued fading channels," *IEEE Transactions on Information Theory*, vol. 48, no. 7, pp. 2086–2091, July 2002.

[105] S. Verdú, *Multiuser Detection.* New York: Cambridge University Press, 1998.

[106] R. R. Müller and L. Cottatellucci, "Joint antenna combining and multiuser detection," in *Smart antennas in Europe–State of the art*, ser. EURASIP books on Signal Processing and Communications, T. Kaiser, A. Bourdoux, H. Boche, J. Fonollosa, J. B. Andersen, and W. Utschik, Eds. Hindawi Publishing Cooperation, 2005.

[107] J. W. Silverstein, "Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices," *Journal of Multivariate Analysis*, vol. 55, pp. 331–339, 1995.

[108] W. Feller, *An Introduction to Probability Theory and Its Applications*, 3rd ed. New York: John Wiley & Sons, 1968, vol. 1.

[109] S. Skiena, *Implementing Discrete Mathematics: Combinatorics and Graph Theory With Mathematica.* Addison–Wesley, 1990.

[110] W. Feller, *An Introduction to Probability Theory and Its Applications*, 2nd ed. New York: John Wiley & Sons, 1971, vol. 2.

[111] R. A. Horn and C. R. Johnson, *Matrix Analysis.* Cambridge, U.K.: Cambridge University Press, 1985.

[112] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York: McGraw–Hill, 1991.