Performance of Reduced-Rank Linear Interference Suppression

Michael L. Honig, *Fellow, IEEE*, and Weimin Xiao, *Member, IEEE*

***Abstract*—The performance of reduced-rank linear filtering is studied for the suppression of multiple-access interference. A reduced-rank filter resides in a lower dimensional space, relative to the full-rank filter, which enables faster convergence and tracking. We evaluate the large system output signal-to-interference plus noise ratio (SINR) as a function of filter rank D for the multistage Wiener filter (MSWF) presented by Goldstein and Reed. The large system limit is defined by letting the number of users K and the number of dimensions N tend to infinity with K/N fixed. For the case where all users are received with the same power, the reduced-rank SINR converges to the full-rank SINR as a *continued fraction*. An important conclusion from this analysis is that the rank D needed to achieve a desired output SINR does not scale with system size. Numerical results show that $D = 8$ is sufficient to achieve near-full-rank performance even under heavy loads $(K/N = 1)$. We also evaluate the large system output SINR for other reduced-rank methods, namely, Principal Components and Cross-Spectral, which are based on an eigendecomposition of the input covariance matrix, and Partial Despreading (PD). For those methods, the large system limit lets $D \to \infty$ with D/N fixed. Our results show that for large systems, the MSWF allows a dramatic reduction in rank relative to the other techniques considered.**

***Index Terms*—Interference suppression, large system analysis, multiuser detection, reduced-rank filters.**

I. Introduction

REDUCED-rank filtering and estimation have been proposed for numerous signal processing applications such as array processing, radar, model order reduction, and quantization (e.g., see [1]–[4] and references therein). A reduced-rank estimator may require relatively little observed data to produce an accurate approximation of the optimal filter. In this paper, we study the performance of reduced-rank linear filters for the suppression of multiple-access interference.

Reduced-rank linear filtering has recently been applied to interference suppression in direct-sequence (DS) code-division multiple access (CDMA) systems [5]–[10]. Although conventional adaptive filtering algorithms can be used to estimate the linear minimum mean-squared error (MMSE) detector, assuming short, or repeated spreading codes [11], the performance may be inadequate when a large number of filter coefficients must be estimated. For example, a conventional implementation of a time-domain adaptive filter which spans three symbols for proposed third-generation wide-band DS-CDMA cellular systems can have over 300 coefficients. Introducing multiple antennas for additional space–time interference suppression capability exacerbates this problem. Adapting such a large number of filter coefficients is hampered by very slow response to changing interference and channel conditions.

In a reduced-rank filter, the received signal is projected onto a lower dimensional subspace, and the filter optimization then occurs within this subspace. This has the advantage of reducing the number of filter coefficients to be estimated. However, by adding this subspace constraint, the overall MMSE may be higher than that achieved by a full-rank filter. Much of the previous work on reduced-rank interference suppression has been based on "Principal Components" in which the received vector is projected onto an estimate of the lower dimensional signal subspace with largest energy (e.g., [8], [12]). This technique can improve convergence and tracking performance when the number of degrees of freedom (e.g., CDMA processing gain) is much larger than the signal subspace. This assumption, however, does not hold in a heavily loaded commercial cellular system.

Our main contribution is to characterize the performance of the reduced-rank multistage Wiener filter (MSWF) presented by Goldstein and Reed [13], [14]. This technique has the important property that the filter rank (i.e., dimension of the projected subspace) can be much less than the dimension of the signal subspace without compromising performance. Furthermore, adaptive estimation of the optimum filter coefficients does not require an eigendecomposition of the input (sample) covariance matrix. Adaptive interference suppression algorithms based on the MSWF are presented in [9].

Our performance evaluation is motivated by the large system analysis for DS-CDMA with random spreading sequences introduced in [15]–[17]. Specifically, let K be the number of users, N be the number of available dimensions (e.g., chips per coded symbol in CDMA, or number of receiver antennas in a narrow-band system), and D be the subspace dimension. We evaluate the signal-to-interference plus noise ratio (SINR) at the output of the MSWF as $K, N \to \infty$ with K/N fixed. For the case where all users are received with the same power, we obtain a closed-form expression for the output SINR as a function of D, K/N, and the background noise variance. As D increases, this expression rapidly converges to the full-rank large system SINR derived in [17] as a *continued-fraction*. The MSWF therefore has the surprising property that the dimension D needed to obtain a target SINR (e.g., within a small ϵ of the

Manuscript received January 10, 2000; revised January 9, 2001. This work was supported by the U.S. Army Research Office under Grant DAAH04-96-1-0378.

The authors are with the Department of Electrical and Computer Engineering, Northwestern University, Evanston, IL 60208 USA (e-mail: mh@ece.nwu.edu; xiao@ece.nwu.edu).

Communicated by U. Madhow, Associate Editor for Detection and Estimation.

Publisher Item Identifier S 0018-9448(01)04413-3.

full-rank SINR) does not scale with the system size (i.e., K and N). Our results show that for moderate to heavy loads, a rank $D = 8$ filter essentially achieves full-rank performance, and the SINR for a rank $D = 4$ filter is within 1 dB of the full-rank SINR.

We also evaluate the large system performance of the reduced-rank MSWF given an arbitrary power distribution. A byproduct of this analysis is a method for computing the full-rank large system SINR which does not explicitly make use of the asymptotic eigenvalue distribution for the class of random matrices derived in [18], and used in [15]–[17].

Finally, we compare the large system performance of the MSWF with the following reduced-rank techniques. 1) Principal Components, 2) Cross-Spectral [19], [20], and 3) Partial-Despreading [21]. (See also [6].) The Cross-Spectral method is based on an eigendecomposition of the input covariance matrix, but unlike Principal Components, selects the basis vectors which minimize MSE. Partial Despreading (PD) refers to a relatively simple reduced-rank technique in which the subspace is spanned by nonoverlapping segments of the matched filter.

In contrast with the MSWF, the large system analysis of the latter techniques lets $D, K, N \to \infty$ with both K/N and D/K fixed. That is, to achieve a target SINR near the full-rank large system limit, $D \to \infty$ as $K, N \to \infty$. For the case where all users are received with the same power, we obtain closed-form expressions which accurately predict output SINR as a function of K/N, D/K, and noise variance.

In the next two sections, we present the system model and the reduced-rank techniques considered. In Section IV, we briefly review large system analysis. Our main results are presented in Section V, and numerical examples are presented in Section VI. Proofs and derivations are given in Section VII.

II. System Model

Let $\boldsymbol{r}(i)$ be the $(N \times 1)$ received vector corresponding to the ith transmitted symbol. For example, the elements of $\boldsymbol{r}(i)$ may be samples at the output of a chip-matched filter (for CDMA) or across an antenna array. We assume that

$$\boldsymbol{r}(i) = \boldsymbol{S}\boldsymbol{A}\boldsymbol{b}(i) + \boldsymbol{n}(i) \tag{1}$$

where

$$\boldsymbol{S} = [\boldsymbol{s}_1, \ldots, \boldsymbol{s}_K] \tag{2}$$

is the $N \times K$ matrix of signature sequences where N is the number of dimensions (e.g., processing gain or number of antennas) and K is the number of users, and $\boldsymbol{s}_k$ is the signature sequence for user k. The amplitude matrix

$$\boldsymbol{A} = \mathrm{diag}(\sqrt{P_1}, \ldots, \sqrt{P_K})$$

where P_k is the power for user k, $\boldsymbol{b}(i)$ is the $(K \times 1)$-vector of symbols across users at time i, and $\boldsymbol{n}(i)$ is the noise vector, which has covariance matrix $\sigma^2\boldsymbol{I}$. We assume that the symbol variance is one for all users, and that all vectors are complex valued.

In what follows we assume that user 1 is the desired user. The MMSE receiver consists of the filter represented by the vector $\boldsymbol{c}$, which is chosen to minimize the mean-square error (MSE)

$$\mathcal{M} = E\{|b_1(i) - \boldsymbol{c}^\dagger\boldsymbol{y}(i)|^2\} \tag{3}$$

where $E\{\cdot\}$ denotes expectation, and $\dagger$ denotes Hermitian transpose. The MMSE solution is [11]

$$\boldsymbol{c} = \boldsymbol{R}^{-1}\boldsymbol{s}_1 \tag{4}$$

where the $N \times N$ covariance matrix

$$\boldsymbol{R} = E[\boldsymbol{r}(i)\boldsymbol{r}^\dagger(i)] = \boldsymbol{S}\boldsymbol{P}\boldsymbol{S}^\dagger + \sigma^2\boldsymbol{I} \tag{5}$$

where $\boldsymbol{P} = \boldsymbol{A}\boldsymbol{A}^\dagger$ and the (full-rank) MMSE is

$$\mathcal{M} = 1 - \boldsymbol{s}_1^\dagger\boldsymbol{R}^{-1}\boldsymbol{s}_1. \tag{6}$$

Let the $N \times (K-1)$ matrix of spreading codes for the interferers be

$$\boldsymbol{S}_I = [\boldsymbol{s}_2, \ldots, \boldsymbol{s}_K] \tag{7}$$

and

$$\boldsymbol{r}_I(i) = \boldsymbol{S}_I\boldsymbol{A}_I\boldsymbol{b}_{2:K}(i) + \boldsymbol{n}(i) \tag{8}$$

where $\boldsymbol{A}_I$ is the diagonal matrix of interference amplitudes and $\boldsymbol{x}_{l:m}$ denotes components l through m of the vector $\boldsymbol{x}$. The interference-plus-noise covariance matrix is

$$\boldsymbol{R}_I = E[\boldsymbol{r}_I(i)\boldsymbol{r}_I^\dagger(i)] = \boldsymbol{S}_I\boldsymbol{A}_I\boldsymbol{A}_I^\dagger\boldsymbol{S}_I^\dagger + \sigma^2\boldsymbol{I}. \tag{9}$$

The output SINR of the MMSE filter is

$$\beta = P_1\boldsymbol{s}_1^\dagger\boldsymbol{R}_I^{-1}\boldsymbol{s}_1 \tag{10}$$

where P_1 is the received power for user 1.

III. Reduced-Rank Linear Filtering

A reduced-rank filter reduces the number of coefficients to be estimated by projecting the received vector onto a lower dimensional subspace [22, Sec. 8.4], [2]. Specifically, let $\boldsymbol{M}_D$ be the $N \times D$ matrix with column vectors forming a basis for a D-dimensional subspace, where $D < N$. The vector of combining coefficients for the ith received vector corresponding to this subspace is given by

$$\tilde{\boldsymbol{r}}(i) = (\boldsymbol{M}_D^\dagger\boldsymbol{M}_D)^{-1}\boldsymbol{M}_D^\dagger\boldsymbol{r}(i). \tag{11}$$

In what follows, a "tilde" denotes a (reduced-rank) D-dimensional vector, or $D \times D$ covariance matrix.

