Asymptotic Normality of Linear Multiuser Receiver Outputs

Dongning Guo, Sergio Verdú, and Lars K. Rasmussen

Abstract— **This paper proves large-system asymptotic normality of the output of a family of linear multiuser receivers that can be arbitrarily well approximated by polynomial receivers. This family of receivers encompasses the single-user matched filter, the decorrelator, the MMSE receiver, the parallel interference cancelers, and many other linear receivers of interest. Both with and without the assumption of perfect power control, we show that the output decision statistic for each user converges to a Gaussian random variable in distribution as the number of users and the spreading factor both tend to infinity with their ratio fixed. Analysis reveals that the distribution conditioned on almost all spreading sequences converges to the same distribution, which is also the unconditional distribution. This normality principle allows the system performance, e.g., the multiuser efficiency, to be completely determined by the output signal-to-interference ratio for large linear systems.**

Index Terms: Code-division multiple access, multiuser detection, multiple access interference, central limit theorem, multiuser efficiency, signal-to-interference ratio.

I. Introduction

Linear multiuser receivers for multiple access channels have been studied extensively during the last two decades due to their performance capabilities and analytical tractability [1]. A linear receiver provides a soft output that can be either hard limited for decision-making or treated as soft reliability information for further processing such as in coded transmission [2], [3]. The conventional matched filter, the decorrelator and the minimum mean square error (MMSE) receiver are among the earliest and most well-known linear receivers. More recently, linear interference cancelers have also been analyzed [4], [5], [6], [7], [8], [9].

The performance, in particular, the uncoded bit-error-rate (BER) of a linear receiver, depends on the cumulative distribution of the multiple access interference (MAI), which is in general a discrete distribution. For all but the decorrelator, the BER is given as a sum of an exponential number of Gaussian error functions (Q-functions[1]), the evaluation of which is infeasible for even moderately sized systems.

To circumvent this difficulty, a Gaussian approximation of the MAI is often used. Weber *et al.* [10] were among the earliest to model interfering users' signals as white Gaussian noise. Since then a large amount of work has been dedicated to the justification of various normality approximations. For a long-code system, the MAI embedded in the output of the matched filter for each user can be well approximated by a Gaussian random variable [1], [11]. For a short-code system, the system size has to be much larger in order for the error in this approximation to normality to be negligible. Nevertheless, the distribution of the MAI in a matched filter output is shown to converge to a Gaussian law in the sense of divergence by Verdú and Shamai [11]. In fact, [11] is one of the first to prove asymptotic normality for an output distribution conditioned on the spreading sequences. Moreover, Poor and Verdú [12] showed that the distribution of the output MAI of the MMSE receiver in a short-code system often has no noticeable difference to a Gaussian law. Recently, it was proved rigorously by Zhang *et al.* that the MAI in the MMSE receiver output is asymptotically Gaussian [13]. Normality is also established for linear blind multiuser receivers [14].

These normality results allow the large-system probability of error of these receivers to be quantified by a single Q-function of the square root of the output signal-to-interference ratio (SIR). It also implies that error-control codes for Gaussian channels are asymptotically optimal if autonomous single-user decoding is to be used. As a result, the receiver and the decoder can be designed and optimized separately.

The normality property of all the above mentioned receivers is not accidental. Indeed, this is a result of the central limit theorem due to the fact that the MAI is a sum of contributions from a large number of users. In this paper, we extend the normality principle to a much wider family of linear receivers, which is defined as a set of matrix filters each of which shares the same eigenvectors as those of the channel correlation matrix and takes eigenvalues given by a function of the eigenvalues of the correlation matrix. Immediately, this family includes the conventional receiver, the decorrelator, and the MMSE receiver. Also, it includes a subset of matrix filters described by polynomial expansion of the correlation matrix, called polynomial receivers [15], which corresponds exactly to the set of linear multistage parallel interference cancelers [16].

Our results are asymptotic in nature, namely, they are large-system limits where the number of users and the spreading factor both tend to infinity with a fixed ratio. Recent work in [17], [18], [11], [13], [19] shows that this approach can average out the dependence on specific spreading sequences and result in simple expressions for system performance. Moreover, asymptotic results provide good approximations for moderately sized systems in many cases of practical interest.

Our results can be summarized as follows. Assuming random spreading sequences, the output decision statistic of the family of linear receivers for each user is asymptotically Gaussian in distribution conditioned on one's own

This work was supported in parts by the National Science Foundation under grant no. CCR0074277 and by the Swedish Research Council for Engineering Sciences under grant no. 217-1997-538.

[1] $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2} dt$

transmitted symbol. Moreover, the asymptotic output distribution of a receiver in this family is the same conditioned on almost all possible spreading sequences, which is also the asymptotic unconditional distribution. The normality principle holds for very general scenarios. The spreading sequences are not limited to binary sequences and the received energies of the users can be different.

Our results follow the work of Zhang *et al.* [13], where asymptotic normality was concluded for the MMSE receiver only. Besides making use of some of the same mathematical tools as in [13], our proof is quite different. We tackle the problem by starting from polynomial receivers. We show that the output decision statistic of every polynomial receiver is asymptotically Gaussian in distribution. We then generalize the normality principle to a large family of receivers that can be approximated arbitrarily by polynomial receivers.

This paper is organized as follows. Section II introduces the system model and notation. Polynomial receivers are studied under the perfect power control assumption in Section III. The normality principle is generalized to a family of receivers in Section IV. Arbitrary energy distribution is considered in Section V. Results on some popular receivers are summarized in Section VI. Some numerical examples are given in Section VII.

II. System Model

A. CDMA Uplink Channels

We assume a symbol-synchronous CDMA system depicted in Fig. 1 where each user's spreading sequence is independently and randomly chosen. For the purpose of large-system analysis, we consider first a K-user system with a spreading factor of $N = N(K)$ and then let K and N go to infinity with their ratio converging to a non-negative number β, i.e.,

$$\lim_{K\to\infty} \frac{K}{N(K)} = \beta. \tag{1}$$

Let $\{P_k \,|\, k = 1, 2, \dots\}$ be a deterministic sequence consisting of the received energies per symbol of all users. The signal-to-interference ratio[2] of user k in absence of interfering users is therefore $\frac{P_k}{\sigma^2}$. By absorbing a common factor into the noise level, we can assume without loss of generality for a K-user system

$$\frac{1}{K}\sum_{k=1}^{K} P_k = 1. \tag{2}$$

In Section III and IV, we study the simplest perfect power control case, where the received energies are equal from all users, i.e., $P_k = 1$ for $k = 1, \dots, K$. In Section V, we allow the received energies to be different. We assume, however, that the empirical distributions of $\{P_1, \dots, P_K\}$ converge as $K \to \infty$ to a distribution F_P, called the *energy distribution*, which has finite moments of any order.

Let $\{\bar{s}_{nk} \,|\, n = 1, 2, \cdots,\ k = 1, 2, \cdots\}$ be an infinite array of real-valued independent identically distributed (i.i.d.) random chips. For convenience we assume that the distribution of the chips has unit-variance and finite higher order moments, and is symmetric, i.e., $-\bar{s}_{nk}$ follows the same distribution as $\bar{s}_{nk}$. Consider a K-user system. The spreading sequence of user k is given as an N-dimensional vector, $\bar{\mathbf{s}}_k^{(K)} = \frac{1}{\sqrt{N}}[\bar{s}_{1k}, \bar{s}_{2k}, \cdots, \bar{s}_{Nk}]^\top$. We define the unnormalized spreading sequence of user k to be

$$\mathbf{s}_k^{(K)} = \frac{1}{\sqrt{N}}[s_{1k}, s_{2k}, \cdots, s_{Nk}]^\top = \sqrt{P_k}\,\bar{\mathbf{s}}_k^{(K)}. \tag{3}$$

We also define the unnormalized correlation matrix $\mathbf{R}^{(K)}$ as a $K \times K$ random matrix with its element on the k^{th} row and the j^{th} column as the crosscorrelation of user k and user j's unnormalized spreading sequences,

$$\left[\mathbf{R}^{(K)}\right]_{kj} = \mathbf{s}_k^\top \mathbf{s}_j = \frac{1}{N}\sum_{n=1}^{N} s_{nk}s_{nj} = \frac{\sqrt{P_k P_j}}{N}\sum_{n=1}^{N} \bar{s}_{nk}\bar{s}_{nj}. \tag{4}$$

Note that we label the correlation matrix with its corresponding system dimension (the number of users). In fact, every variable pertaining to a K-user system has K as its corresponding dimension. For notational convenience, we will often omit the index (K) when the dimension is understood from the context. For instance, $\mathbf{R}^{(K)}$ is simplified to $\mathbf{R}$.

Let $\{d_k \,|\, k = 1, 2, \cdots\}$ be a sequence of independent random antipodal modulated symbols taking the value of $+1$ and -1 equally likely. For a K-user system, let $\mathbf{d} = [d_1, \cdots, d_K]^\top$ be the vector of transmitted symbols.

A set of sufficient decision statistics is obtained by matched filtering using all user's unnormalized spreading sequences[3]

$$\mathbf{y}_{\text{MF}} = \mathbf{R}\mathbf{d} + \mathbf{n} \tag{5}$$

where the correlation matrix $\mathbf{R}$ is determined by (4), and $\mathbf{n}$ is a zero-mean Gaussian noise vector with covariance matrix $\sigma^2\mathbf{R}$, where σ^2 is the noise sample variance.

B. Linear Receivers

A linear receiver assumes knowledge of the spreading sequences, the received energies, as well as the noise variance and makes use of this information in detection. Mathematically, it is a $K \times K$ matrix filter $\mathbf{G}$, dependent on $\mathbf{R}$ and σ^2, applied to the matched-filter output. It outputs a vector of decision statistics expressed as

$$\begin{aligned} \mathbf{y} &= \mathbf{G}\cdot\mathbf{y}_{\text{MF}} && (6)\\ &= \mathbf{G}\cdot(\mathbf{R}\mathbf{d}+\mathbf{n}) && (7)\\ &= (\mathbf{G}\mathbf{R})\cdot\mathbf{d}+\mathbf{z} && (8)\end{aligned}$$

where $\mathbf{z} = \mathbf{G}\mathbf{n}$ is a zero-mean Gaussian random vector with covariance matrix $\sigma^2\mathbf{G}\mathbf{R}\mathbf{G}^\top$. We denote $\mathbf{G}\mathbf{R}$ by $\mathbf{H}$

[2]The SIR is defined as the energy ratio of the useful signal to the noise in the output. In contrast, the signal-to-noise ratio (SNR) of user k is usually defined as the ratio of the input energy and single-sided noise spectral density $P_k/(2\sigma^2)$.

[3]This is in contrast to matched filtering using normalized spreading sequences as in [1].

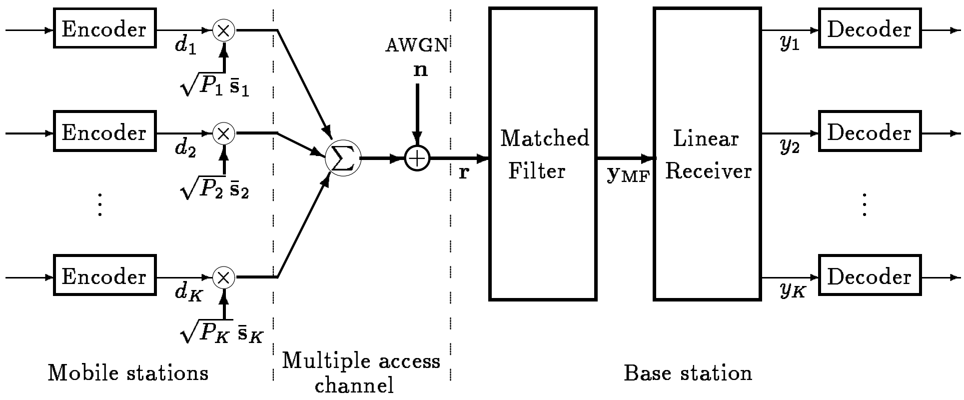

Fig. 1. Discrete time system model.

throughout the paper for notational simplicity. Thus, we have a simple linear system expressed as

$$\mathbf{y} = \mathbf{H} \cdot \mathbf{d} + \mathbf{z}. \tag{9}$$

An advantage of linear receivers is that they can be implemented in a decentralized fashion [1]. Independent single-user decoding is conducted based on the decision statistic sequence produced by the individual linear receiver. From an individual user's point of view, the multiaccess channel collapses to a single-user channel by treating the MAI as noise. The input-to-output characteristic of such a single-user channel is determined by the distribution of the output decision statistic conditioned on the transmitted symbol.

Without loss of generality, we consider user 1. As far as the performance is concerned, we can assume that user 1 always transmits +1.[4] User 1's decision statistic is a scalar

$$y_1 = H_{11} + \sum_{k=2}^{K} H_{1k} d_k + z_1 \tag{10}$$

where H_{kj}, like $[\mathbf{H}]_{kj}$, denotes the element of $\mathbf{H}$ on the k^{th} row and the j^{th} column. Clearly, this decision statistic consists of the transmitted symbol ($d_1 = 1$) scaled by H_{11}, the multiaccess interference aggregated from all the other users, and a Gaussian noise term.

