Asynchronous CDMA Systems with Random Spreading–Part II: Design Criteria

Laura Cottatellucci, Ralf R. Müller, and Merouane Debbah

Abstract

Totally asynchronous code-division multiple-access (CDMA) systems are addressed. In part I, the fundamental limits of asynchronous CDMA systems were analyzed in terms of total capacity and SINR at the output of the optimum linear detector. The focus of Part II, is the design of low-complexity implementations of linear multiuser detectors in systems with many users that admit a multistage representation, e.g. reduced rank multistage Wiener filters, polynomial expansion detectors, weighted parallel interference cancellers.

The effects of the excess bandwidth, chip-pulse shaping, and time delay distribution on CDMA with suboptimum linear receiver structures are investigated. Recursive expressions for universal weight design are given. The performance in terms of SINR is derived in the large-system limit and the performance improvement over synchronous systems is quantified. The considerations distinguish between two ways of forming discrete-time statistics: chip-matched filtering and oversampling.

Index Terms - Asynchronous code division multiple access (CDMA), channel capacity, multiuser detection, random matrix theory, effective interference, linear minimum mean square error (MMSE) detector, multistage detector, random spreading sequences.

This work was presented in part at the IEEE Information Theory Workshop (ITW 2006), Punta de l'Este, Uruguay, March 2006 and at the IEEE Wireless Communications and Networking Conference, Hong Kong, March 2007. It partly appears in Laura Cottatellucci, "Low Complexity Multistage Detectors for Randomly Spread CDMA Systems", Ph.D. thesis, Vienna University of Technology, March 2006.

This work was supported by the French and Australian Science and Technology (FAST) Programme and the Research Council of Norway under grant 171133/V30.

Laura Cottatellucci is with Institute Eurecom, Sophia Antipolis, France (e-mail: laura.cottatellucci@eurecom.fr). She was with Institute of Telecommunications Research, University of South Australia, Adealide, SA, Australia. Ralf Müller is with Norwegian University of Science and Technology, Trondheim, Norway, (e-mail: ralf.mueller@iet.ntnu.no). Merouane Debbah was with Institut Eurecom, Sophia Antipolis, France. He is currently with SUPELEC, Paris, France (e-mail: merouane.debbah@supelec.fr).

I. INTRODUCTION

In part I of this paper [1], we analyzed asynchronous CDMA systems with random spreading sequences in terms of capacity per chip constrained to a given chip pulse waveform and in terms of SINR at the output of an optimum linear multiuser detector. The analysis showed that under realistic conditions, chip-asynchronous CDMA systems significantly outperform chip-synchronous CDMA systems. In order to utilize the benefits from chip-asynchronous CDMA, we need efficient algorithms to cope with multiuser detection for chip-asynchronous users. In part II of this work we, therefore, focus on the generalization of known design rules for low-complexity multiuser detectors to chip-asynchronous CDMA.

A unified framework for the design and analysis of multiuser detectors that admit a multistage representation for synchronous users was given in [2]. The class of multiuser detectors that admit a multistage representation is large and includes popular linear multiuser detectors like linear MMSE detectors (e.g. [3]), reduced rank multistage Wiener filters [4], [5], polynomial expansion detectors [6] or conjugate gradient methods (e.g. [7]), linear parallel interference cancellers (PIC, e.g. [8], [9]), eventually weighted (e.g. [10]), and the single-user matched filters. Multistage detectors are constructed around the matched filter concept. They consist of a projection of the signal onto a subspace of the whole signal space by successive matched filtering and re-spreading followed by a linear filter in the subspace.

Multistage detectors based on universal weights have been proposed in [11], [12] for CDMA systems in AWGN channels and extended to more realistic scenarios in [13], [14], [2]. These references make use of the self-averaging properties of large random matrices to find *universal* weighting coefficients for the linear filter in the subspace. More specifically, the universal weights are obtained by approximating the precise weights designed according to some optimality criterion with asymptotically optimum weights, i.e. the optimum weights for a CDMA system whose number of users and spreading factor tend to infinity with constant ratio. Thanks to the properties of random matrices, asymptotically these weights become independent of the users' spreading sequences and depend only on few macroscopic system parameters, as the system load or number of transmitted symbols per chip, the variance of the noise, and the distribution of the fading. In this way, the weight design for long-code CDMA simplifies considerably, its complexity becomes independent of both the number of users in the system and the spreading factor. Moreover, the weights need updating only when the macroscopic system parameters change.

The fact that users are not received in a time-synchronized manner at the receiver causes two main problems from a signal processing perspective: (i) the need for an infinite observation window to implement a linear MMSE detector and (ii) the potential need for oversampling to form sufficient discrete-time statistics. The need for an infinite observation window is primarily related to asynchronism on the symbol-level, not the chip-level. It was addressed in [15], [16] where it was found that multistage detectors need not have infinite observation windows and can be efficiently implemented without windowing at all. The question of how to form sufficient statistics was addressed in Part I of this paper [1] where a distinction between the following two cases was made:

(A) Sufficient statistics obtained by filtering the received signal by a lowpass filter with bandwidth B_{LOW} larger than the chip-pulse bandwidth and subsequent sampling at rate $2B_{\text{LOW}}$.

(B) Statistics obtained by sampling the output of a filter matched to the chip waveform at the chip rate (*chip rate sampling*). In this case the chip pulses at the output of matched filter need to satisfy the Nyquist criterion. In the following we refer to them as square root Nyquist chip-pulse waveforms.

General results for the design of linear multistage detectors with both kind of statistics are provided in this work. The chip pulse waveforms are assumed to be identical for all users.

For asynchronous CDMA, low-complexity detectors with universal weights are conveniently obtained formed from statistics (A). In fact, these observables enable a joint processing of all users without loss of information. Multistage detectors with universal weights and statistics (A) have a complexity order per bit equal to $\mathcal{O}(rK)$ if the sampling rate is $\frac{r}{T_c}$. On the contrary, discretization scheme (B) provides different observables for each user and does not allow for simultaneous joint detection of all users. An implementation of multistage detectors with universal weights using such statistics implies a complexity order *per bit* equal to $\mathcal{O}(K^2)$. This approach is still interesting from a complexity point of view if detection of a single user is required. However, it suffers from a performance degradation due to the sub-optimality of the statistics.

This work is organized in six additional sections. Section II and III introduce the notation and the system model for asynchronous CDMA, respectively. In Section IV, multistage detectors for asynchronous CDMA are reviewed and a implementation which does not suffer from truncation effects is given. The design of universal weighting is addressed in Section V. Finally, the analytical results are applied to gain further insight into the system in Section VI where methods for pulse-shaping, forming sufficient statistics and synchronization are compared. Conclusions are summed up in Section VII.

II. NOTATION AND SOME USEFUL DEFINITIONS

Throughout Part II we adopt the same notation and definitions already introduced in Part I of this work [1]. In order to make Part II self-contained we repeat here definitions useful in this part. Upper and lower boldface symbols are used respectively for matrices and vectors corresponding to signals spanning a specific symbol interval m. Matrices and vectors describing signals spanning more than a symbol interval are denoted by upper boldface calligraphic letters.

In the following, we utilize *unitary* Fourier transforms both in the continuous time and in the discrete time domain. The unitary Fourier transform of a function $f(t)$ in the continuous time domain is given by $F(j2\pi f) = \frac{1}{\sqrt{2\pi}} \int f(t) \mathrm{e}^{-j2\pi ft} \mathrm{d}t$. The unitary Fourier transform of a sequence $\{\ldots, c_{-1}, c_0, c_1, \ldots\}$ in the discrete time domain is given by $c(\mathrm{e}^{j2\pi x}) = \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{+\infty} c_n \mathrm{e}^{-j2\pi xn}$. We will refer to them shortly as Fourier transform. We denote the argument of a Fourier transform of a continuous function by f and the argument of a Fourier transform of a sequence by x.

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

Definition 1 *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*

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

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

$s = 1 \ldots r$,

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

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

Furthermore, the superscripts $\cdot^T$ and $\cdot^H$ denote the transpose and the conjugate transpose of the matrix argument, respectively. $\boldsymbol{I}_n$ is the identity matrix of size $n \times n$ and $\mathbb{C}$, $\mathbb{Z}$, $\mathbb{Z}^+$, and $\mathbb{R}$ are the fields of complex, integer, nonnegative integers, and real numbers, respectively. $\mathrm{tr}(\cdot)$ is the trace of the matrix argument and $\mathrm{span}(\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_s)$ denotes the vector space spanned by the s vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots \boldsymbol{v}_s$. $\mathrm{diag}(\ldots) : \mathbb{C}^n \rightarrow \mathbb{C}^{n \times n}$ transforms an n-dimensional vector $\boldsymbol{v}$ into a diagonal matrix of size n having as diagonal elements the components of $\boldsymbol{v}$ in the same order. $\mathrm{E}\{\cdot\}$ and $\Pr\{\cdot\}$ are the expectation and probability operators, respectively. δ_{ij} is the Kronecker symbol and $\delta(\lambda)$ is the Dirac's delta function. mod denotes the modulus and $\lfloor \cdot \rfloor$ is the operator that yields the maximum integer not greater than its argument.

III. SYSTEM MODEL

In this section we recall briefly the system model for asynchronous CDMA derived in Section IV of Part I of this work [1]. The reader interested in the details of the derivation can refer to [1].

Let us consider an asynchronous CDMA system with K active users in the uplink (reverse link) channel with spreading factor N. Each user and the base station are equipped with a single antenna. The channel is flat fading and impaired by additive white Gaussian noise with two sided power spectral density N_0. The symbol interval is denoted with T_s and $T_c = \frac{T_s}{N}$ is the chip interval. The modulation of all users is based on the same chip pulse waveform $\psi(t)$ bandlimited with bandwidth B and energy $E_\psi = \int_{-\infty}^{\infty} |\psi(t)|^2 \mathrm{d}t$.

The time delays of the K users are denoted with τ_k, $k = 1, \ldots, K$. Without loss of generality we can assume (i) user 1 as reference user so that $\tau_1 = 0$, (ii) the users ordered according to increasing time delay with respect to the reference user, i.e. $\tau_1 \leq \tau_2 \leq \ldots \leq \tau_K$; (iii) the time delay to be, at most, one symbol interval so that $\tau_k \in [0, T_s)$.[1]

As in Part I of this work [1] we assume the use of one of the following front-ends:

Front-end Type A consists of

- An ideal lowpass filter with cut-off frequency $f = \frac{r}{2Tc}$ where $r \in \mathbb{Z}^+$ satisfies the constraint $B \leq \frac{r}{2T_c}$ such that the sampling theorem applies. The filter is normalized to obtain a unit overall amplification

[1]For a thorough discussion on this assumption the reader can refer to [3].

factor, i.e. the transfer function is

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

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

This front-end satisfies the conditions of the sampling theorem and, thus, provides sufficient discrete-time statistics. For convenience, the sampling rate is an integer multiple of the chip rate. Additionally, the discrete-time noise process is white with zero mean and variance $\sigma^2 = \frac{N_0 r}{E_\psi T_c}$.

Front-end Type B consists of

- A filter $G(f)$ matched to the chip pulse and normalized to the chip pulse energy, i.e. $G(f) = \Psi^*(f) E_\psi^{-\frac{1}{2}}$;
- Subsequent sampling at the symbol rate.

When used with square root Nyquist chip pulses, the discrete time noise process $\{n[p]\}$ is white with variance $\frac{N_0}{E_\psi T_c}$. For synchronous systems with square root Nyquist chip pulses, this front end provides sufficient statistics whereas the observables are not sufficient if the system is asynchronous.

Let us denote with $\boldsymbol{b}[m]$ and $\boldsymbol{y}[m]$ the vectors of transmitted and received signals at time instants $m \in \mathbb{Z}$. The baseband discrete-time asynchronous system is given by

$$\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{B}} + \boldsymbol{\mathcal{N}} \tag{2}$$

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 infinite-dimensional vectors of received and transmitted symbols respectively; $\boldsymbol{\mathcal{N}}$ is an infinite-dimensional noise vector; and $\boldsymbol{\mathcal{H}}$ is a bi-diagonal block matrix of infinite size given by

$$\boldsymbol{\mathcal{H}} = \begin{bmatrix} \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \ldots & \mathbf{0} & \boldsymbol{H}_d[m-1] & \boldsymbol{H}_u[m] & \mathbf{0} & \ldots & \ldots \\ \ldots & \ldots & \mathbf{0} & \boldsymbol{H}_d[m] & \boldsymbol{H}_u[m+1] & \mathbf{0} & \ldots \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}. \tag{3}$$

Here, $\boldsymbol{H}_u[m]$ and $\boldsymbol{H}_d[m]$ are matrices of size $rN \times K$ obtained by the decomposition of the $2rN \times K$ matrix $\boldsymbol{H}[m]$ into two parts such that $\boldsymbol{H}[m] = [\boldsymbol{H}_u^T[m], \boldsymbol{H}_d^T[m]]^T$. For $\boldsymbol{H}[m]$ the relation

$$\boldsymbol{H}[m] = \boldsymbol{S}[m]\boldsymbol{A} \tag{4}$$

holds where $\boldsymbol{A}$ is the $K \times K$ diagonal matrix of the received amplitudes a_{kk} and $\boldsymbol{S}[m]$ is the $2rN \times K$ matrix of virtual spreading. The matrix $\boldsymbol{S}[m]$ of virtual spreading is given by

$$\boldsymbol{S}[m] = (\boldsymbol{\Phi}_1 \boldsymbol{s}_{1m}, \boldsymbol{\Phi}_2 \boldsymbol{s}_{2m}, \dots \boldsymbol{\Phi}_K \boldsymbol{s}_{Km}) \tag{5}$$

where $\boldsymbol{s}_{km}$ is the N-dimensional column vector of the spreading sequence of user k for the transmitted symbol m and $\boldsymbol{\Phi}_k$ is the $2rN \times N$ matrix taking into account the effects of the chip pulse shape and the time delay τ_k of user k. The matrix $\boldsymbol{\Phi}_k$ is of the form

$$\boldsymbol{\Phi}_k = \begin{bmatrix} \boldsymbol{0}_{k,0} \\ \boldsymbol{C}_{\phi,r}\left(\tau_k - \lfloor \frac{\tau_k}{T_c} \rfloor T_c\right) \\ \boldsymbol{0}_{k,1} \end{bmatrix} \tag{6}$$

where $\boldsymbol{0}_{k,0}$ and $\boldsymbol{0}_{k,1}$ are matrices of dimensions $\lfloor \frac{r\tau_k}{T_c} \rfloor \times N$ and $\left(N - \lfloor \frac{r\tau_k}{T_c} \rfloor\right) \times N$, respectively, with zero elements; $\boldsymbol{C}_{\phi,r}(\tau_k)$ is an r-block-wise circulant matrix of order N defined by

$$\boldsymbol{C}_{\phi,r}(\tau) \triangleq \boldsymbol{C}(\phi(x,\tau), \phi(x, \tau - \tfrac{T_c}{r}), \dots, \phi(x, \tau - \tfrac{(r-1)T_c}{r})), \tag{7}$$

with

$$\phi(x,\tau) \triangleq \frac{1}{T_c} \sum_{s=-\infty}^{+\infty} \mathrm{e}^{j2\pi \frac{\tau}{T_c}(x+s)} \Phi^* \left(\tfrac{j2\pi}{T_c}(x+s)\right). \tag{8}$$

Thus, the virtual spreading sequences are the samples of the delayed continuous-time spreading waveforms at sampling rate r/T_c.

Throughout this work we assume that the transmitted symbols are uncorrelated and identically distributed random variables with unitary variance and zero mean, i.e. $\mathrm{E}(\boldsymbol{\mathcal{B}}) = \boldsymbol{\mathcal{O}}$ and $\mathrm{E}(\boldsymbol{\mathcal{B}}\boldsymbol{\mathcal{B}}^H) = \boldsymbol{\mathcal{I}}$ being $\boldsymbol{\mathcal{O}}$ and $\boldsymbol{\mathcal{I}}$ an unlimited zero vector and the unlimited identity matrix, respectively. The elements of the spreading sequences $\boldsymbol{s}_{k,m}$ are assumed to be zero mean i.i.d. random variables over all the users, chips, and symbols with $\mathrm{E}\{\boldsymbol{s}_{km}\boldsymbol{s}_{km}^H\} = \frac{1}{N}\boldsymbol{I}_N$. Finally, $\boldsymbol{h}_{k,m}$ denotes that column of the matrix $\boldsymbol{\mathcal{H}}$ containing the k^{th} column of the matrix $\boldsymbol{H}[m]$. We define the correlation matrices $\boldsymbol{\mathcal{T}} = \boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{H}}^H$ and $\boldsymbol{\mathcal{R}} = \boldsymbol{\mathcal{H}}^H\boldsymbol{\mathcal{H}}$. The system load $\beta = \frac{K}{N}$ is the number of transmitted symbols per chip.

IV. MULTISTAGE STRUCTURES FOR ASYNCHRONOUS CDMA

We consider the large class of linear multistage detectors for asynchronous CDMA. Let $\chi_{L,k}(\boldsymbol{\mathcal{H}})$ be the Krylov subspace [17] of rank $L \in \mathbb{Z}^+$ given by

$$\chi_{L,k,m}(\boldsymbol{\mathcal{H}}) = \mathrm{span}(\boldsymbol{\mathcal{T}}^\ell \boldsymbol{h}_{k,m})|_{\ell=0}^{L-1}. \tag{9}$$

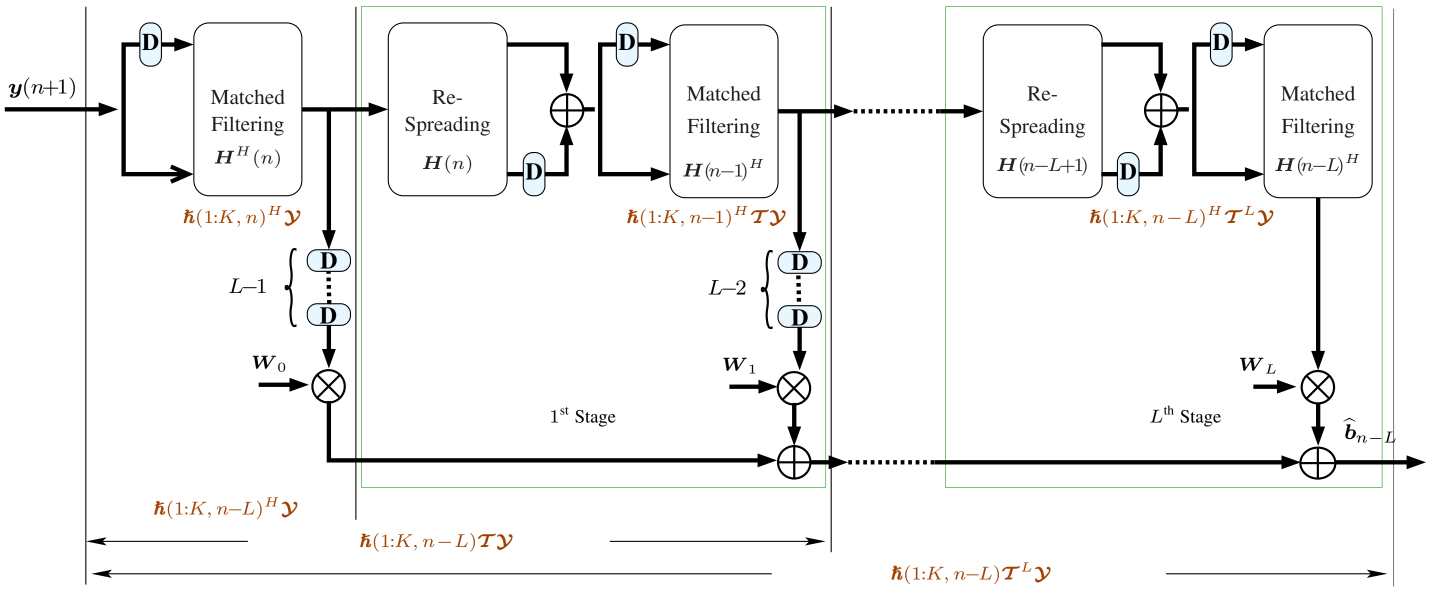

Fig. 1. Multistage detector for asynchronous CDMA systems. Here, $\hbar(1:K,r) = [\boldsymbol{h}_{1r}, \boldsymbol{h}_{2r}, \ldots \boldsymbol{h}_{Kr}]$

A multistage detector of rank $L \in \mathbb{Z}^+$ for user k is given by

$$\widehat{b}_k = \sum_{\ell=0}^{L-1} (\boldsymbol{w}_{k,m})_\ell \boldsymbol{h}_{k,m}^H \boldsymbol{\mathcal{T}}^\ell \boldsymbol{\mathcal{Y}} \tag{10}$$

where $\boldsymbol{w}_{k,m}$ is the L-dimensional vector of weight coefficients.

It has been shown in [16] that, given the weight vector $\boldsymbol{w}_{k,m}$ the detection of the symbol $b_k[m]$ by the multistage detector of rank L in (10) can be performed with finite delay L using the implementation scheme in Figure 1. Although infinite length vectors and infinite dimension matrices appear in (10), the multistage detector in Figure 1 implements exactly (10) and does not suffer from truncation effects. Equivalently, the multistage detector in Figure 1 can be considered as a multistage detector with sliding observation window of size $2L$. The projection of the received vector $\boldsymbol{\mathcal{Y}}$ onto the subspaces $\chi_{L,k,m}(\boldsymbol{\mathcal{H}})$, for $k = 1 \ldots K$, is performed jointly for all users and requires only multiplications between vectors and matrices. The size of those vectors and matrices does not depend on the observation window. For further details the interested reader is referred to [16], [18].

The class of multistage detectors includes many popular multiuser detectors:

- the single-user matched filter for $L = 1$,
- the linear parallel interference canceller (PIC) [19], [20] for weight coefficients chosen irrespective of the properties of the transfer matrix $\boldsymbol{\mathcal{H}}$,
- the polynomial expansion detector [6] and the conjugate gradient method [7], if the weight coefficients are identical for all users and chosen to minimize the mean square error.

- the (reduced rank) multistage Wiener filter [5] if the weight coefficients are chosen to minimize the mean square error, but are allowed to differ from user to user.

Throughout this work we refer to detectors that minimize the MSE in the projection subspace of the user of interest as *optimum detectors in MSE sense*. More specifically this class of multistage detectors includes the linear MMSE detector and the multistage Wiener filter but not the polynomial expansion detector.