The sequence of vectors $\{\tilde{\boldsymbol{r}}(i)\}$ is the input to a tapped-delay line filter, represented by the $(D \times 1)$-vector $\tilde{\boldsymbol{c}}$. The filter output corresponding to the ith transmitted symbol is $z(i) = \tilde{\boldsymbol{c}}^\dagger\tilde{\boldsymbol{r}}(i)$, and the objective is to select $\tilde{\boldsymbol{c}}$ to minimize the reduced-rank MSE

$$\mathcal{M}_D = E\{|b_1(i) - \tilde{\boldsymbol{c}}^\dagger\tilde{\boldsymbol{r}}(i)|^2\}. \tag{12}$$

The solution is

$$\tilde{\boldsymbol{c}} = \tilde{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{s}}_1 \tag{13}$$

where

$$\tilde{\boldsymbol{R}} = (\boldsymbol{M}_D^\dagger\boldsymbol{M}_D)^{-1}(\boldsymbol{M}_D^\dagger\boldsymbol{R}\boldsymbol{M}_D)(\boldsymbol{M}_D^\dagger\boldsymbol{M}_D)^{-1} \tag{14}$$

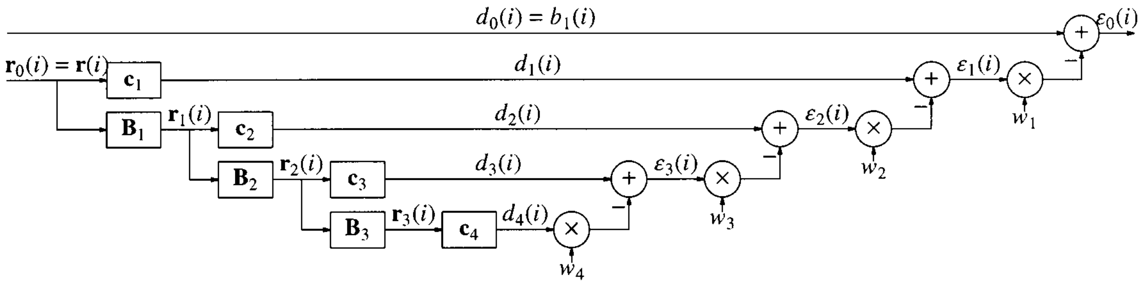

Fig. 1. Multistage Wiener filter (MSWF).

and

$$\tilde{\boldsymbol{s}}_1 = (\boldsymbol{M}_D^\dagger \boldsymbol{M}_D)^{-1} \boldsymbol{M}_D^\dagger \boldsymbol{s}_1. \tag{15}$$

Defining $\tilde{\boldsymbol{R}}_I$ in the obvious way, the output SINR is given by

$$\beta_D = P_1 \tilde{\boldsymbol{s}}_1^\dagger \tilde{\boldsymbol{R}}_I^{-1} \tilde{\boldsymbol{s}}_1 = P_1 \boldsymbol{s}_1^\dagger \boldsymbol{M}_D (\boldsymbol{M}_D^\dagger \boldsymbol{R}_I \boldsymbol{M}_D)^{-1} \boldsymbol{M}_D^\dagger \boldsymbol{s}_1. \tag{16}$$

In what follows, we present the reduced-rank filters of interest. We remark that other reduced-rank methods have been proposed in [5], [10], [23], [24], [4], [25]. (The auxiliary vector method proposed in [10] generates the same D-dimensional subspace as the MSWF.) A simulation study of adaptive versions of the eigendecomposition and PD interference suppression methods described here is presented in [6].

A. *Multistage Wiener Filter (MSWF)*

A block diagram showing four stages of the MSWF is shown in Fig. 1. The stages are associated with the sequence of nested filters $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_D$, where D is the order of the filter. Let $\boldsymbol{B}_m$ denote a *blocking matrix*, i.e.,

$$\boldsymbol{B}_m^\dagger \boldsymbol{c}_m = \boldsymbol{0}. \tag{17}$$

Referring to Fig. 1, let $d_m(i)$ denote the output of the filter $\boldsymbol{c}_m$, and $\boldsymbol{r}_m(i)$ denote the output of the blocking matrix $\boldsymbol{B}_m$, both at time i. Then the filter for the $(m+1)$th stage is determined by

$$\boldsymbol{c}_{m+1} = E[d_m^* \boldsymbol{r}_m] \tag{18}$$

where $*$ denotes complex conjugate. For $m = 0$, we have $d_0(i) = b_1(i)$ (the desired input symbol), $\boldsymbol{r}_0(i) = \boldsymbol{r}(i)$, and $\boldsymbol{c}_1$ is the matched filter $\boldsymbol{s}_1$. Here we assume that each blocking matrix $\boldsymbol{B}_m$ is $N \times N$, so that each vector $\boldsymbol{c}_m$ is $N \times 1$. As in [14], it will be convenient to normalize the filters $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_D$ so that $\|\boldsymbol{c}_m\| = 1$.

The filter output is obtained by linearly combining the outputs of the filters $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_D$ via the weights $w_1, \ldots, w_{D-1}$. This is accomplished stage-by-stage. Referring to Fig. 1, let

$$\epsilon_m(i) = d_m(i) - w_{m+1}\epsilon_{m+1}(i) \tag{19}$$

for $1 \le m \le D$ and $\epsilon_D(i) = d_D(i)$. Then w_{m+1} is selected to minimize $E[|\epsilon_m|^2]$.

The rank D MSWF is given by the following set of recursions.

Initialization:

$$d_0(i) = b_1(i), \qquad \boldsymbol{r}_0(i) = \boldsymbol{r}(i). \tag{20}$$

For $n = 1, \ldots, D$ (*Forward Recursion*)

$$\boldsymbol{c}_n = E[d_{n-1}^* \boldsymbol{r}_{n-1}(i)] / \|E[d_{n-1}^* \boldsymbol{r}_{n-1}]\| \tag{21}$$

$$d_n(i) = \boldsymbol{c}_n^\dagger \boldsymbol{r}_{n-1}(i) \tag{22}$$

$$\boldsymbol{B}_n = \boldsymbol{I} - \boldsymbol{c}_n \boldsymbol{c}_n^\dagger \tag{23}$$

$$\boldsymbol{r}_n = \boldsymbol{B}_n^\dagger \boldsymbol{r}_{n-1}. \tag{24}$$

Decrement $n = D, \ldots, 1$ (*Backward Recursion*)

$$w_n = E[d_{n-1}^*(i)\epsilon_n(i)] / E[|\epsilon_n(i)|^2] \tag{25}$$

$$\epsilon_{n-1}(i) = d_{n-1}(i) - w_n^* \epsilon_n(i) \tag{26}$$

where $\epsilon_D(i) = d_D(i)$.

The MSWF has the following properties.

1) At stage n the filter generates the desired (time) sequence $\{d_n(i)\}$ and the "observation" sequence $\{\boldsymbol{r}_n(i)\}$. The MSWF for estimating the former from the latter has the same structure at each stage. The full-rank MMSE filter can, therefore, be represented as an MSWF with n stages, where $\boldsymbol{c}_n$ is replaced by the MMSE filter for estimating $d_{n-1}(i)$ from $\boldsymbol{r}_{n-1}(i)$.
2) It is shown in [14] that

$$\boldsymbol{c}_{i+1} = \frac{(\boldsymbol{I} - \boldsymbol{c}_i \boldsymbol{c}_i^\dagger)\boldsymbol{R}_{i-1}\boldsymbol{c}_i}{\|(\boldsymbol{I} - \boldsymbol{c}_i \boldsymbol{c}_i^\dagger)\boldsymbol{R}_{i-1}\boldsymbol{c}_i\|} \tag{27}$$

where

$$\boldsymbol{R}_{i+1} = (\boldsymbol{I} - \boldsymbol{c}_{i+1}\boldsymbol{c}_{i+1}^\dagger)\boldsymbol{R}_i(\boldsymbol{I} - \boldsymbol{c}_{i+1}\boldsymbol{c}_{i+1}^\dagger) \tag{28}$$

for $i = 0, 1, 2, \ldots, D-1$, where $\boldsymbol{c}_1 = \boldsymbol{s}_1$ and $\boldsymbol{R}_0 = \boldsymbol{R}$. The following induction argument establishes that $\boldsymbol{c}_i$ is orthogonal to $\boldsymbol{c}_j$ for all $j \neq i$. First, it is easily verified from (27) that $\boldsymbol{c}_2$ is orthogonal to $\boldsymbol{c}_1$. Assume, then, that $\boldsymbol{c}_l$ is orthogonal to $\boldsymbol{c}_m$ for $l, m \le i, l \neq m$. We can rewrite (28) and (27) as

$$\boldsymbol{R}_i = \left(\boldsymbol{I} - \sum_{l=1}^{i} \boldsymbol{c}_l \boldsymbol{c}_l^\dagger\right) \boldsymbol{R} \left(\boldsymbol{I} - \sum_{l=1}^{i} \boldsymbol{c}_l \boldsymbol{c}_l^\dagger\right) \tag{29}$$

and

$$\begin{aligned} \boldsymbol{c}_{i+1} &= \kappa_{i+1}\left(\boldsymbol{I} - \sum_{l=1}^{i} \boldsymbol{c}_l \boldsymbol{c}_l^\dagger\right) \boldsymbol{R} \left(\boldsymbol{I} - \sum_{l=1}^{i-1} \boldsymbol{c}_l \boldsymbol{c}_l^\dagger\right)\boldsymbol{c}_i \\ &= \kappa_{i+1}\left(\boldsymbol{I} - \sum_{l=1}^{i} \boldsymbol{c}_l \boldsymbol{c}_l^\dagger\right) \boldsymbol{R}\boldsymbol{c}_i \\ &= \kappa_{i+1}\left(\boldsymbol{I} - \sum_{l=1}^{i} \boldsymbol{c}_l \boldsymbol{c}_l^\dagger\right) \boldsymbol{R}_I\boldsymbol{c}_i \end{aligned} \tag{30}$$

where κ_{i+1} is a normalization constant, and the last equality holds since $\boldsymbol{R} = \boldsymbol{R}_I + P_1\boldsymbol{s}_1\boldsymbol{s}_1^\dagger$ and $\boldsymbol{c}_1 = \boldsymbol{s}_1$ is orthogonal to c_i. It can then be verified from (30) that c_{i+1} is orthogonal to $c_1, \ldots, c_i$, or, equivalently, that c_l is orthogonal to c_m for $l, m \le i+1, l \ne m$, which establishes the induction step. The relations (29) and (30) will be useful in what follows.