We can interpret the randomness of the decision statistic in two different ways which correspond to short-code and long-code systems respectively. In a short-code system, the spreading sequences are randomly picked at the beginning of transmission and remain the same for every transmitted symbol. For each channel use, the randomness in y_1 includes that of the transmitted symbols and that of the noise. The performance, e.g. the uncoded probability of error, can be easily obtained once we have the distribution of y_1 conditioned on the spreading sequence. In a long-code system, the spreading sequences are randomly and independently chosen symbol-by-symbol. The randomness of y_1 then also includes the randomness of the crosscorrelations reflected in matrix $\mathbf{H}$. The expected performance, for instance the uncoded error probability averaged over all spreading sequences, is now better characterized by the unconditional distribution of y_1. In this paper we address both the conditional distribution and the unconditional distribution as the system size increases without bound. The resulting distributions turn out to be the same for all but a negligible set of spreading sequences.

[4] Accordingly, all distributions we will be considering are implicitly conditioned on $d_1 = 1$.

C. A Family of Linear Receivers

Note that $\mathbf{R}$ is symmetric. It has an eigen-decomposition as

$$\mathbf{R} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^\top \tag{11}$$

where $\mathbf{U}$ is unitary and $\boldsymbol{\Lambda} = \operatorname{diag}(\lambda_1, \cdots, \lambda_K)$ is a diagonal matrix consisting of the eigenvalues of $\mathbf{R}$, which are all nonnegative random variables dependent on $\mathbf{R}$. We limit our study to receivers taking the form of

$$\mathbf{G} = \mathbf{U} \cdot \operatorname{diag}(g(\lambda_1), \cdots, g(\lambda_K)) \cdot \mathbf{U}^\top \tag{12}$$

for some real continuous function g, and refer to them as *the family of linear receivers* throughout this paper. With slight abuse of notation we denote the right hand side of (12) by $g(\mathbf{R})$. Note that $g(\mathbf{R})$ is symmetric and shares the same eigenvectors with $\mathbf{R}$, and the eigenvalues of $g(\mathbf{R})$ are given by the function g evaluated at the eigenvalues of $\mathbf{R}$.

The family of receivers defined in the above represents a subset of linear receivers. It does not include successive interference cancelers, which, unlike (12), treat users unequally. Neither does it include the optimum linear detector in the sense of BER or asymptotic multiuser efficiency [1, Page 288]. Nonetheless, a wide spectrum of linear receivers belong to this family. In particular, if the function g degenerates to a constant 1, the receiver $\mathbf{G}$ is reduced to the single-user matched filter. If g is a polynomial, $\mathbf{G}$ becomes a polynomial receiver (or, equivalently, a parallel interference canceler [16]). If we let

$$g(\lambda) = \begin{cases} \frac{1}{\lambda} & \text{if } \lambda > 0, \\ 0 & \text{if } \lambda = 0, \end{cases} \tag{13}$$

the resulting $\mathbf{G}$ is the decorrelator. If we let $g(\lambda) = (\lambda + \sigma^2)^{-1}$, $\mathbf{G}$ becomes the MMSE receiver.

Let us include the dimension index for the rest of this section to make it clear what we mean by a linear receiver in the large-system analysis. A linear receiver here refers to a sequence of matrix filters $\left\{\mathbf{G}^{(K)}\right\}_{K=1}^{\infty}$, each $\mathbf{G}^{(K)}$ a function of $\mathbf{R}^{(K)}$, for which the vector of output decision statistics is expressed as

$$\mathbf{y}^{(K)} = \mathbf{G}^{(K)} \cdot \mathbf{y}_{\mathrm{MF}}^{(K)}. \tag{14}$$

In particular, a linear receiver in the family of our interest is defined as a sequence of receivers specified by a function g, i.e., for $K = 1, 2, \dots,$

$$\mathbf{G}^{(K)} = \mathbf{U}^{(K)} \cdot \mathrm{diag}\left(g(\lambda_1^{(K)}), \cdots, g(\lambda_K^{(K)})\right) \cdot \left(\mathbf{U}^{(K)}\right)^{\top} \tag{15}$$

where

$$\mathbf{R}^{(K)} = \mathbf{U}^{(K)} \cdot \mathrm{diag}\left(\lambda_1^{(K)}, \cdots, \lambda_K^{(K)}\right) \cdot \left(\mathbf{U}^{(K)}\right)^{\top}. \tag{16}$$

We study the marginal probabilistic law of the decision statistic $\mathbf{y}^{(K)}$, i.e., the distribution of $y_1^{(K)}$, as K goes to infinity.

D. Eigenvalue Distribution

The normality results we will show in the following sections hinge on the intriguing fact that the limiting empirical distribution of the eigenvalues of a large random covariance matrix is deterministic. Denote the limit of the cumulative distribution of the eigenvalues of the random matrix $\mathbf{R}^{(K)}$ by F_Λ. It is dependent on the received energy distribution F_P. In general, this distribution function does not have a closed-form solution. Its Stieltjes transform, defined as

$$m(z) = \int \frac{1}{\lambda - z}\, dF_\Lambda(\lambda) \tag{17}$$

satisfies

$$m(z) = \left[-z + \beta^{-1} \int \frac{P}{1 + Pm(z)}\, dF_P(P)\right]^{-1} \tag{18}$$

where F_P is the energy distribution [20], [21].

In the equal-energy case, a closed-form solution exists [22]

$$F_\Lambda(\lambda) = \max\left(0, 1 - \frac{1}{\beta}\right) \cdot u(\lambda) + \int_{-\infty}^{\lambda} p_\beta(t)\ dt \tag{19}$$

where $u(\lambda)$ is a unit step function, and

$$p_\beta(\lambda) = \begin{cases} \frac{1}{2\pi\beta\lambda}\sqrt{(\lambda - \lambda_{\min})(\lambda_{\max} - \lambda)} & \text{if } \lambda_{\min} < \lambda < \lambda_{\max}, \\ 0, & \text{otherwise,} \end{cases} \tag{20}$$

with $\lambda_{\min} = (1 - \sqrt{\beta})^2$ and $\lambda_{\max} = (1 + \sqrt{\beta})^2$. An expression for the moments of the eigenvalues has been developed in [22]

$$\mathsf{E}\left\{\lambda^i\right\} = \sum_{j=0}^{i-1} \frac{1}{j+1} \binom{i}{j} \binom{i-1}{j} \beta^j. \tag{21}$$

III. Polynomial Receivers Under Perfect Power Control

In this section, we study a subset of linear receivers, namely, the set of polynomial receivers, special cases of which have been considered in [15], [5], [23]. A polynomial receiver $\mathbf{G}$ is of the form

$$\mathbf{G} = \sum_{i=1}^{m} x_i \mathbf{R}^{i-1} \tag{22}$$

where m is an integer, and the weights x_i are arbitrary deterministic real numbers. $\mathbf{G}$ can be expressed as a member of the family of receivers defined by (12) if we let g be a polynomial

$$g(\lambda) = \sum_{i=1}^{m} x_i \lambda^{i-1}. \tag{23}$$

We study the output distribution of the polynomial receiver $\mathbf{G}$ in both the conditional case and the unconditional case.

In this section, we also limit ourselves to the equal-energy case, i.e., $P_k = 1$, $k = 1, \dots, K$, for which we obtain simple expressions for the limiting distribution. The results are extended to unequal-energy case in Section V.

A. Conditional Distribution

Consider the MAI term in the decision statistic y_1 in (10) where $\mathbf{H}$ is given. It is a sum of contributions from all interfering users. Due to the central limit theorem, its distribution becomes closer and closer to a Gaussian law as the system size increases without bound. Precisely, we have the following theorem.

Theorem 1: The decision statistic (10) as a function of $\mathbf{H}$, where the polynomial receiver is given as (22), converges to a Gaussian random variable in distribution with probability 1. The mean value corresponding to the limiting distribution is

$$\mu_1 = \sum_{i=1}^{m} x_i M_i \tag{24}$$

and the variance is

$$\sigma_1^2 = \sum_{i=1}^{m}\sum_{j=1}^{m} x_i x_j [M_{i+j} - M_i M_j + \sigma^2 M_{i+j-1}] \tag{25}$$

where M_i is defined as

$$M_i = \lim_{K\to\infty} \mathsf{E}\left\{\left[\mathbf{R}^i\right]_{11}\right\}. \tag{26}$$

In the special case of perfect power control, M_i is equal to the i^{th}-order moment of the limiting eigenvalue distribution given by (21).

Note that surprisingly the mean and the variance are not dependent on $\mathbf{R}$, since M_i is an average over all spreading sequences. Indeed, the theorem states that the asymptotic conditional distribution is almost surely independent of the spreading sequences. We develop a proof of the theorem starting from showing the existence of M_i.

Lemma 1: Under perfect power control,

$$M_i = \lim_{K\to\infty} \mathsf{E}\left\{\left[\mathbf{R}^i\right]_{11}\right\} \tag{27}$$

exists and is given by (21).

Proof: Under perfect power control, $\left[\mathbf{R}^i\right]_{kk}$'s are identically distributed for all k. Hence

$$\begin{aligned}\mathsf{E}\left\{\left[\mathbf{R}^i\right]_{11}\right\} &= \frac{1}{K}\sum_{k=1}^{K}\mathsf{E}\left\{\left[\mathbf{R}^i\right]_{kk}\right\} &(28)\\ &= \frac{1}{K}\mathsf{E}\left\{\mathrm{tr}\left\{\mathbf{R}^i\right\}\right\} &(29)\\ &= \frac{1}{K}\mathsf{E}\left\{\sum_{k=1}^{K}\lambda_k^i\right\} &(30)\\ &= \mathsf{E}\left\{\lambda^i\right\} &(31)\end{aligned}$$

where λ is a randomly picked eigenvalue of $\mathbf{R}^i$. Taking the large-system limit of both sides of (31), the left hand side is M_i and the right hand side is given by (21). ■

We next present a simple fact that underlies all major results in this paper.

Lemma 2: Let $\{\hat{s}_{nk}\,|\,n=1,2,\cdots,\ k=1,2,\cdots\}$ be an array of independent random variables each taking equally likely values of $+1$ and -1. Let $n_1,\cdots,n_i\in\{1,\cdots,N\}$ and $k_1,\cdots,k_i,k_{i+1}\in\{1,\cdots,K\}$. Then the product

$$\hat{s}_{n_1k_1}\hat{s}_{n_1k_2}\hat{s}_{n_2k_2}\hat{s}_{n_2k_3}\cdots\hat{s}_{n_{i-1}k_{i-1}}\hat{s}_{n_{i-1}k_i}\hat{s}_{n_ik_i}\hat{s}_{n_ik_{i+1}} \tag{32}$$

is a constant $+1$ if the indexes are such that all the indexed $\hat{s}$ variables appear in pairs; otherwise the product is a random variable taking the values of $+1$ and -1 equally likely.

This Lemma holds trivially. For example, $\hat{s}_{21}\hat{s}_{24}\hat{s}_{14}\hat{s}_{14}\hat{s}_{24}\hat{s}_{21}$ consists of 3 pairs of identical binary variables and is equal to $\hat{s}_{21}^2\hat{s}_{24}^2\hat{s}_{14}^2=+1$. On the contrary, $\hat{s}_{12}\hat{s}_{11}\hat{s}_{31}\hat{s}_{32}$ is a fair coin toss of $+1$ and -1, and has a zero mean.

Every entry in the matrix $\mathbf{R}^i$ is a highly structured weighted sum of products of the form of (32),

$$\begin{aligned}&\left[\mathbf{R}^i\right]_{1k}\\ &= \sum_{k_2,\dots,k_i=1}^{K}\mathbf{R}_{1k_2}\mathbf{R}_{k_2k_3}\cdots\mathbf{R}_{k_ik} &(33)\\ &= \frac{1}{N^i}\sum_{k_2,\dots,k_i=1}^{K}\ \sum_{n_1,\dots,n_i=1}^{N}\\ &\qquad s_{n_11}s_{n_1k_2}s_{n_2k_2}s_{n_2k_3}\cdots s_{n_{i-1}k_{i-1}}s_{n_{i-1}k_i}s_{n_ik_i}s_{n_ik} &(34)\end{aligned}$$

Lemma 2, reinforced with combinatorial arguments, powerfully reveals the probabilistic behavior of individual elements in $\mathbf{R}^i$. We have the following proposition, the proof of which is quite lengthy and relegated to Appendix I and II.

Proposition 1: For every positive integer p, and every user index k, the p^{th}-order central moment of $\sqrt{K}\left[\mathbf{R}^i\right]_{1k}$ converges to a deterministic constant as $K\to\infty$.