In the following we focus on the design of multistage Wiener filters implemented as in Figure 1. This reduces the problem to the design of the filter coefficients $\boldsymbol{w}_{k,m}$. The multistage Wiener filter for the detection of the symbol m transmitted by user k reads

$$\boldsymbol{\mathcal{M}}_{k,m} = \sum_{\ell=0}^{L-1} (\boldsymbol{w}_{k,m})_{\ell-1} \boldsymbol{h}_{k,m}^H \boldsymbol{\mathcal{T}}^{\ell}. \tag{11}$$

The weight vector $\boldsymbol{w}_{k,m}$ that minimizes the MSE $\mathrm{E}\{\|\boldsymbol{\mathcal{M}}_{k,m}\boldsymbol{\mathcal{Y}} - b_{k,m}\|^2\}$ is given by

$$\boldsymbol{w}_{k,m} = \underset{\overline{\boldsymbol{w}}_{k,m}}{\operatorname{argmin}} \mathrm{E}\left\{ \left\| \sum_{\ell=0}^{L-1} (\overline{\boldsymbol{w}}_{k,m})_{\ell} \boldsymbol{h}_{k,m}^H \boldsymbol{\mathcal{T}}^{\ell} \boldsymbol{\mathcal{Y}} - b_{k,m} \right\|^2 \right\} \tag{12}$$

$$= \underset{\overline{\boldsymbol{w}}_{k,m}}{\operatorname{argmin}} \mathrm{E}\left\{ \left\| \overline{\boldsymbol{w}}_{k,m}^H \boldsymbol{x}_{k,m} - b_{k,m} \right\|^2 \right\} \tag{13}$$

where $\boldsymbol{x}_{k,m}$ is an L-dimensional vector with j^{th} element $(\boldsymbol{x}_{k,m})_j = \boldsymbol{h}_{k,m}^H \boldsymbol{\mathcal{T}}^{j-1} \boldsymbol{\mathcal{Y}}$. This optimization problem is solved by the Wiener-Hopf theorem [21] and $\boldsymbol{w}_{k,m}$ is given by

$$\boldsymbol{w}_{k,m} = \Xi_{k,m}^{-1} \boldsymbol{\xi}_{k,m} \tag{14}$$

where $\Xi_{k,m} = \mathrm{E}\{\boldsymbol{x}_{k,m}\boldsymbol{x}_{k,m}^H\}$ and $\boldsymbol{\xi} = \mathrm{E}\{b_{k,m}^* \boldsymbol{x}_{k,m}\}$. It is straightforward to verify that in this case

$$\Xi_{k,m} = \begin{pmatrix} (\boldsymbol{\mathcal{R}}^2)_{k,m} + \sigma^2(\boldsymbol{\mathcal{R}})_{k,m} & \cdots & (\boldsymbol{\mathcal{R}}^{L+1})_{k,m} + \sigma^2(\boldsymbol{\mathcal{R}}^{L})_{k,m} \\ (\boldsymbol{\mathcal{R}}^3)_{k,m} + \sigma^2(\boldsymbol{\mathcal{R}}^2)_{k,m} & \cdots & (\boldsymbol{\mathcal{R}}^{L+2})_{k,m} + \sigma^2(\boldsymbol{\mathcal{R}}^{L+1})_{k,m} \\ \vdots & \ddots & \vdots \\ (\boldsymbol{\mathcal{R}}^{L+1})_{k,m} + \sigma^2(\boldsymbol{\mathcal{R}}^{L})_{k,m} & \cdots & (\boldsymbol{\mathcal{R}}^{2L})_{k,m} + \sigma^2(\boldsymbol{\mathcal{R}}^{2L-1})_{k,m} \end{pmatrix}$$

$$\boldsymbol{\xi}_k = \left((\boldsymbol{\mathcal{R}})_{k,m}, (\boldsymbol{\mathcal{R}}^2)_{k,m}, \ldots, (\boldsymbol{\mathcal{R}}^L)_{k,m} \right)^T. \tag{15}$$

where $(\boldsymbol{\mathcal{R}}^s)_{k,m} = \boldsymbol{h}_{k,m}^H \boldsymbol{\mathcal{T}}^{s-1} \boldsymbol{h}_{k,m}$ is the diagonal element of the matrix $\boldsymbol{\mathcal{R}}^s$ corresponding to the m^{th} symbol transmitted by user k.

V. UNIVERSAL WEIGHT DESIGN

Consider the SINR of any linear detector that admits a multistage representation. Let $\overline{\boldsymbol{w}}_{k,m}$ be the weight vector for the detection of the m^{th} symbol transmitted by user k, then the SINR at the output of the multistage detector is given by

$$\mathrm{SINR}_k = \frac{\overline{\boldsymbol{w}}_{k,m}^H \boldsymbol{\xi}_{k,m} \boldsymbol{\xi}_{k,m}^T \overline{\boldsymbol{w}}_{k,m}}{\overline{\boldsymbol{w}}_{k,m}^H (\Xi_{k,m} - \boldsymbol{\xi}_{k,m} \boldsymbol{\xi}_{k,m}^T) \overline{\boldsymbol{w}}_{k,m}^H}. \tag{16}$$

The performance of multistage Wiener filters simplifies to

$$\mathrm{SINR}_k = \frac{\boldsymbol{\xi}_{k,m}^T \Xi_{k,m}^{-1} \boldsymbol{\xi}_{k,m}}{1 - \boldsymbol{\xi}_{k,m}^T \Xi_{k,m}^{-1} \boldsymbol{\xi}_{k,m}}. \tag{17}$$

From (14), (16), and (17) it is apparent that the diagonal elements of the matrix $\boldsymbol{\mathcal{R}}^s$ play a fundamental role in the design and analysis of multistage detectors.

It has been shown in [2] that, if the spreading sequences are random and the CDMA system is synchronous, the diagonal elements of the matrix $\boldsymbol{\mathcal{R}}^s$, $s \in \mathbb{Z}^+$, converge to deterministic values as $K, N \to \infty$ with constant ratio. This asymptotic convergence holds for some classes of random matrices and is a stronger property than the convergence of the eigenvalue distribution. The Stieltjes transform of the asymptotic eigenvalue distribution of $\boldsymbol{\mathcal{R}}$ is related to the SINR at the output of the linear MMSE detector, as pointed out first in [22] for synchronous CDMA systems. The asymptotic eigenvalue moments of $\boldsymbol{\mathcal{R}}$ enable the asymptotic performance analysis of reduced rank multistage Wiener filters [23] and the design of multistage detectors with quadratic complexity order per bit [14], [13]. The convergence of the diagonal elements of $\boldsymbol{\mathcal{R}}^s$ has been utilized in [2] for the design of multistage detectors with linear complexity order per bit in synchronous CDMA systems and for the asymptotic analysis of any multistage detector not necessarily optimum in a MSE sense. In the following we extend the results in [2] to the case of asynchronous CDMA systems making use of the asymptotic properties of the random matrix $\boldsymbol{\mathcal{R}}$ for asynchronous CDMA systems.

The design of low complexity multistage detectors is based on the approximation of the weight vectors $\boldsymbol{w}_{k,m}$ by their asymptotic limit when $K, N \to \infty$ with constant ratio β

$$\boldsymbol{w}_k^\infty = \lim_{K=\beta N \to \infty} \Xi_{k,m}^{-1} \boldsymbol{\xi}_{k,m}. \tag{18}$$

Thanks to the fact that the diagonal elements of $\boldsymbol{\mathcal{R}}^s$ can be computed by a polynomial in few macroscopic system parameters, the computation of the weight vectors becomes independent of the size of $\boldsymbol{\mathcal{R}}$ and independent of m. Thus, the effort for the computation of the weights becomes negligible and the complexity of the detector is dominated by the joint projection of the received signal $\boldsymbol{\mathcal{Y}}$ onto the subspaces $\chi_{k,m}(\boldsymbol{\mathcal{H}})$, $k = 1 \ldots K$ and $m \in \mathbb{Z}$. This projection has linear complexity per bit if the multistage detector in Figure 1 is utilized.

In order to present asymptotic results useful for the design of linear multistage detectors in asynchronous CDMA systems with random spreading we follow the same line as in Section IV of Part I of this article [1].

More specifically, we introduce a matrix $\widetilde{\boldsymbol{\mathcal{H}}}$ as follows. Let $\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau})$, with $\widetilde{\tau} \in [0, T_c]$ be an $rN \times N$ block diagonal matrix with ℓ^{th} block

$$(\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}))_{\ell\ell} = \begin{pmatrix} \phi(\frac{\ell-1}{N}, \widetilde{\tau}) \\ \phi(\frac{\ell-1}{N}, \widetilde{\tau} - \frac{T_c}{r}) \\ \vdots \\ \phi(\frac{\ell-1}{N}, \widetilde{\tau} - \frac{(r-1)T_c}{r}) \end{pmatrix} \tag{19}$$

and $\phi(x, \tau)$ defined in (8). Similarly to the matrix $\boldsymbol{\Phi}_k$ in (6), we define a matrix $\widetilde{\boldsymbol{\Phi}}_k$ as

$$\widetilde{\boldsymbol{\Phi}}_k = \begin{bmatrix} \mathbf{0}_{k,0} \\ \widetilde{\boldsymbol{\Delta}}_{\phi,r}(\tau_k) \\ \mathbf{0}_{k,1} \end{bmatrix} \tag{20}$$

where $\mathbf{0}_{k,0}$ and $\mathbf{0}_{k,1}$ are defined as in (6) and $\widetilde{\boldsymbol{\Delta}}_{\phi,r}(\tau)$ is obtained by the diagonal circular down shift of the diagonal elements of the matrix $\boldsymbol{\Delta}_{\phi,r}\left(\tau - \left\lfloor \frac{\tau}{T_c} \right\rfloor T_c\right)$ by $\left\lfloor \frac{\tau}{T_c} \right\rfloor$ positions, i.e. the $\widetilde{\ell}^{\text{th}}$ diagonal block of $\widetilde{\boldsymbol{\Delta}}_{\phi,r}(\tau)$ is given by $\left(\widetilde{\boldsymbol{\Delta}}_{\phi,r}(\tau)\right)_{\widetilde{\ell}\widetilde{\ell}} = \left(\boldsymbol{\Delta}_{\phi,r}\left(\tau - \left\lfloor \frac{\tau}{T_c} \right\rfloor T_c\right)\right)_{\ell\ell}$ with $\ell = \left(N + \widetilde{\ell} + \left\lfloor \frac{\tau}{T_c} \right\rfloor\right) \mathrm{mod} N + 1$. The matrix $\widetilde{\boldsymbol{\mathcal{H}}}$ is defined as the matrix $\boldsymbol{\mathcal{H}}$ replacing the $2rN \times K$ matrices of virtual spreading $\boldsymbol{S}[m]$ in $\boldsymbol{\mathcal{H}}$ with the matrices of virtual spreading $\widetilde{\boldsymbol{S}}[m] = (\widetilde{\boldsymbol{\Phi}}_1 \widetilde{\boldsymbol{s}}_{1m}, \widetilde{\boldsymbol{\Phi}}_2 \widetilde{\boldsymbol{s}}_{2m} \ldots \widetilde{\boldsymbol{\Phi}}_K \widetilde{\boldsymbol{s}}_{Km})$, where $\widetilde{\boldsymbol{s}}_{km}$ are N-dimensional vectors with the same statistical properties of $\boldsymbol{s}_{km}$.

When the delays of the received signals τ_k, $k = 1, \ldots, K$ are not greater than the chip interval T_c, i.e. $\tau_k \leq T_c$, $k = 1, \ldots, K$, the matrices $\boldsymbol{\Phi}_k$ and $\widetilde{\boldsymbol{\Phi}}_k$ are of the form

$$\boldsymbol{\Phi}_k = \begin{bmatrix} \boldsymbol{C}_{\phi,r}(\tau_k) \\ \mathbf{0}_N \end{bmatrix} \qquad \text{and} \qquad \widetilde{\boldsymbol{\Phi}}_k = \begin{bmatrix} \boldsymbol{\Delta}_{\phi,r}(\tau_k) \\ \mathbf{0}_N \end{bmatrix},$$

respectively, being $\mathbf{0}_N$ an $N \times N$ zero matrix. The matrices $\boldsymbol{\mathcal{H}}$ in (3) and $\widetilde{\boldsymbol{\mathcal{H}}}$ reduce to block diagonal matrices with blocks of dimensions $rN \times K$ and we can focus on the transmission in a single symbol interval. The virtual spreading matrices in the m^{th} symbol interval are given by

$$\overline{\boldsymbol{S}}[m] = [\boldsymbol{C}_{\phi,r}(\tau_1)\boldsymbol{s}_{1m}, \boldsymbol{C}_{\phi,r}(\tau_2)\boldsymbol{s}_{2m} \ldots \boldsymbol{C}_{\phi,r}(\tau_K)\boldsymbol{s}_{Km}] \tag{21}$$

and

$$\widehat{\boldsymbol{S}}[m] = [\boldsymbol{\Delta}_{\phi,r}(\tau_1)\widetilde{\boldsymbol{s}}_{1m}, \boldsymbol{\Delta}_{\phi,r}(\tau_2)\widetilde{\boldsymbol{s}}_{2m} \ldots \boldsymbol{\Delta}_{\phi,r}(\tau_K)\widetilde{\boldsymbol{s}}_{Km}]. \tag{22}$$

Let $\overline{\boldsymbol{H}}[m] = \overline{\boldsymbol{S}}[m]\boldsymbol{A}$ ($\widehat{\boldsymbol{H}}[m] = \widehat{\boldsymbol{S}}[m]\boldsymbol{A}$) be the transfer matrix of the system at time instant m and let $\overline{\boldsymbol{R}} = \overline{\boldsymbol{H}}^H \overline{\boldsymbol{H}}$ ($\widehat{\boldsymbol{R}} = \widehat{\boldsymbol{H}}^H \widehat{\boldsymbol{H}}$). Without ambiguity we can drop the index m in the following. The convergence of the diagonal elements of $\overline{\boldsymbol{R}}^{\ell}$ and of $\widehat{\boldsymbol{R}}^{\ell}$ to deterministic values is established in the following theorem. The definitions and the assumptions in the statement of Theorem 1 summarize and formalize the characteristics of system model (2) for $\widetilde{\tau}_k \in [0, T_c]$.

Theorem 1 *Let $\boldsymbol{A} \in \mathbb{C}^{K\times K}$ be a diagonal matrix with k^{th} diagonal element a_{kk} and T_c a positive real. Given a function $\Phi(j2\pi f) : \mathbb{R} \rightarrow \mathbb{C}$, let $\phi(x,\tau)$ be as in (8). Given $\{\widetilde{\tau}_1, \widetilde{\tau}_2 \ldots \widetilde{\tau}_K\}$ a set of reals in $[0, T_c]$ and a positive integer r, let $\boldsymbol{C}_{\phi,r}(\widetilde{\tau}_k)$, $k = 1,\ldots,K$, be r-block-wise circulant matrices of order N defined in (7) and $\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_k)$, $k = 1,\ldots,K$, be $rN \times N$ block diagonal matrices with ℓ^{th} block defined as in (19). Let $\overline{\boldsymbol{H}} = \overline{\boldsymbol{S}}\boldsymbol{A}$ with $\overline{\boldsymbol{S}} = (\boldsymbol{C}_{\phi,r}(\widetilde{\tau}_1)\boldsymbol{s}_1, \boldsymbol{C}_{\phi,r}(\widetilde{\tau}_2)\boldsymbol{s}_2, \ldots, \boldsymbol{C}_{\phi,r}(\widetilde{\tau}_K)\boldsymbol{s}_K)$ with $\boldsymbol{s}_k \in \mathbb{C}^{N\times 1}$. Furthermore, let $\widehat{\boldsymbol{H}} = \widehat{\boldsymbol{S}}\boldsymbol{A}$ with $\widehat{\boldsymbol{S}} = (\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_1)\boldsymbol{s}_1, \boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_2)\boldsymbol{s}_2, \ldots, \boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_K)\boldsymbol{s}_K)$.*

We assume that the function $\Phi(j2\pi f)$ is upper bounded and has finite support. The vectors $\boldsymbol{s}_k$ are independent with i.i.d. circularly symmetric Gaussian elements. Furthermore, the elements a_{kk} of the matrix $\boldsymbol{A}$ are uniformly bounded for any K. The sequence of the empirical joint distributions $F^{(K)}_{|\boldsymbol{A}|^2,T}(\lambda,\tau) = \frac{1}{K}\sum_{k=1}^{K} 1(\lambda - |a_{kk}|^2) 1(\tau - \widetilde{\tau}_k)$ converges almost surely, as $K \rightarrow \infty$, to a non-random distribution function $F_{|\boldsymbol{A}|^2,T}(\lambda,\tau)$.

Then, conditioned on $(|a_{kk}|^2, \widetilde{\tau}_k)$, the corresponding diagonal elements of the matrices $\overline{\boldsymbol{R}}^{\ell}$ and $\widehat{\boldsymbol{R}}^{\ell}$ converge in probability to the deterministic value

$$\lim_{K=\beta N\rightarrow\infty} (\overline{\boldsymbol{R}}^{\ell})_{kk} = \lim_{K=\beta N\rightarrow\infty} (\widehat{\boldsymbol{R}}^{\ell})_{kk} \stackrel{\mathcal{P}}{=} R_\ell(|a_{kk}|^2, \widetilde{\tau}_k) \tag{23}$$

with $R_\ell(|a_{kk}|^2, \widetilde{\tau}_k)$ determined by the following recursion

$$R_\ell(\lambda,\tau) = \sum_{s=0}^{\ell-1} g(\boldsymbol{T}_{\ell-s-1}, \lambda, \tau) R_s(\lambda,\tau) \tag{24}$$

and

$$\boldsymbol{T}_\ell(x) = \sum_{s=0}^{\ell-1} \mathbf{f}(R_{\ell-s-1}, x)\boldsymbol{T}_s(x) \qquad -\frac{1}{2} \le x \le \frac{1}{2} \tag{25}$$

$$\mathbf{f}(R_\ell, x) = \beta \int \lambda \boldsymbol{\Delta}_{\phi,r}(x,\tau)\boldsymbol{\Delta}^H_{\phi,r}(x,\tau) R_\ell(\lambda,\tau) \mathrm{d}\, F_{|\boldsymbol{A}|^2,T}(\lambda,\tau) \qquad -\frac{1}{2} \le x \le \frac{1}{2} \tag{26}$$

$$g(\boldsymbol{T}_\ell, \lambda, \tau) = \lambda \int_{-\frac{1}{2}}^{\frac{1}{2}} \boldsymbol{\Delta}^H_{\phi,r}(x,\tau)\boldsymbol{T}_\ell(x)\boldsymbol{\Delta}_{\phi,r}(x,\tau) \mathrm{d}\, x \tag{27}$$

with

$$\boldsymbol{\Delta}_{\phi,r}(x,\tau) = \begin{pmatrix} \phi(x,\tau) \\ \phi(x,\tau-\frac{T_c}{r}) \\ \vdots \\ \phi(x,\tau-\frac{T_c(r-1)}{r}) \end{pmatrix}. \tag{28}$$

The recursion is initialized by setting $\boldsymbol{T}_0(x) = \boldsymbol{I}_r$ *and* $R_0(\lambda,\tau) = 1$.

Theorem 1 is proven in Appendix I.

From Theorem 1 we can obtain $m_{\overline{\boldsymbol{R}}}^{(\ell)}$, the asymptotic eigenvalue moment of the matrix $\overline{\boldsymbol{R}}$ of order ℓ by using the relation

$$m_{\overline{\boldsymbol{R}}}^{(\ell)} = \mathrm{E}\{R_\ell(\lambda,\tau)\}$$

where the expectation is taken over the limiting eigenvalue distribution $F_{|\boldsymbol{A}|^2,T}(\lambda,\tau)$. 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, and $\phi(x)$ satisfying the Nyquist criterion the recursive equations (25), (26), and (27) reduce to the recursion in [2] Theorem 1.

This theorem is very general and holds for all chip pulses of practical interest. Furthermore, no constraint is imposed on the time delay distribution and on the front end. The choice of the front end in this work is restricted only by the applicability of (16) or (17), which imply white noise at the front end.

Now, we specialize Theorem 1 to a case of theoretical and practical interest, where sufficient statistics are utilized in the detection, the chip pulse waveform $\phi(t)$ is band-limited, and the empirical distribution function of the time delays converges to a uniform distribution function.

Corollary 1 *Let us adopt the same definitions as in Theorem 1 and let the same assumptions of Theorem 1 be satisfied. Additionally, assume that the random variables* λ *and* τ *in* $F_{|\boldsymbol{A}|^2,T}(\lambda,\tau)$ *are statistically independent and the random variable* τ *is uniformly distributed in* $[0,T_c]$. *Furthermore,* $\Phi(j2\pi f)$ *is bounded in absolute value, and bandlimited with bandwidth* $B \in [\frac{n-1}{2T_c}, \frac{n}{2T_c}]$ *and* $r \geq n$. *Then, given* $(|a_{kk}|^2, \tau_k)$, *the corresponding diagonal element of the matrix* $\overline{\boldsymbol{R}}^\ell$ *or of the matrix* $\widehat{\boldsymbol{R}}^\ell$ *converges in probability to a deterministic value, conditionally on* $|a_{kk}|^2$,

$$\lim_{K=\beta N\to\infty} (\overline{\boldsymbol{R}})_{kk}^\ell = \lim_{K=\beta N\to\infty} (\widehat{\boldsymbol{R}})_{kk}^\ell \stackrel{\mathcal{P}}{=} R_\ell(|a_{kk}|^2)$$

with $R_\ell(\lambda)|_{\lambda=|a_{kk}|^2}$ *determined by the following recursion:*

$$R_\ell(\lambda) = \sum_{s=0}^{\ell-1} \lambda R_s(\lambda)\nu_{\ell-s-1}$$

and

$$\begin{aligned} T_\ell(x) &= \frac{r}{T_c}\sum_{s=0}^{\ell-1} \beta f(R_{\ell-s-1})\frac{1}{T_c}\left|\Phi\left(j2\pi\frac{x}{T_c}\right)\right|^2 T_s(x) \qquad\qquad -n \le x \le n \\ f(R_\ell) &= \int \lambda R_\ell(\lambda)\mathrm{d}\, F_{|\boldsymbol{A}|^2}(\lambda) \\ \nu_\ell &= \frac{r}{T_c}\int_{-n}^{n} \frac{1}{T_c}\left|\Phi\left(j2\pi\frac{x}{T_c}\right)\right|^2 T_\ell(x)\mathrm{d}\, x. \end{aligned}$$

The recursion is initialized by setting $T_0(x) = 1$ *and* $R_0(\lambda) = 1$.

Corollary 1 is derived in Appendix II.

The eigenvalue moments of $\overline{\boldsymbol{R}}$ or of $\widehat{\boldsymbol{R}}$ can be expressed in terms of the auxiliary quantities $f(R_s)$ and ν_s in the recursion of Corollary 1 by the following expression:

$$m_{\overline{\boldsymbol{R}}}^{(\ell)} = \mathrm{E}\{R_\ell(\lambda)\} = \sum_{s=0}^{\ell-1} f(R_s)\nu_{\ell-s-1}.$$

Applying Corollary 1 we obtain the following algorithm to compute the asymptotic limits of the diagonal elements of $\overline{\boldsymbol{R}}$ ($\widehat{\boldsymbol{R}}$) and the eigenvalue moments.

Algorithm 1

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

l^{th} step:

- *Define* $u_{\ell-1}(y) = \frac{r}{T_c} y\mu_{\ell-1}(y)$ *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*

$$\mathcal{E}_s = \frac{1}{T_c^s}\int_{-B}^{B} T_c|\Phi(j2\pi f)|^{2s}\mathrm{d}\, f \tag{29}$$

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* $m^{(1)}_{|\boldsymbol{A}|^2}$, $m^{(2)}_{|\boldsymbol{A}|^2}$,..., $m^{(\ell)}_{|\boldsymbol{A}|^2}$, *respectively. Denote the result by* $V_{\ell-1}$.
- *Calculate*

$$\rho_{\ell}(z) = \sum_{s=0}^{\ell-1} zU_{\ell-s-1}\rho_s(z)$$

$$\mu_{\ell}(y) = \frac{r}{T_c}\sum_{s=0}^{\ell-1} \beta y V_{\ell-s-1}\mu_s(y).$$

- *Assign* $\rho_{\ell}(\lambda)$ *to* $R^{\ell}(\lambda)$.

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

Algorithm 1 is derived in Appendix III.

Interestingly, the recursive equations in Corollary 1 do not depend on the time delay τ_k of the signal of user k, i.e. the performance of a CDMA system with multistage detection is independent of the sampling instants and time delays if the assumptions of Corollary 1 on the chip waveforms and on the time delays are satisfied.

Additionally, the dependence of $R^{\ell}(\lambda)$ on the chip pulse waveforms becomes clear from Algorithm 1: $R^{\ell}(\lambda)$ depends on $\Phi(j2\pi f)$ through the quantities $\mathcal{E}_s$, $s = 1, 2, \ldots$, defined in (29).

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

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

In general, the eigenvalue moments of $\overline{\boldsymbol{R}}$ ($\widehat{\boldsymbol{R}}$) depend only on the system load β, the sampling rate $\frac{r}{T_c}$, 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 $\phi(t)$. The asymptotic limits of the diagonal elements of the matrix $\overline{\boldsymbol{R}}^\ell$ corresponding to user k depends also on $|a_{kk}|^2$ but not on the time delay τ_k.

In the special case of chip pulse waveforms $\psi(t)$ having bandwidth not greater than the half of the chip rate, i.e. $B \leq \frac{T_c}{2}$ the result of Corollary 1 holds for any sets of time delays included synchronous systems.

In Theorem 2, chip pulse waveforms with bandwidth $B \leq \frac{1}{2T_c}$ are considered and the diagonal elements of $\boldsymbol{\mathcal{R}}^s$, or equivalently $\widetilde{\boldsymbol{\mathcal{R}}}$, are shown to be independent of the time delays of the active users.

Theorem 2 *Let the definitions of Theorem 1 hold.*

We assume that the function $\Phi(j2\pi f)$ is bounded in absolute value and has support $\Omega \subseteq \left[-\frac{1}{2T_c}, \frac{1}{2T_c}\right]$. The vectors $\boldsymbol{s}_k$ are independent with i.i.d. Gaussian elements $s_{nk} \in \mathbb{C}$ such that $\mathrm{E}\{s_{nk}\} = 0$ and $\mathrm{E}\{|s_{nk}|^2\} = \frac{1}{N}$. Furthermore, the elements a_{kk} of the matrix $\boldsymbol{A}$ are uniformly bounded for any K. The sequence of the empirical distributions $F^{(K)}_{|\boldsymbol{A}|^2}(\lambda) = \frac{1}{K}\sum_{k=1}^K 1(\lambda - |a_{kk}|^2)$ converges in law almost surely, as $K \to \infty$, to a non-random distribution function $F_{|\boldsymbol{A}|^2}(\lambda)$.