From Fig. 1 it is easily seen that the matrix of basis vectors for the MSWF is given by

$$\begin{aligned} \boldsymbol{M}_D &= \left[\boldsymbol{c}_1\ \boldsymbol{B}_1\boldsymbol{c}_2\ \boldsymbol{B}_1\boldsymbol{B}_2\boldsymbol{c}_3\ \cdots\ \prod_{n=1}^{D-1}\boldsymbol{B}_n\boldsymbol{c}_D\right] \\ &= [\boldsymbol{c}_1\ \boldsymbol{c}_2\ \cdots\ \boldsymbol{c}_D] \end{aligned} \tag{31}$$

where the last equality is due to the fact that the c_i's are orthogonal. (This implies that the blocking matrices in Fig. 1 can be replaced by the identity matrix without affecting the variables $d_n(i)$ and $\epsilon_n(i), n = 0, \ldots, D$.) An alternate set of nonorthogonal basis vectors is given in the next section.

3) It is easily shown that each c_m is contained in the signal subspace, hence K stages are needed to form the full-rank filter.

4) It is shown in [14] that

$$E[d_n^*(i)d_{n+m}(i)] = \boldsymbol{c}_n^\dagger \boldsymbol{R}\boldsymbol{c}_{n+m} = 0 \tag{32}$$

for $|m| > 1$ and $0 \le n, n+m \le D-1$. It follows that $\boldsymbol{M}_D^\dagger \boldsymbol{R}\boldsymbol{M}_D$ is *tridiagonal*.

5) The blocking matrix $\boldsymbol{B}_m$ is not unique. (In [14], $\boldsymbol{B}_m$ is assumed to be an $[N-(m-1)] \times (N-m)$ matrix, so that c_n is $[N-(n-1)] \times 1$.) Although any rank $N-m$ matrix that satisfies (18) achieves the same performance (MMSE), this choice can affect the performance for a specific data record. In particular, a poor choice of blocking matrix can lead to numerical instability.

6) Computation of the MMSE filter coefficients does not require an estimate of the signal subspace, as do the eigendecomposition techniques to be described. Successive filters are determined by "residual correlations" of signals in the preceding stage. Adaptive algorithms based on this technique are presented in [9].

B. Eigendecomposition Methods

The reduced-rank technique which has probably received the most attention is "Principal Components" (PC), which is based on the following eigendecomposition of the covariance matrix

$$\boldsymbol{R} = \boldsymbol{V}\Lambda\boldsymbol{V}^\dagger \tag{33}$$

where $\boldsymbol{V}$ is the orthonormal matrix of eigenvectors of $\boldsymbol{R}$ and Λ is the diagonal matrix of eigenvalues. Suppose that the eigenvalues are ordered as $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_N$. For a given subspace dimension D, the projection matrix for PC is $\boldsymbol{M}_D = \boldsymbol{V}_{1:D}$, the first D columns of $\boldsymbol{V}$.

For $K < N$, the eigenvalues $\lambda_1, \ldots, \lambda_K$ are associated with the signal subspace, and the remaining eigenvalues are associated with the noise subspace, i.e., $\lambda_m = \sigma^2$ for $K < m \le N$. Consequently, by selecting $D \ge K$, PC retains full-rank MMSE performance (e.g., see [8], [26]). However, the performance can degrade quite rapidly as D decreases below K, since there is no guarantee that the associated subspace will retain most of the desired signal energy. This is especially troublesome in a near–far scenario, since for small D, the subspace which contains most of the energy will likely correspond to the interference, and not the desired signal. We remark that in a heavily loaded cellular system, the dimension of the signal subspace may be near, or even exceed the number of dimensions available, in which case PC does not offer much of an advantage relative to conventional full-rank adaptive techniques.

An alternative to PC is to choose a set of D eigenvectors for the projection matrix which minimizes the MSE. Specifically, we can rewrite the MSE (6) in terms of reduced-rank variables as

$$\mathcal{M} = 1 - \|\Lambda^{-1}\tilde{\boldsymbol{s}}_1\|^2. \tag{34}$$

The subspace that minimizes the MSE has basis vectors which are the eigenvectors of $\boldsymbol{R}$ associated with the D largest values of $|\tilde{s}_{1,k}|^2/\lambda_k$, where $\tilde{s}_{1,k}$ is the kth component of $\tilde{\boldsymbol{s}}_1$, and is given by $\boldsymbol{v}_k^\dagger \boldsymbol{s}_1$, where $\boldsymbol{v}_k$ is the kth column of $\boldsymbol{V}$. (Note the inverse weighting of λ_k in contrast with PC.)

This technique, called "Cross-Spectral" (CS) reduced-rank filtering, was proposed in [19] and [20]. This technique can perform well for $D < K$ since it takes into account the energy in the subspace contributed by the desired user. Unlike PC, the projection subspace for CS requires knowledge of the desired user's spreading code $\boldsymbol{s}_1$. A disadvantage of eigendecomposition techniques in general is the complexity associated with estimation of the signal subspace.

C. Partial Despreading (PD)

In this method, proposed for DS-CDMA in [21], the received signal is *partially despread* over consecutive segments of m chips, where m is a parameter. The partially despread vector has dimension $D = \lceil N/m \rceil$, and is the input to the D-tap filter. Consequently, $m = 1$ corresponds to the full-rank MMSE filter, and $m = N$ corresponds to the matched filter. The columns of $\boldsymbol{M}_D$ in this case are nonoverlapping segments of $\boldsymbol{s}_1$, the signature for user 1, where each segment is of length m.

Specifically, if $N/m = D$, the jth column of $\boldsymbol{M}_D$ is

$$[\boldsymbol{M}_D]_j' = [0 \cdots 0 \boldsymbol{s}_{1,[(j-1)m+1:jm]}' 0 \cdots 0] \tag{35}$$

where $1 \le j \le D$, prime ($'$) denotes transpose, and there are $(j-1)m$ zeros on the left and $(D-j)m$ zeros on the right. This is a simple reduced-rank technique that allows the selection of MSE performance between the matched and full-rank MMSE filters by adjusting the number of adaptive filter coefficients.

IV. Large System Analysis

Our main results, presented in the next section, are motivated by the large system results for synchronous CDMA with random signature sequences presented in [15]–[17]. Specifically, we evaluate the large system limit of the output SINR for the reduced-rank filters described in the preceding section when the signatures are chosen randomly. This limit is defined by letting the number of dimensions N and number of users K tend to infinity with $K/N = \alpha$ held constant.

The large system results presented in [15]–[17], as well as some of the results presented here, make use of the limiting eigenvalue distribution of a class of random matrices. Let C_{ij} be an infinite matrix of independent and identically distributed (i.i.d.) complex-valued random variables with variance 1, and P_i be a sequence of real-valued random variables (corresponding to user powers). Let $\boldsymbol{S}$ be an $N \times K$ matrix, whose (i, j)th entry is $\frac{C_{ij}}{\sqrt{N}}$. Let $\boldsymbol{P}$ be a $K \times K$ diagonal matrix with diagonal entries $P_1, \ldots, P_K$. As $K \to \infty$, we assume that the empirical distribution function of these entries converges almost surely in distribution to a deterministic limit $F(\cdot)$.

Let $G_N(\lambda)$ denote the empirical distribution function of the eigenvalues of the Hermitian matrix $\boldsymbol{SPS}^\dagger$. It is shown in [18] that as $N, K \to \infty$, and for $\frac{K}{N} \to \alpha > 0$, G_N converges almost surely to a deterministic limit G. Let $m_G(z)$ denote the Stieltjes transform of the limit distribution G

$$m_G(z) = \int \frac{1}{\lambda - z}\, dG(\lambda) \tag{36}$$

for $z \in \mathcal{C}^+ \equiv \{z \in \mathcal{C}: \Im(z) > 0\}$. It is shown in [18] that

$$m_G(z) = \frac{1}{-z + \alpha \int \frac{\tau dF(\tau)}{1+\tau m_G(z)}} \tag{37}$$

for all $z \in \mathcal{C}^+$. In what follows, we will denote the range of λ for which $G(\lambda)$ is nonzero as $[a(\alpha), b(\alpha)]$.

For an arbitrary distribution F, closed-form expressions for G, $a(\alpha)$, and $b(\alpha)$ do not exist. However, a closed-form expression for G is given in [27] for the case where $F(x) = 1, x > P$, and $F(x) = 0, x < P$ (i.e., all users are received with the same power). This will be used in Section VII to derive some of our results.

The preceding result was used in [17] to derive the large system limit of the output SINR for the linear MMSE filter in a synchronous CDMA system. Specifically, let P_k denote the received power of user k, and P denote the received power of a random user, which has the limit distribution $F(P)$. Let user 1 be the user of interest. It is shown in [17] that as $K = \alpha N \to \infty$, the (random) output SINR of the linear MMSE receiver for user 1 converges in probability to the deterministic limit

$$\beta^\infty = \frac{P_1}{\sigma^2 + \alpha \int_0^\infty I(P, P_1, \beta^\infty)\, dF(P)} \tag{38}$$

where

$$I(P, P_1, \beta^\infty) = \frac{P_1 P}{P_1 + P\beta^\infty} \tag{39}$$

is the "effective interference" associated with an interferer received with power P.

For the case where all users are received with the same power, (38) becomes

$$\beta^\infty = \frac{P}{\sigma^2 + \frac{\alpha P}{1+\beta^\infty}} \tag{40}$$

which yields a closed-form solution for β^∞ [17]. It will be convenient to denote this solution as

$$\beta^\infty = \mathcal{B}_U\left(\alpha, \frac{P}{\sigma^2}\right). \tag{41}$$

Similary, we will denote the solution to (38) for an arbitrary power distribution as

$$\beta^\infty = \mathcal{B}_F[\alpha, P_1, \sigma^2, F(\cdot)]. \tag{42}$$

Finally, we remark that the analogous large system limit for the matched filter is

$$\beta^{\text{MF}} = \frac{P_1}{\sigma^2 + \alpha \int_0^\infty P\, dF(P)}. \tag{43}$$

V. Main Results

In this section, we present the large system limits of output SINR for the reduced-rank filters presented in Section III. Proofs and derivations are given in Section VII. For finite K and N, the output SINR is a random variable due to the assignment of random signature sequences. For the MSWF and PD, we are able to show, in analogy with the full-rank MMSE receiver, that the output SINR converges to a deterministic limit as $K = \alpha N \to \infty$. We conjecture that this is also true for the PC and CS methods.

A. Multistage Wiener Filter (MSWF)

We first state the large system SINR for the MSWF assuming that all users are received with the same power.