As a consequence, the asymptotic distribution of

$$\sqrt{K}\left(\left[\mathbf{R}^i\right]_{1k}-\mathsf{E}\left\{\left[\mathbf{R}^i\right]_{1k}\right\}\right) \tag{35}$$

has finite moments of any order. This allows us to have a clear picture of the probabilistic property of the elements in $\mathbf{R}^i$. By first subtracting its mean value (the off-diagonal elements have zero-mean) each entry diminishes as $K\to\infty$. This vanishing rate is quite fast. In fact, if we amplify it by $\sqrt{K}$, each entry converges to some random variable, the moments of which are the finite limits of the central moments of (35).

A weaker boundedness result is sufficient for our study, namely, the p^{th} order central moment of $\sqrt{K}\left[\mathbf{R}^i\right]_{1k}$ is bounded by some number for all K, due to its convergence. The following is immediate.

Corollary 1: For every positive integer p, the p^{th} order central moment of $\left[\mathbf{R}^i\right]_{1k}$ is upper bounded by $\gamma K^{-\frac{p}{2}}$ for all p and K where γ is a positive number independent of K.

Corollary 1 leads to the following almost sure convergence.

Lemma 3: $\left[\mathbf{R}^i\right]_{11}$ converges to M_i with probability 1 as $K\to\infty$.

Proof: Define

$$v_K = \left[\mathbf{R}^i\right]_{11}-\mathsf{E}\left\{\left[\mathbf{R}^i\right]_{11}\right\}. \tag{36}$$

The system dimension (K) is explicit in v_K. By Corollary 1, there exists $\gamma>0$,

$$\mathsf{E}\left\{v_K^4\right\} < \frac{\gamma}{K^2},\quad \forall K. \tag{37}$$

By the Markov Inequality [24], for every $\epsilon>0$,

$$\begin{aligned}\mathsf{P}\left(|v_K|>\epsilon\right) &\le \frac{\mathsf{E}\left\{v_K^4\right\}}{\epsilon^4} &(38)\\ &< \frac{\gamma}{\epsilon^4K^2}. &(39)\end{aligned}$$

Clearly,

$$\sum_{K=1}^{\infty}\mathsf{P}\left(|v_K|>\epsilon\right)<\infty. \tag{40}$$

By the Borell-Cantelli Lemma [25], v_K converges to 0 with probability 1. Therefore,

$$\left[\mathbf{R}^i\right]_{11} = \mathsf{E}\left\{\left[\mathbf{R}^i\right]_{11}\right\}+v_K \tag{41}$$

converges with probability 1 to M_i by Lemma 1. ■

The following is immediate from Lemma 3.

Corollary 2: The coefficient $H_{11}=\sum_{i=1}^{m}x_i\left[\mathbf{R}^i\right]_{11}$ converges to $\sum_{i=1}^{m}x_iM_i$ with probability 1.

Also, we have the following lemma about the MAI term in (10).

Lemma 4: The distribution of $\sum_{k=2}^{K}H_{1k}d_k$, conditioned on the spreading sequences, converges to a Gaussian law with probability 1.

To prove Lemma 4, we use the Lindeberg-Feller central limit theorem [26, page 448], which is restated here in a form convenient for this paper.

Theorem 2 (Lindeberg-Feller) For each K, let $\{X_{K,1},\dots,X_{K,K}\}$ be independent zero-mean random variables with finite variance. Suppose that

$$\sum_{k=1}^{K} \mathsf{E}\left\{X_{K,k}^2\right\} \to 1, \tag{42}$$

and that the Lindeberg condition is satisfied, i.e., for all $\epsilon > 0$,

$$\lim_{K\to\infty} \sum_{k=1}^{K} \mathsf{E}\left\{X_{K,k}^2 \cdot 1_{\{|X_{K,k}|>\epsilon\}}\right\} = 0 \tag{43}$$

where $1_{\{\cdot\}}$ is the indicator function which takes the value of 1 if the condition in the braces is satisfied and 0 otherwise. Then $\sum_{k=1}^{K} X_{K,k}$ converges to a standard Gaussian random variable in distribution.

Lemma 4 can be proved as follows.

Proof: **[Lemma 4]** Assume that the spreading sequences are given so that $\mathbf{H}$ is determined. We study the set of random variables $\{H_{12}d_2,\dots,H_{1K}d_K\}$ for each K and show that for almost all possible spreading sequence assignments, the conditions required by Theorem 2 are satisfied, namely,

(L4.1) $\{H_{1k}d_k \mid k=2,\cdots,K\}$ is a set of independent zero-mean random variables;

(L4.2) $\sum_{k=2}^{K} \mathsf{E}\left\{ (H_{1k}d_k)^2 \middle| \mathbf{H}\right\}$ converges as $K\to\infty$; and

(L4.3) The Lindeberg condition

$$\lim_{K\to\infty} \sum_{k=2}^{K} \mathsf{E}\left\{ (H_{1k}d_k)^2 \cdot 1_{\{|H_{1k}d_k|>\epsilon\}} \middle| \mathbf{H}\right\} = 0, \quad \forall \epsilon > 0. \tag{44}$$

Condition (L4.1) holds for all $\mathbf{H}$ by independence of d_k's. The sum in condition (L4.2) can be obtained as

$$\sum_{k=2}^{K} H_{1k}^2 = \sum_{k=1}^{K} H_{1k}H_{k1} - H_{11}^2 \tag{45}$$

$$= [\mathbf{H}^2]_{11} - H_{11}^2. \tag{46}$$

Notice that

$$\mathbf{H}^2 = \sum_{i=1}^{m}\sum_{j=1}^{m} x_i x_j \mathbf{R}^{i+j}. \tag{47}$$

The right hand side of (46) converges to

$$\sum_{i=1}^{m}\sum_{j=1}^{m} x_i x_j M_{i+j} - \left(\sum_{i=1}^{m} x_i M_i\right)^2 \tag{48}$$

with probability 1 by Corollary 2.

To examine condition (L4.3) we define

$$W_K(\mathbf{H},\epsilon) = \sum_{k=2}^{K} \mathsf{E}\left\{ (H_{1k}d_k)^2 \cdot 1_{\{|H_{1k}d_k|>\epsilon\}} \middle| \mathbf{H}\right\}. \tag{49}$$

Since $d_k = \pm 1$, and noting that

$$1_{\{x>\epsilon\}} \le \left(\frac{x}{\epsilon}\right)^n, \quad \forall x,\epsilon,n > 0, \tag{50}$$

we have

$$W_K(\mathbf{H},\epsilon) = \sum_{k=2}^{K} H_{1k}^2 \cdot 1_{\{|H_{1k}|>\epsilon\}} \tag{51}$$

$$\le \sum_{k=2}^{K} H_{1k}^2 \cdot \frac{H_{1k}^4}{\epsilon^4} \tag{52}$$

Hence for every $\delta > 0$,

$$\mathsf{P}\left(W_K(\mathbf{H},\epsilon) > \delta\right) \le \frac{1}{\delta}\cdot \mathsf{E}\left\{W_K(\mathbf{H},\epsilon)\right\} \tag{53}$$

$$\le \frac{1}{\delta\epsilon^4}\cdot \mathsf{E}\left\{\sum_{k=2}^{K} H_{1k}^6\right\} \tag{54}$$

$$= \frac{K-1}{\delta\epsilon^4}\cdot \mathsf{E}\left\{H_{12}^6\right\} \tag{55}$$

$$\le \frac{1}{\delta\epsilon^4 K^2}\cdot \mathsf{E}\left\{(\sqrt{K}H_{12})^6\right\} \tag{56}$$

where (53) is by the Markov Inequality and (55) holds because the H_{1k}'s, $k=2,\cdots,K$, are identically distributed. Since every moment of $\sqrt{K}H_{1k} = \sum_{i=1}^{m} x_i\sqrt{K}\left[\mathbf{R}^i\right]_{1k}$ is bounded due to Corollary 1, the probability $\mathsf{P}\left(W_K(\mathbf{H},\epsilon) > \delta\right)$ is bounded by some γK^{-2} which is summable over K. Analogously to Lemma 3, the Borell-Cantelli Lemma leads to

$$\lim_{K\to\infty} W_K(\mathbf{H},\epsilon) = 0 \text{ with probability 1,} \tag{57}$$

i.e., the Lindeberg condition (L4.3) is satisfied with probability 1.

In all, conditions (L4.1-3) are satisfied with probability 1. Invoking Theorem 2, we obtain that conditioned on almost all spreading sequences, the MAI is asymptotically Gaussian. ■

With the distribution of all three components of the decision statistic (10) known (the noise is trivially Gaussian), we can now prove its asymptotic normality.

Proof: **[Theorem 1]** The first term on the right hand side of (10) converges with probability 1 to a deterministic value by Corollary 2. The MAI term is asymptotically Gaussian with probability 1 by Lemma 4. The noise is zero-mean Gaussian with variance

$$\mathsf{E}\left\{z_1^2\right\} = [\mathbf{GRG}]_{11}\,\sigma^2, \tag{58}$$

which can be easily shown to converge to

$$\nu_\infty^2 = \sigma^2 \sum_{i=1}^{m}\sum_{j=1}^{m} x_i x_j M_{i+j-1} \tag{59}$$

with probability 1. Conditioned on the spreading sequences, the MAI and the noise are independent. The distribution of the decision statistic is therefore asymptotically Gaussian with mean value and variance given by (24) and (25) respectively. ■

B. Unconditional Distribution

It is also desirable to know the distribution of the decision statistic when the spreading sequences are allowed to vary symbol-by-symbol as in a long-code system. In every symbol interval, the MAI is a large sum of contributions from individual interfering users, whose spreading sequences are chosen independently from that of the desired user. It is expected that the MAI approaches a Gaussian probability law in distribution as the number of users increases. The resulting distribution is trivially the same as in the conditional case since the unconditional distribution is a mixture of the conditional ones, the limit of which are all the same except for a negligible set in probabilistic sense by virtue of Theorem 1. Precisely, we have the following theorem.

Theorem 3: The unconditional distribution of the decision statistic given by (10), where the polynomial receiver is given as (22), converges to a Gaussian law with mean value and variance given by (24) and (25) respectively.

Proof: We present a direct proof similar to that of the conditional case rather than citing Theorem 1. The proof reveals some dependence subtleties not present in the conditional case.

First, $H_{11} \to \sum_{i=1}^{m} x_i M_i$ with probability 1 by Corollary 2.

Second, we need to show that the distribution of $\sum_{k=2}^{K} H_{1k} d_k$ converges to a Gaussian law. This is not as simple as in the conditional case, since the H_{1k}'s are now dependent random variables. We resort to a more general central limit theorem in [27]. We show that the following 3 conditions are satisfied.

(T3.1) $\sum_{k=2}^{K} H_{1k} d_k$ is a martingale for every K;

(T3.2) $\mathsf{E}\left\{\left(\sum_{k=2}^{K} H_{1k} d_k\right)^2\right\}$ converges as $K \to \infty$; and

(T3.3) The Lindeberg condition is satisfied, i.e.,

$$\sum_{k=2}^{K} \mathsf{E}\left\{(H_{1k} d_k)^2 \cdot 1_{\{|H_{1k} d_k| > \epsilon\}}\right\} \to 0, \quad \forall \epsilon > 0. \tag{60}$$

For every K, and an arbitrary user index $k > 1$, the conditional expectation

$$\mathsf{E}\left\{H_{1k} d_k | H_{12} d_2, \cdots, H_{1\,k-1} d_{k-1}\right\} = 0 \tag{61}$$

by independence of the data symbols. By definition, H_{1k} is an absolutely fair sequence, and therefore $\sum_{k=2}^{K} H_{1k} d_k$ is a martingale [25, p. 209]. Hence (T3.1) is true.

Also by the independence of the antipodal symbols,

$$\begin{aligned} \mathsf{E}\left\{\left(\sum_{k=2}^{K} H_{1k} d_k\right)^2\right\} &= \sum_{k=2}^{K} \mathsf{E}\left\{H_{1k}^2\right\} & (62) \\ &= \mathsf{E}\left\{[\mathbf{H}^2]_{11}\right\} - \mathsf{E}\left\{H_{11}^2\right\}. & (63) \end{aligned}$$

The right hand side of (63) converges by Corollary 2. Thus we have (T3.2).

To verify (T3.3), we find, for every $\epsilon > 0$

$$\begin{aligned} & \sum_{k=2}^{K} \mathsf{E}\left\{(H_{1k} d_k)^2 \cdot 1_{\{|H_{1k} d_k| > \epsilon\}}\right\} \\ = \; & (K-1) \cdot \mathsf{E}\left\{H_{12}^2 \cdot 1_{\{|H_{12}| > \epsilon\}}\right\} & (64) \\ \le \; & (K-1) \cdot \frac{1}{\epsilon^2} \cdot \mathsf{E}\left\{H_{12}^4\right\} & (65) \\ \le \; & \frac{1}{K\epsilon^2} \cdot \mathsf{E}\left\{(\sqrt{K} H_{12})^4\right\} & (66) \\ \to \; & 0 & (67) \end{aligned}$$

as $K \to \infty$ by Corollary 1.

With conditions (T3.1)–(T3.3) verified, the sum $\sum_{k=2}^{K} H_{1k} d_k$ converges to a Gaussian law, with zero mean and a variance given by (48), following a dependent central limit theorem for martingales in [27].