Then, given $|a_{kk}|^2$, the k^{th} diagonal elements of the matrix $\overline{\boldsymbol{R}}^\ell$ or of the matrix $\widehat{\boldsymbol{R}}^\ell$ converges in probability to a deterministic value, conditionally on $|a_{kk}|^2$,

$$\lim_{K=\beta N \to \infty} (\overline{\boldsymbol{R}}^\ell)_{kk} = \lim_{K=\beta N \to \infty} (\widehat{\boldsymbol{R}}^\ell)_{kk} \overset{a.s.\mathcal{P}}{=} R_\ell(|a_{kk}|^2)$$

with $R_\ell(|a_{kk}|^2)$ determined by the following recursion

$$R_\ell(\lambda) = \sum_{s=0}^{\ell-1} \lambda R_s(\lambda) \nu_{\ell-s-1} \tag{30}$$

and

$$T_\ell(x) = \frac{r}{T_c}\sum_{s=0}^{\ell-1} \beta f(R_{\ell-s-1}) \frac{1}{T_c} |\Phi(j2\pi x/T_c)|^2 T_s(x) \qquad -\frac{1}{2} \leq x \leq \frac{1}{2} \tag{31}$$

$$f(R_\ell) = \int \lambda R_\ell(\lambda) \mathrm{d}\, F_{|\boldsymbol{A}|^2}(\lambda) \tag{32}$$

$$\nu_\ell = \frac{r^2}{T_c^2}\int_{-\frac{1}{2}}^{\frac{1}{2}} |\Phi(j2\pi x/T_c)|^2 T_\ell(x) \mathrm{d}\, x. \tag{33}$$

The recursion is initialized by setting $T_0(x) = \frac{T_c}{r}$ *and* $R_0(\lambda) = 1$.

Theorem 2 is shown in Appendix IV. It is straightforward to verify that Algorithm 1 can be applied to determine $R_\ell(\lambda)$, the asymptotic limit of the diagonal elements and the eigenvalue moments of matrices $\overline{\boldsymbol{R}}$ and $\widehat{\boldsymbol{R}}$ satisfying the conditions of Theorem 2.

As for linear MMSE detectors, we can show that the diagonal elements of the matrix $\widetilde{\boldsymbol{\mathcal{R}}}^s = (\widetilde{\boldsymbol{\mathcal{H}}}^H\widetilde{\boldsymbol{\mathcal{H}}})^s$ with time delays $\tau_k \in [0, T_s]$ converges to the same asymptotic limit as the diagonal elements of the matrix $\widehat{\boldsymbol{R}}^s$ ($\overline{\boldsymbol{R}}^s$) for a symbol quasi synchronous but chip asynchronous CDMA system with time delays $\widetilde{\tau}_k = \tau_k - \left\lfloor \frac{\tau_k}{T_c} \right\rfloor T_c$. This property is established in the following theorem.

Theorem 3 *Given* $\{\tau_1, \tau_2, \dots \tau_K\}$ *a set of reals in* $[0, T_s]$ *let us consider the set of reals in* $[0, T_c]$ *defined as* $\left\{\widetilde{\tau}_k :\ \widetilde{\tau}_k = \tau_k - \left\lfloor \frac{\tau_k}{T_c} T_c \right\rfloor,\ k = 1, \dots K\right\}$. *Let* $\boldsymbol{A}$, $\Phi(j2\pi f)$, $\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_k)$, $\widehat{\boldsymbol{S}}$, *and* $\widehat{\boldsymbol{H}}$ *be defined as in Theorem 1. Furthermore* $\widetilde{\boldsymbol{\Phi}}_k$, $k = 1 \dots K$ *are* $2rN \times N$ *matrices such that* $\widetilde{\boldsymbol{\Phi}}_k = [\boldsymbol{0}_{k,0}^T, \widetilde{\boldsymbol{\Delta}}_{\phi,r}^T(\tau_k), \boldsymbol{0}_{k,1}^T]^T$ *being* $\widetilde{\boldsymbol{\Delta}}_{\phi,r}^T(\tau_k)$ *an* $rN \times N$ *block diagonal matrix with* $\widetilde{\ell}^{\rm th}$ *diagonal block* $\left(\widetilde{\boldsymbol{\Delta}}_{\phi,r}(\tau_k)\right)_{\widetilde{\ell}\widetilde{\ell}} = (\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_k))_{\ell\ell}$, $\ell = \left(N + \widetilde{\ell} + \left\lfloor \frac{\tau_k}{T_c} \right\rfloor\right) \mathrm{mod} N + 1$, $\boldsymbol{0}_{k,0}$ *and* $\boldsymbol{0}_{k,1}$ *zero matrices of dimensions* $r \left\lfloor \frac{\tau_k}{T_c} \right\rfloor \times N$ *and* $\left(N - r \left\lfloor \frac{\tau_k}{T_c} \right\rfloor\right) \times N$, *respectively. Let* $\widetilde{\boldsymbol{S}}[m] = \left(\widetilde{\boldsymbol{\Phi}}_1 \boldsymbol{s}_{1m}, \widetilde{\boldsymbol{\Phi}}_2 \boldsymbol{s}_{2m} \dots \widetilde{\boldsymbol{\Phi}}_K \boldsymbol{s}_{Km}\right)$, $\widetilde{\boldsymbol{H}}[m] = [\widetilde{\boldsymbol{H}}_u^T[m], \widetilde{\boldsymbol{H}}_d^T[m]]^T = \widetilde{\boldsymbol{S}}\boldsymbol{A}$ *and* $\widetilde{\boldsymbol{\mathcal{H}}}$ *the infinite block row and block column matrix of the same form as in (3). Furthermore,* $\widetilde{\boldsymbol{h}}_{km}$ *is the infinite length column vector with non zero elements given by* $|a_{kk}|^2 \widetilde{\boldsymbol{\Phi}}_k \boldsymbol{s}_{km}$. *Let the same assumptions as in Theorem 1 hold.*

Then, given $(|a_{kk}|^2, \widetilde{\tau}_k)$, *the diagonal elements* $\widetilde{\boldsymbol{h}}_{km}^H \widetilde{\boldsymbol{\mathcal{T}}}^\ell \widetilde{\boldsymbol{h}}_{km}$ *of the matrix* $\widetilde{\boldsymbol{\mathcal{R}}}^\ell$ *converges in probability to the deterministic value* $R_\ell(|a_{kk}|^2, \widetilde{\tau}_k)$ *in (24).*

Theorem 3 is proven in Appendix V.

The results of Corollary 1 and Theorem 2 can be readily extended to CDMA systems with transfer matrix $\widetilde{\boldsymbol{\mathcal{H}}}$.

As in Part I Section IV of this article [1], we conjecture the equivalence between a CDMA system with transfer matrix $\boldsymbol{\mathcal{H}}$ and a CDMA system with transfer matrix $\widetilde{\boldsymbol{\mathcal{H}}}$. 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 (thus, chip pulse waveforms that are not square root Nyquist), QPSK modulation, and random spreading sequences with

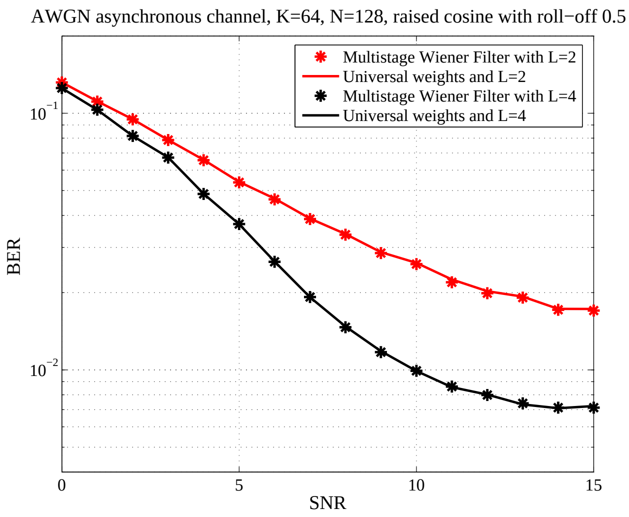

Fig. 2. BER of MSWF with universal weights (solid lines), and weight optimized for the finite system (markers) with $L = 2$ and 4 versus SNR. CDMA systems with equal received powers, raised cosine chip waveforms (roll-off $\gamma = 0.5$), sampling rate $\frac{2}{T_c}$, and system load $\beta = \frac{1}{2}$ are considered.

$N = 128$. Perfect power control was applied, i.e. all users were received with the same power, and sampled at rate $\frac{2}{T_c}$. At the receiver, detection was performed by multistage Wiener filters with either universal weights or weights optimized for the finite system, cf. Figure 2. There is clearly no loss due to the approximation of weights by their asymptotic limits.

The mathematical results presented in this section have important implications on the design and analysis of asynchronous CDMA systems and linear detectors for asynchronous CDMA systems. We elaborate on them in the following section.

VI. Effects of Asynchronism, Chip Pulse Waveforms, and Sets of Observables

The theoretical framework developed in Section V enables the analysis and design of linear multistage detectors for CDMA systems using optimum and suboptimum statistics and possibly non ideal chip pulse waveforms. In this section we focus on the following aspects:

1) Analysis of the effects of chip pulse waveforms and time delay distributions when the multistage detectors are fed by sufficient statistics.
2) Impact of the use of sufficient and suboptimum statistics on the complexity and the performance of multistage detectors.

A. *Sufficient Statistics*

Sufficient statistics impaired by discrete additive Gaussian noise are obtained as output of detector Type A. For chip pulse waveforms with bandwidth $B \leq \frac{1}{2T_c}$ and any set of time delays, Theorem 2 applies. For $B > \frac{1}{2T_c}$ and uniform time delay distribution, Corollary 1 holds. In both cases, as $K, N \to \infty$ with constant ratio the diagonal elements of the matrix $\overline{\boldsymbol{R}}^{\ell}$ and the eigenvalue moments $m_{\overline{\boldsymbol{R}}}^{(\ell)}$ can be obtained from Algorithm 1. As consequence of (16) the performance of the large class of multiuser detectors that admit a representation as multistage detectors depends only on the diagonal elements $\overline{\boldsymbol{R}}_{\ell}(|a_{kk}|^2)$ and the variance of the noise. In large CDMA systems, the SINR depends on the system load β, the sampling rate $\frac{r}{T_c}$, the limit distribution of the received powers $F_{|\boldsymbol{A}|^2}(\lambda)$, the variance of the noise σ^2, the coefficients $\mathcal{E}_{\ell}$, $\ell \in \mathbb{Z}^+$ and the received powers $|a_{kk}|^2$, but it is independent of the time delay τ_k, in general. For $B \leq \frac{1}{2T_c}$, the SINR is also independent of the time delay distribution. Therefore we can state the following corollary

Corollary 2 *If the bandwidth of the chip pulse waveform satisfies the constraint $B \leq \frac{1}{2T_c}$, large synchronous and asynchronous CDMA systems have the same performance in terms of SINR when a linear detector that admits a representation as multistage detector is used at the receiver.*

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, synchronous and asynchronous CDMA systems are have the same performance. In [24] is established the equivalence between synchronous and asynchronous CDMA systems using an ideal Nyquist sinc waveform ($B = \frac{1}{2T_c}$) and linear MMSE detector. Corollary 2 generalize that equivalence to any kind of chip pulse waveforms with bandwidth $B \leq \frac{1}{2T_c}$ and any linear multiuser detector with a multistage representation.

By inspection of Algorithm 1 we can verify that the dependence of $R_{\ell}(|a_{kk}|^2)$ and $m_{\overline{\boldsymbol{R}}}^{(\ell)}$ on the sampling rate $\frac{r}{T_c}$ can be expressed by the following relations

$$R_{\ell}(|a_{kk}|^2) = \left(\frac{r}{T_c}\right)^{\ell} R_{\ell}^{*}(|a_{kk}|^2) \tag{34}$$

and

$$m_{\overline{\boldsymbol{R}}}^{(\ell)} = \left(\frac{r}{T_c}\right)^{\ell} m_{\overline{\boldsymbol{R}}}^{*(\ell)} \tag{35}$$

where $R^{*}_{\ell}(|a_{kk}|^2)$ and $m^{*(\ell)}_{\overline{\boldsymbol{R}}}$ are independent of the sampling rate $\frac{r}{T_c}$. Thanks to this particular dependence and the fact that $\sigma^2 = \frac{r}{T_c}N_0$ the quadratic forms $\boldsymbol{\xi}^H_{k,m}\Xi^{-1}_{k,m}\boldsymbol{\xi}_{k,m}$, $\boldsymbol{\xi}^H_{k,m}\Xi^{-1}\boldsymbol{\xi}$, and $\boldsymbol{\xi}^H\Xi^{-1}\Xi_{k,m}\Xi^{-1}\boldsymbol{\xi}$ are independent of the sampling rate for large systems. Thus, the large system performance of (1) linear multistage detectors optimum in a mean square sense (see (17)), (2) of the polynomial expansion detectors and (3) the matched filters is independent of the sampling rate. This property is not general. Detectors that are not designed to benefit at the best from the available sufficient statistics may improve their performance using different sets of sufficient statistics. Therefore, the large system performance of other multistage detectors like PIC detectors depends on the sampling rate and can be improved increasing the oversampling factor r.

Given a positive real γ, let us consider the chip pulse

$$\Phi(j2\pi f) = \begin{cases} \sqrt{\frac{T_c}{\gamma}} & \text{for } |f| \leq \frac{\gamma}{2T_c}, \\ 0 & \text{otherwise.} \end{cases} \tag{36}$$

corresponding to a sinc waveform with bandwidth $B = \frac{\gamma}{2T_c}$ and unit energy. For waveform (36) with $\gamma = 1$, $T_c = 1$, and $r = 1$ Algorithm 1 reduces to Algorithm 1 in [18] for synchronous systems. Let us denote by $R^{(\text{syn})}_{\ell}(|a_{kk}|^2, \beta)$ and $m^{(\ell)}_{\overline{\boldsymbol{R}}^{(\text{syn})}}(\beta)$ the values of $R_{\ell}(|a_{kk}|^2)$ and $m^{(\ell)}_{\overline{\boldsymbol{R}}}$ for such a case and system load β. Then, in general case for chip pulse waveform (36) Algorithm 1 yields

$$R^{(\text{sinc})}_{\ell}(|a_{kk}|^2) = \left(\frac{r}{T_c}\right)^{\ell} \mathcal{R}^{(\text{syn})}_{\ell}\left(|a_{kk}|^2, \frac{\beta}{\gamma}\right) \tag{37}$$

and

$$m^{(\ell)}_{\overline{\boldsymbol{R}}^{(\text{sinc})}} = \left(\frac{r}{T_c}\right)^{\ell} m^{(\ell)}_{\overline{\boldsymbol{R}}^{(\text{syn})}}\left(\frac{\beta}{\gamma}\right). \tag{38}$$

Therefore, the same property pointed out in part I of this paper [1] for linear MMSE detectors holds for several multistage detectors (namely, multistage Wiener filters, polynomial expansion detectors, matched filters): In a large asynchronous CDMA systems using a sinc function with bandwidth $\frac{\gamma}{2T_c}$ as chip pulse waveform and system load β any multistage detector whose performance is independent of the sampling rate performs as well as in a large synchronous CDMA system with modulation based on square root Nyquist chip pulses and system load $\beta' = \frac{\beta}{\gamma}$.

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 time delays . We elaborate on these aspects focusing on square root raised cosine chip-pulse waveforms with roll-off $\vartheta \in [0, 1]$ and on chip pulse waveforms (36) with $\gamma \in [1, 2]$. To simplify the notation we assume $T_c = 1$. Let

$$S(x) = \begin{cases} 1 & 0 \leq |x| \leq \frac{1-\vartheta}{2} \\ \frac{1}{2}\left(1 - \sin\frac{\pi}{\vartheta}\left(|x| - \frac{1}{2}\right)\right) & \frac{1-\gamma}{2} \leq |x| \leq \frac{1+\vartheta}{2} \\ 0 & |x| \geq \frac{1+\vartheta}{2}. \end{cases}$$

The energy frequency spectrum of a square root raised cosine waveform with unit energy is given by $|\Psi_{\mathrm{sqrc}}(j2\pi x)|^2 = S(x)$. The large system analysis of an asynchronous CDMA system using square root raised cosine chip pulse waveform is obtained applying Algorithm 1. The corresponding coefficients $\mathcal{E}_{\mathrm{sqrc},s}$, $s = \mathbb{Z}^+$, are given by

$$\mathcal{E}_{\mathrm{sqrt},s}{=}2^s(1-\gamma) + 2\int_{\frac{1-\gamma}{2}}^{\frac{1+\gamma}{2}} \sin^s\left(\frac{\pi}{\gamma}\left(\frac{1}{2}{-}x\right)\right)\mathrm{d}x.$$

It is well known that in a synchronous CDMA system the performance is maximized using square root Nyquist waveforms. In this case the performance is independent of the specific waveform and the bandwidth and equals the performance of a large synchronous system using the sinc function with bandwidth $\frac{1}{2T_c}$ as chip pulse. Since the square root raised cosine pulses are square root Nyquist waveforms they attain the maximum SINR in synchronous systems. The large system performance of multistage Wiener filters for synchronous CDMA systems with a square root raised cosine waveform is obtained making use of (17) and Algorithm 1 with $r = 1$ and $\mathcal{E}_s = 1$, $s \in \mathbb{Z}^+$.

In general, chip pulse waveform (36) is not a square root Nyquist waveform. For this reason the performance analysis of linear multistage Wiener filters for synchronous CDMA sytems [14], [18] is not applicable. In this case characterized by interchip interference we can still apply Theorem 1, sampling at rate $\frac{2}{T_c}$ and assuming a Dirac function $f_T(\tau) = \delta(\tau)$ as probability density function of the time delays. For the chip pulse waveform (36) the matrix $\boldsymbol{Q}(x) = \boldsymbol{\Delta}_{\Phi,2}(x,0)\boldsymbol{\Delta}_{\Phi,2}^H(x,0)$ used in the recursion of Theorem 1 is given by

$$\boldsymbol{Q}(x) = \begin{cases} \frac{1}{\gamma}\begin{pmatrix} 1 & \mathrm{e}^{-j\pi x} \\ \mathrm{e}^{j\pi x} & 1 \end{pmatrix} & |x| \leq 1 - \frac{\gamma}{2} \\ \frac{1}{\gamma}\begin{pmatrix} 4 & 0 \\ 0 & 0 \end{pmatrix} & 1 - \frac{\gamma}{2} \leq |x| \leq \frac{1}{2}. \end{cases}$$

The large system analysis in the asynchronous case with chip pulse (36) can be readily performed making use of (17) and (37).

In Figure 3 the large system SINR at the output of a multistage Wiener filter with $L = 4$ is plotted as function of the bandwidth for synchronous and asynchronous CDMA systems based on modulation by square root raised cosine or by pulse (36). We assume perfect power control, i.e. $\boldsymbol{A} = \boldsymbol{I}$, system load $\beta = 0.5$, and input $\mathrm{SNR} = 10$ dB.

Consistently with the theoretical knowledge on synchronous CDMA systems, if the modulation is based on square root raised cosine a synchronous CDMA system outperforms a similar CDMA system with modulation based on (36). Asynchronous CDMA systems with both chip pulse waveforms widely outperform the corresponding synchronous systems. In contrast to the synchronous case, chip pulse waveforms (36) exploits better than square root cosine the additional degrees of freedom introduced by increasing the bandwidth and an asynchronous CDMA system with modulation based on (36) outperforms considerably a system using square root raised cosine pulses. Interestingly, while in synchronous systems avoidance of interchip interference at the output of a chip matched filter is a sensible criterion for chip pulse design and yields optimum chip pulse waveforms, in asynchronous systems the detrimental effects of the interchip interference can be eventually compensated by a reduction of MAI due to a careful chip design and the avoidance of interchip interference is not the driving criterion for the chip pulse optimization.

In Figure 4 the SINR at the output of a multistage Wiener filter with $L = 8$ is plotted as a function of the system load, parametric in the bandwidth, for $\mathrm{SNR} = 10$ dB. The improvement achievable by asynchronous systems over synchronous systems increases as the the system load increases.

B. Chip Rate Sampling

Chip rate sampling is a widely used approach to generate statistics for asynchronous CDMA systems. It implies the use of square root Nyquist chip pulses and makes use of front end Type B. Hereafter, we refer to these CDMA systems as systems B, while we refer to the systems that use sufficient statistics from a front end Type A as systems A.

A bound on the performance of systems B with linear MMSE detectors is in [25]. The performance analysis of linear multistage detectors as $K, N \to \infty$ with $\frac{K}{N} \to \beta$ can be performed applying Theorem 1 to the chip pulse waveform at the output of the chip matched filter $\Phi(j2\pi f) = \frac{1}{\sqrt{E_\psi}} |\Psi(j2\pi f)|^2$ and assuming $r = 1$. In order to elaborate further on systems B we focus on the square root raised cosine chip pulse with roll-off θ [26]

$$\psi(t) = \frac{4\theta(\frac{t}{T_c}) \cos(\pi(1+\theta)\frac{t}{T_c}) + \sin(\pi(1-\theta)\frac{t}{T_c})}{\pi t(1 - (4\theta\frac{t}{T_c})^2)} \qquad \theta \in [0, 1]. \tag{39}$$

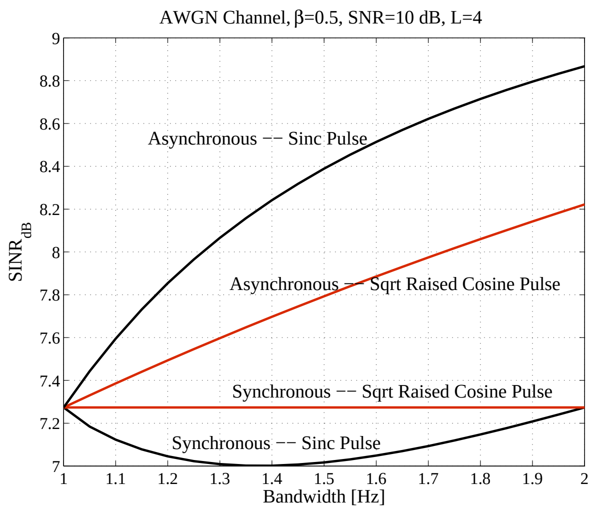

Fig. 3. Output SINR of a multistage Wiener filter with $L = 4$ versus the bandwidth. CDMA systems with equal received powers, square root raised cosine chip waveforms or sinc pulses, system load $\beta = \frac{1}{2}$ and input $\mathrm{SNR} = 10$ dB are considered.

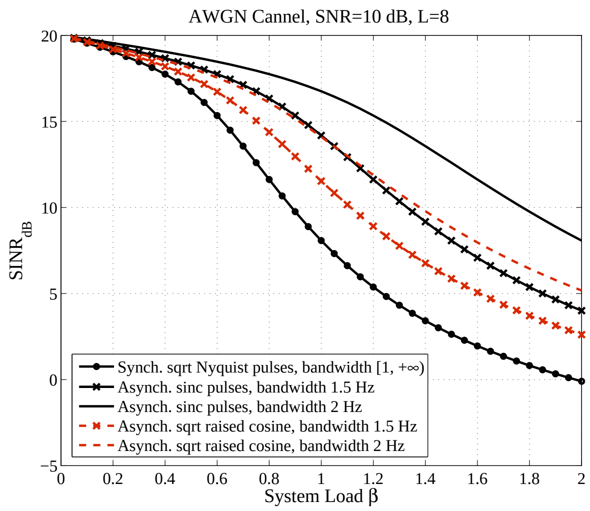

Fig. 4. Output SINR of a multistage Wiener filter with $L = 8$ versus the system load. Asynchronous CDMA systems with equal received powers, square root raised cosine chip waveforms or sinc pulses with bandwidth $B = 1.5, 2$ Hz, input $\mathrm{SNR} = 10$ dB are compared to synchronous CDMA systems with square root Nyquist chip pulses.