Theorem 1: As $K = \alpha N \to \infty$, the output SINR of the rank D MSWF converges in probability to the limit β_D^{MS}, which satisfies

$$\beta_{D+1}^{MS} = \frac{P}{\sigma^2 + \alpha \frac{P}{1+\beta_D^{MS}}} \qquad \text{for } D \geq 0 \tag{44}$$

where P is the received power for each user, $\beta_0^{MS} = 0$, and $\beta_1^{MS} = P/(\sigma^2 + \alpha P)$ is the large system limit of the output SINR for the matched filter.

The proof is given in Section VII-A.

According to this theorem, for finite D the output SINR of the MSWF can be expressed as a *continued fraction*. For example,

$$\beta_2^{MS} = \frac{P}{\sigma^2 + \alpha \frac{P}{1+\frac{P}{\sigma^2+\alpha P}}}. \tag{45}$$

As D increases, this continued fraction converges to the full-rank MMSE given by (40). Two important consequences of this result are as follows.

1) The dimension D needed to achieve a target SINR within some small ϵ of the full-rank SINR does not scale with the system size (K and N). This is in contrast with the other techniques considered, for which the large system output SINR is determined by the ratio D/N.

2) As D increases, β_D^{MS} converges rapidly to the full-rank MMSE. Specifically, consider the case without background noise, $\sigma^2 = 0$. It can be shown that

$$\beta_D^{MS} = \sum_{n=1}^{D} \alpha^{-n} = \begin{cases} \frac{\alpha^{-D}-1}{1-\alpha}, & \text{for } \alpha < 1 \\ D, & \text{for } \alpha = 1 \\ \frac{1-\alpha^{-D+1}}{1-\alpha}, & \text{for } \alpha > 1. \end{cases} \tag{46}$$

In particular, β_D^{MS} increases exponentially with D for $\alpha < 1$ and linearly with D for $\alpha = 1$. If $\alpha > 1$, then the gap between β_i^{MS} and the full-rank performance $\beta^\infty = \frac{1}{\alpha-1}$ decreases exponentially. Numerical results to be presented in the next section indicate that for signal-to-noise ratios (SNRs) (P/σ^2) and loads (K/N) of interest, the full-rank MMSE performance is essentially achieved with $D = 8$.

We now consider the MSWF with an arbitrary power distribution. In this case, we do not have a closed-form result for the large system SINR, although we can compute it numerically. In analogy with the uniform power case, we also have the *approximation*

$$\beta_{D+1}^{MS} \approx \frac{P_1}{\sigma^2 + \alpha\int_0^\infty \frac{P_1 P}{P_1 + P\beta_D^{MS}}\, dF(P)} \tag{47}$$

where $\beta_0^{MS} = 0$, β_1^{MS} is the asymptotic SINR of the matched filter given by (43), and $F(P)$ is an arbitrary power distribution. This approximation is accurate for many cases we have considered; however, in Appendix A we show that it is not exact.

To compute the output SINR for the MSWF with an arbitrary power distribution, we first give an alternate representation of the subspace spanned by the basis vectors, or columns of $\boldsymbol{M}_D$. Let $\mathcal{S}_D$ denote the D-dimensional subspace associated with the rank D MSWF, which is spanned by the set of basis vectors given by (31), and let $\boldsymbol{R}_I$ denote the interference-plus-noise covariance matrix given by (9).

Theorem 2: The subspace $\mathcal{S}_D$ is spanned by the vectors $\boldsymbol{y}_0, \ldots, \boldsymbol{y}_{D-1}$ where $\boldsymbol{y}_n = \boldsymbol{R}_I^n \boldsymbol{s}_1$.

The proof is given in Section VII-B.

The matrix of basis vectors for the MSWF can, therefore, be written as

$$\boldsymbol{M}_D = [\boldsymbol{s}_1\ \boldsymbol{R}_I\boldsymbol{s}_1\ \boldsymbol{R}_I^2\boldsymbol{s}_1\ \cdots\ \boldsymbol{R}_I^{D-1}\boldsymbol{s}_1]. \tag{48}$$

It is straightforward to show that $\boldsymbol{R}_I$ can be replaced by $\boldsymbol{R} = E[\boldsymbol{r}(i)\boldsymbol{r}^\dagger(i)]$ in Theorem 2. Approximations of the full-rank MMSE filter in terms of powers of the covariance matrix have also been considered in [28] and [29].

Let

$$\gamma_m = \boldsymbol{s}_1^\dagger \boldsymbol{y}_m = \boldsymbol{s}_1^\dagger \boldsymbol{R}_I^m \boldsymbol{s}_1 \tag{49}$$

$$\gamma_{l:m} = [\gamma_l\, \gamma_{l+1} \cdots \gamma_m]' \tag{50}$$

$$\Gamma_{l:l+m} = [\gamma_{l:l+m}\, \gamma_{l+1:l+m+1} \cdots \gamma_{l+m:l+2m}]. \tag{51}$$

Note that $\Gamma_{l:l+m}$ is an $(m+1)\times(m+1)$ matrix. From (13)–(15), the reduced-rank MSWF is

$$\tilde{\boldsymbol{c}}_D = (\boldsymbol{M}_D^\dagger \boldsymbol{M}_D)(\boldsymbol{M}_D^\dagger \boldsymbol{R} \boldsymbol{M}_D)^{-1}\boldsymbol{M}_D^\dagger \boldsymbol{s}_1 \tag{52}$$

$$= \Gamma_{0:D-1}(\Gamma_{1:D} + P_1\gamma_{0:D-1}\gamma_{0:D-1}^\dagger)^{-1}\gamma_{0:D-1} \tag{53}$$

$$= \frac{1}{1+\beta_D}(\Gamma_{0:D-1}\Gamma_{1:D}^{-1})\gamma_{0:D-1} \tag{54}$$

$$= \frac{1}{1+\beta_D}\begin{bmatrix} \beta_D/P_1 \\ --- \\ \gamma_{0:D-1} \end{bmatrix} \tag{55}$$

where

$$\beta_D = P_1 \gamma_{0:D-1}^\dagger \Gamma_{1:D}^{-1} \gamma_{0:D-1} \tag{56}$$

is the output SINR from (16). To compute the large system limit, we, therefore, need to compute the large system limit of γ_n, which is given by the following lemma, where G is the asymptotic eigenvalue distribution of $\boldsymbol{S}_I \boldsymbol{P} \boldsymbol{S}_I^\dagger$.

Lemma 1: As $K = \alpha N \to \infty$, γ_n converges in probability to the limit

$$\gamma_n^\infty(\alpha, \sigma^2) = \int (\lambda + \sigma^2)^n\, dG(\lambda) \tag{57}$$

provided that this moment is finite.

Proof: From (49) we have

$$\gamma_n = \sum_{k=1}^N (\lambda_k + \sigma^2)^n |\boldsymbol{s}_1^\dagger \boldsymbol{v}_{k;I}|^2 \tag{58}$$

where $\boldsymbol{v}_{k;I}$ is the kth eigenvector of $\boldsymbol{R}_I$, and the sum can be restricted to $\boldsymbol{v}_{k;I}$ in the signal subspace since otherwise $\boldsymbol{s}_1^\dagger \boldsymbol{v}_{k;I} = 0$. It is shown in [17] that $|\boldsymbol{s}_1^\dagger \boldsymbol{v}_{k;I}|^2$ is $O(1/N)$, and the Lemma follows from the same argument used to prove [17, Lemma 4.3].

When all users have the same power P and $\sigma = 0$, the limit can be evaluated explicitly as [27]

$$\gamma_n^\infty(\alpha, 0) = P^n \sum_{k=0}^{n-1} \frac{1}{k+1}\binom{n}{k}\binom{n-1}{k}\alpha^{k+1}, \qquad n = 1, 2, \ldots. \tag{59}$$

To compute $\gamma_n^\infty(\alpha, \sigma)$ for $\sigma > 0$, we observe that

$$\frac{d\gamma_n^\infty}{d(\sigma^2)} = n\int (\lambda + \sigma^2)^{n-1}\, dG(\lambda) = n\gamma_{n-1}^\infty \tag{60}$$

which leads to a recursive method for computing the sequence $\{\gamma_n^\infty\}$. For a nonuniform power distribution, the large system limit γ_n^∞ can be computed directly from (57). In Appendix A, we give two other methods for computing γ_n, one of which does not make explicit use of the asymptotic eigenvalue distribution G.

We can now state the large system SINR for the rank D MSWF where the limiting power distribution is $F(\cdot)$. The superscript ∞ indicates the large system limit of the associated variable(s).

Theorem 3: As $K = \alpha N \to \infty$, the output SINR of the rank D MSWF converges in probability to

$$\beta_D^{MS} = P_1(\gamma_{0:D-1}^\infty)^\dagger (\Gamma_{1:D}^\infty)^{-1}\gamma_{0:D-1}^\infty. \tag{61}$$

Proof: This follows directly from Lemma 1 and the fact that β_D, given by (56), is a continuous and bounded function of $\gamma_1, \ldots, \gamma_{2D-1}$.

As for the uniform power case, the dimension D needed to achieve a target SINR within some small constant ϵ of the full-rank SINR does not scale with K and N. Because this representation for the output SINR is not as transparent as that for the uniform power case, it is difficult to see how fast the SINR given by (61) converges to the full-rank value as D increases. Numerical examples are presented in Section VI, and indicate that, as for the uniform power case, full-rank performance is achieved for $D < 10$.

For large D, Theorem 3 gives an alternative method to (38) for computing the full-rank MMSE. This method does not require knowledge of the asymptotic eigenvalue distribution if the second method for computing the moments $\{\gamma_n^\infty\}$ presented in Appendix B is used.

Finally, we remark that for a uniform power distribution, the SINR shown in Theorem 3 must be the same as the continued fraction representation in Theorem 1. Finding a direct proof of this equivalence appears to be an open problem.

B. Eigendecomposition Methods

We now state our results for the PC and CS reduced-rank methods presented in Section III-B. In what follows, the large system limit is defined as $D, K, N \to \infty$ where $K/N = \alpha$ and $D/N = \delta$. In particular, D now increases proportionally with K and N. As stated earlier, we conjecture that the SINR converges to a deterministic large system limit. The technical difficulty in proving this is characterization of the large system limit of $|\boldsymbol{v}_n^\dagger \boldsymbol{s}_1|^2$ where $\boldsymbol{v}_n$ is the nth eigenvector of the covariance matrix $\boldsymbol{R}$ corresponding to the particular ordering given in Section III-B. Note, in particular, that $\boldsymbol{v}_n$ and $\boldsymbol{s}_1$ are correlated.

In order to proceed, we assume that the conjecture is true, and evaluate the corresponding large system limit. The numerical results in Section VI show that the large system results are nearly identical to the corresponding simulation results. For PC, we are only able to evaluate the large system limit for the uniform (equal) power case. In what follows, G is the limit distribution for the eigenvalues of $\boldsymbol{SS}^\dagger$ (see Section VII-C).