Moreover, the noise converges trivially to a Gaussian random variable in mean square sense. Note that both the MAI and the noise are dependent on the spreading sequences. Given the output noise variance, however, the noise and the MAI are mutually independent. Consider a slight modification of the decision statistic, where we introduce a scalar multiplier to the noise term to remove dependence,

$$y_1' = H_{11} + \sum_{k=2}^{K} H_{1k} d_k + \frac{\nu_\infty}{\sqrt{\mathsf{E}\left\{z_1^2\right\}}} \cdot z_1 \tag{68}$$

where ν_∞ is defined in (59). The standard Gaussian random variable $z_1/\sqrt{\mathsf{E}\left\{z_1^2\right\}}$ can be easily shown to be independent of the spreading sequences and hence of the MAI. Since the multiplier converges to 1, $(y_1' - y_1)$ converges to 0 in mean square sense, and hence y_1 and y_1' share the same asymptotic distribution.

In all, the distribution of y_1 is asymptotically Gaussian with a mean value as that of the limit of the first term, and a variance as the sum of the limits of those of the MAI and the noise. They are given in (24) and (25), respectively. ■

IV. Asymptotic Normality

We have shown above that, under perfect power control, every polynomial receiver yields asymptotically Gaussian outputs. By the Weierstrass Theorem, the set of polynomials is dense in the space of continuous functions defined on a finite interval [28]. For this reason, we can show that every receiver of the form $\mathbf{G} = g(\mathbf{R})$ can be arbitrarily well approximated by a sequence of polynomial receivers. As a consequence its output is also asymptotically Gaussian in distribution. Formally, we have the following theorem.

Theorem 4: Assume perfect power control. For every function g continuous on $\left(\max^2(0, 1-\sqrt{\beta}), (1+\sqrt{\beta})^2\right)$, the output decision statistic of any linear receiver $\mathbf{G} = g(\mathbf{R})$ defined in (12) is asymptotically Gaussian in distribution conditioned on almost all spreading sequences. The mean value corresponding to the limiting distribution is

$$\mu_g = \int g(\lambda) \cdot \lambda \, dF_\Lambda(\lambda) \tag{69}$$

and the variance is

$$\sigma_g^2 = \int g^2(\lambda) \cdot \lambda \cdot (\lambda + \sigma^2) \, dF_\Lambda(\lambda) - \mu_g^2. \tag{70}$$

The following lemma is useful for proving Theorem 4.

Lemma 5: Let $\mathbf{G} = g(\mathbf{R})$. For all $\epsilon > 0$, there exists a polynomial receiver $\tilde{\mathbf{G}} = \tilde{g}(\mathbf{R})$ such that the mean square difference of every user's output decision statistic of $\mathbf{G}$ and $\tilde{\mathbf{G}}$ is less than ϵ for sufficiently large K for almost all spreading sequences.

Lemma 5, proved in Appendix III, establishes that a linear receiver can be arbitrarily well approximated by polynomial receivers in the large-system limit. The fact that the output of every polynomial receiver converges to a Gaussian random variable leads to the asymptotic normality of linear receiver outputs. The proof of Theorem 4 also requires the following lemma.

Lemma 6: Let $\left\{ Y_m^{(K)} \middle| K = 1, 2, \cdots, m = 1, 2, \cdots \right\}$ be an array of continuous random variables. Let $\left\{ Y^{(K)} \right\}_{K=1}^{\infty}$ and $\{Y_m\}_{m=1}^{\infty}$ be two sequences of random variables. Denote the cumulative distribution functions of $Y_m^{(K)}$, $Y^{(K)}$ and Y_m as $F_m^{(K)}$, $F^{(K)}$ and F_m respectively. Suppose that

$$\text{(L6.1)} \quad \lim_{m\to\infty} \lim_{K\to\infty} \mathsf{E} \left\{ \left| Y^{(K)} - Y_m^{(K)} \right|^2 \right\} = 0; \text{ and} \tag{71}$$

$$\text{(L6.2)} \quad \lim_{K\to\infty} \left| F_m^{(K)}(a) - F_m(a) \right| = 0, \ \forall a, m, \tag{72}$$

then both $F^{(K)}$ and F_m converge pointwise to the same distribution.

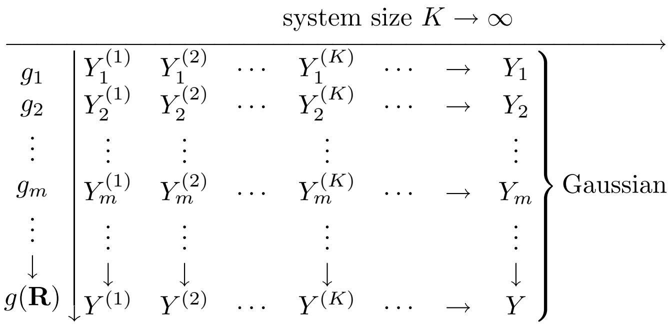

Fig. 2. Convergence of output distributions.

The proof of Lemma 6 is in Appendix IV. The idea is illustrated in Fig. 2. Each row corresponds to a particular polynomial receiver, whose output converges to a Gaussian random variable in distribution. Each column corresponds to a sequence of polynomial receivers for a particular system size. The sequences of $\{Y^{(K)}\}$ and $\{Y_m\}$ converge in distribution to the same Gaussian law.

We are now equipped to prove the asymptotic normality for linear detectors of the form $\mathbf{G} = g(\mathbf{R})$.

Proof: **[Theorem 4]** For clarity, we explicitly label relevant variables with their corresponding system dimension. Take a sequence $\epsilon_m \to 0$. For each m, by Lemma 5, there exists a polynomial receiver $\left\{ \mathbf{G}_m^{(K)} \right\}$ determined by a polynomial g_m such that for sufficiently large K and almost all spreading sequences,

$$\mathsf{E} \left\{ \left| y_1^{(K)} - y_{m1}^{(K)} \right|^2 \right\} < \epsilon_m \tag{73}$$

where $y_1^{(K)}$ and $y_{m1}^{(K)}$ are the output of $\mathbf{G}$ and $\mathbf{G}_m^{(K)}$ for user 1 respectively.

Let $Y^{(K)} = y_1^{(K)}$ and $Y_m^{(K)} = y_{m1}^{(K)}$. Due to the presence of noise, $Y_m^{(K)}$ are continuous random variables. Equation (73) implies (71). Also, by Theorem 1, for every m, $Y_m^{(K)}$ converge as $K \to \infty$ to some Gaussian random variable, defined as Y_m. Hence (72) is satisfied. Note that a sequence of Gaussian distribution functions converge also to a Gaussian law. Invoking Lemma 6, we have that $Y^{(K)}$ converge in distribution to a Gaussian law. Equation (69) is straightforward by noting that for $g_m(\lambda) = \sum_{i=1}^m x_i \lambda^{i-1}$,

$$\begin{aligned} \sum_{i=1}^m x_i M_i &= \mathsf{E} \left\{ \sum_{i=1}^m x_i \lambda^i \right\} &(74)\\ &= \mathsf{E} \{ g_m(\lambda) \cdot \lambda \} &(75)\\ &= \int g_m(\lambda) \cdot \lambda \mathrm{d} F_\Lambda(\lambda). &(76) \end{aligned}$$

Equation (70) can be obtained analogously. ■

V. Asymptotic Normality for Unequal Powers

We have established asymptotic normality of linear receiver outputs assuming perfect power control. In fact, the normality principle holds for very general scenarios. In this section we generalize the normality results to the case where the received energies from all users are not equal. We assume that the energies are independent of the spreading sequences and that in the large-system limit, the empirical distributions of the energies converge to the energy distribution F_P.

It is important to note that Proposition 1 still holds in this case, as is proved in Appendix II. Therefore, Theorem 1 and 3 also hold, since the proof still applies in principle. The complication here is that (28) is no longer true and M_i does not have a simple expression. We can follow the approach in [29] to obtain each moment by exploiting the structure of $\left[\mathbf{R}^i\right]_{11}$ as a sum of products of random chips as in (34). For instance, assuming binary spreading, the first 4 M_i's are

$$\begin{aligned} M_1 &= P_1, &(77)\\ M_2 &= P_1 \left[P_1 + \beta \right], &(78)\\ M_3 &= P_1 \left[P_1^2 + 2\beta P_1 + \beta \mathsf{E}\left\{P^2\right\} + \beta^2 \right], &(79)\\ M_4 &= P_1 \left[P_1^3 + 3\beta P_1^2 + (2\beta \mathsf{E}\left\{P^2\right\} + 3\beta^2) P_1 \right.\\ &\quad \left. + (\beta \mathsf{E}\left\{P^3\right\} + 3\beta^2 \mathsf{E}\left\{P^2\right\} + \beta^3) \right], &(80) \end{aligned}$$

where the expectations are taken over the energy distribution. In this way, the mean and variance of the limiting output distribution of a polynomial receiver can be determined.

Furthermore, we can still approximate a linear receiver determined by a continuous function g by a series of polynomial receivers. Indeed, the asymptotic normality is true for an arbitrary receiver $\mathbf{G} = g(\mathbf{R})$ without the assumption that all users are received at the same energy. Unfortunately, the mean and the variance of the limiting distribution do not allow simple expression as in Theorem 4. In principle, the mean and variance can be well-approximated by that of a polynomial receiver output, which can be obtained by (24)–(25). It is often easier, however, to find the mean and the variance using properties of the particular receiver of interest. Some useful results on popular linear receivers are listed in Section VI.

In summary, we have the following theorem.

Theorem 5: For every continuous function g, the output decision statistic of linear receiver $\mathbf{G} = g(\mathbf{R})$ has asymptotically the same Gaussian distribution conditioned on almost all spreading sequences.

This theorem states that the output of a large family of linear receivers has the same asymptotic Gaussian distribution conditioned on almost every spreading sequence assignment, which is nothing but the asymptotic unconditional distribution. This somewhat surprising result may be understood as follows. First, under mild conditions, all interfering users have "comparable" and uniformly small contributions in interference to the desired user as the system size gets larger and larger. So the total contribution turns out to be Gaussian in the limit. Second, a large system is self-averaging, i.e., a particular realization of the covariance matrix is almost surely "sufficiently representative" of the whole ensemble. In other words, empirical averaging is the same as ensemble averaging in the large-system limit. As the system size increases, a short-code system where the spreading sequences are randomly chosen behaves more and more like a long-code system. Interestingly, the fundamental law underlying this principle is statistical physics. A multiuser system is equivalent to a thermodynamic system, whose fluctuation vanishes and the emerging stable macroscopic properties dominate in the large-system limit [19], [30], [31].

VI. Multiuser Efficiency

The asymptotic normality of linear receiver outputs allows the multiuser efficiency [1, page 121], which uniquely characterizes the uncoded bit-error-rate for an arbitrary noise level, to be completely determined by the signal-to-interference ratio in the large-system limit.

For a receiver of the form in (12) determined by g, we have its large-system limit of the SIR expressed as

$$\gamma = \frac{\mu_g^2}{\sigma_g^2} \tag{81}$$

where μ_g and σ_g^2 are the mean value and the variance of the limiting distribution respectively. Assuming threshold detection, we have the uncoded probability of error expressed as a single Q-function of the square root of the SIR, i.e.,

$$\mathsf{P} = Q\left(\sqrt{\gamma}\right). \tag{82}$$

The multiuser efficiency, defined as the ratio between the energy that a user would require to achieve the same BER in absence of interfering users and the actual energy, is then

$$\eta = \frac{\mu_g^2 \sigma^2}{P_1 \sigma_g^2}. \tag{83}$$

For a polynomial receiver given as (22) we have

$$\eta^{(\mathrm{p})} = \frac{\sigma^2}{P_1} \cdot \frac{\left(\sum_{i=1}^m x_i M_i\right)^2}{\sum_{i=1}^m \sum_{j=1}^m x_i x_j [M_{i+j} - M_i M_j + \sigma^2 M_{i+j-1}]} \tag{84}$$

where M_i, defined in (26), can be obtained by (21) for the equal-energy case, or as (77)–(80) by following the approach in [29] otherwise.[5] A trivial example of a polynomial receiver is the single-user matched filter, whose large-system multiuser efficiency is obtained as

$$\eta^{(\mathrm{mf})} = \frac{\sigma^2}{P_1} \frac{M_1^2}{M_2 - M_1^2 + \sigma^2 M_1} \tag{85}$$

$$= \frac{1}{1 + \frac{\beta}{\sigma^2}}. \tag{86}$$

Another popular receiver, the decorrelator, is determined by (13). In case of $\beta < 1$, the MAI term is 0 with probability 1 and hence

$$\eta^{(\mathrm{dec})} = 1 - \beta, \quad \beta < 1. \tag{87}$$

In case of $\beta > 1$, the MAI is nontrivial but the asymptotic normality is still true. The multiuser efficiency is obtained in [19], which, in the equal-energy case can be simplified to [32]

$$\eta^{(\mathrm{dec})} = \frac{(\beta - 1)\sigma^2}{(\beta - 1)^2 + \sigma^2 \beta}, \quad \beta > 1. \tag{88}$$

The MMSE receiver is determined by $g(\lambda) = (\lambda + \sigma^2)^{-1}$. It is hard to obtain μ_g and σ_g^2 directly. Applying the known Stieltjes transform of the eigenvalue distribution function, the multiuser efficiency can be obtained as the positive solution to the Tse-Hanly fixed-point equation [18], [13], [33]

$$\eta + \beta\, \mathsf{E}\left\{\frac{P\,\eta}{P\,\eta + \sigma_0^2}\right\} = 1, \tag{89}$$

where the expectation is taken over the random variable P drawn according to the energy distribution.