In this case, the matrix function $\boldsymbol{Q}(x,\tau) = \boldsymbol{\Delta}_{\phi,1}(x,\tau)\boldsymbol{\Delta}_{\phi,1}^H(x,\tau)$ occurring in Theorem 1 reduces to the scalar function

$$\boldsymbol{Q}(x,\tau) = \begin{cases} \frac{1}{2} + \frac{1}{2}\sin^2\left(\frac{\pi}{\theta}(x+\frac{1}{2})\right) + \frac{\cos 2\pi\tau}{2}\left(1 - \sin^2\left(\frac{\pi}{\theta}(x+\frac{1}{2})\right)\right) & -\frac{1}{2} \le x \le -\frac{1-\theta}{2} \\ 1 & -\frac{1-\theta}{2} \le x \le \frac{1-\theta}{2} \\ \frac{1}{2} + \frac{1}{2}\sin^2\left(\frac{\pi}{\theta}(x-\frac{1}{2})\right) + \frac{\cos 2\pi\tau}{2}\left(1 - \sin^2\left(\frac{\pi}{\theta}(x-\frac{1}{2})\right)\right) & \frac{1-\theta}{2} \le x \le \frac{1}{2}. \end{cases}$$

due to the fact that $r = 1$. Equal received powers, system load $\beta = \frac{1}{2}$, multistage Wiener filters with $L = 3$ define the scenario we consider for the asymptotic analysis.

The analysis shows a strong dependence of the performance on the time delays. As expected, it is possible to verify that the best SINR is obtained when the sampling instants coincide with the time delays of the user of interest.

In Figure 5 we compare the performance of system B with square root raised cosine chip pulse to the SINR of a system A with the same modulating pulse. In the comparison we consider the best SINR for system B obtained when the sampling times coincide with the time delays of the user of interest. The curves represent the output SINR as a function of the roll-off θ parameterized with respect to SNR. The parameter (SNR) varies from 0 dB to 20 dB in steps of 5 dB. As reference we also plot the performance of synchronous

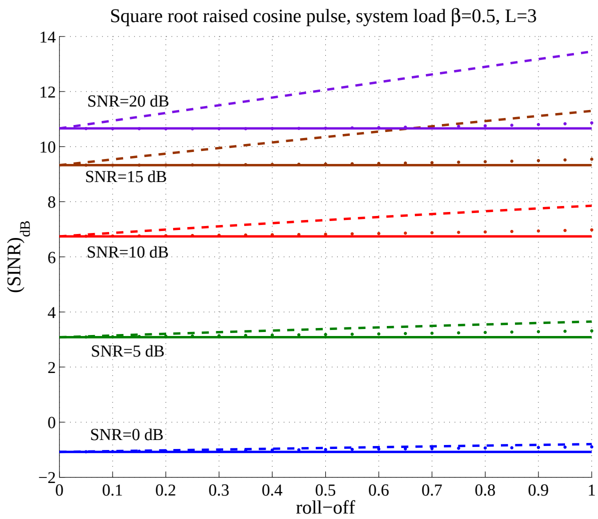

Fig. 5. Asymptotic output SINR of a multistage Wiener filter with $L = 3$ versus the roll-off θ as front-end A (dashed lines) and front-end B (dots) are in use in an asynchronous CDMA system. The solid lines show the reference performance in synchronous CDMA systems. The curves are parametric in the input SNR with SNR varying between 0 dB and 20 dB in steps of 5 dB.

CDMA systems. As expected, multistage detectors with front-end A outperform the corresponding multistage detectors with front-end B.

Interestingly, while linear multistage detectors and asynchronism in system A can compensate to some extent for the loss in spectral efficiency caused by the increasing roll-off and typical of synchronous CDMA systems such a compensation is not possible in systems B. Systems B behave similarly to synchronous CDMA systems. In fact, the SINR for system B is very close to the performance of synchronous systems for any SNR level.

Since the SINR in system B heavily depends on the sampling instants with respect to τ_k, different statistics are needed for the detection of different users in order to obtain good performance. As consequence, joint detection is not feasible and each user has to be detected independently. This is a significant drawback when several or all users have to be detected (e.g. uplink) and has a relevant impact on the complexity of the system. For example, the complexity order per bit of a multistage Wiener filter or polynomial expansion detector is linear in rK in system A while the complexity order per bit of the same detectors is quadratic in K in system B. A similar increase in complexity can be noticed also for other detectors (e.g. linear MMSE detectors, or any multistage detector).

VII. CONCLUSIONS

In Part II of this work we provided guidelines for the design of asynchronous CDMA systems via the analysis of the effects of chip pulse waveforms, time delay distributions, sufficient and suboptimum observables on the complexity and performance of the broad class of multiuser detectors with multistage representation.

Similarly to the results obtained in part I of this article [1], i.e. the chip-pulse constrained capacity and the performance of linear MMSE detectors, multistage detectors show performances independent of the time delays of the active users if the bandwidth of the chip pulse waveform is not greater than half of the chip rate, i.e. $B \leq \frac{1}{2T_c}$. Above that threshold the performances of linear multistage detectors depend on the time delay distributions and asynchronous CDMA systems outperform synchronous CDMA systems.

The framework presented here enabled the analysis of optimum and suboptimum multistage detectors using statistics (A), which are sufficient, or observables (B), which are suboptimum. In the two cases of (i) chip pulses with bandwidth $B \leq \frac{1}{2T_c}$ and (ii) chip pulses with bandwidth $B > \frac{1}{2T_c}$, sufficient statistics, and uniform distribution, the effects of the chip pulse waveforms on the detector performance are described by the coefficients $\mathcal{E}_s = \frac{1}{T_c^s} \int_{-B}^{B} T_c |\Psi(j2\pi f)|^{2s} \mathrm{d}f$. The output SINR of linear MMSE detectors, multistage Wiener filters, polynomial expansion detectors, and matched filters is independent of the sampling rate. In contrast, the output SINR of other multistage detectors like PIC detectors depends on the sampling rate and increases with it.

Comparing the performance of synchronous and asynchronous CDMA systems with modulation based on square root Nyquist pulses, namely square root raised cosine waveforms, and modulation based on sinc functions with increasing bandwidth, it becomes apparent that the chip pulse design for synchronous CDMA systems follows the same guidelines as the chip pulse design for single user systems. In contrast, chip pulse design for asynchronous CDMA systems is governed by entirely different rules. In fact, CDMA systems we found to perform well if the spectrum of the received signal is as white as possible.

The asymptotic analysis of asynchronous CDMA systems using statistics (B) shows that the performance of multistage Wiener filters is close to the SINR of the corresponding synchronous CDMA systems for any bandwidth and level of SNR. Therefore, this kind of front-end is not capable of exploiting the benefits of asynchronous CDMA.

The universal weights proposed for the design of low complexity detectors account for the effects of asynchronism, sub-optimality of the statistics, and non-ideality of pulse-shapers. They depend on the sampling rate although the large system performance do not.

From the asymptotic analysis and design performed in this work we can draw the following conclusions:

- Multistage detectors with front end Type B and universal weights are asymptotically suboptimal and have the same complexity order per bit $\mathcal{O}(K^2)$ in uplink as the linear MMSE detector.
- Multistage Wiener filters and polynomial expansion detectors with statistics A and universal weights are asymptotically optimum and have the same complexity order per bit as the matched filter, i.e. $\mathcal{O}(rK)$ with $r \ll K$.
- If only a user has to be detected, multistage detectors using statistics (B) have slightly lower complexity than multistage detectors with statistics (A), namely they have a complexity per bit $\mathcal{O}(K^2)$ while in the later case the complexity per bit is $\mathcal{O}(rK^2)$. However, they perform almost as the multistage detectors for synchronous systems at any SNR and do not provide the gain in performance due to asynchronism in contrast to statistics (A).

ACKNOWLEDGMENT

The authors thank Alex Grant and Dirk Slock for useful discussions.

APPENDIX I

PROOF OF THEOREM 1

By applying the same approach as in part I Theorem 1 of this paper [1] it can be shown that the eigenvalue moments of the matrix $\overline{\boldsymbol{R}} = \boldsymbol{A}^H \overline{\boldsymbol{S}}^H \overline{\boldsymbol{S}} \boldsymbol{A} = \overline{\boldsymbol{H}}^H \overline{\boldsymbol{H}}$ and $\widehat{\boldsymbol{R}} = \boldsymbol{A}^H \widehat{\boldsymbol{S}}^H \widehat{\boldsymbol{S}} \boldsymbol{A} = \widehat{\boldsymbol{H}}^H \widehat{\boldsymbol{H}}$ coincide. The same property holds for the diagonal elements of the matrices $\overline{\boldsymbol{R}}^\ell$ and $\widehat{\boldsymbol{R}}^\ell$ with $\ell \in \mathbb{Z}^+$.

In the following we focus on the asymptotic analysis of the diagonal elements of the matrices $\widehat{\boldsymbol{R}}^\ell = (\widehat{\boldsymbol{H}}^H \widehat{\boldsymbol{H}})^\ell$ with $\widehat{\boldsymbol{H}} = (\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_1)\boldsymbol{s}_1, \boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_2)\boldsymbol{s}_2, \ldots, \boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_K)\boldsymbol{s}_K)\boldsymbol{A}$.

Throughout this proof we adopt the following notation. For $k = 1, \ldots, K$ and $n = 1, \ldots, N$

- $\widehat{\boldsymbol{h}}_k$ is the k^{th} column of the matrix $\widehat{\boldsymbol{H}}$;
- $\widehat{\boldsymbol{h}}_{nk}$ is the n^{th} $r \times 1$ block of the vector $\widehat{\boldsymbol{h}}_k$ and $\widehat{\boldsymbol{h}}_{nk} = a_{kk}(\boldsymbol{\Delta}_{\phi,r})_{nn} s_{nk}$;
- $\widehat{\boldsymbol{\delta}}_n$ is the n^{th} block row of $\widehat{\boldsymbol{H}}$ of dimensions $r \times K$;
- $\widehat{\boldsymbol{H}}_{\vDash n}$ is the matrix obtained from $\widehat{\boldsymbol{H}}$ by suppressing $\widehat{\boldsymbol{\delta}}_n$;
- $\widehat{\boldsymbol{H}}_{\sim k}$ is the matrix obtained from $\widehat{\boldsymbol{H}}$ by suppressing $\widehat{\boldsymbol{h}}_k$.
- $\widehat{\boldsymbol{T}}_{\sim k} = \widehat{\boldsymbol{H}}_{\sim k} \widehat{\boldsymbol{H}}_{\sim k}^H$;

- $\widehat{\boldsymbol{R}}_{\vDash n} = \widehat{\boldsymbol{H}}^H_{\vDash n} \widehat{\boldsymbol{H}}_{\vDash n}$;
- $\widehat{\boldsymbol{\sigma}}_n = (s_{n1}, s_{n2}, \dots, s_{nK})$.
- $\boldsymbol{\nabla}_{n,t}$, for $t = 1, \dots, r$ and $n = 1, \dots, N$, is the $K \times K$ diagonal matrix with the k^{th} element equal to $\phi\left(\frac{n-1}{N}, \widetilde{\tau}_k - \frac{(t-1)T_c}{r}\right)$. Note that $\widehat{\boldsymbol{\sigma}}_n \boldsymbol{\nabla}_{n,t} \boldsymbol{A}$ coincides with the $(t + (n-1)r)^{\text{th}}$ row of the matrix $\widehat{\boldsymbol{H}}$.
- $\widehat{\boldsymbol{T}}^s_{[nn]}$ is the n^{th} diagonal block of $\widehat{\boldsymbol{T}}^s$ of dimensions $r \times r$.

Furthermore, since the channel gains a_{kk} are bounded, we denote by a_{MAX} their upper bound, i.e. $|a_{kk}| < a_{\text{MAX}}, \forall k$. Finally, thanks to the assumption that $\Phi(j2\pi f)$ is bounded in absolute value with finite support also $\phi(x, \tau)$ is upper bounded for any x and τ. We denote by Φ_{MAX} its bound.

Let us observe first that the eigenvalue moments of the matrix $\widehat{\boldsymbol{R}}$ (or equivalently of $\widehat{\boldsymbol{T}}$) are almost surely upper bounded by a finite positive values $C^{(s)}$,

In fact,

$$\begin{aligned}
\frac{1}{N}\text{tr}\widehat{\boldsymbol{R}}^s &= \frac{1}{N} \sum_{k_1,\dots k_s=1}^{K} \sum_{n_1,\dots n_s=1}^{N} \widehat{\boldsymbol{h}}^H_{n_1,k_1} \widehat{\boldsymbol{h}}_{n_1,k_2} \widehat{\boldsymbol{h}}^H_{n_2,k_2} \widehat{\boldsymbol{h}}_{n_2,k_3} \dots \widehat{\boldsymbol{h}}^H_{n_s,k_s} \widehat{\boldsymbol{h}}_{n_s,k_1} \\
&= \frac{1}{N} \sum_{k_1,\dots k_s=1}^{K} |a_{k_1k_1}|^2 \dots |a_{k_sk_s}|^2 \sum_{n_1,\dots n_s=1}^{N} (\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_1)^H_{n_1n_1} (\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_2)_{n_1n_1} \dots (\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_s)^H_{n_sn_s} (\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_1)_{n_sn_s} \times \\
&\times s^*_{n_1,k_1} s_{n_1,k_2} s^*_{n_2,k_2} s_{n_2,k_3} \dots s^*_{n_s,k_s} s_{n_s,k_1}
\end{aligned}$$

Applying the approach of non-crossing partitions [27], [28], it is possible to recognize that the factors $s^*_{n_1,k_1} s_{n_1,k_2} s^*_{n_2,k_2} s_{n_2,k_3} \dots s^*_{n_s,k_s} s_{n_s,k_1}$ which do not vanish asymptotically, correspond to the ones having nonzero non-crossing partitions. Correspondingly, also the remaining factors

$$(\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_1)^H_{n_1n_1} (\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_2)_{n_1n_1} \dots (\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_s)^H_{n_sn_s} (\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_1)_{n_sn_s}$$

are positive and bounded by

$$|(\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_1)^H_{n_1n_1} (\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_2)_{n_1n_1} \dots (\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_s)^H_{n_sn_s} (\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_1)_{n_sn_s}| \leq \frac{r^{2s} \Delta^{2s}_{\text{MAX}}}{T_c^{2s}}.$$

Therefore,

$$\frac{1}{N}\text{Tr}\widehat{\boldsymbol{R}}^s \leq \frac{r^{2s} \Delta_{\text{MAX}} a^{2s}_{\text{MAX}}}{T_c^{2s}} \left(\frac{1}{N} \sum_{k_1,\dots k_s=1}^{K} \sum_{n_1,\dots n_s=1}^{N} s^*_{n_1,k_1} s_{n_1,k_2} s^*_{n_2,k_2} s_{n_2,k_3} \dots s^*_{n_s,k_s} s_{n_s,k_1} \right). \tag{40}$$

The last factor in (40) is the s-th eigenvalue moment of a central Wishart matrix with zeromean i.i.d Gaussian entries having variance $\frac{1}{N}$. Well established results of random matrix theory [29], [28], [12] show that the eigenvalue moments of such a matrix converge almost surely to finite values. More specifically,

$$\frac{1}{N}\sum_{n_1,\dots n_s=1}^{N} s^*_{n_1,k_1}s_{n_1,k_2}s^*_{n_2,k_2}s_{n_2,k_3}\dots s^*_{n_s,k_s}s_{n_s,k_1} \xrightarrow{a.s.} \sum_{i=0}^{s-1}\begin{pmatrix} s \\ i \end{pmatrix}\begin{pmatrix} s \\ i+1 \end{pmatrix}\frac{\beta^i}{s}. \tag{41}$$

Then, appealing to (40) and (41), the eigenvalue moments of the matrices $\widehat{\boldsymbol{R}}$ and $\widehat{\boldsymbol{T}}$ are upper bounded almost surely by

$$C^{(s)} = \frac{r^{2s}\Delta^{2s}_{\mathrm{MAX}a_{\mathrm{MAX}}}}{T_c^{2s}}\sum_{i=0}^{s-1}\begin{pmatrix} s \\ i \end{pmatrix}\begin{pmatrix} s \\ i+1 \end{pmatrix}\frac{\beta^i}{s}. \tag{42}$$

The proof of Theorem 1 is based on strong induction. In the first step we prove the following facts:

1) The diagonal elements of the matrix $\widehat{\boldsymbol{R}}$ converge in probability, as $N \rightarrow \infty$, to deterministic values $R_1(|a_{kk}|^2, \widetilde{\tau}_k)$, conditionally on $(|a_{kk}|^2, \widetilde{\tau}_k)$. Furthermore, $\forall \varepsilon > 0$ and large $K = \beta N$

$$\Pr\{|\widehat{\boldsymbol{R}}_{kk} - R_1(|a_{kk}|^2, \widetilde{\tau}_k)| > \varepsilon\} \leq o\left(N^{-2}\right).$$

2) $\widehat{\boldsymbol{T}}_{[nn]}$, the $r \times r$ block diagonal elements of the matrix $\widehat{\boldsymbol{T}} = \widehat{\boldsymbol{H}}\widehat{\boldsymbol{H}}^H$, converge in probability to deterministic blocks $\boldsymbol{T}_1(x)$, with $x = \lim_{N\rightarrow\infty}\frac{n}{N}$. Additionally, $\forall \varepsilon > 0$, large $K = \beta N$ and $u, v = 1, \dots r$,

$$\Pr\{|(\widehat{\boldsymbol{T}}_{[nn]})_{uv} - (\boldsymbol{T}_1(x))_{uv}| > \varepsilon\} \leq o\left(N^{-2}\right).$$

Then, in the recursion step, we use the following induction assumptions:

1) For $s = 1, \dots, \ell-1$, the diagonal elements of the matrix $\widehat{\boldsymbol{R}}^s$, converge in probability, as $K = \beta N \rightarrow \infty$, to deterministic values $R_s(|a_{kk}|^2, \widetilde{\tau}_k)$, conditionally on $(|a_{kk}|^2, \widetilde{\tau}_k)$. Additionally, $\forall \varepsilon > 0$ and large $K = \beta N$, $\Pr\{|(\widehat{\boldsymbol{R}}^s)_{kk} - R_s(|a_{kk}|^2, \widetilde{\tau}_k)| > \varepsilon\} \leq o\left(N^{-2}\right)$.
2) For $s = 1, \dots, \ell-1$, $\widehat{\boldsymbol{T}}^s_{[nn]}$, the $r \times r$ block diagonal elements of the matrix $\widehat{\boldsymbol{T}}^s$ converge in probability to deterministic blocks $\boldsymbol{T}_s(x)$, with $x = \lim_{N\rightarrow\infty}\frac{n}{N}$. Additionally, $\forall \varepsilon > 0$, large $K = \beta N$, and $u, v = 1, \dots r$, $\Pr\{|(\widehat{\boldsymbol{T}}^s_{[nn]})_{uv} - (\boldsymbol{T}_s(x))_{uv}| > \varepsilon\} \leq o\left(N^{-2}\right)$.

We prove:

1) The diagonal elements of the matrix $\widehat{\boldsymbol{R}}^\ell$, converge in probability, as $K = \beta N \rightarrow \infty$, to deterministic values $R^\ell(|a_{kk}|^2, \widetilde{\tau}_k)$, conditionally on $(|a_{kk}|^2, \widetilde{\tau}_k)$. Furthermore, $\forall \varepsilon > 0$ and large $K = \beta N$

$$\Pr\{|(\widehat{\boldsymbol{R}}^\ell)_{kk} - R_\ell(|a_{kk}|^2, \widetilde{\tau}_k)| > \varepsilon\} \leq o\left(N^{-2}\right). \tag{43}$$

2) The blocks $\widehat{\boldsymbol{T}}^\ell_{[nn]}$, converge in probability to deterministic blocks $\boldsymbol{T}^\ell(x)$ with $\lim_{N\rightarrow\infty}\frac{n}{N}$. Additionally, $\forall \varepsilon > 0$, large N and $u, v = 1, \dots r$,

$$\Pr\{|(\widehat{\boldsymbol{T}}^\ell_{[nn]})_{uv} - (\boldsymbol{T}_\ell(x))_{uv}| > \varepsilon\} \leq o\left(N^{-2}\right). \tag{44}$$

First step: Consider $\widehat{\boldsymbol{R}}_{kk} = \widehat{\boldsymbol{h}}_k^H \widehat{\boldsymbol{h}}_k = |a_{kk}|^2 \boldsymbol{s}_k^H \boldsymbol{\Delta}_{\phi,r}^H(\widetilde{\tau}_k)\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_k)\boldsymbol{s}_k$. Thanks to the bound $|\phi(x,\tau)| < \Phi_{\mathrm{MAX}}$ which holds for any x and τ, also the eigenvalues of the matrix $\boldsymbol{\Delta}_{\phi,r}^H(\widetilde{\tau})\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau})$ are upper bounded. In fact, they are given by $\sum_{t=1}^{r}\left|\phi\left(\frac{n-1}{N}, \widetilde{\tau}_k - \frac{(t-1)T_c}{r}\right)\right|^2$ for $n = 1, \ldots, N$. Therefore, the limit eigenvalue distribution of the matrix $\boldsymbol{\Delta}_{\phi,r}^H(\widetilde{\tau})\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau})$ has upper bounded support Δ_{MAX}. Then, by appealing to Lemma 9 in part I [1] with $p = 4$ and by making use of the bound for any Hermitian matrix $\boldsymbol{C} \in \mathbb{C}^{N\times N}$ $(\mathrm{tr}\boldsymbol{C})^2 \leq N\mathrm{tr}(\boldsymbol{C}^2)$ we obtain

$$\begin{aligned}
\zeta_1 &= \mathrm{E}\left||a_{kk}|^2 \boldsymbol{s}_k^H \boldsymbol{\Delta}_{\phi,r}^H(\widetilde{\tau}_k)\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_k)\boldsymbol{s} - \frac{|a_{kk}|^2}{N}\mathrm{tr}(\boldsymbol{\Delta}_{\phi,r}^H(\widetilde{\tau}_k)\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_k))\right|^4 \\
&\leq \frac{K_4|a_{kk}|^4}{N^3}\mathrm{tr}(\boldsymbol{\Delta}_{\phi,r}^H(\widetilde{\tau}_k)\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_k))^4 \\
&\leq \frac{K_4|a_{kk}|^4}{N^2}\Delta_{\mathrm{MAX}}^4.
\end{aligned}$$

Since $|a_{kk}| \leq a_{\mathrm{MAX}} < +\infty$, the Bienaymé inequality yields $\forall \varepsilon > 0$

$$\begin{aligned}
\Pr\left\{\left|\widehat{\boldsymbol{R}}_{kk} - \frac{|a_{kk}|^2}{N}\mathrm{tr}(\boldsymbol{\Delta}_{\phi,r}^H(\widetilde{\tau}_k)\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_k))\right| \geq \varepsilon\right\} &\leq \frac{\mathrm{E}\left|\widehat{\boldsymbol{R}}_{kk} - \frac{|a_{kk}|^2}{N}\mathrm{tr}(\boldsymbol{\Delta}_{\phi,r}^H(\widetilde{\tau}_k)\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_k))\right|^4}{\varepsilon^4} \\
&\leq \frac{K_4|a_{kk}|^4\Delta_{\mathrm{MAX}}^4}{N^2\varepsilon^4}
\end{aligned} \tag{45}$$

This bound implies the following convergence in probability[2]

$$\begin{aligned}
R_1(\lambda,\tau)|_{(\lambda,\tau)=(|a_{kk}|^2,\tau_k)} &= \lim_{K=\beta N\to\infty}\widehat{\boldsymbol{R}}_{kk} \\
&= \lim_{K=\beta N\to\infty}\frac{|a_{kk}|^2}{N}\mathrm{tr}(\boldsymbol{\Delta}_{\phi,r}^H(\widetilde{\tau}_k)\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_k)) \\
&= \lim_{K=\beta N\to\infty}\frac{|a_{kk}|^2}{N}\sum_{\ell=1}^{N}(\boldsymbol{\Delta}_{\phi,r}^H(\widetilde{\tau}_k))_{\ell,\ell}(\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_k))_{\ell,\ell} \\
&= \lambda\int_0^1 \boldsymbol{\Delta}_{\phi,r}^H(x,\tau)\boldsymbol{\Delta}_{\phi,r}(x,\tau)\mathrm{d}\,x\bigg|_{(\lambda,\tau)=(|a_{kk}|^2,\widetilde{\tau}_k)}.
\end{aligned} \tag{46}$$

Furthermore, thanks to the bound (45) $\forall\varepsilon > 0$

$$\Pr\left\{\left|\widehat{\boldsymbol{R}}_{kk} - R_1(|a_{kk}|^2,\widetilde{\tau}_k)\right| \geq \varepsilon\right\} \leq o(N^{-2}).$$

Let us now consider the block matrix $\widehat{\boldsymbol{T}}_{[nn]}$ whose (u,v) element $(\widehat{\boldsymbol{T}}_{[nn]})_{uv}$ is given by

$$(\widehat{\boldsymbol{T}}_{[nn]})_{uv} = \widehat{\boldsymbol{\sigma}}_n\boldsymbol{A}\boldsymbol{\nabla}_{n,u}\boldsymbol{\nabla}_{n,v}^H\boldsymbol{A}^H\widehat{\boldsymbol{\sigma}}_n^H.$$

[2] In this case it is easy to show also the convergence with probability one or almost sure convergence.