For both the PC and CS methods, the output SINR for a rank D filter can be written as

$$\beta_D = \frac{v_D}{1 - v_D} \tag{62}$$

where

$$v_D = P \sum_{n=1}^{D} \frac{|\boldsymbol{v}_n^\dagger \boldsymbol{s}_1|^2}{\lambda_n + \sigma^2}, \tag{63}$$

and where λ_n and $\boldsymbol{v}_n$ are the nth eigenvalue and eigenvector, respectively, of the covariance matrix $\boldsymbol{R}$. We first consider PC, for which $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_N$. We show in Section VII-C that

$$\begin{aligned}
&\lim_{\substack{K=\alpha N \to \infty \\ D=\delta N \to \infty}} E(v_D) \\
&= v^{\rm PC}(\alpha, \delta) = \frac{1}{\alpha} \int_c^{b(\alpha)} \frac{\lambda}{\lambda + \sigma^2} \, dG(\lambda) \\
&= -\frac{1}{2\pi\alpha} \sqrt{[c - a(\alpha)][b(\alpha) - c]} + \frac{a(\alpha) + b(\alpha) + 2\sigma^2}{4\pi\alpha} \\
&\quad \cdot \left[\frac{\pi}{2} - \arcsin \frac{2c - a(\alpha) - b(\alpha)}{b(\alpha) - a(\alpha)} \right] \\
&\quad - \frac{1}{2\pi} \sqrt{[a(\alpha) + \sigma^2][b(\alpha) + \sigma^2]} \\
&\quad \cdot \left[\frac{\pi}{2} - \arcsin \frac{(c - \sigma^2)[a(\alpha) + b(\alpha)] - 2a(\alpha)b(\alpha) + 2\sigma^2 c}{(c + \sigma^2)[b(\alpha) - a(\alpha)]} \right]
\end{aligned} \tag{64}$$

where

$$a(\alpha) = \left(1 - \sqrt{1/\alpha}\right)^2 P \qquad b(\alpha) = \left(1 + \sqrt{1/\alpha}\right)^2 P \tag{65}$$

and c is defined by

$$\begin{aligned}
G(c) &= \sqrt{[c - a(\alpha)][b(\alpha) - c]} + \frac{a(\alpha) + b(\alpha)}{2} \\
&\quad \cdot \left[\arcsin \frac{2c - a(\alpha) - b(\alpha)}{b(\alpha) - a(\alpha)} + \frac{\pi}{2} \right] \\
&\quad + \sqrt{a(\alpha) b(\alpha)} \\
&\quad \cdot \left[\arcsin \frac{[a(\alpha) + b(\alpha)]c - 2a(\alpha)b(\alpha)}{c[b(\alpha) - a(\alpha)]} + \frac{\pi}{2} \right] \\
&\quad + (1 - \alpha) \mathbf{1}\{\alpha < 1\} \\
&= 1 - \delta.
\end{aligned} \tag{66}$$

If v_D in (63) converges to a deterministic large system limit, then this limit must be $v^{\rm PC}(\alpha, \delta)$. The large system limit for output SINR is then

$$\beta^{\rm PC} = \frac{v^{\rm PC}(\alpha, \delta)}{1 - v^{\rm PC}(\alpha, \delta)}. \tag{67}$$

Numerical examples are presented in the next section, and show that, as expected, when $\delta = \alpha$, the PC algorithm achieves full-rank performance. As δ decreases below α, the performance degrades substantially.

Although we do not have an analogous result for an arbitrary power distribution, we observe that for $\delta \ge \alpha$, the PC algorithm again achieves full-rank performance, since the eigenvectors chosen for the projection include the signal subspace. As δ decreases below α, in a near–far situation the performance can be substantially worse than that with equal received powers. For example, assume that there are two groups of users where users in each group have the same power, but users in the first group transmit with much more power than users in the second group. (The groups may correspond to different services, such as voice and data.) In this situation, the eigenvalues corresponding to the signal subspace can also be roughly divided into two sets corresponding to the two user groups. If the desired user belongs to the second group, then its energy is mostly contained in the subspace spanned by eigenvectors associated with the small eigenvalues. Consequently, PC will choose a subspace which contains little energy from the desired user, resulting in poor performance.

The CS method, described in Section III-B, performs better than the PC method for $\delta < \alpha$ since it accounts for the projection of the desired user spreading sequence onto the selected subspace. The output SINR is again given by (62) and (63) where the ordering of eigenvalues and eigenvectors corresponds to decreasing values of $|\boldsymbol{v}_n^\dagger \boldsymbol{s}_1|^2 / (\lambda_n + \sigma^2)$. Let

$$\xi_n = \sqrt{K} \frac{\boldsymbol{v}_n^\dagger \boldsymbol{s}_1}{\sqrt{\lambda_n + \sigma^2}} \tag{68}$$

so that

$$v_D = \frac{1}{K} \sum_{n=1}^{D} |\xi_n|^2. \tag{69}$$

As $K = \alpha N \to \infty$, numerical results indicate that the sequence $\{\xi_n\}$ converges to a deterministic distribution $H(\xi)$. Assuming this is true, it follows that as $K = \alpha N \to \infty$ and $D = \delta N \to \infty$

$$v_D \to v^{\rm CS}(\alpha, \delta) = 2 \int_c^{\infty} |\xi|^2 \, dH(\xi) \tag{70}$$

where c satisfies $H(c) = 1 - \frac{\delta}{2}$.

In what follows, we assume that $H(\cdot)$ is zero-mean Gaussian. Justification for this assumption stems from the analysis in [30], where it is shown that if $\boldsymbol{v}_{n;I}$ is a randomly chosen eigenvector of the interference-plus-noise covariance matrix, then $\boldsymbol{v}_{n;I}^\dagger \boldsymbol{s}_1$ is zero-mean Gaussian. (In that case, $\boldsymbol{v}_{n;I}$ and $\boldsymbol{s}_1$ are independent.)

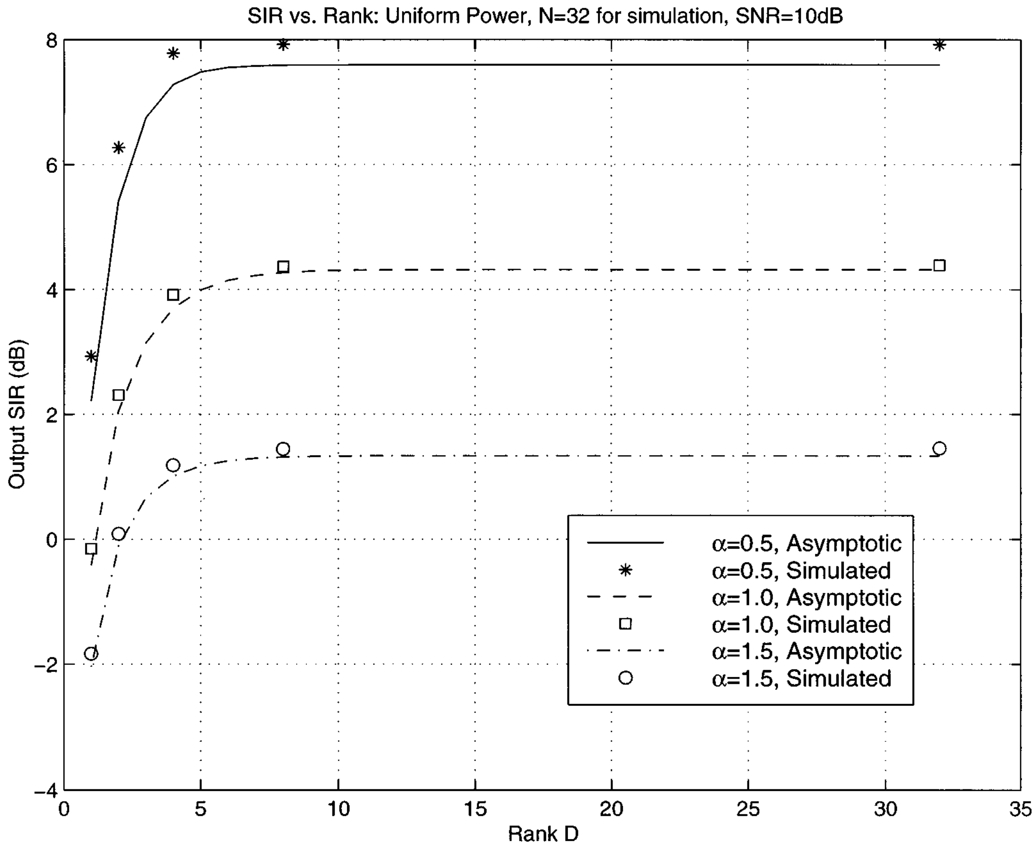

Fig. 2. Output SINR versus rank D for the MSWF with different loads α.

It is shown in Section VII-C that $\sqrt{K}\boldsymbol{v}_n^\dagger \boldsymbol{s}_1$ has variance $E[\lambda_n]$. (Note that λ_n converges to a deterministic large system limit for fixed n/K.) Consequently, at high SNRs ($\sigma^2 \ll \lambda_n$ for $n < D$), $E[|\xi_n|^2] \approx 1$, independent of n. Further justification for this assumption is the close agreement between the large system analytical and simulation results shown in the next section.

With the Gaussian assumption for $H(\cdot)$

$$v^{\mathrm{CS}}(\alpha, \delta) = 2\int_c^\infty x^2 \frac{1}{\sqrt{2\pi\sigma_\xi^2}}\, e^{-x^2/(2\sigma_\xi^2)}\, dx \tag{71}$$

where c satisfies

$$Q(c/\sigma_\xi) = \int_{c/\sigma_\xi}^\infty 1/\sqrt{2\pi} e^{-x^2/2}\, dx = 1 - \frac{\delta}{2}$$

and

$$\sigma_\xi^2 = \int_{-\infty}^\infty x^2\, dH(x) \tag{72}$$

which is the large system limit of $E[|\xi_n|^2]$ where n is chosen randomly according to a uniform distribution between one and K. In Section VII-D, it is shown that

$$\sigma_\xi^2 = \frac{\beta^\infty}{1+\beta^\infty} \tag{73}$$

where β^∞ is the full-rank SINR. When all users have the same power P

$$\sigma_\xi^2 = \frac{1}{2\sqrt{\alpha}}\left(\sqrt{a(\alpha)+\sigma^2} - \sqrt{b(\alpha)+\sigma^2}\right)^2 \tag{74}$$

and $Q(c/\sigma_\xi) = 1 - \delta/2$.