VII. Numerical Results

Fig. 3 shows convergence to a Gaussian distribution of the output decision statistics. Chips and symbols are BPSK modulated. Perfect power control and an SNR of 5 dB are

[5] M_i can also be obtained using other dedicated numerical approaches.

assumed for all users. We plot the histograms of the output statistics of a linear polynomial receiver

$$\mathbf{G}(\mathbf{R}) = 2.24\mathbf{I} - 1.61\mathbf{R} + 0.345\mathbf{R}^2, \tag{90}$$

which is the 3-stage parallel interference canceler that gives asymptotically the least achievable output mean square error [16]. The number of users considered are $K = 1$, 2, 16 and 256 respectively, and $K/N = 1/2$ is assumed in all cases. The predicted asymptotic Gaussian distribution for an infinite number of users is also plotted for reference. It is evident that the distribution of the output decision statistics converges to the Gaussian distribution as the system size increases. For sixteen users or more, the approximation is excellent. Note that the area under each curve on the left half plane (< 0) corresponds to the uncoded bit-error-rate of the receiver.

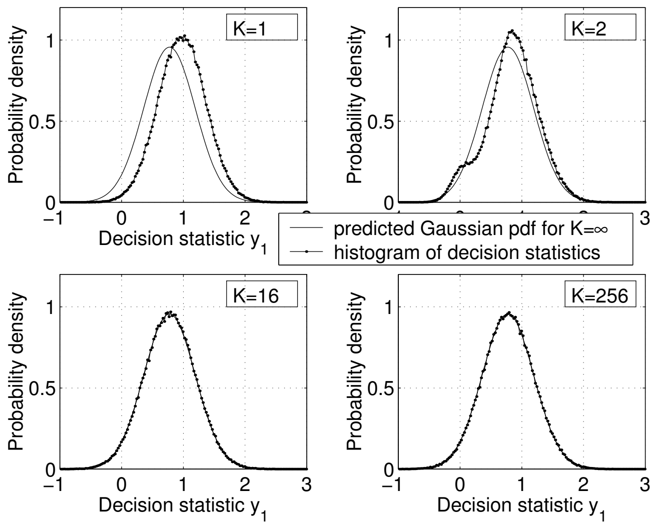

Fig. 3. Distribution of output decision statistics. The 2-user, 16-user and 256-user cases are shown as well as the single-user case ($K = 1$). They are compared with the asymptotic Gaussian distribution.

In Fig. 4 we plot the BER of various linear receivers averaged over spreading sequences and observe the trend as the system size increases. This corresponds to long-code system performance. BERs of the single-user matched filter, the decorrelator, the MMSE receiver, and two polynomial receivers of order 3 and 6 respectively are obtained through Monte Carlo simulation. As in the previous figure, the polynomial receivers are chosen as the ones that give asymptotically the least achievable mean square error as suggested in [16]. The ratio K/N is always 1/2 but the SNR is assumed to be 10 dB for all users. Asymptotic estimates using the results in Section VI are marked by solid triangles on the right border of the figure for reference. It is clear that for all receivers the BERs converge to the asymptotic estimates.

For Fig. 5 we do the simulation in the same setting as in Fig. 4 except for that a particular choice of spreading sequences is used to simulate a short-code system. The BER varies much more than in Fig. 4, depending on whether the particular set of spreading sequences is favorable or unfavorable for the system size. Nevertheless, as the system

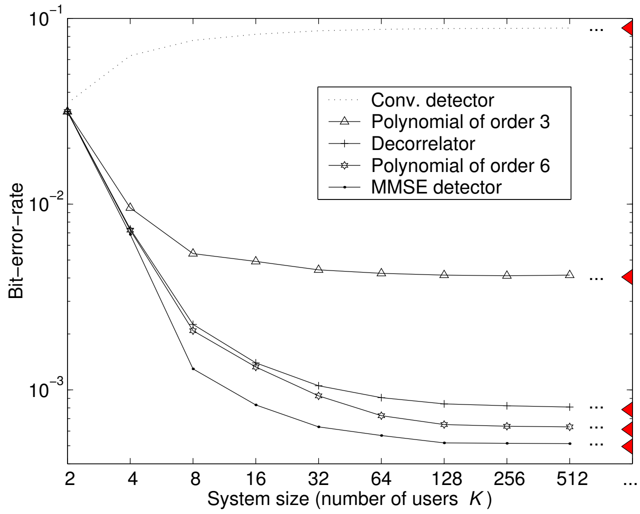

Fig. 4. BER vs. the number of users (long sequences). The asymptotic estimates are marked by solid triangles on the right border of the plot.

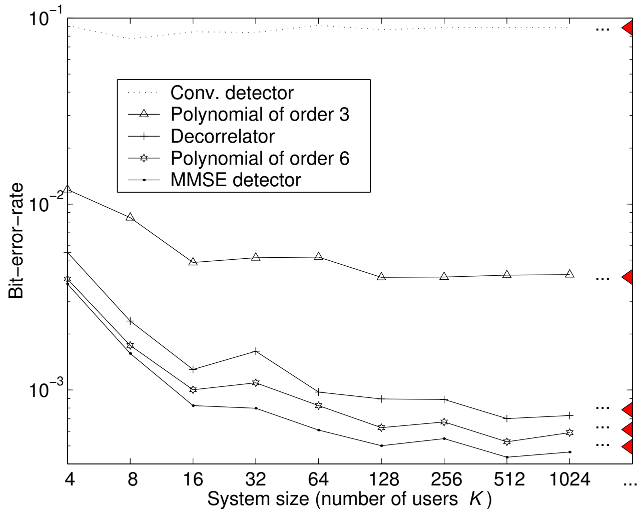

Fig. 5. BER vs. the number of users (short sequences). The asymptotic estimates are marked by solid triangles on the right border of the plot.

size increases, the BERs converge to asymptotic predictions, which are marked by the solid triangles. The convergence speed to a Gaussian law is much slower than for long sequences under the same system setting.

VIII. Conclusion

In this paper, we have proved asymptotic normality of the output decision statistics of a large family of linear receivers, which can be arbitrarily well approximated by polynomial receivers. The limiting output distribution of the decision statistics conditioned on almost all choices of spreading sequences is asymptotically the same as the unconditional distribution. The normality principle shows that the signal-to-interference ratio is a decisive index of uncoded system performance for linear receivers. The large-system limits of the multiuser efficiency of the single-user matched filter, the decorrelator, the MMSE receiver as well as the polynomial receivers are determined by way of evaluating the large-system SIR. We can further conclude that, if single-user decoding is used, error-control codes that are optimal for Gaussian channels will also be asymptotically optimal for a multiuser channel.

Appendices

I. Proof of Proposition 1

We develop a combinatorial proof for Proposition 1 based on the simple fact of Lemma 2.

We first introduce some notation. Define a random variable

$$S(i) = \sum_{k_1=1}^{K} \cdots \sum_{k_i=1}^{K} \cdot \sum_{n_1=1}^{N} \cdots \sum_{n_i=1}^{N} s_{n_1k_1}s_{n_1k_2}s_{n_2k_2}s_{n_2k_3}\cdots s_{n_{i-1}k_{i-1}}s_{n_{i-1}k_i}s_{n_ik_i}s_{n_ik_1}. \tag{91}$$

Let $\mathcal{I}(i)$ denote a vector of indexes, $[k_1, \cdots, k_i, n_1, \cdots, n_i]$, and $A(i) = \{1, \cdots, K\}^i \times \{1, \cdots, N\}^i$. $S(i)$ can then be written as a single summation

$$S(i) = \sum_{\mathcal{I}(i)\in A(i)} s_{n_1k_1}s_{n_1k_2}s_{n_2k_2}s_{n_2k_3}\cdots s_{n_{i-1}k_{i-1}}s_{n_{i-1}k_i}s_{n_ik_i}s_{n_ik_1}. \tag{92}$$

We are also interested in cases when some of the indexes in $\mathcal{I}(i)$ are fixed. We define for $k = 1, 2$,

$$S_k(i) = \sum_{\mathcal{J}(i)\in B(i)} s_{n_1k}s_{n_1k_2}s_{n_2k_2}\cdots s_{n_{i-1}k_i}s_{n_ik_i}s_{n_ik} \tag{93}$$

where $\mathcal{J}(i) = [k_2, \cdots, k_i, n_1, \cdots, n_i]$ and $B(i) = \{1, \cdots, K\}^{i-1} \times \{1, \cdots, N\}^i$. Clearly, this is equivalent to evaluating the sum of $S(i)$ with k_1 in $\mathcal{I}(i)$ forced to take the value of $k = 1$ (or 2). Similarly, we define

$$T(i) = \sum_{\mathcal{J}(i)\in B(i)} s_{n_11}s_{n_1k_2}s_{n_2k_2}\cdots s_{n_{i-1}k_i}s_{n_ik_i}s_{n_i2}. \tag{94}$$

Interestingly, each of the above defined variables is a sum of products of the form in Lemma 2. For $S(i)$ and $S_k(i)$, the indexes in each product term always form a closed loop, whereas the indexes in $T(i)$ form an open loop.

To keep the presentation clear, we introduce an equivalent notation (see [34] for a similar technique). Instead of writing $S(i)$ as a sum of products as in (92), we record the topology of the indexes in a two-row array:

$$\left\langle \begin{array}{cccc} n_1 & n_2 & \cdots & n_i \\ k_1 & k_2 & \cdots & k_i \end{array} \right\rangle. \tag{95}$$

The way to understand it is to look upon it as a sum of a product of indexed s variables with their indexes zigzagging through the array, i.e., $s_{n_1k_1}$, $s_{n_1k_2}$, $s_{n_2k_2}$, $s_{n_2k_3}$, $\cdots$, $s_{n_ik_i}$ and finally $s_{n_ik_1}$ to close the loop. It would be helpful to take the indexes as vertices and the s_{nk} variables as edges of a graph. Also, $S_1(i)$ is denoted by

$$\left\langle \begin{array}{cccc} n_1 & n_2 & \cdots & n_i \\ 1 & k_2 & \cdots & k_i \end{array} \right\rangle, \tag{96}$$

$S_2(i)$ denoted similarly, and $T(i)$ denoted by

$$\left\langle \begin{array}{ccccc} n_1 & n_2 & \cdots & n_i & \\ 1 & k_2 & \cdots & k_i & 2 \end{array} \right\rangle. \tag{97}$$

$T(i)$ has one more element in the second row since it is an open loop. Its last variable in its corresponding product is s_{n_i2} instead of s_{n_i1}. This angle bracket notation is very illustrative and greatly simplifies our task of estimating the size of interesting variables.

The reader may have noticed that

$$\left[\mathbf{R}^i\right]_{11} = N^{-i}S_1(i) \tag{98}$$

and

$$\left[\mathbf{R}^i\right]_{12} = N^{-i}T(i). \tag{99}$$

Hence studying the statistical properties of $S_1(i)$ and $T(i)$ is sufficient for Proposition 1.

In the following we prove Proposition 1 assuming perfect power control and antipodal spreading, i.e., all s_{nk}'s are independent chips that take on values ± 1 only. The proof is generalized to non-binary spreading with no power control in Appendix II.

The following intermediate result is useful.

Lemma 7: For $p \geq 1$ and integers $i_1, \cdots, i_p \geq 1$,

$$\mathsf{E}\left\{\prod_{w=1}^{p} S(i_w)\right\} \tag{100}$$

is a polynomial in K of degree $D = \sum_{w=1}^{p}(i_w + 1)$.

Proof: Let Θ denote the product of the $S(i_w)$'s inside the expectation brackets in (100). Let $\sum_{w=1}^{p}(i_w + 1)$ be the size of Θ. This lemma is equivalent to saying that the degree of $\mathsf{E}\{\Theta\}$ is equal to the size of Θ, or, in other words, each factor $S(i_w)$ in Θ contributes to the overall degree of $\mathsf{E}\{\Theta\}$ by $(i_w + 1)$. The proof is by induction on the size of Θ.

The smallest size Θ may take is 2, where $p = 1$ and $i_1 = 1$. Thus

$$\begin{aligned} \mathsf{E}\{\Theta\} &= \mathsf{E}\{S(1)\} && (101)\\ &= \mathsf{E}\left\{\sum_{n=1}^{N}\sum_{k=1}^{K} s_{nk}s_{nk}\right\} && (102)\\ &= \beta^{-1}K^2. && (103) \end{aligned}$$

It is a polynomial of degree $D = 2$ as predicted.