Thanks to the assumption of Theorem 1 that the support of $F_{|\boldsymbol{A}|^2,T}(\lambda,\tau)$ is bounded and $\phi(x,\tau)$ is bounded in absolute value, the diagonal elements of the diagonal matrix $\boldsymbol{A}\boldsymbol{\nabla}_{n,u}\boldsymbol{\nabla}_{n,v}^H\boldsymbol{A}^H$ are upper bounded in absolute value by a positive constant T_{MAX}. Then, by appealing to Lemma 9 in part I [1] we obtain

$$\begin{aligned}\mathrm{E}\left(\left|(\widehat{\boldsymbol{T}}_{[nn]})_{u,v}-\frac{1}{N}\mathrm{tr}\boldsymbol{A}\boldsymbol{\nabla}_{n,u}\boldsymbol{\nabla}_{n,v}^H\boldsymbol{A}^H\right|^4\right) &\leq \frac{K_4}{N^3}\mathrm{tr}(\boldsymbol{A}\boldsymbol{\nabla}_{n,u}\boldsymbol{\nabla}_{n,v}^H\boldsymbol{A}^H)^4\\ &\leq \frac{K_4}{N^2}T_{\mathrm{MAX}}^4. \end{aligned}\tag{47}$$

By appealing again to the Bienaymé inequality and by making use of the bound (47) we obtain $\forall\varepsilon>0$

$$\begin{aligned}\Pr\left\{\left|(\widehat{\boldsymbol{T}}_{[nn]})_{u,v}-\frac{1}{N}\mathrm{tr}(\boldsymbol{A}\boldsymbol{\nabla}_{n,u}\boldsymbol{\nabla}_{n,v}^H\boldsymbol{A}^H)\right|>\varepsilon\right\} &\leq \frac{1}{\varepsilon^4}\mathrm{E}\left(\left|(\widehat{\boldsymbol{T}}_{[nn]})_{u,v}-\frac{1}{N}\mathrm{tr}(\boldsymbol{A}\boldsymbol{\nabla}_{n,u}\boldsymbol{\nabla}_{n,v}^H\boldsymbol{A}^H)\right|^4\right)\\ &\leq \frac{K_4T_{\mathrm{MAX}}^4}{\varepsilon^4N^2}.\end{aligned}$$

Thus, the following convergence in probability holds

$$\begin{aligned}\lim_{K=\beta N\to\infty}(\widehat{\boldsymbol{T}}_{[nn]})_{u,v} &= \lim_{K=\beta N\to\infty}\frac{1}{N}\mathrm{tr}\boldsymbol{A}\boldsymbol{\nabla}_{n,u}\boldsymbol{\nabla}_{n,v}^H\boldsymbol{A}^H\\ &= \lim_{K=\beta N\to\infty}\frac{\beta}{K}\sum_{k=1}^{K}|a_{kk}|^2\phi\left(\frac{n-1}{N},\widetilde{\tau}_k-\frac{u-1}{r}T_c\right)\phi^*\left(\frac{n-1}{N},\widetilde{\tau}_k-\frac{v-1}{r}T_c\right)\\ &= \beta\int\lambda\phi\left(x,\tau-\frac{u-1}{r}T_c\right)\phi\left(x,\tau-\frac{v-1}{r}T_c\right)\mathrm{d}\,F_{|\boldsymbol{A}|^2,T}(\lambda,\tau),\end{aligned}\tag{48}$$

with $x=\lim_{N\to\infty}\frac{n}{N}$ and $0\leq x\leq 1$. Therefore, the block matrix $\widehat{\boldsymbol{T}}_{[nn]}$ converges in probability and in mean square sense to the $r\times r$ matrix

$$\begin{aligned}\boldsymbol{T}_1(x) &= \lim_{K=\beta N\to\infty}\widehat{\boldsymbol{T}}_{[nn]}\\ &= \beta\int\lambda\boldsymbol{\Delta}_{\phi,r}(x,\tau)\boldsymbol{\Delta}_{\phi,r}^H(x,\tau)\mathrm{d}\,F_{|\boldsymbol{A}|^2,T}(\lambda,\tau)\end{aligned}$$

with $0\leq x\leq 1$. Thanks to the bound (47) for large $K=\beta N$ and $\forall\varepsilon>0$ the bound

$$\Pr\left\{\left|(\widehat{\boldsymbol{T}}_{[nn]})_{u,v}-(\boldsymbol{T}(x))_{u,v}\right|<\varepsilon\right\}\leq o(N^{-2})$$

holds. This concludes the proof of the first step.

Step ℓ: First of all, let us introduce some properties of the convergence in probability.

Property A: Let us consider a finite number q of random sequences $\{a_n^{(1)}\},\dots,\{a_n^{(q)}\}$ that converge in probability to deterministic limits $a_1,\dots,a_q$, respectively. Then, any linear combination of such sequences converges in probability to the linear combination of the limits. Furthermore, if $|a_n^{(s)}-a_s|\xrightarrow{\mathcal{P}}o(N^{-i_s})$, with $i_s \in \mathbb{R}^+$, and $s = 1, \dots q$, then any linear combination of the random sequences converges as $o(N^{-\min_{s=1,\dots q}(i_s)})$, at worst.

Property B: Let $\{a_n\}$ and $\{b_n\}$ two random sequences that converge in probability to a and b, respectively. Then, the sequence $\{a_n b_n\}$ converges in probability to ab. In fact, $\forall \varepsilon > 0$

$$\begin{aligned}
\Pr\{|a_n b_n - ab| > \varepsilon\} &= \Pr\{|(a_n - a)(b_n - b) + a(b_n - b) + b(a_n - a)| > \varepsilon\} \\
&\leq \Pr\left\{|(a_n - a)(b_n - b)| > \frac{\varepsilon}{3}\right\} + \Pr\left\{|a(b_n - b)| > \frac{\varepsilon}{3}\right\} + \Pr\left\{|b(a_n - a)| > \frac{\varepsilon}{3}\right\} \\
&\leq \Pr\left\{|(a_n - a)| > \sqrt{\frac{\varepsilon}{3}}\right\} + \Pr\left\{|(b_n - b)| > \sqrt{\frac{\varepsilon}{3}}\right\} + \Pr\left\{|(b_n - b)| > \frac{\varepsilon}{3|a|}\right\} \\
&\quad + \Pr\left\{|(a_n - a)| > \frac{\varepsilon}{3|b|}\right\}. \qquad (49)
\end{aligned}$$

Because of the convergence in probability of a_n and b_n the right hand side in (49) vanishes as $n \to \infty$ and this proves the convergence in probability of the random sequence $\{a_n b_n\}$.

Property C: If for large n, $\Pr\{|a_n - a| > \varepsilon\} \leq o(n^{-s})$ and $\Pr\{|b_n - b| > \varepsilon\} \leq o(n^{-t})$, with $s, t \in \mathbb{R}^+$, then also $\Pr\{|(a_n - a)(b_n - b)| > \varepsilon\} \leq o(n^{-\min(s,t)})$, at worst.

Thanks to the convergence of the diagonal elements of $\widehat{\boldsymbol{R}}^s$ and of the diagonal $r \times r$ blocks of $\widehat{\boldsymbol{T}}^s$, for $s = 1, \dots, \ell - 1$ the following convergence in probability holds:

$$\begin{aligned}
\lim_{K=\beta N \to \infty} \frac{\mathrm{tr} \boldsymbol{A} \boldsymbol{\nabla}_{n,u} \widehat{\boldsymbol{R}}^s_{\vDash n} \boldsymbol{\nabla}^H_{n,v} \boldsymbol{A}^H}{N} &= \lim_{K=\beta N \to \infty} \sum_{k=1}^{K} \frac{|a_{kk}|^2}{N} \phi\left(\frac{n-1}{N}, \widetilde{\tau}_k - \frac{u-1}{r} T_c\right) \phi^*\left(\frac{n-1}{N}, \widetilde{\tau}_k - \frac{v-1}{r} T_c\right) (\widehat{\boldsymbol{R}}^s_{\vDash n})_{kk} \\
&= \beta \int \lambda \phi\left(x, \tau - \frac{u-1}{r} T_c\right) \phi^*\left(x, \tau - \frac{v-1}{r} T_c\right) R_s(\lambda, \tau) \mathrm{d} F_{|\boldsymbol{A}|^2, T}(\lambda, \tau)
\end{aligned} \qquad (50)$$

with $x = \lim_{N \to \infty} \frac{n-1}{N}$, $s = 1, \dots \ell - 1$ and

$$R_s(\lambda, \tau)|_{(\lambda,\tau)=(|a_{kk}|^2, \widetilde{\tau}_k)} = \lim_{K=\beta N \to \infty} (\widehat{\boldsymbol{R}}^s)_{kk} + o(N^{-2}) \qquad (51)$$

as from the recursion assumptions. Furthermore,

$$\begin{aligned}
\lim_{K=\beta N \to \infty} \frac{|a_{kk}|^2}{N} \mathrm{tr} \boldsymbol{\Delta}^H_{\phi,r}(\widetilde{\tau}_k) \widehat{\boldsymbol{T}}^s_{\sim k} \boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_k) &= \lim_{K=\beta N \to \infty} \frac{|a_{kk}|^2}{N} \sum_{n=1}^{N} (\boldsymbol{\Delta}^H_{\phi,r}(\widetilde{\tau}_k))_{nn} (\widehat{\boldsymbol{T}}^s)_{nn} (\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_k))_{nn} \\
&= \lambda \int_0^1 \boldsymbol{\Delta}^H_{\phi,r}(x, \tau) \boldsymbol{T}_s(x) \boldsymbol{\Delta}_{\phi,r}(x, \tau) \mathrm{d}\, x \bigg|_{(\lambda,\tau)=(|a_{kk}|^2, \widetilde{\tau}_k)}
\end{aligned} \qquad (52)$$

with $s = 1, \dots \ell - 1$ and

$$\boldsymbol{T}_s(x) = \lim_{K=\beta N \to \infty} (\widehat{\boldsymbol{T}}^s)_{nn}. \qquad (53)$$

In fact, for (50) we can write

$$\zeta_2 = \Pr\left\{\left|\frac{1}{N}\mathrm{tr}\boldsymbol{A}\boldsymbol{\nabla}_{n,u}\widehat{\boldsymbol{R}}^s_{\vDash n}\boldsymbol{\nabla}^H_{n,v}\boldsymbol{A}^H - \frac{1}{N}\sum_{k=1}^{K}|a_{kk}|^2\phi\left(\frac{n-1}{N},\widetilde{\tau}_k-\frac{u-1}{r}T_c\right)\phi^*\left(\frac{n-1}{N},\widetilde{\tau}_k-\frac{v-1}{r}T_c\right)R_s(|a_{kk}|^2,\widetilde{\tau}_k)\right| > \varepsilon\right\}$$
$$\leq \zeta_{2a}+\zeta_{2b}$$

where

$$\zeta_{2a} = \Pr\left\{\left|\frac{1}{N}\mathrm{tr}\boldsymbol{A}\boldsymbol{\nabla}_{n,u}(\widehat{\boldsymbol{R}}^s - \widehat{\boldsymbol{R}}^s_{\vDash n})\boldsymbol{\nabla}^H_{n,v}\boldsymbol{A}^H\right| > \frac{\varepsilon}{2}\right\}$$

and

$$\zeta_{2b} = \Pr\left\{\left|\frac{1}{N}\sum_{k=1}^{K}|a_{kk}|^2\phi\left(\frac{n-1}{N},\widetilde{\tau}_k-\frac{u-1}{r}T_c\right)\phi^*\left(\frac{n-1}{N},\widetilde{\tau}_k-\frac{v-1}{r}T_c\right)\left((\widehat{\boldsymbol{R}}^s)_{kk}-R_s(|a_{kk}|^2,\widetilde{\tau}_k)\right)\right| > \frac{\varepsilon}{2}\right\}.$$

Note that

$$\zeta_{2a} \leq \Pr\left\{\left|\frac{1}{K}\mathrm{tr}(\widehat{\boldsymbol{R}}^s - \widehat{\boldsymbol{R}}^s_{\vDash n})\right| > \frac{\varepsilon}{2\beta a^2_{\mathrm{MAX}}\phi^2_{\mathrm{MAX}}}\right\}.$$

The expansion of the matrix $\widehat{\boldsymbol{R}}^s = (\widehat{\boldsymbol{R}}_{\vDash n} + \widehat{\boldsymbol{\delta}}^H_n\widehat{\boldsymbol{\delta}}_n)^s$ yields

$$\mathrm{tr}\widehat{\boldsymbol{R}}^s = \mathrm{tr}\widehat{\boldsymbol{R}}^s_{\vDash n} + \sum_{\substack{(i_0,i_1,\dots i_{s-1})\\ i_0+\sum_{j=1}^{s-1}(j+1)i_j=s_0}} \varphi(i_0,i_1,\dots i_{s-1})\prod_{u=0}^{s-1}\left(\widehat{\boldsymbol{\delta}}^H_n\widehat{\boldsymbol{R}}^u_{\vDash n}\widehat{\boldsymbol{\delta}}_n\right)^{i_u}$$

where $\varphi(i_0,i_1,\dots i_{s-1}) \leq 2^s$ is the number of the terms of the expansion of $\widehat{\boldsymbol{R}}^s$ whose trace equals $\prod_{u=0}^{s-1}\left(\widehat{\boldsymbol{\delta}}^H_n\widehat{\boldsymbol{R}}^u_{\vDash n}\widehat{\boldsymbol{\delta}}_n\right)^{i_u}$. Then,

$$\zeta_{2a} \leq 2^s \sum_{\substack{(i_0,i_1,\dots i_{s-1})\\ i_0+\sum_{j=1}^{s-1}(j+1)i_j=s_0}} \Pr\left\{\frac{1}{N}\prod_{u=0}^{s-1}\left(\widehat{\boldsymbol{\delta}}^H_n\widehat{\boldsymbol{R}}^u_{\vDash n}\widehat{\boldsymbol{\delta}}_n\right)^{i_u} > \frac{\varepsilon}{\beta a^4_{\mathrm{MAX}}\phi^4_{\mathrm{MAX}}2^{s+1}}\right\}$$

Thanks to Property B on the convergence in probability, ζ_{2a} converges in probability with rate $o(N^{-2-\frac{4}{s}})$ if $\forall\varepsilon>0$,

$$\lim_{K=\beta N\to\infty}\Pr\left\{\frac{\prod_{u=0}^{s-1}\widehat{\boldsymbol{\delta}}^H_n\widehat{\boldsymbol{R}}^u_{\vDash n}\widehat{\boldsymbol{\delta}}_n}{N} > \sqrt[s]{\frac{\varepsilon}{\beta 2^{s+1}a^4_{\mathrm{MAX}}\phi^4_{\mathrm{MAX}}}}\right\} = o\left(\frac{1}{N^{2+\frac{4}{s}}}\right) \quad (54)$$

In fact, for $\varepsilon' = \frac{\varepsilon}{\beta 2^{s+1} a_{\mathrm{MAX}}^4 \phi_{\mathrm{MAX}}^4}$

$$
\begin{aligned}
\Pr\left\{ \frac{\prod_{u=0}^{s-1} (\widehat{\boldsymbol{\delta}}_n^H \widehat{\boldsymbol{R}}_{\vDash n}^u \widehat{\boldsymbol{\delta}}_n)^{i_u}}{N} > \varepsilon' \right\} &\leq \sum_{u=0}^{s-1} \Pr\left\{ \widehat{\boldsymbol{\delta}}_n^H \widehat{\boldsymbol{R}}_{\vDash n}^u \widehat{\boldsymbol{\delta}}_n > \sqrt[s]{\varepsilon' N} \right\} \\
&\stackrel{(a)}{\leq} \sum_{u=0}^{s-1} \Pr\left\{ \left| \widehat{\boldsymbol{\delta}}_n^H \widehat{\boldsymbol{R}}_{\vDash n}^u \widehat{\boldsymbol{\delta}}_n - \frac{\mathrm{tr} \widehat{\boldsymbol{R}}_{\vDash n}^u}{N} \right| > \sqrt[s]{\varepsilon' N} - \frac{\mathrm{tr} \widehat{\boldsymbol{R}}_{\vDash n}^u}{N} \right\} \\
&\stackrel{(b)}{\leq} \sum_{u=0}^{s-1} \frac{\mathrm{E}\left\{ \left| \widehat{\boldsymbol{\delta}}_n^H \widehat{\boldsymbol{R}}_{\vDash n}^u \widehat{\boldsymbol{\delta}}_n - \frac{\mathrm{tr} \widehat{\boldsymbol{R}}_{\vDash n}^u}{N} \right|^4 \right\}}{\sqrt[s]{(\varepsilon' N)^4}} \\
&\stackrel{(c)}{\leq} \frac{K_4 C^{(u)}}{N^2 ((N\varepsilon')^{\frac{1}{s}} - C^{(u)})^4} \qquad (55)
\end{aligned}
$$

where inequality (a) holds for N sufficiently large, inequality (b) follows from the Bienaymé inequality, and inequality (c) is a consequence of Lemma 9 in part I [1] and the bound on the eigenvalues moments of the matrix $\widehat{\boldsymbol{R}}$.

Let us consider now the probability ζ_{2b},

$$
\begin{aligned}
\zeta_{2b} &\leq \Pr\left\{ \frac{1}{N} \sum_{k=1}^{K} |(\widehat{\boldsymbol{R}}^s)_{kk} - R_s(|a_{kk}|^2, \widetilde{\tau}_k)| > \frac{\varepsilon}{a_{\mathrm{MAX}}^2 \phi_{\mathrm{MAX}}^2} \right\} \\
&\leq \Pr\left\{ \max_k |(\widehat{\boldsymbol{R}}^s)_{kk} - R_s(|a_{kk}|^2, \widetilde{\tau}_k)| > \frac{\varepsilon}{\beta a_{\mathrm{MAX}}^2 \phi_{\mathrm{MAX}}^2} \right\} \qquad (56)
\end{aligned}
$$

for $s = 1, \ldots \ell - 1$. Thanks to the assumption of the recursive step that $\forall \varepsilon' > 0$ and large $K = \beta N$, $\Pr\{|(\widehat{\boldsymbol{R}}^s)_{kk} - R_s(|a_{kk}|^2, \widetilde{\tau}_k)| > \varepsilon'\} \leq o(N^{-2})$, $\zeta_{2b} \to o(N^{-2})$, i.e. it vanishes asymptotically as $N, K \to \infty$ with constant ratio with the same converge rate as $o(N^{-2})$ at worst. Therefore, (50) converges in probability as $o(N^{-2})$ for $N \to +\infty$, at worst.

The proof of the convergence (52) in probability follows along similar lines.

Following the same approach as in the proof of Theorem 1 in [2], we can expand $(\widehat{\boldsymbol{R}}^\ell)_{kk}$ and $\widehat{\boldsymbol{T}}^\ell_{[nn]}$ as follows:

$$
(\widehat{\boldsymbol{R}}^\ell)_{kk} = \sum_{s=0}^{\ell-1} \widehat{\boldsymbol{h}}_k^H \widehat{\boldsymbol{T}}_{\sim k}^{\ell-s-1} \widehat{\boldsymbol{h}}_k (\widehat{\boldsymbol{R}}^s)_{kk} \qquad \ell = 1, 2, \ldots \qquad (57)
$$

$$
\widehat{\boldsymbol{T}}^\ell_{[nn]} = \sum_{s=0}^{\ell-1} \widehat{\boldsymbol{\delta}}_n \widehat{\boldsymbol{R}}_{\vDash n}^{\ell-s-1} \widehat{\boldsymbol{\delta}}_n^H \widehat{\boldsymbol{T}}^s_{[nn]}. \qquad \ell = 1, 2, \ldots \qquad (58)
$$

being $\widehat{\boldsymbol{T}}^0$ and $\widehat{\boldsymbol{R}}^0$ the identity matrices of dimensions $rN \times rN$ and $K \times K$, respectively.

Thanks to Property A and Property B of the convergence in probability of random sequences and the induction assumptions, the convergence in probability of the sequences $\{(\widehat{\boldsymbol{R}}^\ell)_{kk}\}$ and $\{\widehat{\boldsymbol{T}}^\ell_{[nn]}\}$ reduces to show the convergence in probability of $\widehat{\boldsymbol{h}}_k^H \widehat{\boldsymbol{T}}_{\sim k}^s \widehat{\boldsymbol{h}}_k$ and $\widehat{\boldsymbol{\delta}}_n \widehat{\boldsymbol{R}}_{\vDash n}^s \widehat{\boldsymbol{\delta}}_n^H$ to a deterministic limit, respectively. Let us define

$$\zeta_3 = \widehat{\boldsymbol{h}}_k^H \widehat{\boldsymbol{T}}_{\sim k}^s \widehat{\boldsymbol{h}}_k - \frac{|a_{kk}|^2}{N} \mathrm{tr} \boldsymbol{\Delta}_{\phi,r}^H(\widetilde{\tau}_k) \widehat{\boldsymbol{T}}_{\sim k}^s \boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_k).$$

Lemma 9 in part I [1] applied to the quadratic form $\widehat{\boldsymbol{h}}_k^H \widehat{\boldsymbol{T}}_{\sim k}^s \widehat{\boldsymbol{h}}_k$ with $p=4$ yields

$$\begin{aligned} \mathrm{E}\,|\zeta_3|^4 &< \frac{K_4 |a_{kk}|^4}{N^3} \mathrm{E}\left(\mathrm{tr}(\boldsymbol{\Delta}_{\phi,r}^H(\widetilde{\tau}_k) \widehat{\boldsymbol{T}}_{\sim k}^s \boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_k))^4 \right) \\ &\leq \frac{K_4}{N^3} a_{\mathrm{MAX}}^8 \phi_{\mathrm{MAX}}^8 \mathrm{tr}(\widehat{\boldsymbol{T}}_{\sim k}^{4s}). \end{aligned} \tag{59}$$

Thanks to the bound on the eigenvalues moments of the matrix $\widehat{\boldsymbol{T}}$, $\lim_{K=\beta N\to\infty} \frac{1}{N}\mathrm{E}(\mathrm{tr}\widehat{\boldsymbol{T}}_{\sim k}^{4s})$ is almost sure upper bounded $\forall s$ as $N=\beta K \to +\infty$. Therefore, $\mathrm{E}|\zeta_3|^4 \to 0$ as $K,N\to\infty$ with $\frac{K}{N}\to\beta$ and $\widehat{\boldsymbol{h}}_k^H \widehat{\boldsymbol{T}}_{\sim k}^s \widehat{\boldsymbol{h}}_k$ converges in mean square sense, and thus in probability. Furthermore, the Bienaymé inequality implies that $\Pr\{|\zeta_3|>\varepsilon\} \leq o(N^{-2})$ as $N\to+\infty$. Thanks to (52)

$$\begin{aligned} \lim_{N=\beta K\to\infty} \frac{|a_{kk}|^2}{N} \mathrm{tr} \boldsymbol{\Delta}_{\phi,r}^H(\widetilde{\tau}_k) \widehat{\boldsymbol{T}}_{\sim k}^s \boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_k) &= \lambda \int_0^1 \boldsymbol{\Delta}_{\phi,r}^H(x,\tau) \boldsymbol{T}_s(x) \boldsymbol{\Delta}_{\phi,r}(x,\tau) \mathrm{d}\, x \bigg|_{(\lambda,\tau)=(|a_{kk}|^2,\widetilde{\tau}_k)} + o(N^{-2}) \\ &= g(\boldsymbol{T}_s,\lambda,\tau) + o(N^{-2}). \end{aligned} \tag{60}$$

then

$$\Pr\{|\widehat{\boldsymbol{h}}_k^H \widehat{\boldsymbol{T}}_{\sim k}^s \widehat{\boldsymbol{h}}_k - g(\boldsymbol{T}_s,\lambda,\tau)| > \varepsilon\} \to o(N^{-2}) \tag{61}$$

for property A.