If v_D converges to a deterministic large system limit, then it must converge to $v^{\mathrm{CS}}(\alpha, \delta)$ in which case β_D converges in probability to the corresponding limit

$$\beta^{\mathrm{CS}} = \frac{v^{\mathrm{CS}}(\alpha, \delta)}{1 - v^{\mathrm{CS}}(\alpha, \delta)}. \tag{75}$$

The numerical results shown in the next section are generated according to these assumptions.

C. Partial Despreading (PD)

The output SINR for PD can be expressed in terms of the full-rank MMSE expressions $\mathcal{B}_U$ and $\mathcal{B}_F$ given by (41) and (42). The large system limit is obtained by despreading over $M \geq 1$ chips, where M is held constant, so that $N/D = M = 1/\delta$.

Theorem 4: Assume that the elements of $\boldsymbol{S}$ are i.i.d., zero-mean, and are selected from either a binary or Gaussian distribution. As $K = \alpha N \to \infty$ and $D = N/M \to \infty$, the output SINR of the PD MMSE filter converges in probability to the limit

$$\beta^{\mathrm{PD}}(\alpha, M) = \mathcal{B}_F(M\alpha, MP_1, M\sigma^2, F(\cdot)) \tag{76}$$

where $M = 1/\delta$.

The proof is given in Section VII-E.

If $\delta = 0$ ($M \to \infty$), then β^{PD} is the large system limit for the matched-filter output SINR given by (43). The large system limit of the output SINR for the MMSE PD filter with a uniform power distribution is

$$\beta^{\mathrm{PD}}(\alpha, \delta) = M\mathcal{B}_U\left[M\alpha, \frac{P}{M\sigma^2}\right]. \tag{77}$$

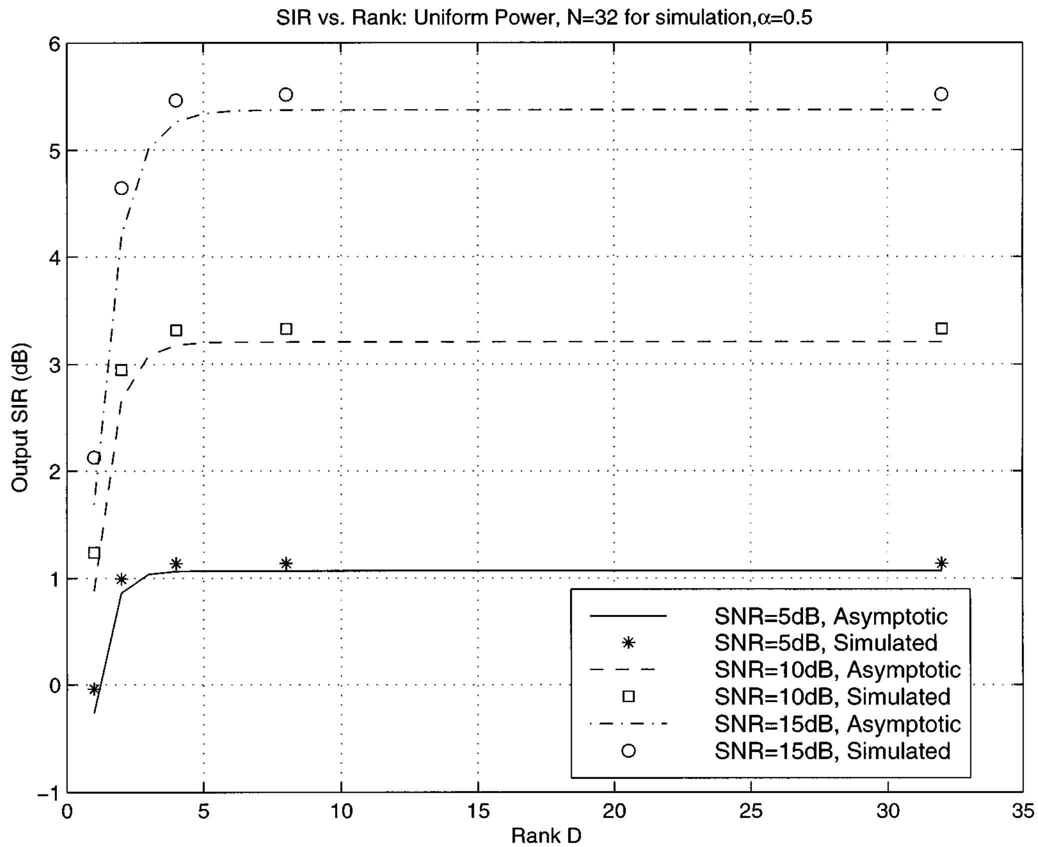

Fig. 3. Output SINR versus rank D for the MSWF with different SNRs.

VI. Numerical Results

In this section, we present numerical results, which illustrate the performance of the reduced-rank techniques considered. Simulation results for a CDMA system with finite N and random binary signature sequences are included for comparison with the large system limit. The latter results are averaged over random binary signature sequences and the received power distribution.

Fig. 2 shows plots of output SINR versus rank D for the MSWF with different loads α, assuming the background SNR is $1/\sigma^2 = 10$ dB and that all users are received with equal power. Also included are simulation results corresponding to $N = 32$. Fig. 3 shows the analogous results for different SNRs and fixed load $\alpha = 1/2$. These results show that the large system limit accurately predicts the simulated values. In all cases shown, the MSWF achieves essentially full-rank performance for $D < 10$. Furthermore, the SINR for $D = 4$ is within 1 dB of the full-rank SINR, and the SINR for $D = 2$ is approximately midway between the SINRs for the matched-filter and full-rank MMSE receivers.

Fig. 4 shows simulated output SINR for the MSWF as a function of normalized rank D/N for $N = 32$, 64, and 128. This illustrates the convergence to the large system limit, which is the full-rank performance for all values of D/N (shown as the solid line in the figure).

Fig. 5 shows output SINR versus normalized rank D/K for the reduced-rank filters considered assuming uniform (equal) power, SNR $= 10$ dB, and $\alpha = 1/2$. For all four methods considered, the large system analysis accurately predicts the simulation results, which are shown for $N = 32$. As discussed previously, the large system SINR for the MSWF as $D \to \infty$ is the full-rank SINR for any $D/N > 0$. (Large system results for the MSWF corresponding to finite D are not shown.) Consequently, there is a large gap between the curve for the MSWF and the curves for the other methods for small D/N. The CS and PC reduced-rank filters can achieve the full-rank performance only when $D \geq K$. For $D < K$, these results show that the CS filter performs much better than the PC filter.

The PD filter can achieve the full-rank performance only when $D = N$, since for any $D < N$, the selected subspace $\mathcal{S}_D$ does not generally contain the MMSE solution. For small D/K, the PD filter performs close to the matched filter, which is significantly better than the eigendecomposition methods. This is because for the latter methods, the desired signal energy is spread over many eigenvectors, so that for small D, relatively little desired signal energy is retained in the selected subspace.

Performance results for nonuniform power distributions are shown in Figs. 6–8. Two distributions are considered: log-normal, and discrete with two powers. In the former case, the desired user has power $P_1 = 1$, and the log-variance of the log-normal distribution is 6 dB. In the latter case, $P^{(2)}/P^{(1)} =$ 10 dB, where $P^{(j)}$ is the power associated with users in group $j = 1, 2$, and the fraction of high-power users is 0.2. The desired user is assumed to be in group one with an SNR of 10 dB. The first case applies to the reverse link of an isolated cell

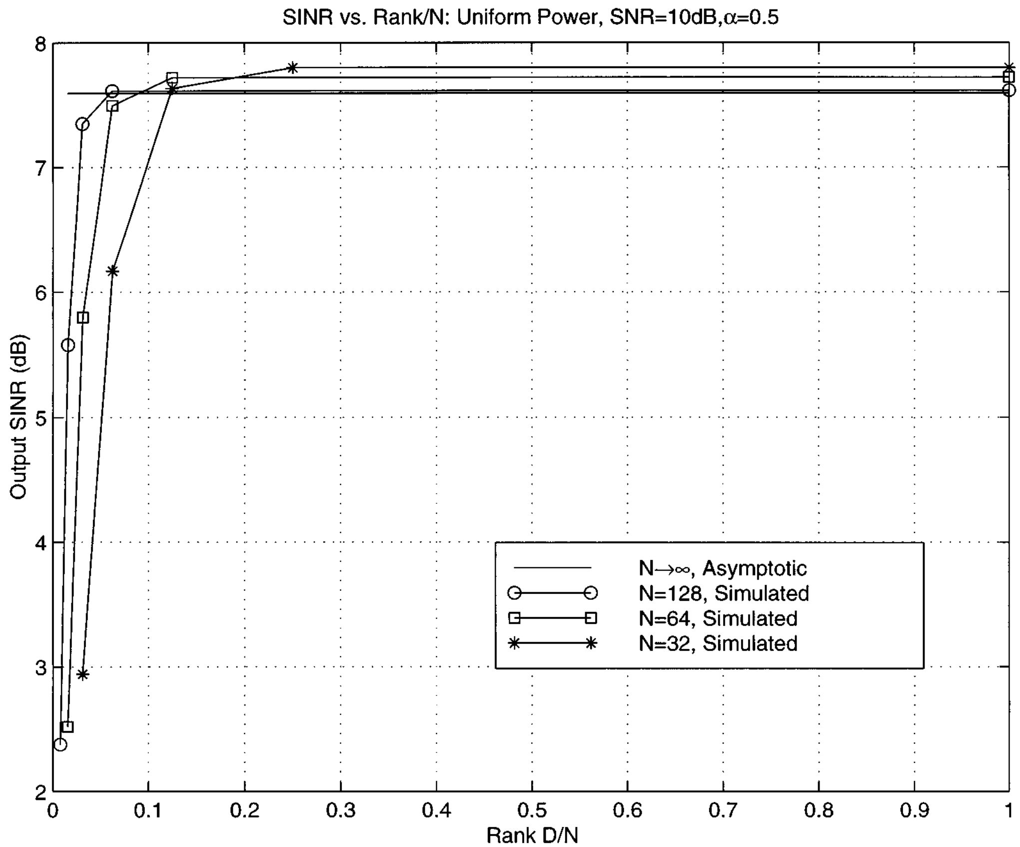

Fig. 4. Output SINR versus normalized rank D/N for the MSWF with different spreading gains.

where power control is used to compensate for shadowing. The second case corresponds to two service categories, such as data and voice, with perfect power control.

Fig. 6 compares the large system output SINR of the MSWF computed via the approximation (47) with the exact SINR computed from (61). Figs. 7 and 8 show output SINR versus normalized rank for the different reduced-rank filters considered. Simulation results are shown for $N = 32$ (and $N = 512$ in Fig. 6). In these figures, $\alpha = 0.5$. Figs. 7 and 8 show that all methods perform approximately the same as for the uniform power case, except for PC, which performs significantly worse for $D < K$.