Suppose now the lemma is true for the sizes of $2, \cdots, D-1$. We show that it must also be true for a size of D.

Let $\mathcal{I}^w = (n_1^w, \cdots, n_{i_w}^w, k_1^w, \cdots, k_{i_w}^w)$. $\mathsf{E}\{\Theta\}$ can be written as

$$\mathsf{E}\left\{\prod_{w=1}^{p}\sum_{\mathcal{I}^w\in A(i_w)} s_{n_1^wk_1^w}s_{n_1^wk_2^w}\cdots s_{n_{i_w}^wk_{i_w}^w}s_{n_{i_w}^wk_1^w}\right\}$$
$$= \sum_{\mathcal{I}^1\in A(i_1)}\cdots\sum_{\mathcal{I}^p\in A(i_p)} \mathsf{E}\left\{s_{n_1^1k_1^1}\cdots s_{n_{i_1}^1k_1^1}\cdots s_{n_1^pk_1^p}\cdots s_{n_{i_p}^pk_1^p}\right\}. \tag{104}$$

By the simple fact of Lemma 2, for each product to have nonzero expectation, the indexes must be such that the s variables form complete pairs, in which case the product term is 1. The value of $\mathsf{E}\{\Theta\}$ is therefore the number of occurrences of such cases.

Consider adding an extra constraint on the indexes: for each $w \in \{1, \cdots, p\}$, $n_1^w = n_2^w = \cdots = n_{i_w}^w$. The s variables then trivially form complete pairs. The number of such occurrences is $\beta^{-p}K^D$. Hence (104) is lower bounded by a polynomial of degree D. Surprisingly, even if all possible combinations of the indexes are counted, which appears to significantly increase the number of terms, this sum is still a polynomial of degree D.

To show this, we study the matching problem of a variable $s_{n_1k_1}$, the first term of some $S(i)$. It must be paired with some other s variable, which either comes from $S(i)$ itself, or from another $S(i')$. In either case, the two indexes, n_1 and k_1 are replaced by another two and can be dropped, but with the topology of the summation still in the form of Θ, i.e., a product of $S(i_w)$'s. However, the size of the problem can be reduced by one and therefore solved by the induction hypothesis. We develop this idea in full in the following.

Suppose this other s variable comes from $S(i)$ itself. Four possibilities arise:
(1) $n_1 = n_i$ so that $s_{n_1k_1} = s_{n_ik_1}$, $\forall k_1$;
(2) $k_1 = k_2$ so that $s_{n_1k_1} = s_{n_1k_2}$, $\forall n_1$;
(3) $n_1 = n_u$, $k_1 = k_u$ for some $u \neq 2, i$, so that $s_{n_1k_1} = s_{n_uk_u}$;
(4) $n_1 = n_u$, $k_1 = k_{u+1}$ for some $u \neq 1, i$, so that $s_{n_1k_1} = s_{n_uk_{u+1}}$.
We show that with each of the above four constraints Θ can be reduced to a variable of the same topology as the unconstrained Θ but of a smaller size.

In case of (1), $s_{n_1k_1}$ is paired with $s_{n_ik_1}$ for every choice of k_1. The product of the two is 1 and can be dropped. k_1 becomes an index of complete freedom, by summing over which a multiplicative factor K is contributed to the remaining sum. With this constraint, $S(i)$ becomes

$$K \cdot \left\langle \begin{array}{ccc} n_2 & \cdots & n_i \\ k_2 & \cdots & k_i \end{array} \right\rangle. \tag{105}$$

It is easily identified as $K \cdot S(i-1)$. Consequently, Θ is reduced to $K \cdot \Theta'$ where Θ' has the same topology as Θ but is one less in size than Θ. By the induction hypothesis, $\mathsf{E}\{\Theta'\}$ has a degree of $(D-1)$, so $\mathsf{E}\{\Theta\}$ with this constraint is a polynomial in degree D.

Case (2) is similar to case (1) and results in $N \cdot S(i-1)$, which is also a polynomial of degree D.

In case of (3), $s_{n_1k_1}$ couples with $s_{n_uk_u}$ and both terms are dropped. Indexes n_1 and k_1 are replaced by n_u and k_u respectively. The $S(i)$ becomes

$$\left\langle \begin{array}{ccccccc} n_2 & \cdots & n_{u-1} & n_i & n_{i-1} & \cdots & n_u \\ k_2 & \cdots & k_{u-1} & k_u & k_i & \cdots & k_{u+1} \end{array} \right\rangle = S(i-1). \tag{106}$$

Clearly, with this constraint, Θ is reduced to the same topology but with a size of $(D-1)$. $\mathsf{E}\{\Theta\}$ is a polynomial of degree $(D-1)$ in this case.

In case of (4), $s_{n_1k_1}$ couples with $s_{n_uk_{u+1}}$ and both terms are dropped. Indexes n_1 and k_1 are replaced by n_u and k_{u+1} respectively. $S(i)$ becomes

$$\begin{aligned} &\left\langle \begin{array}{ccc} n_2 & \cdots & n_u \\ k_2 & \cdots & k_u \end{array} \right\rangle \times \left\langle \begin{array}{ccc} n_{u+1} & \cdots & n_i \\ k_{u+1} & \cdots & k_i \end{array} \right\rangle \\ =&S(u-1) \cdot S(i-u). \end{aligned} \tag{107}$$

Thus under this constraint, $S(i)$ is split into two unconstrained S variables of the same form. Consequently Θ is reduced to some Θ' of the same topology. Although two indexes are dropped, the size of the resulting Θ' is the same as that of Θ, so the induction hypothesis does not directly apply in this case. However, we can take $S(u-1)$ in Θ' as $S(i)$ and go over the above reduction procedure recursively. One of the other cases must happen after some iterations since each splitting eliminates two of the less than $2p$ indexes. Hence, the induction applies indirectly in this case.

We now consider the situation that $s_{n_1k_1}$ is coupled to an s variable from $S(i')$. Two possibilities arise:
(5) $n_1 = n'_u$, $k_1 = k'_u$ for some u, so that $s_{n_1k_1} = s_{n'_uk'_u}$;
(6) $n_1 = n'_u$, $k_1 = k'_{u+1}$ for some u, so that $s_{n_1k_1} = s_{n'_uk'_{u+1}}$.

In case of (5), $s_{n_1k_1}$ and $s_{n'_uk'_u}$ are both dropped. Indexes n_1 and k_1 are replaced by n'_u and k'_u. Then,

$$S(i) \cdot S(i') = \left\langle \begin{array}{ccc} n_1 & \cdots & n_i \\ k_1 & \cdots & k_i \end{array} \right\rangle \times \left\langle \begin{array}{ccc} n'_1 & \cdots & n'_{i'} \\ k'_1 & \cdots & k'_{i'} \end{array} \right\rangle \tag{108}$$

becomes

$$\begin{aligned} &\left\langle \begin{array}{ccccccccc} n_2 & \cdots & n_i & n'_u & \cdots & n'_1 & n'_{i'} & \cdots & n'_{u+1} \\ k_2 & \cdots & k_i & k'_u & \cdots & k'_1 & k'_{i'} & \cdots & k'_{u+1} \end{array} \right\rangle \\ =&S(i+i'-1). \end{aligned} \tag{109}$$

With this constraint, Θ is reduced in size by two while maintaining the same topology. The resulting expectation is a polynomial of degree $(D-2)$ in degree, two less than it could if the two variables from different $S(\cdot)$'s were not forced to be paired.

Case (6) is similar to case (5) and the resulting contribution is also a polynomial of degree $(D-2)$.

The overall value of $\mathsf{E}\{\Theta\}$ is the sum of the above six cases. Some subtlety arises since the six cases overlap. With some patience the overlapping parts can be identified as polynomials of smaller degrees so there is essentially no over-counting by ignoring the overlap. The reader may have noticed that cases (3), (4), (5) and (6) may happen for multiple choices of u. However, u may take at most p different values in each case, which is fixed and not dependent on K. The sum over all possible choices of u is still a polynomial in K of the given degree. We can in fact neglect case (3), (5) and (6) when estimating $\mathsf{E}\{\Theta\}$ since they contribute a degree less than D. We can also conclude that the cases of matching variables from different $S(i)$'s can be neglected.

In all, the overall sum is a polynomial of degree D. The induction holds and the proof is complete. ■

Equipped with the techniques developed above, we solve a harder problem where some of the indexes in the sum are forced to take fixed values. We have the following result.

Lemma 8: Let $p, q, r, t \geq 0$ be integers. Let $i_1, \cdots, i_p$, $j_1, \cdots, j_q$, $l_1, \cdots, l_{2r}$ and $m_1, \cdots, m_t$ be positive integers. Then

$$\mathsf{E}\left\{\prod_{w=1}^{p} S(i_w) \prod_{w=1}^{q} S_1(j_w) \prod_{w=1}^{2r} T(l_w) \prod_{w=1}^{t} S_2(m_w)\right\} \quad (110)$$

is a polynomial in K of degree

$$D = \sum_{w=1}^{p}(i_w + 1) + \sum_{w=1}^{q} j_w + \sum_{w=1}^{2r}(l_w - \frac{1}{2}) + \sum_{w=1}^{t} m_w. \quad (111)$$

Proof: Let Θ denote the product in the expectation in (110). The size of Θ is defined as the right hand side of (111). The lemma is equivalent to saying that the degree of $\mathsf{E}\{\Theta\}$ is equal to the size of Θ. In other words, each $S(i)$ in Θ contributes $(i+1)$ to the degree of $\mathsf{E}\{\Theta\}$, each $S_1(j)$ contributes j, each $T(l)$ contributes $(l - \frac{1}{2})$, and each $S_2(m)$ contributes m. Note that we have an even number of $T(l_w)$'s so the size is always an integer. The proof is also by induction on the size of Θ.

We take a variable s_{nk} in the expansion of Θ and discuss all possible ways of matching it with another s variable. Upon a match, coinciding indexes merge into one. Equivalently, under the constraint that some indexes coincide, the s variables match and can be dropped. The resulting unconstrained sum, denoted as Θ', remains in the same topology with the same or a reduced size. If Θ' is of the same size as Θ, it may go through the reduction recursively until the size is reduced. In particular, if s_{nk} is matched to an immediate neighbor in the topology, a free index is produced, which contributes a degree of one to the overall sum, and the resulting unconstrained sum is one less in size. By the induction hypothesis, $\mathsf{E}\{\Theta'\}$ can be obtained as a polynomial and the degree of $\mathsf{E}\{\Theta\}$ can be deduced.

Consider the basis case of the induction: Θ has a size of one. Three possibilities arise:
(1) $q = 1$. Then $p = r = t = 0$ and $j_1 = 1$. Trivially,

$$\mathsf{E}\{\Theta\} = \mathsf{E}\{S_1(1)\} = \beta^{-1}K. \quad (112)$$

(2) $t = 1$. Similar to case (1). $\mathsf{E}\{\Theta\}$ takes the same value $\beta^{-1}K$.
(3) $r = 1$. Then $p = q = t = 0$ and $m_1 = m_2 = 1$. Hence,

$$\begin{aligned} \mathsf{E}\{\Theta\} &= \mathsf{E}\{T(1)T(1)\} && (113)\\ &= \sum_{n=1}^{N}\sum_{n'=1}^{N} \mathsf{E}\{s_{n2}s_{n1}s_{n'2}s_{n'1}\} && (114)\\ &= \beta^{-1}K. && (115)\end{aligned}$$

The lemma is therefore true for a size of one.

Suppose that the lemma is true for sizes of $1, \cdots, D-1$. We show that it is also true for a size of D. Take arbitrarily an indexed s variable, say s_{nk} from the expansion of Θ. The variable must be in one of three forms: s_{n2}, s_{n1}, or s_{nk}, with n and k as variable indexes. If no s variable of the first two forms exists, (110) is reduced to the form of (100) and the lemma holds trivially by Lemma 7. We assume that there exists a variable $s_{n_1 1}$ (if not, there must exist an $s_{n_1 2}$, which has no statistical difference to $s_{n_1 1}$ by homogeneity of the users).

$s_{n_1 1}$ is either from some $S_1(j)$ or from some $T(l)$. We consider both cases. Whichever the case, for $s_{n_1 1}$ to contribute nonzero to the expectation, it must be coupled with another s variable.