The convergence in probability of the diagonal blocks $\widehat{\boldsymbol{T}}_{[nn]}^{\ell}$ can be proven in a similar way. More specifically, it can be shown that the $r\times r$ block $\widehat{\boldsymbol{\delta}}_n \widehat{\boldsymbol{R}}_{\vDash n}^s \widehat{\boldsymbol{\delta}}_n^H$ converges to the $r\times r$ deterministic matrix

$$\mathbf{f}(R_s,x) = \beta \int \lambda \boldsymbol{\Delta}_{\phi,r}(x,\tau) \boldsymbol{\Delta}_{\phi,r}(x,\tau)^H R_s(\lambda,\tau) \mathrm{d}\, F_{|\boldsymbol{A}|^2,T}(\lambda,\tau). \tag{62}$$

such that $\Pr\left\{\left|(\widehat{\boldsymbol{\delta}}_n)_u \widehat{\boldsymbol{R}}_{\vDash n}^s (\widehat{\boldsymbol{\delta}}_n^H)_u - (\mathbf{f}(R_s,x))_u\right| > \varepsilon\right\} \to o(N^{-2})$.

Finally, by making use of equations (57) and (58) and the definitions (51), (53), (62), and (60) we obtain

$$R_\ell(\lambda,\tau) = \sum_{s=0}^{\ell-1} g(\boldsymbol{T}_{\ell-s-1},\lambda,\tau) R_s(\lambda,\tau) \qquad \ell=1,2,\ldots \tag{63}$$

and

$$\boldsymbol{T}_\ell(x) = \sum_{s=0}^{\ell-1} \mathbf{f}(R_{\ell-s-1},x) \boldsymbol{T}_s(x) \qquad \ell=1,2,\ldots. \tag{64}$$

with $g(\boldsymbol{T}_s, \lambda, \tau)$ and $\mathbf{f}(R_s, x)$ given in (60) and (62), respectively. Consistently to the definitions of $\widehat{\boldsymbol{T}}^0$ and $\widehat{\boldsymbol{R}}^0$, $\boldsymbol{T}_0(x) = \boldsymbol{I}_r$, being $\boldsymbol{I}_r$ the $r \times r$ identity matrix and $R_0(\lambda) = 1$.

Then, $g(R_0, \lambda, \tau) = \lambda \int_{-\frac{1}{2}}^{\frac{1}{2}} \boldsymbol{\Delta}_{\phi,r}^H(x,\tau)\boldsymbol{\Delta}_{\phi,r}(x,\tau)\mathrm{d}x$ and $\boldsymbol{f}(\boldsymbol{T}_0, x) = \beta \int \lambda \boldsymbol{\Delta}_{\phi,r}(x,\tau)\boldsymbol{\Delta}_{\phi,r}^H(x,\tau)\mathrm{d}F_{|\boldsymbol{A}|^2,T}(\lambda,\tau)$ and (63) and (64) reduce to the asymptotic limits $R_1(\lambda,\tau)$ and $\boldsymbol{T}_1(x)$ already derived in *step 1*. Therefore, we can begin the recursion with $\ell = 0$, $R_0(\lambda,\tau) = 1$ and $\boldsymbol{T}_0(x) = \boldsymbol{I}_r$.

Properties A, B, and C, the induction assumptions, relations (57) and (63), the convergence rates $\zeta_2 \rightarrow o(N^{-2})$ and $\Pr\{\zeta_3 > \varepsilon\} \leq\rightarrow o(N^{-2})$ yield (43). The proof of (44) follows immediately along similar lines.

This concludes the proof of Theorem 1.

APPENDIX II

PROOF OF COROLLARY 1

Corollary 1 is derived by specializing Theorem 1 to a unitary Fourier transform $\Phi(j2\pi f)$ with bandwidth $B \leq \frac{r}{2T_c}$. Let us recall here that the unitary Fourier transform in the discrete time domain is given by

$$\phi(x,\tau) = \frac{1}{T_c}\mathrm{e}^{j2\pi\frac{\tau}{T_c}x} \sum_{s=-\mathrm{sign}(x)\left\lfloor\frac{r-1}{2}\right\rfloor}^{\mathrm{sign}(x)\left\lfloor\frac{r}{2}\right\rfloor} \mathrm{e}^{j2\pi\frac{\tau}{T_c}s}\Phi^*\left(\frac{j2\pi}{T_c}(x+s)\right) \quad \text{for} \quad |x| \leq \frac{1}{2}. \tag{65}$$

The matrix $\boldsymbol{Q}(x,\tau) = \boldsymbol{\Delta}_{\phi,r}(x,\tau)\boldsymbol{\Delta}_{\phi,r}(x,\tau)^H$, with $\boldsymbol{\Delta}_{\phi,r}(x,\tau)$ defined in (28), can be decomposed as $\boldsymbol{Q}(x,\tau) = \boldsymbol{Q}(x) + \overline{\boldsymbol{Q}}(x,\tau)$ with the elements of $\boldsymbol{Q}(x)$ and $\overline{\boldsymbol{Q}}(x,\tau)$ defined by

$$(\boldsymbol{Q}(x))_{k,\ell} = \frac{1}{T_c^2} \sum_{s=-\mathrm{sign}(x)\left\lfloor\frac{r-1}{2}\right\rfloor}^{\mathrm{sign}(x)\left\lfloor\frac{r}{2}\right\rfloor} \left|\Phi\left(\frac{j2\pi}{T_c}(x+s)\right)\right|^2 \mathrm{e}^{-j2\pi\frac{k-\ell}{r}(x+s)} \quad \text{for} \quad |x| \leq \frac{1}{2}, \tag{66}$$

and

$$(\overline{\boldsymbol{Q}}(x,\tau))_{k,\ell} = \frac{1}{T_c^2} \sum_{\substack{s,u=-\mathrm{sign}(x)\left\lfloor\frac{r-1}{2}\right\rfloor \\ s\neq u}}^{\mathrm{sign}(x)\left\lfloor\frac{r}{2}\right\rfloor} \Phi\left(\frac{j2\pi}{T_c}(x+u)\right)\Phi^*\left(\frac{j2\pi}{T_c}(x+s)\right)\mathrm{e}^{-j2\pi\frac{\tau}{T_c}(s-u)}\mathrm{e}^{-j2\pi\left(\frac{k-1}{r}(x-s)-\frac{\ell-1}{r}(x-u)\right)} \quad \text{for} \quad |x| \leq \frac{1}{2}, \tag{67}$$

respectively.

Equations (26) and (27) can be rewritten as

$$\begin{aligned}\boldsymbol{f}(R_s,x) &= \beta\boldsymbol{Q}(x)\int \lambda R_s(\lambda,\tau)\mathrm{d}F_{|\boldsymbol{A}|^2,T}(\lambda,\tau)\\ &\quad+\beta\int \lambda R_s(\lambda,\tau)\overline{\boldsymbol{Q}}(x,\tau)\mathrm{d}F_{|\boldsymbol{A}|^2,T}(\lambda,\tau), \qquad -\frac{1}{2}\leq x\leq\frac{1}{2} \qquad (68)\\ g(\boldsymbol{T}_s,\lambda,\tau) &= \lambda\int_{-\frac{1}{2}}^{\frac{1}{2}}\mathrm{tr}(\boldsymbol{T}_s(x)\boldsymbol{Q}(x))\mathrm{d}x+\lambda\int_{-\frac{1}{2}}^{\frac{1}{2}}\mathrm{tr}(\boldsymbol{T}_s(x)\overline{\boldsymbol{Q}}(x,\tau))\mathrm{d}x, \qquad (69)\end{aligned}$$

respectively. If the conditions of Corollary 1 are satisfied, i.e. if $B\leq\frac{r}{2T_c}$ and τ is uniformly distributed in $[0,T_c]$, it can be shown that

- $R_\ell(\lambda,\tau)$, $\ell\in\mathbb{Z}^+$, are independent of τ and
- $\boldsymbol{T}_\ell(x)$ is a matrix of the form (70).

These properties can be proven by strong induction. It is straightforward to verify that they are satisfied for $s=0$. In fact, $R_0(\lambda,\tau)=1$ is independent of τ and $\boldsymbol{T}_0(x)=\boldsymbol{I}$ is of the form (70) with $b_0=1$ and $b_i(x)=0$ with $i=1,\dots r-1$. By appealing to Lemma 1 in part I [1] Appendix I $\mathrm{tr}(\overline{\boldsymbol{Q}}(x,\tau))=0$ and $g(\boldsymbol{T}_0,\lambda,\tau)=\lambda\int_{-\frac{1}{2}}^{\frac{1}{2}}\mathrm{tr}(\boldsymbol{Q}(x))\mathrm{d}x$. Hence, $g(\boldsymbol{T}_0,\lambda,\tau)$ is independent of τ.

The induction step is proven using the following induction assumptions:

- For $s=0,1,\dots\ell-1$, $R_s(\lambda,\tau)$ is independent of τ;
- For $s=0,1,\dots\ell-1$, $\boldsymbol{T}_s(x)$ is of the form

$$\boldsymbol{B}=\boldsymbol{B}(x)=\begin{bmatrix} b_0 & b_1\mathrm{e}^{j\frac{2\pi}{r}x} & \dots & \dots & b_{r-1}\mathrm{e}^{j\frac{2\pi(r-1)}{r}x}\\ b_{r-1}\mathrm{e}^{-j\frac{2\pi}{r}x} & b_0 & b_1\mathrm{e}^{j\frac{2\pi}{r}x} & \dots & b_{r-2}\mathrm{e}^{j\frac{2\pi(r-2)}{r}x}\\ \dots & \ddots & \ddots & \ddots & \ddots\\ b_1\mathrm{e}^{-j\frac{2\pi(r-1)}{r}x} & \ddots & \ddots & b_{r-1}\mathrm{e}^{-j\frac{2\pi}{r}x} & b_0\end{bmatrix}, \qquad (70)$$

being $b_0=b_0(x), b_1=b_1(x),\dots b_{r-1}=b_{r-1}(x)$, eventually functions of x.

Thanks to the form (70) of $\boldsymbol{T}_s(x)$, $s=1,\dots\ell-1$, given by the induction assumptions and by applying Lemma I in part I Appendix I we have $\mathrm{tr}(\boldsymbol{T}_s(x)\overline{\boldsymbol{Q}}(x,\tau))=0$, for $s=0,1,\dots,\ell-1$. Then, (69) reduces to $g(\boldsymbol{T}_s,\lambda,\tau)=\lambda\int_{-\frac{1}{2}}^{\frac{1}{2}}\mathrm{tr}\left(\boldsymbol{T}_s(x)\boldsymbol{Q}(x)\right)\mathrm{d}x$ and $g(\boldsymbol{T}_s,\lambda,\tau)$ is independent of τ for $s=0,1,\dots,\ell-1$. Therefore, all quantities that appear in the right hand side of (24) are independent of τ and $R_\ell(\lambda,\tau)$ is also independent of τ. In the following we will shortly write $R_\ell(\lambda)$ and $g(\boldsymbol{T}_s,\lambda)$ instead of $R_\ell(\lambda,\tau)$ and $g(\boldsymbol{T}_s,\lambda,\tau)$. Thanks to the fact that $R_s(\lambda,\tau)$ is independent of τ and λ and τ are statistically independent with τ uniformly distributed, (68) can be rewritten as

$$f(R_s,x)=\beta\int\lambda R_s(\lambda)\mathrm{d}F_{|\boldsymbol{A}|^2}\left(\boldsymbol{Q}(x)+\frac{1}{T_c}\int_0^{T_c}\overline{\boldsymbol{Q}}(x,\tau)\mathrm{d}\tau\right). \qquad (71)$$

It is straightforward to verify that $\int_0^{T_c} \overline{\boldsymbol{Q}}(x,\tau)\mathrm{d}\tau = 0$ from the definition of $\overline{\boldsymbol{Q}}(x,\tau)$ in (67). Then,

$$\begin{aligned} \boldsymbol{f}(R_s, x) &= \beta \boldsymbol{Q}(x) \int \lambda R_s(\lambda) \mathrm{d}F_{|\boldsymbol{A}|^2}(\lambda) \\ &= f(R_s)\boldsymbol{Q}(x) \end{aligned} \tag{72}$$

with $f(R_s) = \beta \int \lambda R_s(\lambda)\mathrm{d}F_{|\boldsymbol{A}|^2}(\lambda)$. Substituting (72) in (25) yields

$$\boldsymbol{T}_\ell(x) = \sum_{s=0}^{\ell-1} f(R_{\ell-s-1})\boldsymbol{Q}(x)\boldsymbol{T}_s(x), \qquad -\frac{1}{2} \le x \le \frac{1}{2}. \tag{73}$$

Since $\boldsymbol{T}_s(x)$ is of form (70), the conditions of Lemma 2 in part I Appendix I are satisfied for $\boldsymbol{B} = \boldsymbol{T}_s(x)$. This implies that $\boldsymbol{Q}(x)\boldsymbol{T}_s(x)$ is also of the form (70). Since $\boldsymbol{T}_\ell(x)$ is a linear combination of matrices of the form (70), $\boldsymbol{T}_\ell(x)$ is also a matrix of the form (70). Then, the statement of the strong induction is proven .

Thanks to the properties shown by strong induction the recursive equations in Theorem (1) reduce to the following set of recursive equations:

$$R_\ell(\lambda) = \sum_{s=0}^{\ell-1} g(\boldsymbol{T}_{\ell-s-1}, \lambda) R_s(\lambda) \tag{74}$$

$$\boldsymbol{T}_\ell(x) = \sum_{s=0}^{\ell-1} f(R_{\ell-s-1})\boldsymbol{Q}(x)\boldsymbol{T}_s(x) \qquad -\frac{1}{2} \le x \le \frac{1}{2} \tag{75}$$

$$f(R_s) = \beta \int \lambda R_s(\lambda) \mathrm{d}\, F_{|\boldsymbol{A}|^2}(\lambda), \qquad -\frac{1}{2} \le x \le \frac{1}{2} \tag{76}$$

$$g(\boldsymbol{T}_s, \lambda) = \lambda \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathrm{tr}(\boldsymbol{T}_s(x)\boldsymbol{Q}(x))\mathrm{d}\, x \tag{77}$$

with $\boldsymbol{T}_0(x) = \boldsymbol{I}_r$ and $R_0(\lambda) = 1$.

Then, applying again Theorem 1 we obtain the following convergence in probability

$$\lim_{K=\beta N \to \infty} (\widehat{\boldsymbol{R}}^\ell)_{kk} = R_\ell(\lambda)|_{\lambda=|a_{kk}|^2}.$$

From (75) and $\boldsymbol{T}_0(x) = \boldsymbol{I}_r$ it is apparent that $\boldsymbol{T}_\ell(x)$ is a polynomial in $\boldsymbol{Q}^s(x)$, for $s = 0, 1, \ldots \ell$. Then, $\boldsymbol{T}_\ell(x)$ has the same eigenvectors as $\boldsymbol{Q}(x)$ and it can written as $\boldsymbol{T}_\ell(x) = \boldsymbol{U}(x)\boldsymbol{\Lambda}_\ell(x)\boldsymbol{U}^H(x)$ where $\boldsymbol{\Lambda}_\ell(x)$ is a diagonal matrix with diagonal elements $t_{\ell,1}, t_{\ell,2}, \ldots t_{\ell,r}$ and

$$\boldsymbol{U}(x) = \left(\boldsymbol{e}\left(x - \mathrm{sign}(x)\left\lfloor \frac{r-1}{2} \right\rfloor\right), \ldots \boldsymbol{e}\,(x) \ldots \boldsymbol{e}\left(x + \mathrm{sign}(x)\left\lfloor \frac{r}{2} \right\rfloor\right)\right) \tag{78}$$

with $\boldsymbol{e}\,(x)$ r-dimensional column vector defined by

$$\boldsymbol{e}\,(x) = \frac{1}{\sqrt{r}}\left(1, \mathrm{e}^{-j2\pi\frac{1}{r}x}, \ldots \mathrm{e}^{-j2\pi\frac{r-1}{r}x}\right)^T.$$

By making use of the eigenvalue decomposition of the matrix $\boldsymbol{Q}(x)$ in part I Appendix I Lemma 3 the matrix equation (75) reduces to r scalar equations

$$t_{\ell,u}(x) = \sum_{s=0}^{\ell-1} f(R_{\ell-s-1}) \frac{r}{T_c^2} \left| \Phi\left(j\frac{2\pi}{T_c}\left(x - \mathrm{sign}(x)\left(\left\lfloor \frac{r-1}{2} \right\rfloor - u + 1 \right)\right)\right)\right|^2 t_{s,u}(x) \qquad u = 1, \ldots r \quad \text{and} \quad |x| \le \frac{1}{2}.$$

By substituting $y = x - \mathrm{sign}(x)\left(\left\lfloor \frac{r-1}{2} \right\rfloor - u + 1\right)$ for $|x| \le \frac{1}{2}$ we obtain

$$t_{\ell,u}\left(y + \left\lfloor \frac{r-1}{2} \right\rfloor - u + 1 \right) = \sum_{s=0}^{\ell-1} f(R_{\ell-s-1}) \frac{r}{T_c^2} \left| \Phi\left(j\frac{2\pi}{T_c} y \right) \right|^2 t_{s,u}\left(y + \left\lfloor \frac{r-1}{2} \right\rfloor - u + 1 \right) \tag{79}$$

for $0 \le y + \left\lfloor \frac{r-1}{2} \right\rfloor - u + 1 \le \frac{1}{2}$ and

$$t_{\ell,u}\left(y - \left\lfloor \frac{r-1}{2} \right\rfloor + u - 1 \right) = \sum_{s=0}^{\ell-1} f(R_{\ell-s-1}) \frac{r}{T_c^2} \left| \Phi\left(j\frac{2\pi}{T_c} y \right) \right|^2 t_{s,u}\left(y - \left\lfloor \frac{r-1}{2} \right\rfloor + u - 1 \right) \tag{80}$$

for $\frac{1}{2} \le y - \left\lfloor \frac{r-1}{2} \right\rfloor + u - 1 \le 0$. Then, for $u = 1, \ldots r$ the r functions (79) and (80) defined in not overlapping intervals in $[-r, r]$ can be combined in a unique scalar functions $T_\ell(y)$ in the interval $|y| \le r$ satisfying the recursive equation

$$T_\ell(y) = \sum_{s=0}^{\ell-1} \frac{r}{T_c^2} f(R_{\ell-s-1}) \left| \Phi\left(j\frac{2\pi}{T_c} y \right) \right|^2 T_s(y).$$

Similar arguments applied to (77) yield

$$g(T_s, \lambda) = \lambda \int_r^r \frac{r}{T_c^2} T_s(y) \left| \Phi\left(j\frac{2\pi}{T_c} y \right) \right|^2 \mathrm{d}y.$$

This concludes the derivation of Corollary 1 from Theorem 1.

APPENDIX III

DERIVATION OF ALGORITHM 1

Algorithm 1 can be derived from the recursive equations of Corollary 1 by using the following substitutions[3]:

$$
\begin{array}{lcl}
\lambda & \rightarrow & z \\
R_s(\lambda) & \rightarrow & \rho_s(z) \\
\lambda R_s(\lambda) & \rightarrow & v_s(z) \\
f(R_s) = \mathrm{E}(\lambda R_s(\lambda)) & \rightarrow & V_s \\
\frac{1}{T_c}\left|\Phi\left(\frac{j2\pi x}{T_c}\right)\right|^2 & \rightarrow & y \\
T_s(\cdot) & \rightarrow & \mu_s(y) \\
\frac{r}{T_c^2}\left|\Phi\left(\frac{j2\pi x}{T_c}\right)\right|^2 T_s(x) & \rightarrow & u_s(y) \\
\nu_s = \frac{r}{T_c}\int_{-r}^{r}\frac{1}{T_c}\left|\Phi\left(\frac{j2\pi x}{T_c}\right)\right|^2 T_s(x)\mathrm{d}x & \rightarrow & U_s.
\end{array}
$$

Then, the initial step is obtained by defining $\mu_0(y) = 1$ and $\rho_0(z) = 1$. The recursive equations in step ℓ are obtained by using the previous substitutions. In order to derive U_s let us observe that $\frac{1}{T_c}\left|\Phi\left(\frac{j2\pi x}{T_c}\right)\right|^2 T_s(x)$ is a polynomial in $y = \left|\Phi\left(\frac{j2\pi x}{T_c}\right)\right|^2$ of degree $s+1$. Then, U_s is a linear combination of $\mathcal{E}_n$ where

$$
\mathcal{E}_n = \frac{1}{T_c^n}\int_{-r}^{r}\left|\Phi\left(\frac{j2\pi x}{T_c}\right)\right|^{2n}\mathrm{d}x = \frac{1}{T_c^n}\int_{-B}^{B} T_c\left|\Phi\left(j2\pi f\right)\right|^{2n}\mathrm{d}f.
$$

The coefficients of the linear combination are obtained by expanding $u_s(y)$ as a polynomial in y.

We conclude the derivation of Algorithm 1 by summarizing the previous considerations and substitutions:

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

[3]Note that the substitution of λ with z is redundant. It is used to obtain polynomials in the commonly used variable z.

- expanding $u_s(y)$ and $v_s(z)$ as polynomials in y and z, respectively,
- replacing the monomials y^n and z^n, $n \in \mathbb{Z}^+$ with $\mathcal{E}_n$ and $m_{|\boldsymbol{A}|^2}^{(s)}$, respectively.

Then, $R_\ell(\lambda) = \rho_\ell(\lambda)$ and the eigenvalue moment $m_{\boldsymbol{R}}^{(\ell)} = \mathrm{E}\{R_\ell(\lambda)\}$ is obtained by replacing all monomials $z, z^2, \ldots, z^\ell$ in the polynomial $\rho_\ell(z)$ by the moments $m_{|\boldsymbol{A}|^2}^1, m_{|\boldsymbol{A}|^2}^2, \ldots, m_{|\boldsymbol{A}|^2}^\ell$, respectively.

APPENDIX IV

PROOF OF THEOREM 2

The proof of Theorem 2 follows along the line of the proof of Theorem 1.

For a signal with bandwidth $B \leq \frac{1}{2T_c}$,

$$\underline{\phi}(x,\tau) = \frac{1}{T_c}\mathrm{e}^{j2\pi\frac{\tau x}{T_c}}\Phi^*\left(j\frac{2\pi}{T_c}x\right) \qquad |x| \leq \frac{1}{2}$$

and $\phi(x,\tau) = \underline{\phi}(x - \lfloor 2x \rfloor, \tau)$ for any x. Correspondingly, we define

$$\underline{\boldsymbol{\Delta}}_{\phi,r}(x,\tau) = \frac{1}{T_c}\Phi(\frac{j2\pi x}{T_c})\mathrm{e}^{-\frac{j2\pi\tau x}{T_c}}\boldsymbol{e}(x), \qquad |x| \leq \frac{1}{2}$$

with $\boldsymbol{e}(x) = (1, \mathrm{e}^{j2\pi\frac{x}{r}}, \ldots \mathrm{e}^{j2\pi\frac{(r-1)}{r}x})$ and

$$\boldsymbol{\Delta}_{\phi,r}(x,\tau) = \underline{\boldsymbol{\Delta}}_{\phi,r}(x - \lfloor 2x \rfloor, \tau) \qquad \text{for any } \ x.$$

We adopt here the same notation as in the proof of Theorem 1. Then, the $K \times K$ diagonal matrix $\boldsymbol{\nabla}_{nt}$, for $t = 1, \ldots r$ and $n = 1, \ldots N$ is given by

$$\boldsymbol{\nabla}_{nt} = \frac{1}{T_c}\Phi^*\left(\frac{j2\pi}{T_c}\underline{n}\right)\mathrm{e}^{-\frac{j2\pi\underline{n}(t-1)}{r}}\mathrm{diag}\left(\mathrm{e}^{\frac{j2\pi\underline{n}\widetilde{\tau}_1}{T_c}}, \mathrm{e}^{\frac{j2\pi\underline{n}\widetilde{\tau}_2}{T_c}}, \ldots \mathrm{e}^{\frac{j2\pi\underline{n}\widetilde{\tau}_K}{T_c}}\right)$$

with $\underline{n} = \frac{n-1}{N} - \left\lfloor 2\frac{n-1}{N}\right\rfloor$ and $\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_k)$ is the $rN \times N$ block diagonal matrix with n diagonal block $\boldsymbol{\Delta}_{\phi,r}(\underline{n}, \widetilde{\tau}_k)$. We develop the proof by strong induction as in Theorem 1 with similar initial step and similar induction step.