VII. Proofs and Derivations

A. Theorem 1: MSWF with Uniform Power

The proof is based on an induction argument in which the full-rank MSWF is partitioned into two component filters. The first filter consists of the first $i-1$ stages and the second filter consists of stages i through K (i.e., the full-rank filter which estimates d_{i-1} from $\boldsymbol{r}_{i-1}$). We first consider the case $i = 2$ and prove that: i) the theorem is valid for $D = i = 2$; ii) the large system SINR associated with the second component filter is the full-rank large system SINR β^{∞}; and iii) the large system SINR associated with the filter $\boldsymbol{c}_2$ (with appropriately defined desired signal and interference components) is equal to the large system SINR for the matched filter. For the induction step, we make the analogous assumptions i)–iii) for some i where $2 \le i \le K$, and prove that i)–iii) hold for $i+1$.

The rank one MSWF is the matched filter $\boldsymbol{c}_1 = \boldsymbol{s}_1$, which has output

$$d_1 = \boldsymbol{c}_1^{\dagger}(b_1\boldsymbol{s}_1 + \boldsymbol{S}_I\boldsymbol{b}_I + \boldsymbol{n}) = S_1 + I_1 + N_1 \tag{78}$$

where S_1, I_1, and N_1 denote the corresponding desired signal, interference, and noise terms. The SINR at the output of $\boldsymbol{c}_1$ is

$$\beta_1 = \frac{E[|S_1|^2]}{E[|N_1|^2] + E[|I_1|^2]} \underset{K\to\infty}{\longrightarrow} \frac{P}{\sigma^2 + \alpha P} \tag{79}$$

where $K \to \infty$ denotes the large system limit $(K/N = \alpha)$, the expectation is with respect to the transmitted symbols and noise, and the limit follows from the fact that $|I_1|^2 \to \alpha P$.

Let $\boldsymbol{c}_1^{\perp}$ be a vector which is orthogonal to $\boldsymbol{c}_1$. The output of $\boldsymbol{c}_1^{\perp}$ is

$$d_1^{\perp} = (\boldsymbol{c}_1^{\perp})^{\dagger}\boldsymbol{r} = (\boldsymbol{c}_1^{\perp})^{\dagger}\boldsymbol{S}_I\boldsymbol{b}_I + (\boldsymbol{c}_1^{\perp})^{\dagger}\boldsymbol{n} \tag{80}$$

and it is easily shown that

$$E[(d_1^{\perp})^* S_1] = 0 \tag{81}$$

$$E[(d_1^{\perp})^* N_1] = 0 \tag{82}$$

$$E[(d_1^{\perp})^* I_1] = \boldsymbol{c}_1^{\dagger}(\boldsymbol{S}_I\boldsymbol{S}_I^{\dagger})\boldsymbol{c}_1^{\perp} = \boldsymbol{c}_1^{\dagger}\boldsymbol{R}\boldsymbol{c}_1^{\perp}. \tag{83}$$

We now express $d_1^{\perp}$ as the sum of a desired signal component, interference, and noise

$$d_1^{\perp} = S_1^{\perp} + I_1^{\perp} + N_1^{\perp}. \tag{84}$$

We define the desired signal as

$$S_1^{\perp} = aI_1 \tag{85}$$

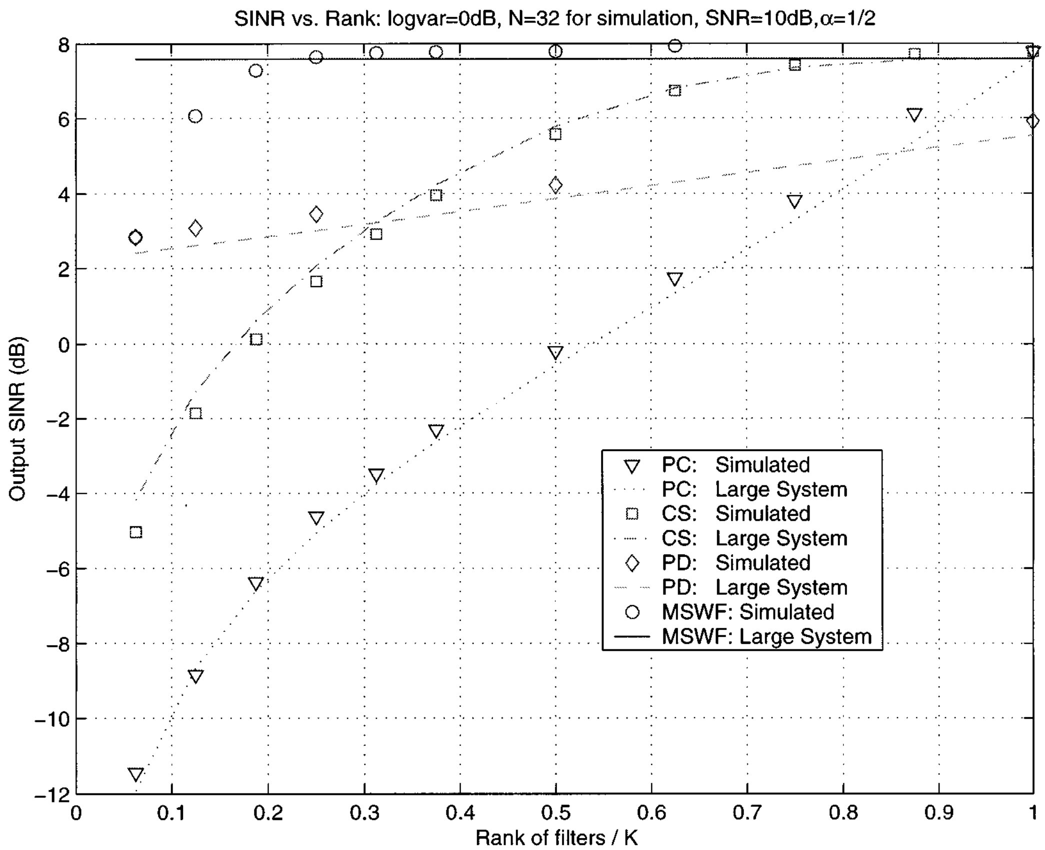

Fig. 5. Output SINR versus normalized rank D/K for reduced-rank filters with equal received powers.

where a minimizes $E[|d_1^\perp - aI_1|^2]$. That is,

$$S_1^\perp = \frac{E[(d_1^\perp)^* I_1]}{E[|I_1|^2]} I_1 = \frac{c_1^\dagger R c_1^\perp}{E[|I_1|^2]} I_1 \tag{86}$$

is the MMSE estimate of $d_1^\perp$ given I_1, so that by the orthogonality principle

$$E[(S_1^\perp)^*(I_1^\perp + N_1^\perp)] = 0. \tag{87}$$

Given these definitions of the signal, interference, and noise, we associate an (output) SINR with the filter $c_1^\perp$, which is given by

$$\beta_1^\perp = \frac{E[|S_1^\perp|^2]}{E[|I_1^\perp + N_1^\perp|^2]} = \frac{E[|S_1^\perp|^2]}{E[|d_1^\perp|^2] - E[|S_1^\perp|^2]} \tag{88}$$

$$= \frac{\frac{|c_1^\dagger R c_1^\perp|^2}{E[|I_1|^2]}}{(c_1^\perp)^\dagger R c_1^\perp - \frac{|c_1^\dagger R c_1^\perp|^2}{E[|I_1|^2]}} \tag{89}$$

where the expectation is again with respect to the transmitted symbols and noise. From (88) and (89) we have that

$$\frac{\beta_1^\perp}{1+\beta_1^\perp} = \frac{E[|S_1^\perp|^2]}{E[|d_1^\perp|^2]} = \frac{1}{E[|I_1|^2]} \frac{|c_1^\dagger R c_1^\perp|^2}{(c_1^\perp)^\dagger R c_1^\perp}. \tag{90}$$

Now consider the filter $c = c_1 + w_2 c_1^\perp$, which has output $d = d_1 + w_2 d_1^\perp$. Since the output contains the desired signal S_1, the SINR associated with c is

$$\beta_c = \frac{P}{E[|d|^2] - P}. \tag{91}$$

Choosing w_2 to maximize the SINR, or equivalently, minimize the output energy $E[|d|^2]$ gives

$$w_2 = -\frac{E[(d_1^\perp)^* d_1]}{E[|d_1^\perp|^2]} = -\frac{E[(d_1^\perp)^* I_1]}{(c_1^\perp)^\dagger R c_1^\perp} \tag{92}$$

$$= -\frac{c_1^\dagger R c_1^\perp}{(c_1^\perp)^\dagger R c_1^\perp} \tag{93}$$

and

$$\begin{aligned} E[|d|^2] &= E[|d_1|^2] - \frac{|E[(d_1^\perp)^* d_1]|^2}{E[|d_1^\perp|^2]} \\ &= c_1^\dagger R c_1 - \frac{|c_1^\dagger R c_1^\perp|^2}{(c_1^\perp)^\dagger R c_1^\perp} \\ &= c_1^\dagger R c_1 - E[|I_1|^2] \frac{\beta_1^\perp}{1+\beta_1^\perp} \\ &= P + \sigma^2 + E[|I_1|^2] \frac{1}{1+\beta_1^\perp} \\ &\xrightarrow[K\to\infty]{} P + \sigma^2 + \alpha P \frac{1}{1+(\beta_1^\perp)^\infty}. \end{aligned} \tag{94}$$

Combining (91)–(94) gives

$$\beta_c = \frac{P}{\sigma^2 + E[|I_1|^2] \frac{1}{1+\beta_1^\perp}} \tag{95}$$

and letting $K = \alpha N \to \infty$ gives

$$\beta_c^\infty = \frac{P}{\sigma^2 + \alpha P \frac{1}{1+(\beta_1^\perp)^\infty}}. \tag{96}$$