Suppose first that $s_{n_1 1}$ is from some $S_1(j)$. Six possibilities arise:
(1) $s_{n_1 1}$ is paired with a variable from $S_1(j)$ itself. Similar to discussions in the proof of Lemma 7, the size of Θ remains unchanged or is reduced.
(2) $s_{n_1 1}$ is paired with $s_{n'_1 1}$ in some $S_1(j')$. Under this constraint, $S_1(j) \cdot S_1(j')$ is reduced to

$$\begin{aligned}&\left\langle \begin{matrix} n'_{j'} & n'_{j'-1} & \cdots & n'_1 & n_2 & \cdots & n_j \\ 1 & k'_{j'} & \cdots & k'_2 & k_2 & \cdots & k_j \end{matrix} \right\rangle \\ =&S_1(j + j' - 1).\end{aligned} \quad (116)$$

The size of Θ is reduced by one.
(3) $s_{n_1 1}$ is paired with $s_{n'_u k'_u}$ in some $S_1(j')$. Under this constraint, $S_1(j) \cdot S_1(j')$ is reduced to

$$\begin{aligned}&\left\langle \begin{matrix} n'_1 & n'_2 & \cdots & n'_{u-1} \\ 1 & k'_2 & \cdots & k'_{u-1} \end{matrix} \right\rangle \times \\ &\left\langle \begin{matrix} n'_{j'} & n'_{j'-1} & \cdots & n'_u & n_2 & \cdots & n_j \\ 1 & k'_{j'} & \cdots & k'_{u+1} & k_2 & \cdots & k_j \end{matrix} \right\rangle \\ =&S_1(u-1) \cdot S_1(j + j' - u).\end{aligned} \quad (117)$$

The size of Θ is reduced by one.
(4) $s_{n_1 1}$ is paired with $s_{n'_u k'_u}$ in some $S(i)$. Under this constraint, $S_1(j) \cdot S(i)$ is reduced to

$$\begin{aligned}&\left\langle \begin{matrix} n'_{u-1} & n'_{u-2} & \cdots & n'_1 & n'_i & n'_{i-1} & \cdots & n'_u & n_2 & \cdots & n_j \\ 1 & k'_{u-1} & \cdots & k'_2 & k'_1 & k'_i & \cdots & k'_{u+1} & k_2 & \cdots & k_j \end{matrix} \right\rangle \\ =&S_1(j + j' - 1).\end{aligned} \quad (118)$$

The size of Θ is reduced by two.
(5) $s_{n_1 1}$ is paired with $s_{n'_u k'_u}$ in some $T(l)$. Under this constraint, $S_1(j) \cdot T(l)$ is reduced to

$$\begin{aligned}&\left\langle \begin{matrix} n_j & n_{j-1} & \cdots & n_2 & n'_u & n'_{u+1} & \cdots & n'_l & \\ 1 & k_j & \cdots & k_3 & k_2 & k'_{u+1} & \cdots & k'_l & 2 \end{matrix} \right\rangle \times \\ &\left\langle \begin{matrix} n'_1 & \cdots & n'_{u-1} \\ 1 & \cdots & k'_{u-1} \end{matrix} \right\rangle \\ =&T(j + l - u) \cdot S_1(u-1).\end{aligned} \quad (119)$$

The size of Θ is reduced by one.
(6) $s_{n_1 1}$ is paired with $s_{n'_u k'_u}$ in some $S_2(m)$. Under this

constraint, $S_1(j) \cdot S_2(m)$ is reduced to

$$\begin{aligned}&\left\langle \begin{matrix} n'_{u-1} & n'_{u-2} & \cdots & n'_1 & \\ 1 & k'_{u-1} & \cdots & k'_2 & 2 \end{matrix} \right\rangle \times \\ &\left\langle \begin{matrix} n_j & n_{j-1} & \cdots & n_2 & n'_u & n'_{u+1} & \cdots & n'_m & \\ 1 & k_j & \cdots & k_3 & k_2 & k'_{u+1} & \cdots & k'_m & 2 \end{matrix} \right\rangle \\ =&T(u-1) \cdot T_1(j+m-u). \end{aligned} \tag{120}$$

The size of Θ is reduced by two.

For similar reasons as in the proof of Lemma 7, contributions of (2)–(6) can be neglected since case (1) has a higher degree in K and hence dominates. Note that case (1) becomes a trivial after a finite number of splitting and results in a degree of j. The induction is then verified for the case that $s_{n_1 1}$ comes from some $S_1(j)$.

Assume that $s_{n_1 1}$ is from some $T(l)$ for some l. Six possibilities arise:

(1) $s_{n_1 1}$ is paired with a variable from $T(l)$ itself. Similar to discussions in possibility (1) in the previous case study. The size of Θ is either of the same or a reduced size.

(2) $s_{n_1 1}$ is paired with $s_{n'_1 1}$ from some $T(l')$. Under this constraint, $T(l) \cdot T(l')$ is reduced to

$$\begin{aligned}&\left\langle \begin{matrix} n_l & n_{l-1} & \cdots & n_2 & n'_1 & n'_2 & \cdots & n'_{l'} \\ 2 & k_l & \cdots & k_3 & k_2 & k'_2 & \cdots & k'_{l'} \end{matrix} \right\rangle \\ =\ & S_2(l+l'-1). \end{aligned} \tag{121}$$

The size of Θ remains unchanged.

(3) $s_{n_1 1}$ is paired with $s_{n'_u k'_u}$ from some $T(l')$. Under this constraint, $T(l) \cdot T(l')$ is reduced to

$$\begin{aligned}&\left\langle \begin{matrix} n'_1 & n'_2 & \cdots & n'_{u-1} \\ 1 & k'_2 & \cdots & k'_{u-1} \end{matrix} \right\rangle \times \\ &\left\langle \begin{matrix} n'_{l'} & n'_{l'-1} & \cdots & n'_u & n_2 & \cdots & n_l \\ 2 & k'_{l'} & \cdots & k'_{u+1} & k_2 & \cdots & k_l \end{matrix} \right\rangle \\ =&S_1(u-1) \cdot S_2(l+l'-u). \end{aligned} \tag{122}$$

The size of Θ remains unchanged.

(4) $s_{n_1 1}$ is paired with $s_{n'_u k'_u}$ from some $S(i)$. Under this constraint, $T(l) \cdot S(i)$ is reduced to

$$\begin{aligned}&\left\langle \begin{matrix} n'_{u-1} & n'_{u-2} & \cdots & n'_1 & n'_i & n'_{i-1} & \cdots & n'_u & n_2 & \cdots & n_l & \\ 1 & k'_{u-1} & \cdots & k'_2 & k'_1 & k'_i & \cdots & k'_{u+1} & k_2 & \cdots & k_l & 2 \end{matrix} \right\rangle \\ =&T(l+i-1) \end{aligned} \tag{123}$$

The size of Θ is reduced by one.

(5) $s_{n_1 1}$ is paired with $s_{n'_u k'_u}$ from some $S_1(j)$. Under this constraint, $T(l) \cdot S_1(j)$ is reduced to

$$\begin{aligned}&\left\langle \begin{matrix} n'_1 & n'_2 & \cdots & n'_{u-1} \\ 1 & k'_2 & \cdots & k'_{u-1} \end{matrix} \right\rangle \times \\ &\left\langle \begin{matrix} n'_j & n'_{j-1} & \cdots & n'_u & n_2 & \cdots & n_l & \\ 1 & k'_j & \cdots & k'_{u+1} & k_2 & \cdots & k_l & 2 \end{matrix} \right\rangle \\ =&S_1(u-1) \cdot T(l+j-u) \end{aligned} \tag{124}$$

The size of Θ is reduced by one.

(6) $s_{n_1 1}$ is paired with $s_{n'_u k'_u}$ from some $S_2(m)$. Under this constraint, $T(l) \cdot S_2(m)$ is reduced to

$$\begin{aligned}&\left\langle \begin{matrix} n'_{u-1} & n'_{u-2} & \cdots & n'_1 & \\ 1 & k'_{u-1} & \cdots & k'_2 & 2 \end{matrix} \right\rangle \times \\ &\left\langle \begin{matrix} n'_m & n'_{m-1} & \cdots & n'_u & n_2 & \cdots & n_l \\ 2 & k'_m & \cdots & k'_{u+1} & k_2 & \cdots & k_l \end{matrix} \right\rangle \\ =&T(u-1) \cdot S_2(l+m-u). \end{aligned} \tag{125}$$

The size of Θ is reduced by one.

Again, in none of the above six cases, the size of Θ is increased. Cases (4)–(6) can be neglected and cases (1)–(3) can be reduced to trivial cases after a finite number of iterations. It is not difficult to see by the induction hypothesis that the overall expectation of Θ is a polynomial of degree D. The induction is therefore also verified for the case that $s_{n_1 1}$ comes from some $T(l)$.

The induction holds, thus the proof is complete. ■

We need a further result on the constrained sum where certain patterns of the indexes are prohibited. Define

$$S'_1(i) = \sum_{\mathcal{J} \in B(i)-C(i)} s_{n_1 1} s_{n_1 k_2} s_{n_2 k_2} \cdots s_{n_{i-1} k_i} s_{n_i k_i} s_{n_i 1} \tag{126}$$

where $C(i) = \{\mathcal{J} | \mathcal{J}$ is such that the indexed s variables in (126) form complete pairs$\}$.

Lemma 9: Let $p, q, r \geq 0$ be integers. Let $i_1, \cdots, i_p$, $j_1, \cdots, j_q$ and $l_1, \cdots, l_{2r}$ be positive integers. Then

$$\mathsf{E}\left\{ \prod_{w=1}^{p} S(i_w) \prod_{w=1}^{q} S_1(j_w) \prod_{w=1}^{2r} S'_1(l_w) \right\} \tag{127}$$

is a polynomial in K of degree

$$D = \sum_{w=1}^{p} (i_w + 1) + \sum_{w=1}^{q} j_w + \sum_{w=1}^{2r} (l_w - \frac{1}{2}). \tag{128}$$

Proof: If $r = 0$, the lemma is trivial by Lemma 8. Suppose $r \geq 1$. Take any $S'_1(l)$ and convert it to the form of $S(i)$ or $S_1(j)$. Since the s variables in $S'_1(l)$ are not allowed to form complete pairs, there must be a s_{nk} from it that is paired with either a variable from $S(i)$ for some i, $S_1(j)$ from some j, or $S'_1(l')$ for some l'. In the first two cases we can look upon $S'_1(l)$ as if it is a larger $S_1(l)$. Note that by adding more terms to $S'_1(l)$ we will only increase the overall expectation. By similar arguments as in the proof of Lemma 8, the resulting contribution has one degree less. Consequently, these two cases can be neglected. In the last case, $S'_1(l) \cdot S'_1(l')$ is reduced to $S_1(u-1) \cdot S_1(l+l'-u)$ for some u. By the same method, all $S'_1(l)$ variables can be converted to the form of $S_1(i)$. The resulting problem is solved by Lemma 8 and we have the desired result. ■

With the above lemmas established, we can prove Proposition 1.

Proof: [**Proposition 1**] For an odd p, the p^{th}-order moment is always zero by symmetry of the distribution of the random chips. Note also that $\left\{ \left[\mathbf{R}^i\right]_{1k} \right\}_{k=2}^{K}$ is a set of identically distributed random variables. Hence it suffices to show the finiteness of the moments for $k = 1, 2$ and all even p.

Consider first the case of $k = 2$. By (99),

$$\mathsf{E}\left\{\left(\sqrt{K}\left[\mathbf{R}^i\right]_{12}\right)^p\right\} = \left(\beta^i K^{\frac{1}{2}} K^{-i}\right)^p \mathsf{E}\left\{T^p(i)\right\}. \quad (129)$$

By Lemma 8, $\mathsf{E}\{T^p(i)\}$ is a polynomial in K of degree $p(i - \frac{1}{2})$. Equation (129) is therefore a polynomial fraction whose numerator and denominator have the same degree. Its convergence as $K \to \infty$ is evident.

Consider now the case of $k = 1$. Clearly, $\left[\mathbf{R}^i\right]_{11} = N^{-i} S_1(i)$. Note that

$$\begin{aligned}
&\left[\mathbf{R}^i\right]_{11} - \mathsf{E}\left\{\left[\mathbf{R}^i\right]_{11}\right\} \\
=\ & \beta^{pi} K^{-i} \sum_{\mathcal{J} \in B(i) - C(i)} s_{n_1 1} s_{n_1 k_2} \cdots s_{n_i k_i} s_{n_i 1} \quad (130) \\
=\ & \beta^i K^{-i} S_1'(i) \quad (131)
\end{aligned}$$

since for the product terms of $\left[\mathbf{R}^i\right]_{11}$ to have nonzero expectation, the indexes must be such that the s variables form complete pairs. Therefore,

$$\begin{aligned}
&\mathsf{E}\left\{\left[\sqrt{K}\left(\left[\mathbf{R}^i\right]_{11} - \mathsf{E}\left\{\left[\mathbf{R}^i\right]_{11}\right\}\right)\right]^p\right\} \\
=&\beta^i K^{-p(i-\frac{1}{2})} \mathsf{E}\left\{\left[S_1'(i)\right]^p\right\}
\end{aligned} \quad (132)$$

converges as $K \to \infty$ by Lemma 9. ■

II. Non-binary Spreading With No Power Control

We now drop the perfect power control assumption and allow non-binary spreading sequences. Define a set of antipodal spreading sequences induced from the non-binary chips $\{s_{nk}\}$ by taking the sign of each chip,

$$\hat{s}_{nk} = \operatorname{sgn}(s_{nk}). \quad (133)$$

Therefore,

$$s_{nk} = \sqrt{P_k} \cdot a_{nk} \cdot \hat{s}_{nk} \quad (134)$$

where $a_{nk} = |\bar{s}_{nk}|$ is the amplitude of a chip of the normalized spreading sequence. Note that the three variables on the right hand side of (134) are independent.