Step 1: In this case

$$\widehat{\boldsymbol{R}}_{kk} = |a_{kk}|^2 \boldsymbol{s}_k^H \boldsymbol{\Delta}_{\phi,r}^H(\widetilde{\tau}_k)\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_k)\boldsymbol{s}_k = |a_{kk}|^2 \boldsymbol{s}_k^H \boldsymbol{\Phi}\boldsymbol{s}_k$$

where $\boldsymbol{\Phi}$ is a matrix independent of $\widetilde{\tau}_k$ and the n^{th} element is given by $\boldsymbol{\Phi}_{nn} = \frac{r}{T_c}\left|\Phi\left(\frac{j2\pi\underline{n}}{T_c}\right)\right|^2$.

By following the same approach as in Theorem 1 it results $\forall \varepsilon > 0$

$$\Pr\left\{\left|\widehat{\boldsymbol{R}}_{kk} - \frac{r|a_{kk}|^2}{T_c N}\sum_{n=0}^{N-1}\left|\Phi\left(\frac{j2\pi\underline{n}}{T_c}\right)\right|^2\right| > \varepsilon\right\} \leq \frac{K_4|a_{kk}|^4\Delta_{\mathrm{MAX}}^4}{N^2\varepsilon^4}$$

being $\Delta_{\mathrm{MAX}} = \max_{x\in\left[-\frac{1}{2},\frac{1}{2}\right]} \left|\Phi\left(\frac{j2\pi x}{T_c}\right)\right|^2$ and

$$\begin{aligned} R_1(\lambda)|_{\lambda=|a_{kk}|^2} &= \lim_{K=\beta N\to\infty} \frac{|a_{kk}|^2}{N} \sum_{\ell=0}^{N-1} \left|\Phi\left(\frac{j2\pi}{T_c}\left(\frac{n}{N} - \left\lfloor\frac{2n}{N}\right\rfloor\right)\right)\right|^2 \\ &= \lambda \int_{-\frac{1}{2}}^{\frac{1}{2}} \left|\Phi\left(\frac{j2\pi x}{T_c}\right)\right|^2 \mathrm{d}x \bigg|_{\lambda=|a_{kk}|^2} . \end{aligned} \tag{81}$$

Furthermore, $\Pr\left\{|\widehat{\boldsymbol{R}}_{kk} - R_1(|a_{kk}|^2)| > \varepsilon\right\} \leq o\left(N^{-2}\right)$.

Similarly, $(\widehat{\boldsymbol{T}}_{[nn]})_{uv}$, the (u,v)-element of the matrix $\widehat{\boldsymbol{T}}_{[nn]}$ is given by

$$\begin{aligned} \widehat{\boldsymbol{T}}_{[nn]} &= \widehat{\boldsymbol{\sigma}}_n \boldsymbol{A} \boldsymbol{\nabla}_{n,u} \boldsymbol{\nabla}_{n,v}^H \boldsymbol{A}^H \widehat{\boldsymbol{\sigma}}_n^H \\ &= \frac{1}{T_c}\left|\Phi\left(\frac{j2\pi \underline{n}}{T_c}\right)\right| \mathrm{e}^{-j2\pi\underline{n}\frac{v-u}{r}} \widehat{\boldsymbol{\sigma}}_n \boldsymbol{A}\boldsymbol{A}^H \widehat{\boldsymbol{\sigma}}_n^H . \end{aligned} \tag{82}$$

As in Theorem 1 it can been shown that

$$\Pr\left\{\left|(\widehat{\boldsymbol{T}}_{[nn]})_{uv} - \frac{1}{NT_c}\left|\Phi\left(\frac{j2\pi\underline{n}}{T_c}\right)\right|^2 \mathrm{e}^{-j2\pi\underline{n}\frac{v-u}{r}} \mathrm{tr}(\boldsymbol{A}\boldsymbol{A}^H)\right| > \varepsilon\right\} \leq \frac{K_4 T_{\mathrm{MAX}}^4}{N^2\varepsilon^4}$$

with $T_{\mathrm{MAX}} = \max_{x\in\left[-\frac{1}{2},\frac{1}{2}\right]} \left|\Phi\left(\frac{j2\pi\underline{n}}{T_c}\right)\right|^2 \sup_K \max_k |a_{kk}|^2$ and the following convergence in probability holds

$$\begin{aligned} \lim_{K=\beta N\to\infty} (\widehat{\boldsymbol{T}}_{[nn]})_{uv} &= \lim_{K=\beta N\to\infty} \frac{\beta}{T_c K}\left|\Phi\left(\frac{j2\pi\underline{n}}{T_c}\right)\right|^2 \mathrm{e}^{-j2\pi\underline{n}\frac{v-u}{r}} \sum_{k=1}^{K} |a_{kk}|^2 \\ &= \frac{\beta}{T_c}\left|\Phi\left(\frac{j2\pi x}{T_c}\right)\right|^2 \mathrm{e}^{-j2\pi\underline{n}\frac{v-u}{r}} \int \lambda \mathrm{d}F_{|\boldsymbol{A}|^2}(\lambda) \end{aligned}$$

with $x = \lim_{N\to\infty} \underline{n}$ and $|x| \leq \frac{1}{2}$. Thus, the diagonal block converges in probability as follows

$$\begin{aligned} \boldsymbol{T}_1(x) &= \lim_{K=\beta N\to\infty} (\widehat{\boldsymbol{T}}_{[nn]})_{uv} \\ &= \frac{\beta}{T_c}\left|\Phi\left(\frac{j2\pi x}{T_c}\right)\right|^2 \int \lambda \mathrm{d}F_{|\boldsymbol{A}|^2}(\lambda) \boldsymbol{e}(x)\boldsymbol{e}^H(x) \end{aligned} \tag{83}$$

Furthermore,

$$\Pr\left\{\left|(\widehat{\boldsymbol{T}}_{[nn]})_{uv} - (\boldsymbol{T}_1(x))_{uv}\right| > \varepsilon\right\} \leq o(N^{-2}).$$

This concludes the first step of the induction.

Step ℓ: Let us observe that

$$\begin{aligned} \vartheta_1 &= \frac{1}{N}\mathrm{tr}\boldsymbol{A}\boldsymbol{\nabla}_{n,u} \widehat{\boldsymbol{R}}^s_{\vDash n} \boldsymbol{\nabla}_{n,u}^H \boldsymbol{A}^H \\ &= \frac{\mathrm{e}^{-j2\pi\underline{n}\frac{u-v}{r}}}{N} \sum_{k=1}^{K} \frac{|a_{kk}|^2}{T_c^2}\left|\Phi\left(\frac{j2\pi\underline{n}}{T_c}\right)\right|^2 (\widehat{\boldsymbol{R}}^s_{\vDash n})_{kk} \end{aligned}$$

and

$$\vartheta_2 = \frac{|a_{kk}|^2}{N}\mathrm{tr}\boldsymbol{\Delta}^H_{\Phi,r}(\widetilde{\tau}_k)\widehat{\boldsymbol{T}}^s_{\sim k}\boldsymbol{\Delta}_{\Phi,r}(\widetilde{\tau}_k)$$
$$= \frac{|a_{kk}|^2}{N}\sum_{n=1}^{N}\frac{1}{T_c^2}\left|\Phi\left(\frac{j2\pi\underline{n}}{T_c}\right)\right|^2 \boldsymbol{e}^H(\underline{n})(\widehat{\boldsymbol{T}}^s_{\sim k})_{nn}\boldsymbol{e}(\underline{n}).$$

By following the same approach as in Theorem 1 it can be shown that ϑ_1 and ϑ_2 converge in probability to the following limits

$$\lim_{K=\beta N\to\infty}\vartheta_1 = \frac{\beta}{T_c^2}\mathrm{e}^{-j2\pi\underline{n}\frac{u-v}{r}}\left|\Phi\left(\frac{j2\pi x}{T_c}\right)\right|^2\int \lambda R_s(\lambda)\mathrm{d}F_{|\boldsymbol{A}|^2}(\lambda)$$

and

$$\lim_{K=\beta N\to\infty}\vartheta_2 = \frac{\lambda}{T_c^2}\int_{-\frac{1}{2}}^{\frac{1}{2}}\left|\Phi\left(\frac{j2\pi x}{T_c}\right)\right|^2\boldsymbol{e}^H(x)\boldsymbol{T}_s(x)\boldsymbol{e}(x)\mathrm{d}x\Bigg|_{\lambda=|a_{kk}|^2}$$

with $R_s(\lambda)|_{\lambda=|a_{kk}|^2} = \lim_{K=\beta N\to\infty}(\widehat{\boldsymbol{R}}^s)_{kk}$ and $\boldsymbol{T}_s(x)| = \lim_{K=\beta N\to\infty}\widehat{\boldsymbol{T}}^s_{[nn]}$ given by the recursion assumptions.

Additionally, it can be shown that the following convergence in probability holds

$$g(\boldsymbol{T}_s,\lambda)|_{\lambda=|a_{kk}|^2} = \lim_{K=\beta N\to\infty}\widehat{\boldsymbol{h}}_k^H\widehat{\boldsymbol{T}}^s_{\sim k}\widehat{\boldsymbol{h}}_k$$
$$= \frac{\lambda}{T_c}\int_{-\frac{1}{2}}^{\frac{1}{2}}\left|\Phi\left(\frac{j2\pi x}{T_c}\right)\right|^2\boldsymbol{e}^H(x)\boldsymbol{T}_s(x)\boldsymbol{e}(x)\mathrm{d}x\Bigg|_{\lambda=|a_{kk}|^2} \quad (84)$$

and

$$\boldsymbol{f}(R_s,x) = \lim_{K=\beta N\to\infty}\widehat{\boldsymbol{\delta}}_n\widehat{\boldsymbol{R}}^s_{\vDash n}\widehat{\boldsymbol{\delta}}_n^H$$
$$= \frac{\lambda}{T_c^2}\int_{-\frac{1}{2}}^{\frac{1}{2}}\left|\Phi\left(\frac{j2\pi x}{T_c}\right)\right|^2\boldsymbol{e}^H(x)\boldsymbol{T}_s(x)\boldsymbol{e}(x)\mathrm{d}x\Bigg|_{\lambda=|a_{kk}|^2} \quad (85)$$

The convergence in probability satisfies the bounds

$$\Pr\left\{|\widehat{\boldsymbol{h}}_k^H\widehat{\boldsymbol{T}}^s_{\sim k}\widehat{\boldsymbol{h}}_k - g(\boldsymbol{T}_s,|a_{kk}|^2)| > \varepsilon\right\} < o(N^{-2})$$

and

$$\Pr\left\{|(\widehat{\boldsymbol{\delta}}_n)_u\widehat{\boldsymbol{R}}^s_{\vDash n}(\widehat{\boldsymbol{\delta}}_n^H)_v - (\boldsymbol{f}(R_s,x))_{u,v}| > \varepsilon\right\} < o(N^{-2})$$

for large N and $\forall\varepsilon$.

The recursion assumptions and the limits (84) and (85) in (57) and (58) yield

$$\begin{aligned} R_\ell(\lambda)|_{\lambda=|a_{kk}|^2} &= \sum_{s=0}^{\ell-1} g(\boldsymbol{T}_{\ell-s-1},\lambda) R_s(\lambda) \\ &= \sum_{s=0}^{\ell-1} R_s(\lambda) \frac{\lambda}{T_c^2} \int_{-\frac{1}{2}}^{\frac{1}{2}} \left| \Phi\left(\frac{j2\pi x}{T_c}\right)\right|^2 \operatorname{tr}\left(\boldsymbol{T}_s(x)\boldsymbol{e}(x)\boldsymbol{e}^H(x)\right) \mathrm{d}x \Bigg|_{\lambda=|a_{kk}|^2} \end{aligned} \tag{86}$$

and

$$\begin{aligned} \boldsymbol{T}_\ell &= \sum_{s=0}^{\ell-1} \boldsymbol{f}(R_{\ell-s-1},x)\boldsymbol{T}_s(x) \\ &\sum_{s=0}^{\ell-1} \frac{\beta}{T_c^2} \left|\Phi\left(\frac{j2\pi x}{T_c}\right)\right|^2 \int \lambda R_s(\lambda) \mathrm{d}F_{|\boldsymbol{A}|^2}(\lambda) \boldsymbol{e}(x)\boldsymbol{e}^H(x)\boldsymbol{T}_s(x) \end{aligned} \tag{87}$$

where $R_0(\lambda) = 1$ and $\boldsymbol{T}_0(x) = \boldsymbol{I}_r$. With a similar approach as in Theorem 1 it can be proven that for large N and $\forall\varepsilon > 0$

$$\Pr\left\{\left|\widehat{\boldsymbol{R}}_{kk}^{\ell} - R_\ell(|a_{kk}|^2)\right| > \varepsilon\right\} \leq o(N^{-2})$$

and

$$\Pr\left\{\left|(\widehat{\boldsymbol{T}}_{[nn]}^{\ell})_{uv} - (\boldsymbol{T}_\ell(x))_{uv}\right| > \varepsilon\right\} \leq o(N^{-2}).$$

In contrast to Theorem 1 the recursive equations (86), (87), (84), and (85) are independent of the time delay $\widetilde{\tau}_k$.

The recursive equations can be further simplified by observing that $(\boldsymbol{e}(x)\boldsymbol{e}^H(x))^m = r^{m-1}\boldsymbol{e}(x)\boldsymbol{e}^H(x)$. Then, it is straightforward to verify by recursion that the matrix $\boldsymbol{T}_s(x)$, $s = 1, 2, \ldots, \ell-1$, is proportional to the matrix $\boldsymbol{e}(x)\boldsymbol{e}^H(x)$ and we can express it as $\boldsymbol{T}_s(x) = T_s(x)\boldsymbol{e}(x)\boldsymbol{e}^H(x)$, $s = 1, 2, \ldots$. Thus, the recursive equations can be rewritten as

$$\begin{aligned} R_\ell(\lambda) &= \sum_{s=0}^{\ell-1} g(\boldsymbol{T}_{\ell-s-1},\lambda)R_s(\lambda) \\ T_\ell(x)\boldsymbol{e}(x)\boldsymbol{e}^H(x) &= \sum_{s=1}^{\ell-1} \boldsymbol{f}(R_{\ell-s-1},x)T_s(x)\boldsymbol{e}(x)\boldsymbol{e}^H(x) + \boldsymbol{f}(R_{\ell-1},x)\boldsymbol{T}_0(x) \qquad \ell = 1, 2, \ldots \end{aligned} \tag{88}$$

$$\boldsymbol{f}(R_s,x) = f(R_s,x)\boldsymbol{e}(x)\boldsymbol{e}^H(x) \tag{89}$$

$$f(R_s,x) = \frac{\beta}{T_c^2}\left|\Phi\left(\frac{j2\pi x}{T_c}\right)\right|^2 \int \lambda R_s(\lambda)\mathrm{d}\,F_{|\boldsymbol{A}|^2}(\lambda) \qquad -\frac{1}{2} \leq x \leq \frac{1}{2}$$

$$g(T_s,\lambda) = \begin{cases} \frac{r^2\lambda}{T_c^2}\int_{-1/2}^{1/2}\left|\Phi\left(\frac{j2\pi x}{T_c}\right)\right|^2 \overline{T}^s(x)\mathrm{d}\,x & s = 1, 2, \ldots \\ \frac{r\lambda}{T_c^2}\int_{-1/2}^{1/2}\left|\Phi\left(\frac{j2\pi x}{T_c}\right)\right|^2 \mathrm{d}\,x & s = 0. \end{cases}$$

with $\boldsymbol{T}_0(x) = \boldsymbol{I}_r$ and $R_0(\lambda) = 1$.

Substituting (89) in (88) we obtain

$$\begin{aligned} T_\ell(x)\boldsymbol{e}(x)\boldsymbol{e}^H(x) &= \sum_{s=1}^{\ell-1} f(R_{\ell-s-1}, x) T_s(x)\boldsymbol{e}(x)\boldsymbol{e}^H(x)\boldsymbol{e}(x)\boldsymbol{e}^H(x) + f(R_{\ell-1}, x)\boldsymbol{T}_0(x)\boldsymbol{e}(x)\boldsymbol{e}^H(x) \\ &= r\sum_{s=1}^{\ell-1} f(R_{\ell-s-1}, x) T_s(x)\boldsymbol{e}(x)\boldsymbol{e}^H(x) + f(R_{\ell-1}, x) T_0'(x)\boldsymbol{e}(x)\boldsymbol{e}^H(x) \end{aligned} \tag{90}$$

Recalling that $\boldsymbol{T}_0(x) = \boldsymbol{I}_r$ and defining $T_0'(x) = \frac{1}{r}$, we obtain from (90) the scalar $T_\ell(x)$:

$$T_\ell(x) = r\left(\sum_{s=1}^{\ell-1} f(R_{\ell-s-1}, x) T_s(x) + f(R_{\ell-1}, x) T_0'(x)\right). \tag{91}$$

The following equations summarize the recursion in terms of only scalar functions.

$$\begin{aligned} R_\ell(\lambda) &= \sum_{s=0}^{\ell-1} g(T_{\ell-s-1}, \lambda) R_s(\lambda) \\ T_\ell(x) &= r\sum_{s=0}^{\ell-1} f(R_{\ell-s-1}, x) T_s(x) \\ f(R_s, x) &= \frac{\beta}{T_c^2}\left|\Phi\left(\frac{j2\pi x}{T_c}\right)\right|^2 \int \lambda R_s(\lambda)\mathrm{d}\, F_{|\boldsymbol{A}|^2}(\lambda) & -\frac{1}{2} \le x \le \frac{1}{2} \\ g(T_s, \lambda) &= \left(\frac{r}{T_c}\right)^2 \lambda \int_{-1/2}^{1/2} \left|\Phi\left(\frac{j2\pi x}{T_c}\right)\right|^2 T_s(x)\mathrm{d}\, x & s = 0, 1, \ldots \end{aligned}$$

with $T_0(x) = \frac{T_c}{r}$ and $R_0(\lambda) = 1$. Let us observe that the different expressions of $g(T_s, \lambda)$ for $s = 0, 1, \ldots$ could be absorbed in a unified expression by initialize the recursion with $T_0(x) = \frac{T_c}{r}$ instead of using $T_0'(x) = \frac{1}{r}$.

The recursion in the statement of Theorem 2 is obtained by defining

$$f(R_s) = \int \lambda R_s(\lambda)\mathrm{d}F_{|\boldsymbol{A}|^2}(\lambda)$$

and

$$\nu(T_s) = \left(\frac{r}{T_c}\right)^2 \int_{-1/2}^{1/2} \left|\Phi\left(\frac{j2\pi x}{T_c}\right)\right|^2 T_s(x)\mathrm{d}\, x$$

and by expressing $R_\ell(\lambda)$ and $T_\ell(x)$ as recursive functions of $f(R_s)$ and $\nu(T_s)$.

APPENDIX V

PROOF OF THEOREM 3

Theorem 3 in part I Appendix V [1] jointly with Lemma 8 i part I Appendix I [1] implies that the eigenvalues moments of the matrix $\widetilde{\boldsymbol{\mathcal{T}}}$ and, thus, the eigenvalue moments of the matrix $\widetilde{\boldsymbol{\mathcal{R}}}$ converge to the eigenvalue moments of the matrix $\widehat{\boldsymbol{T}}$ and $\widehat{\boldsymbol{R}}$. However, in Theorem 3 we show a stronger result: asymptotically, the diagonal elements of the matrix $\widetilde{\boldsymbol{\mathcal{R}}}$ are a periodical repetition of the diagonal elements of the matrix $\widehat{\boldsymbol{R}}$. The proof follows the same approach as for Theorem 1.

Throughout this proof we adopt the following notation. For $k = 1, \ldots, K$ and $n = 1, \ldots, N$

- $\widetilde{\boldsymbol{h}}_{k,m}$ is the unlimited column vector of the matrix $\widetilde{\boldsymbol{\mathcal{H}}}$ containing $\widetilde{\boldsymbol{\Phi}}_k \boldsymbol{s}_{km}$;
- $\widetilde{\boldsymbol{\delta}}_{n,m}$ is the unlimited block row of $\widetilde{\boldsymbol{H}}$ containing $\boldsymbol{\Delta}_{\Phi,r}\left(\frac{n-1}{N}, \widetilde{\tau}_1\right)(\boldsymbol{s}_{1m})_n$, the n^{th} $r \times 1$ block of the vector $\widetilde{\boldsymbol{\Phi}}_1 \boldsymbol{s}_{1m}$ of the reference user. Since we assume that the users are ordered according to an increasing time delay with respect to the reference user, the vector $\widetilde{\boldsymbol{\delta}}_{n,m}$ has K consecutive $r \times 1$ blocks that are nonzero and the remaining equal to zero. More specifically, $\widetilde{\boldsymbol{\delta}}_{n,m}$ is of the form

$$\widetilde{\boldsymbol{\delta}}_{n,m} = \left(\ldots \mathbf{0}, |a_{k_n+1,k_n+1}|^2 \boldsymbol{\Delta}\left(\frac{n-1}{N}, \widetilde{\tau}_{k_n+1}\right) s_{n,m-1,k_n-1}, \ldots, |a_{K,K}|^2 \boldsymbol{\Delta}\left(\frac{n-1}{N}, \widetilde{\tau}_K\right) s_{n,m-1,K}, \right.$$
$$\left. |a_{1,1}|^2 \boldsymbol{\Delta}\left(\frac{n-1}{N}, \widetilde{\tau}_1\right) s_{n,m,1}, \ldots, |a_{k_n,k_n}|^2 \boldsymbol{\Delta}\left(\frac{n-1}{N}, \widetilde{\tau}_{k_n}\right) s_{n,m,k_n}, \mathbf{0} \ldots \right),$$

  where k_n is the user with highest time delay lower than nT_c, i.e. $k_n = \max(\tau_i : \tau_i < nT_c, i = 1, \ldots K)$ and $s_{n,m,\ell}$ is the $\left(\left(N + n - \left\lfloor \frac{\tau_\ell}{T_c} \right\rfloor - 1\right) \mathrm{mod} N + 1\right)^{\mathrm{th}}$ component of the vector $\boldsymbol{s}_{\ell,m}$;
- $\widetilde{\boldsymbol{\mathcal{H}}}_{\vDash n,m}$ is the matrix obtained from $\widetilde{\boldsymbol{\mathcal{H}}}$ by suppressing $\widetilde{\boldsymbol{\delta}}_{n,m}$;
- $\widetilde{\boldsymbol{\mathcal{H}}}_{\sim k,m}$ is the matrix obtained from $\widetilde{\boldsymbol{\mathcal{H}}}$ by suppressing $\widetilde{\boldsymbol{h}}_{k,m}$.
- $\widetilde{\boldsymbol{\mathcal{T}}}_{\sim k,m} = \widetilde{\boldsymbol{\mathcal{H}}}_{\sim k,m} \widetilde{\boldsymbol{\mathcal{H}}}^H_{\sim k,m}$;
- $\widetilde{\boldsymbol{\mathcal{R}}}_{\vDash n,m} = \widetilde{\boldsymbol{\mathcal{H}}}^H_{\vDash n,m} \widetilde{\boldsymbol{\mathcal{H}}}_{\vDash n,m}$;
- $\widetilde{\boldsymbol{\sigma}}_{n,m} = (s_{n,m-1,k_n+1}, \ldots, s_{n,m-1,K}, s_{n,m,1} \ldots, s_{n,m,k_n})$;
- $\widetilde{\boldsymbol{A}}_n$ is the diagonal matrix obtained by circular diagonal down shift of the elements of the matrix $\boldsymbol{A}$ by $K - k_n$ steps, i.e. $\widetilde{\boldsymbol{A}}_n = \mathrm{diag}(a_{k_n+1,k_n+1}, \ldots, a_{K,K}, a_{1,1} \ldots, a_{k_n,k_n})$.
- $\widetilde{\boldsymbol{\nabla}}_{n,t}$, for $t = 1, \ldots, r$ and $n = 1, \ldots, N$, is the $K \times K$ diagonal matrix obtained by circular diagonal down shift of the elements of the matrix $\boldsymbol{\nabla}_{n,t}$, defined in the proof of Theorem 1, by $K - k_n$ steps, i.e.

$$\widetilde{\boldsymbol{\nabla}}_{n,t} = \mathrm{diag}\left(\phi\left(\frac{n-1}{N}, \widetilde{\tau}_{k_n+1} - \frac{t-1}{r} T_c\right), \ldots, \phi\left(\frac{n-1}{N}, \widetilde{\tau}_K - \frac{t-1}{r} T_c\right), \phi\left(\frac{n-1}{N}, \widetilde{\tau}_1 - \frac{t-1}{r} T_c\right), \ldots \phi\left(\frac{n-1}{N}, \widetilde{\tau}_{k_n} - \frac{t-1}{r} T_c\right)\right);$$

  Note that $\widetilde{\boldsymbol{\sigma}}_{n,m} \boldsymbol{\nabla}_{n,m,t} \boldsymbol{A}_n$ coincides with the vector of nonzero elements of t^{th} row of $\widetilde{\boldsymbol{\delta}}_{n,m}$.
- $\widetilde{\boldsymbol{\mathcal{T}}}^s_{[nn],m}$ is the diagonal block of the matrix $\widetilde{\boldsymbol{\mathcal{T}}}^s$ of dimensions $r \times r$ given by

$$\widetilde{\boldsymbol{\mathcal{T}}}^s_{[nn],m} = \widetilde{\boldsymbol{\delta}}_{n,m} \widetilde{\boldsymbol{\mathcal{R}}}^{s-1} \widetilde{\boldsymbol{\delta}}^H_{n,m};$$

- $\widetilde{\boldsymbol{\mathcal{R}}}^s_{k,m}$ is the diagonal element of the matrix $\widetilde{\boldsymbol{\mathcal{R}}}^s$ given by

$$\widetilde{\boldsymbol{\mathcal{R}}}^s_{k,m} = \widetilde{\boldsymbol{h}}^H_{k,m} \widetilde{\boldsymbol{\mathcal{T}}}^{s-1} \widetilde{\boldsymbol{h}}_{k,m}.$$

As the proof of Theorem 1, the proof of Theorem 3 is based on strong induction. In the first step we prove the following facts:

1) Conditionally on $(|a_{kk}|^2, \widetilde{\tau}_k)$, the diagonal elements of the matrix $\widetilde{\boldsymbol{\mathcal{R}}}$ converges in probability, as $N \to \infty$, to the deterministic value $R_1(|a_{kk}|^2, \widetilde{\tau}_k)$ given by the recursive equations (24)–(27) of Theorem 1. Since asymptotically $\widetilde{\boldsymbol{\mathcal{R}}}_{k,m}$ depends only on k via $(|a_{kk}|^2, \widetilde{\tau}_k)$ but not on m, asymptotically $\widetilde{\boldsymbol{\mathcal{R}}}_{k,m}$ is a periodic repetition of the diagonal elements of $\widehat{\boldsymbol{R}}$.