Let $S(i)$ be defined as in (92). Consider the expectation of each summand,

$$\begin{aligned}
&\mathsf{E}\left\{s_{n_1 k_1} s_{n_1 k_2} s_{n_2 k_2} \cdots s_{n_{i-1} k_i} s_{n_i k_i} s_{n_i k_1}\right\} \\
=&\mathsf{E}\left\{\sqrt{P_{k_1} P_{k_2} \cdots P_{k_i}}\right\} \times \\
&\quad \mathsf{E}\left\{a_{n_1 k_1} a_{n_1 k_2} a_{n_2 k_2} \cdots a_{n_{i-1} k_i} a_{n_i k_i} a_{n_i k_1}\right\} \times \\
&\quad\quad \mathsf{E}\left\{\hat{s}_{n_1 k_1} \hat{s}_{n_1 k_2} \hat{s}_{n_2 k_2} \cdots \hat{s}_{n_{i-1} k_i} \hat{s}_{n_i k_i} \hat{s}_{n_i k_1}\right\}.
\end{aligned} \quad (135)$$

Note that the first expectation on the right hand side of (135), i.e., the expectation of the product of energies, is a product of moments of P_k's dependent on the values of $k_1, \ldots, k_i$. It is a deterministic number independent of K. Similarly, the second expectation is a product of moments of $|\bar{s}_{nk}|$'s, or equivalently of $\bar{s}_{nk}$'s, for every possible combination of $n_1, \ldots, n_i, k_1, \ldots, k_i$. For the final expectation to be nonzero, these moments are non-negative since the variables appear in pairs. Hence it is also some known number independent of K under our assumption of finite moments.

It is easy to show that Corollary 1 still holds by noting that the moments of the energies as well as the moments of the non-binary chips are bounded by some numbers independent of K, so that the problem is reduced to the case of binary spreading with perfect power control.

Moreover, we can exploit the structure of $S(i)$ in the same way as in Lemma 7. The result is nonetheless a polynomial of the same degree in K. For similar reasons, Lemma 8 and 9 also hold. Therefore, Proposition 1 can be proved for this case, i.e., the central moments of $\sqrt{K}\left[\mathbf{R}^i\right]_{1k}$ converge to deterministic constants as $K \to \infty$, even with non-binary spreading and no power control.

III. Proof of Lemma 5

Proof: **[Lemma 5]** Fix the dimension K. Let $\tilde{\mathbf{G}} = \tilde{g}(\mathbf{R})$ be a polynomial receiver. Let $\mathbf{y}$ and $\tilde{\mathbf{y}}$ denote the vector of output decision statistics of receiver $\mathbf{G}$ and $\tilde{\mathbf{G}}$ respectively. Then

$$\mathbf{y} - \tilde{\mathbf{y}} = (\mathbf{G} - \tilde{\mathbf{G}}) \cdot \mathbf{y}_{\text{MF}}. \quad (136)$$

Focusing on user 1,

$$\begin{aligned}
&\mathsf{E}\left\{(y_1 - \tilde{y}_1)^2 \,\middle|\, \mathbf{R}\right\} \\
=\ & \mathsf{E}\left\{\left[(\mathbf{y} - \tilde{\mathbf{y}})(\mathbf{y} - \tilde{\mathbf{y}})^\top\right]_{11} \,\middle|\, \mathbf{R}\right\} \quad (137) \\
=\ & \mathsf{E}\left\{\left[\left(\mathbf{G} - \tilde{\mathbf{G}}\right) \mathbf{y}_{\text{MF}} \mathbf{y}_{\text{MF}}^\top \left(\mathbf{G} - \tilde{\mathbf{G}}\right)\right]_{11} \,\middle|\, \mathbf{R}\right\} \quad (138) \\
=\ & \left[\left(\mathbf{G} - \tilde{\mathbf{G}}\right)(\mathbf{R} + \sigma^2)\mathbf{R}\left(\mathbf{G} - \tilde{\mathbf{G}}\right)\right]_{11} \quad (139) \\
=\ & \left[\mathbf{U}\left(g(\mathbf{\Lambda}) - \tilde{g}(\mathbf{\Lambda})\right)^2 (\mathbf{\Lambda} + \sigma^2\mathbf{I})\mathbf{\Lambda}\mathbf{U}^\top\right]_{11} \quad (140) \\
=\ & \sum_{k=1}^{K} |u_{1k}|^2 \left(g(\lambda_k) - \tilde{g}(\lambda_k)\right)^2 (\lambda_k + \sigma^2)\lambda_k \quad (141)
\end{aligned}$$

where u_{1k} is the k^{th} element on the first row of $\mathbf{U}$. By choosing the polynomial $\tilde{g}$ to be within a distance of ϵ to g on the support of F_Λ, we have

$$\begin{aligned}
\mathsf{E}\left\{(y_1 - \tilde{y}_1)^2 \mid \mathbf{R}\right\} &< \epsilon^2 \cdot \sum_{k=1}^{K} |u_{1k}|^2 (\lambda_k + \sigma^2)\lambda_k \quad (142) \\
&= \epsilon^2 \cdot \left[\mathbf{R}(\mathbf{R} + \sigma^2\mathbf{I})\right]_{11} \quad (143) \\
&\to \epsilon^2 \cdot (M_2 + \sigma^2 M_1) \quad (144)
\end{aligned}$$

with probability 1 as $K \to \infty$. Hence the mean square error of $(y_1 - \tilde{y}_1)$ can be made arbitrarily small for sufficiently large K for almost all spreading sequences. ■

IV. Proof of Lemma 6

Proof: **[Lemma 6]** We first show that $F_m(a)$ is a Cauchy sequence for every a. Since mean square convergence implies convergence in distribution at all points of continuity, we have

$$\lim_{m \to \infty} \lim_{K \to \infty} \left|F^{(K)}(a) - F_m^{(K)}(a)\right| = 0, \quad \forall a \quad (145)$$

by (L6.1). Also, $\forall a$, $\forall \epsilon > 0$, $\forall p > 0$, and $\forall K$,

$$\begin{aligned} &|F_{m+p}(a) - F_m(a)| \\ &\le \left|F_{m+p}(a) - F^{(K)}_{m+p}(a)\right| + \left|F^{(K)}_{m+p}(a) - F^{(K)}(a)\right| \\ &\quad + \left|F^{(K)}(a) - F^{(K)}_m(a)\right| + \left|F^{(K)}_m(a) - F_m(a)\right|. \end{aligned} \tag{146}$$

Taking the limit $K \to \infty$ and then $m \to \infty$, we have all four terms converging to 0 and so

$$\lim_{m\to\infty} |F_{m+p}(a) - F_m(a)| = 0, \quad \forall a, \forall p > 0. \tag{147}$$

Therefore F_m converge pointwise. Furthermore, for every a,

$$\begin{aligned} &\left|F^{(K)}(a) - F_m(a)\right| \\ &\le \left|F^{(K)}(a) - F^{(K)}_m(a)\right| + \left|F^{(K)}_m(a) - F_m(a)\right|. \end{aligned} \tag{148}$$

Taking the limit $K \to \infty$ and then $m \to \infty$, we have

$$\lim_{K\to\infty} F^{(K)}(a) = \lim_{m\to\infty} F_m(a). \tag{149}$$

■

References

[1] Sergio Verdú, *Multiuser Detection*, Cambridge University Press, 1998.

[2] P. D. Alexander, A. J. Grant, and M. C. Reed, "Iterative detection in code-division multiple-access with error control coding," *European Trans. Telecommun.*, vol. 9, pp. 419–426, Sept./Oct. 1998.

[3] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Commun.*, vol. 47, pp. 1046–1060, July 1999.

[4] L. K. Rasmussen, T. J. Lim, and A.-L. Johansson, "A matrix-algebraic approach to successive interference cancellation in CDMA," *IEEE Trans. Commun.*, vol. 48, pp. 145–151, Jan. 2000.

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

[6] H. Elders-Boll, H. D. Schotten, and A. Busboom, "Efficient implementation of linear multiuser detectors for asynchronous CDMA systems by linear interference cancellation," *European Trans. on Telecommun.*, vol. 9, no. 4, pp. 427–438, Sept./Oct. 1998.

[7] R. M. Buehrer, S. P. Nicoloso, and S. Gollamudi, "Linear versus nonlinear interference cancellation," *J. Commun. Networks*, vol. 1, pp. 118–133, June 1999.

[8] P. H. Tan and L. K. Rasmussen, "Linear interference cancellation in CDMA based on iterative techniques for linear equation systems," *IEEE Trans. Commun.*, vol. 48, pp. 2099–2108, Dec. 2000.

[9] A. Grant and C. Schlegel, "Convergence of linear interference cancellation multiuser receivers," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1824–1834, Oct. 2001.

[10] C. L. Weber, G. K. Huth, and B. H. Batson, "Performance considerations of code division multiple-access systems," *IEEE Trans. Veh. Technol.*, vol. 30, no. 1, pp. 3–10, Feb. 1981.

[11] S. Verdú and S. Shamai, "Spectral efficiency of CDMA with random spreading," *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 622–640, March 1999.

[12] H. Vincent Poor and Sergio Verdú, "Probability of error in MMSE multiuser detection," *IEEE Trans. Inform. Theory*, vol. 43, no. 3, pp. 858–871, May 1997.

[13] Junshan Zhang, Edwin K. P. Chong, and David N. C. Tse, "Output MAI distribution of linear MMSE multiuser receivers in DS-CDMA systems," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1128–1144, March 2001.

[14] Junshan Zhang and Xiaodong Wang, "Large-system performance analysis of blind and group-blind multiuser receivers," *IEEE Trans. Inform. Theory*, vol. 48, pp. 2507–2523, Sept. 2002.

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

[16] D. Guo, L. K. Rasmussen, and T. J. Lim, "Linear parallel interference cancellation in long-code CDMA multiuser detection," *IEEE J. Selected Areas Commun.*, vol. 17, pp. 2074–2081, Dec. 1999.

[17] Alex J. Grant and Paul D. Alexander, "Random sequence multisets for synchronous code-division multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 44, no. 7, pp. 2832–2836, Nov. 1998.

[18] David N. C. Tse and Stephen V. Hanly, "Linear multiuser receivers: effective interference, effective bandwidth and user capacity," *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 622–640, March 1999.

[19] D. Guo and S. Verdú, "Multiuser detection and statistical mechanics," in *Communications, Information and Network Security* (V. Bhargava, H. V. Poor, V. Tarokh, and S. Yoon, eds.), ch. 13, pp. 229–277, Kluwer Academic Publishers, 2002. [Online]. Available: http://www.princeton.edu/˜dguo/research/.

[20] V. A. Marcenko and L. A. Pastur, "Distribution of eigenvalues for some sets of random matrices," *Math. USSR-Sbornik 1*, pp. 457–483, 1967.

[21] J. W. Silverstein and Z. D. Bai, "On the empirical distribution of eigenvalues of a class of large dimensional random matrices," *Journal of Multivariate Analysis*, vol. 54, no. 2, pp. 175–192, Aug. 1995.

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

[23] R. Mueller and S. Verdú, "Design and analysis of low-complexity interference mitigation on vector channels," *IEEE J. Selected Areas Commun.*, vol. 19, no. 8, pp. 1429–1441, Aug. 2001.

[24] A. Papoulis, *Probability, Random Variable, and Stochastic Processes*, McGraw-Hill, Inc., 3rd edition, 1991.

[25] William Feller, *An Introduction to Probability Theory and Its Applications*, vol. I, John Wiley & Sons, Inc., 3rd edition, 1968.

[26] A. Rényi, *Probability Theory*, Budapest, Hungary: North Holland – Academiai Kiado, 1970.

[27] D. L. McLeish, "Dependent central limit theorems and invariance principles," *Annals of Probability*, vol. 2, no. 4, pp. 620–628, 1974.

[28] Walter Rudin, *Principles of Mathematical Analysis*, McGraw-Hill, Inc., 3rd edition, 1976.

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

[30] Toshiyuki Tanaka, "A statistical mechanics approach to large-system analysis of CDMA multiuser detectors," *submitted to IEEE Trans. Inform. Theory*, 2002.

[31] H. Nishimori, *Statistical Physics of Spin Glasses and Information Processing: An Introduction*, Number 111 in International Series of Monographs on Physics. Oxford University Press, 2001.

[32] Y. C. Eldar and A. M. Chan, "The independence of wishart matrix eigenvalues and eigenvectors, and its application to the asymptotic performance of the decorrelator," *submitted to IEEE Trans. Inform. Theory*, 2001.

[33] S. Shamai and S. Verdú, "The impact of frequency-flat fading on the spectral efficiency of CDMA," *IEEE Trans. Inform. Theory*, vol. 47, no. 4, pp. 1302–1327, May 2001.

[34] U. Grenander and J. W. Silverstein, "Spectral analysis of networks with random topologies," *SIAM J. Appl. Math.*, vol. 32, no. 2, pp. 499–519, March 1977.