   Furthermore, as in Theorem 1, $\forall \varepsilon > 0$ and large $K = \beta N$

$$\Pr\{|\widetilde{\boldsymbol{\mathcal{R}}}_{kk} - R_1(|a_{kk}|^2, \widetilde{\tau}_k)| > \varepsilon\} \leq o\left(N^{-2}\right). \tag{92}$$

2) $\widetilde{\boldsymbol{\mathcal{T}}}_{[nn],m}$, the $r \times r$ block diagonal elements of the matrix $\widetilde{\boldsymbol{\mathcal{T}}}$, converge in probability to deterministic blocks $\boldsymbol{T}_1(x)$, with $x = \lim_{N\to\infty} \frac{n}{N}$ given by the recursive equations (24)–(27) of Theorem 1. Also in this case, asymptotically, the diagonal blocks of $\widetilde{\boldsymbol{\mathcal{T}}}$ are a periodic repetition of the diagonal $r \times r$ blocks of the matrix $\widehat{\boldsymbol{T}}$. Additionally, $\forall \varepsilon > 0$, large $K = \beta N$ and $u, v = 1, \dots r$,

$$\Pr\{|(\widetilde{\boldsymbol{\mathcal{T}}}_{[nn],m})_{uv} - (\boldsymbol{T}_1(x))_{uv}| > \varepsilon\} \leq o\left(N^{-2}\right). \tag{93}$$

Then, in the recursion step, we use the following induction assumptions:

1) For $s = 1, \dots, \ell - 1$, the diagonal elements of the matrix $\widetilde{\boldsymbol{\mathcal{R}}}^s$, $k = 1, \dots, K$ and $m \in \mathbb{Z}$, converge in probability, as $K = \beta N \to \infty$, to deterministic values $R_s(|a_{kk}|^2, \widetilde{\tau}_k)$ given by the recursive equations (24)–(27) of Theorem 1. Asymptotically they are a periodical repetition of the diagonal elements of $\widehat{\boldsymbol{R}}^s$. Furthermore, $\forall \varepsilon > 0$ and large $K = \beta N$ $\Pr\{|(\widetilde{\boldsymbol{\mathcal{R}}}^s)_{kk} - R_s(|a_{kk}|^2, \widetilde{\tau}_k)| > \varepsilon\} \leq o\left(N^{-2}\right)$.
2) For $s = 1, \dots, \ell - 1$, $\widetilde{\boldsymbol{\mathcal{T}}}^s_{[nn],m}$, the $r \times r$ block diagonal elements of the matrix $\widetilde{\boldsymbol{\mathcal{T}}}^s$ converge in probability to deterministic blocks $\boldsymbol{T}_s(x)$ given by the recursive equations the recursive equations (24)–(27) of Theorem 1. Asymptotically, they are a periodical repetition of the diagonal blocks of the matrix $\widehat{\boldsymbol{T}}^s$. Furthermore, $\forall \varepsilon > 0$, large $K = \beta N$ and $u, v = 1, \dots r$, $\Pr\{|(\widetilde{\boldsymbol{\mathcal{T}}}^s_{[nn],m})_{uv} - (\boldsymbol{T}_s(x))_{uv}| > \varepsilon\} \leq o\left(N^{-2}\right)$.

We prove:

1) The diagonal elements of the matrix $\widetilde{\boldsymbol{\mathcal{R}}}^\ell$, converge in probability, as $K = \beta N \to \infty$, to deterministic values $R^\ell(|a_{kk}|^2, \widetilde{\tau}_k)$, conditionally on $(|a_{kk}|^2, \widetilde{\tau}_k)$. Asymptotically, the diagonal elements of the matrix $\widetilde{\boldsymbol{\mathcal{R}}}^\ell$ are a periodical repetition of the limiting diagonal elements of the matrix $\widehat{\boldsymbol{R}}^\ell$. Furthermore, $\forall \varepsilon > 0$ and large $K = \beta N$

$$\Pr\{|(\widetilde{\boldsymbol{\mathcal{R}}}^s)_{kk} - R_s(|a_{kk}|^2, \widetilde{\tau}_k)| > \varepsilon\} \leq o\left(N^{-2}\right). \tag{94}$$

2) The blocks $\widetilde{\boldsymbol{\mathcal{T}}}^{\ell}_{[nn],m}$, converge in probability to deterministic blocks $\boldsymbol{T}^{\ell}(x)$ with $\lim_{N\to\infty}\frac{n}{N}$ given by the recursive equations (24)–(27) of Theorem 1. Asymptotically they are a periodical repetition of the limit diagonal blocks of the matrix $\widehat{\boldsymbol{\mathcal{T}}}^{s}$. Additionally, $\forall\varepsilon>0$, large N and $u,v=1,\dots r$,

$$\Pr\{|(\widetilde{\boldsymbol{\mathcal{T}}}^{\ell}_{[nn],m})_{uv}-(\boldsymbol{T}_{\ell}(x))_{uv}|>\varepsilon\}\leq o\left(N^{-2}\right). \tag{95}$$

First step: Consider

$$\begin{aligned}\widetilde{\boldsymbol{\mathcal{R}}}_{k,m} &= \widetilde{\boldsymbol{h}}^{H}_{k,m}\widetilde{\boldsymbol{h}}_{k,m}\\ &|a_{kk}|^{2}\boldsymbol{s}^{H}_{k,m}\widetilde{\boldsymbol{\Phi}}_{k}\widetilde{\boldsymbol{\Phi}}^{H}_{k}\boldsymbol{s}_{k,m}\\ &|a_{kk}|^{2}\boldsymbol{s}^{H}_{k,m}\widetilde{\boldsymbol{\Delta}}^{H}_{\phi,r}(\tau_k)\widetilde{\boldsymbol{\Delta}}_{\phi,r}(\tau_k)\boldsymbol{s}_{k,m}.\end{aligned} \tag{96}$$

Applying the same approach as in Theorem 1 we can prove that

$$\begin{aligned}\lim_{K=\beta N\to\infty}\widetilde{\boldsymbol{\mathcal{R}}}_{k,m} &= \lim_{K=\beta N\to\infty}\frac{|a_{kk}|^{2}}{N}\mathrm{tr}\widetilde{\boldsymbol{\Delta}}^{H}_{\phi,r}(\tau_k)\widetilde{\boldsymbol{\Delta}}_{\phi,r}(\tau_k)\\ &= \lim_{K=\beta N\to\infty}\frac{|a_{kk}|^{2}}{N}\mathrm{tr}\boldsymbol{\Delta}^{H}_{\phi,r}(\widetilde{\tau}_k)\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_k)\\ &= R_1(|a_{kk}|^{2},\widetilde{\tau}_k).\end{aligned} \tag{97}$$

The convergence rate (92) can be shown by applying the same arguments as in Theorem 1.

Let us now consider the block matrix $\widetilde{\boldsymbol{\mathcal{T}}}_{[nn],m}$ whose (u,v) element $(\widetilde{\boldsymbol{\mathcal{T}}}_{[nn],m})_{u,v}$ is given by

$$\begin{aligned}(\widetilde{\boldsymbol{\mathcal{T}}}_{[nn],m})_{u,v} &= (\widetilde{\boldsymbol{\delta}}_{n,m})_u(\widetilde{\boldsymbol{\delta}}^{H}_{n,m})_v\\ &= \widetilde{\boldsymbol{\sigma}}_{n,m}\widetilde{\boldsymbol{A}}_n\widetilde{\boldsymbol{\nabla}}_{n,m,u}\widetilde{\boldsymbol{\nabla}}^{H}_{n,m,v}\widetilde{\boldsymbol{A}}^{H}_{n}\widetilde{\boldsymbol{\sigma}}^{H}_{n,m}\end{aligned} \tag{98}$$

where $(\widetilde{\boldsymbol{\delta}}_{n,m})_j$ denotes the j^{th} row of $\widetilde{\boldsymbol{\delta}}_{n,m}$. By applying the same approach as in Theorem 1 we can prove that

$$\begin{aligned}\lim_{K=\beta N\to\infty}(\widetilde{\boldsymbol{\mathcal{T}}}_{[nn],m})_{u,v} &= \lim_{K=\beta N\to\infty}\frac{1}{N}\mathrm{tr}\left(\widetilde{\boldsymbol{A}}_n\widetilde{\boldsymbol{\nabla}}_{n,m,u}\widetilde{\boldsymbol{\nabla}}^{H}_{n,m,v}\widetilde{\boldsymbol{A}}^{H}_{n}\right)\\ &= \lim_{K=\beta N\to\infty}\frac{1}{N}\mathrm{tr}\left(\boldsymbol{A}\boldsymbol{\nabla}_{n,u}\boldsymbol{\nabla}^{H}_{n,v}\boldsymbol{A}^{H}\right) \qquad (99)\\ &= (\boldsymbol{T}_1(x))_{u,v} \qquad (100)\end{aligned}$$

Equation (99) follows from the fact that $\widetilde{\boldsymbol{A}}_n$ and $\widetilde{\boldsymbol{\nabla}}_{n,m,u}$ are obtained from $\boldsymbol{A}$ and $\boldsymbol{\nabla}_{n,u}$, respectively, by circular diagonal down shift of $K-k_n$ steps. The asymptotic periodical behaviour of the $\widetilde{\boldsymbol{\mathcal{T}}}_{[nn],m}$ is a direct consequence of the fact that $\lim_{K=\beta N\to\infty}\widetilde{\boldsymbol{\mathcal{T}}}_{[nn],m}$ does not depend on m. The convergence rate (93) can be shown by appealing to the same arguments as in Theorem 1. This concludes the proof of the first step.

Step ℓ: As in Theorem 1, we can expand $(\widetilde{\boldsymbol{\mathcal{R}}}^{\ell})_{k,m}$ and $\widetilde{\boldsymbol{\mathcal{T}}}^{\ell}_{[nn],m}$ as follows:

$$(\widetilde{\boldsymbol{\mathcal{R}}}^{\ell})_{k,m}=\sum_{s=0}^{\ell-1}\widetilde{\boldsymbol{h}}^{H}_{k,m}\widetilde{\boldsymbol{\mathcal{T}}}^{\ell-s-1}_{\sim k,m}\widetilde{\boldsymbol{h}}_{k,m}(\widetilde{\boldsymbol{\mathcal{R}}}^{s})_{k,m} \qquad \ell=1,2,\dots \tag{101}$$

$$\widetilde{\boldsymbol{\mathcal{T}}}^{\ell}_{[nn],m}=\sum_{s=0}^{\ell-1}\widetilde{\boldsymbol{\delta}}_{n,m}\widetilde{\boldsymbol{\mathcal{R}}}^{\ell-s-1}_{\vDash n,m}\widetilde{\boldsymbol{\delta}}^{H}_{n,m}\widetilde{\boldsymbol{\mathcal{T}}}^{s}_{[nn],m}. \qquad \ell=1,2,\dots \tag{102}$$

being $\widetilde{\boldsymbol{\mathcal{T}}}^{0}$ and $\widetilde{\boldsymbol{\mathcal{R}}}^{0}$ identity matrices of unlimited dimensions.

By applying the same arguments as in Theorem 1 the convergence in probability of the sequences $\{(\widetilde{\boldsymbol{\mathcal{R}}}^{\ell})_{k,n}\}$ and $\{\widetilde{\boldsymbol{\mathcal{T}}}^{\ell}_{[nn],m}\}$ reduces to show the convergence in probability of $\widetilde{\boldsymbol{h}}^{H}_{k,m}\widetilde{\boldsymbol{\mathcal{T}}}^{s}_{\sim k,m}\widetilde{\boldsymbol{h}}_{k,m}$ and $\widetilde{\boldsymbol{\delta}}_{n,m}\widetilde{\boldsymbol{\mathcal{R}}}^{s}_{\vDash n,m}\widetilde{\boldsymbol{\delta}}^{H}_{n,m}$, for $s=0,\dots,\ell-1$, respectively.

As in Theorem 1 it can be shown that

$$\begin{aligned}\lim_{K=\beta N\to\infty}(\widetilde{\boldsymbol{\delta}}_{n,m})_u\widetilde{\boldsymbol{\mathcal{R}}}^{s}_{\vDash n,m}(\widetilde{\boldsymbol{\delta}}^{H}_{n,m})_v &= \lim_{K=\beta N\to\infty}\frac{1}{N}\sum_{j=1}^{k_n}|a_{jj}|^2\phi(\frac{n-1}{N},\widetilde{\tau}_j-\frac{u-1}{r}T_c)\phi^*(\frac{n-1}{N},\widetilde{\tau}_j-\frac{v-1}{r}T_c)\widetilde{\boldsymbol{\mathcal{R}}}^{s}_{j,m}\\ &+\lim_{K=\beta N\to\infty}\frac{1}{N}\sum_{j=k_n+1}^{K}|a_{jj}|^2\phi(\frac{n-1}{N},\widetilde{\tau}_j-\frac{u-1}{r}T_c)\phi^*(\frac{n-1}{N},\widetilde{\tau}_j-\frac{v-1}{r}T_c)\widetilde{\boldsymbol{\mathcal{R}}}^{s}_{j,m-1}\\ &=\lim_{N=\beta K\to\infty}\lim_{K=\beta N\to\infty}\frac{1}{N}\sum_{j=1}^{K}|a_{jj}|^2\phi(\frac{n-1}{N},\widetilde{\tau}_j-\frac{u-1}{r}T_c)\phi^*(\frac{n-1}{N},\widetilde{\tau}_j-\frac{v-1}{r}T_c)\widehat{\boldsymbol{R}}^{s}_{jj}\\ &=(\boldsymbol{f}(R_s,x))_{u,v}.\end{aligned}$$

Note that also the $r\times r$ blocks $\widetilde{\boldsymbol{\delta}}_{n,m}\widetilde{\boldsymbol{\mathcal{R}}}^{s}_{\vDash n,m}\widetilde{\boldsymbol{\delta}}^{H}_{n,m}$ are independent of m, asymptotically, and equal to the corresponding blocks $\boldsymbol{\delta}\widehat{\boldsymbol{R}}^{s}_{\vDash n}\boldsymbol{\delta}^{H}$. Thus, they are a periodical repetition of the blocks $\boldsymbol{\delta}\widehat{\boldsymbol{R}}^{s}_{\vDash n}\boldsymbol{\delta}^{H}$. By making use of the induction assumptions and the recursive equation (102) we obtain

$$\lim_{K=\beta N\to\infty}\widetilde{\boldsymbol{\mathcal{T}}}^{\ell}_{[nn],m}\overset{\mathcal{P}}{=}\boldsymbol{T}^{\ell}(x)=\sum_{s=0}^{\ell-1}\boldsymbol{f}(R_{\ell-s-1},x)\boldsymbol{T}_s(x).$$

Applying the same arguments as in Theorem 1 we can show that (95) is satisfied.

Similarly, for $\widetilde{\boldsymbol{h}}_{k,m}^H \widetilde{\boldsymbol{\mathcal{T}}}_{\sim k,m}^s \widetilde{\boldsymbol{h}}_{k,m}$, $s = 0, \dots, \ell - 1$ it can be shown that

$$\begin{aligned}
\lim_{K=\beta N\to\infty} \widetilde{\boldsymbol{h}}_{k,m}^H \widetilde{\boldsymbol{\mathcal{T}}}_{\sim k,m}^s \widetilde{\boldsymbol{h}}_{k,m} &= \lim_{K=\beta N\to\infty} \frac{|a_{kk}|^2}{N} \sum_{n=\left\lfloor \frac{\tau_k}{T_c} \right\rfloor+1}^{N} (\widetilde{\boldsymbol{\Phi}}_k^H)_{[nn]} \widetilde{\boldsymbol{\mathcal{T}}}_{[nn],m}^s (\widetilde{\boldsymbol{\Phi}}_k)_{[nn]} \\
&+ \lim_{K=\beta N\to\infty} \frac{|a_{kk}|^2}{N} \sum_{n=1}^{\left\lfloor \frac{\tau_k}{T_c} \right\rfloor} (\widetilde{\boldsymbol{\Phi}}_k^H)_{[nn]} \widetilde{\boldsymbol{\mathcal{T}}}_{[nn],m+1}^s (\widetilde{\boldsymbol{\Phi}}_k)_{[nn]} \\
&= \lim_{K=\beta N\to\infty} \frac{|a_{kk}|^2}{N} \sum_{n=1}^{N} \boldsymbol{\Delta}_{\phi,r}^H \left(\frac{n-1}{N}, \widetilde{\tau}_k \right) \widehat{\boldsymbol{T}}_{[nn]}^s \boldsymbol{\Delta}_{\phi,r} \left(\frac{n-1}{N}, \widetilde{\tau}_k \right) \\
&= g(\boldsymbol{T}_s, |a_{kk}|^2, \widetilde{\tau}_k)
\end{aligned} \tag{103}$$

where $(\boldsymbol{X})_{[nn]}$, for $\boldsymbol{X} = \widetilde{\boldsymbol{\Phi}}, \widetilde{\boldsymbol{\Delta}}_{\phi,r}(\tau), \boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau})$ denotes the n block diagonal elements of size $r \times 1$ of the matrix $\boldsymbol{X}$. In (103) we make use of the fact that for $n = \left\lfloor \frac{\tau_k}{T_c} \right\rfloor + 1, \dots \left\lfloor \frac{\tau_k}{T_c} \right\rfloor + N + 1$

$$\begin{aligned}
(\widetilde{\boldsymbol{\Phi}}_k)_{[nn]} &= \left(\widetilde{\boldsymbol{\Delta}}_{\phi,r}(\tau_k) \right)_{\left[n-\left\lfloor \frac{\tau_k}{T_c} \right\rfloor-1, n-\left\lfloor \frac{\tau_k}{T_c} \right\rfloor-1\right]} \\
&= \left(\boldsymbol{\Delta}_{\phi,r}(\widetilde{\tau}_k) \right)_{[(n-1)\mathrm{mod}N+1, (n-1)\mathrm{mod}N+1]} \\
&= \boldsymbol{\Delta}_{\phi,r} \left(\frac{(n-1)\mathrm{mod}N}{N}, \widetilde{\tau}_k \right).
\end{aligned} \tag{104}$$

The induction assumptions and the recursive equation (101) yield

$$\lim_{K=\beta N\to\infty} (\widetilde{\boldsymbol{\mathcal{R}}})_{k,m} = R_\ell(|a_{kk}|^2, \widetilde{\tau}_k) = \sum_{s=0}^{\ell-1} g(\boldsymbol{T}_{\ell-s-1}, \widetilde{\tau}_k) R_s(|a_{kk}|^2, \widetilde{\tau}_k). \tag{105}$$

From the previous equation it becomes apparent the periodical asymptotic behaviour of the diagonal elements of the matrix $\widetilde{\boldsymbol{\mathcal{R}}}^\ell$. The inequality (94) follows along the same lines as in the proof of Theorem 1. Furthermore, the limit of $(\widetilde{\boldsymbol{\mathcal{R}}}^\ell)_{k,m}$ and $(\widetilde{\boldsymbol{\mathcal{T}}}^\ell)_{[nn],m}$ coincide with the equations for the limits of $(\widehat{\boldsymbol{R}}^\ell)_{kk}$ and $(\widehat{\boldsymbol{T}}^\ell)_{[nn]}$. This concludes the proof of Theorem 3.

REFERENCES

[1] L. Cottatellucci, R. R. Müller, and M. Debbah, "Asynchronous CDMA systems with random spreading–part i: Fundamental limits," *Submitted to IEEE Transactions on Information Theory*, Feb. 2007.

[2] L. Cottatellucci and R. R. Müller, "A systematic approach to multistage detectors in multipath fading channels," *IEEE Transactions on Information Theory*, vol. 51, no. 9, pp. 3146–3158, Sept. 2005.

[3] S. Verdú, *Multiuser Detection*. New York: Cambridge University Press, 1998.

[4] 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.

[5] 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.

[6] 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.

[7] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 3rd ed. Baltimore and London: The Johns Hopkins University Press, 1996.

[8] D. Divsalar and M. K. Simon, "Improved CDMA performance using parallel interference cancellation," in *Proc. of Military Communications Conference*, Oct. 1994, pp. 911–917.

[9] 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.

[10] 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.

[11] 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.

[12] ——, "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.

[13] 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.

[14] 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.

[15] Laura Cottatellucci, Merouane Debbah, and R. R. Müller, "Asymptotic analysis of linear detectors for asynchronous CDMA systems," in *Proc. of IEEE International Symposium on Information Theory (ISIT)*, Chicago, Illinois, June/July 2004.

[16] L. Cottatellucci, R. R. Müller, and M. Debbah, "Efficient implementation of multiuser detectors for asynchronous CDMA," in *Proc. 42nd Allerton Conf. on Communication, Control and Computing*, Monticello, Illinois, Sept./Oct. 2004, pp. 357–366.

[17] H. A. van der Vorst, *Iterative Krylov Methods for Large Linear Systems*. Cambridge, U.K.: Cambridge University Press, 2003.

[18] L. Cottatellucci, "Low complexity multiuser detectors with random spreading," Ph.D. dissertation, TU Wien, Vienna, Austria, Mar. 2006.

[19] D. R. Brown, M. Motani, V. V. Veravalli, V. Poor, and C. R. Johnson, "On the performance of linear parallel interference cancellation," *IEEE Transactions on Information Theory*, pp. 1957–1970, July 2001.

[20] D. Guo, L. K. Rasmussen, and T. J. Lim, "Linear parallel interference cancellation in long–code CDMA multiuser detection," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 12, pp. 2074–2081, Dec. 1999.

[21] S. Kay, *Fundamentals of Statistical Signal Processing, Estimation Theory*, ser. Prentice Hall Signal Processing Series. Prentice Hall, 1993, vol. 1.

[22] 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.

[23] 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.

[24] 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.

[25] 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.

[26] J. Huber, *Trelliscodierung*. Berlin, Germany: Springer–Verlag, 1992.

[27] R. Speicher, "Freie Wahrscheinlichkeitstheorie," Lecture Notes, Heidelberg, Germany, 1997/98.

[28] Y. Q. Yin, "Limiting spectral distribution for a class of random matrices," *Journal of Multivariate Analysis*, vol. 20, pp. 50–68, 1986.

[29] K. W. Wachter, "The strong limits of random matrix spectra for sample matrices of independent elements," *The Annals of Probability*, vol. 6, no. 1, pp. 1–18, 1978.