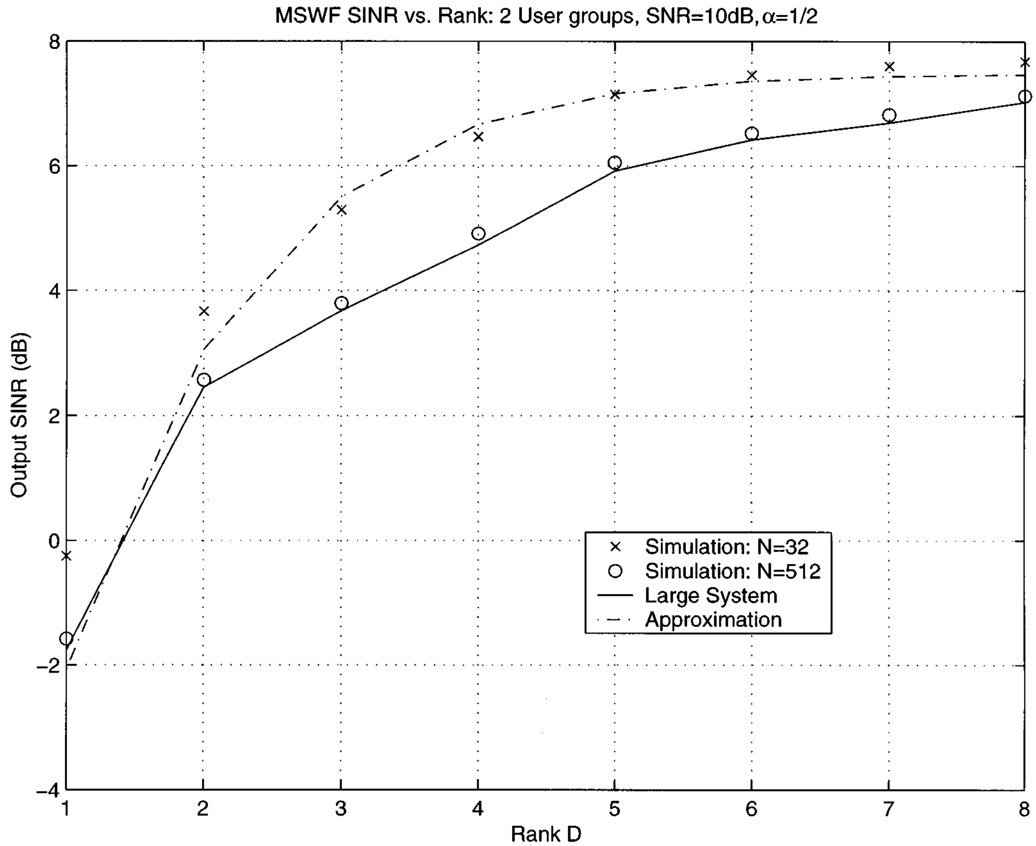

Fig. 6. Output SINR versus D for the MSWF with two groups of high- and low-power users. The large system approximation (47) is compared with the exact large system SINR computed from (61).

Now suppose that we choose $\boldsymbol{c}_1^{\perp} = \boldsymbol{c}_2$, so that $\beta_c^{\infty} = \beta_2^{MS}$. To prove the theorem for $D = 2$, we must show that the large system SINR associated with $\boldsymbol{c}_1^{\perp} = \boldsymbol{c}_2$ is

$$(\beta_1^{\perp})^{\infty} = \beta_1^{MS} = \frac{P}{\alpha P + \sigma^2} \tag{97}$$

which is the large system SINR for the matched filter. We, therefore, let

$$\boldsymbol{c}_1^{\perp} = \boldsymbol{c}_2 = \kappa_2 \mathcal{P}_{\boldsymbol{c}_1}^{\perp}(\boldsymbol{R}\boldsymbol{c}_1) \tag{98}$$

$$= \kappa_2 \mathcal{P}_{\boldsymbol{c}_1}^{\perp}[(\boldsymbol{S}_I \boldsymbol{S}_I^{\dagger})\boldsymbol{c}_1] \tag{99}$$

where $\mathcal{P}_{\boldsymbol{x}}^{\perp}(\boldsymbol{y}) = \boldsymbol{y} - \|\boldsymbol{x}\|^{-2}(\boldsymbol{x}^{\dagger}\boldsymbol{y})\boldsymbol{x}$ is the orthogonal projection of $\boldsymbol{y}$ onto $\boldsymbol{x}$, and κ_2 is the normalization constant. Now from (99)

$$\|\mathcal{P}_{\boldsymbol{c}_1}^{\perp}(\boldsymbol{R}\boldsymbol{c}_1)\|^2 = (\boldsymbol{R}\boldsymbol{c}_1)^{\dagger}[\mathcal{P}_{\boldsymbol{c}_1}^{\perp}(\boldsymbol{R}\boldsymbol{c}_1)] \tag{100}$$

$$= [(\boldsymbol{S}_I \boldsymbol{S}_I^{\perp})\boldsymbol{c}_1]^{\dagger} \mathcal{P}_{\boldsymbol{c}_1}^{\perp}[(\boldsymbol{S}_I \boldsymbol{S}_I^{\dagger})\boldsymbol{c}_1] \tag{101}$$

$$= \boldsymbol{c}_1^{\dagger}(\boldsymbol{S}_I \boldsymbol{S}_I^{\dagger})^2 \boldsymbol{c}_1 - (\boldsymbol{c}_1^{\dagger} \boldsymbol{S}_I \boldsymbol{S}_I^{\dagger} \boldsymbol{c}_1)^2$$

$$\overrightarrow{K\to\infty} (\alpha^2 + \alpha)P^2 - (\alpha P)^2 = \alpha P^2. \tag{102}$$

(See Appendix B, which discusses the computation of the large system limit of the moments $\boldsymbol{c}_1^{\dagger}(\boldsymbol{S}_I \boldsymbol{S}_I^{\dagger})^k \boldsymbol{c}_1$.) From (86) and (102) we have

$$|S_1^{\perp}|^2 = \frac{|\boldsymbol{c}_2^{\dagger}\boldsymbol{R}\boldsymbol{c}_1|^2}{E[|I_1|^2]^2}|I_1|^2 = \frac{\|\mathcal{P}_{\boldsymbol{c}_1}^{\perp}(\boldsymbol{R}\boldsymbol{c}_1)\|^2}{E[|I_1|^2]^2}|I_1|^2$$

$$\overrightarrow{K\to\infty} P \tag{103}$$

and it is shown in Appendix D that

$$|d_1^{\perp}|^2 = \boldsymbol{c}_2^{\dagger}\boldsymbol{R}\boldsymbol{c}_2 \overrightarrow{K\to\infty} P + \alpha P + \sigma^2. \tag{104}$$

Combining (88) with (103) and (104) gives (97), and substituting into (96) gives

$$\beta_2^{MS} = \frac{P}{\sigma^2 + \alpha \frac{P}{1+\beta_1^{MS}}}. \tag{105}$$

Before proceeding to the induction step, we need to make an additional observation. Suppose that we choose $\boldsymbol{c}_1^{\perp} = \boldsymbol{c}_{2:K}$, which consists of stages two through K of the MSWF. Referring to Fig. 1, this filter has input $r_1(i)$ and output $w_2\epsilon_2(i)$. Let $\boldsymbol{c}_{m:K}$ denote stages m through K of the MSWF (input $r_{m-1}(i)$ and output $w_m\epsilon_m(i)$). From Fig. 1 and (26) we can write

$$\boldsymbol{c}_{m:K} = w_m(\boldsymbol{c}_m - \boldsymbol{c}_{m+1:K}) \tag{106}$$

where we have used the fact that $\prod_{m=1}^{n-1} \boldsymbol{B}_m \boldsymbol{c}_n = \boldsymbol{c}_n$. We can, therefore, express $\boldsymbol{c}_{2:K}$ as a linear combination of the MSWF filters $\boldsymbol{c}_2, \ldots, \boldsymbol{c}_K$, i.e.,

$$\boldsymbol{c}_{2:K} = \sum_{i=2}^{K} w_{i;(2:K)}\boldsymbol{c}_i \tag{107}$$

where the $w_{i;(2:K)}$'s, $i = 2, \ldots, K$, are the corresponding combining coefficients, and depend on the filter indices $2 : K$. Since the $\boldsymbol{c}_i$'s are orthogonal, $\boldsymbol{c}_{2:K}$ is orthogonal to $\boldsymbol{c}_1$, as required.

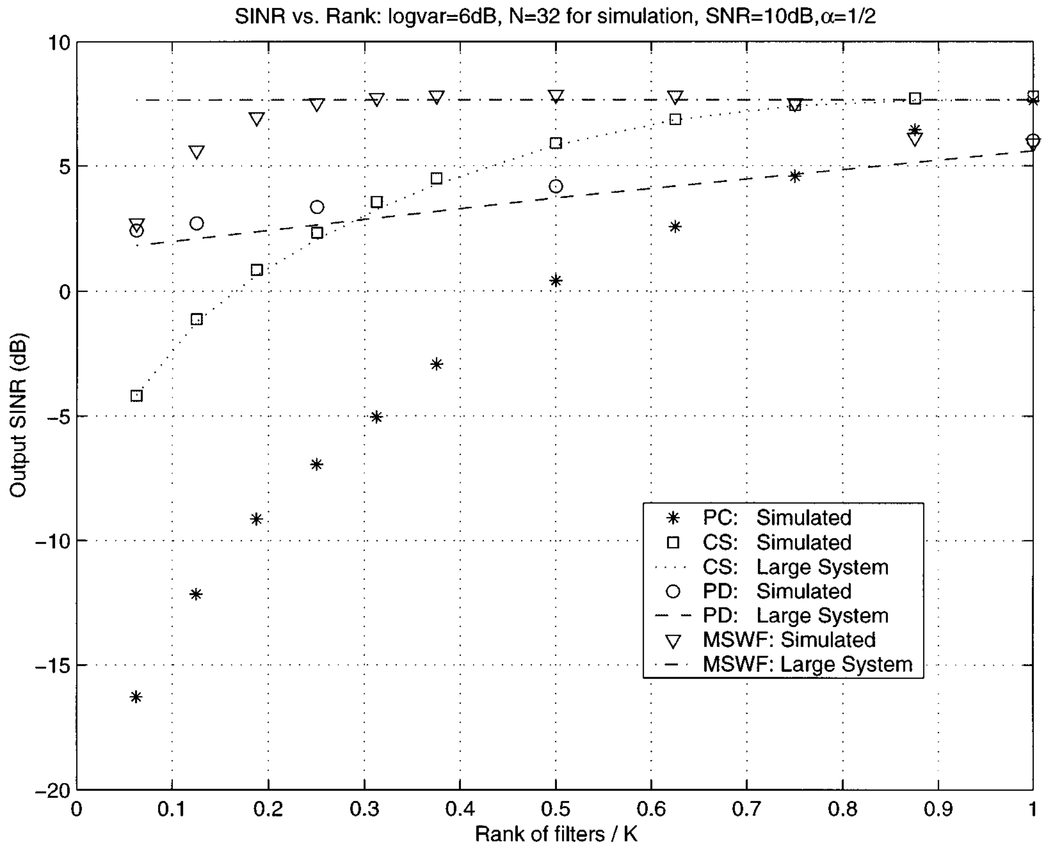

Fig. 7. Output SINR versus normalized rank D/K for reduced-rank filters with a log-normal received power distribution.

From the discussion in Section III, $\boldsymbol{c}_{2:K}$ is the MMSE filter for estimating d_1 from $\boldsymbol{r}_1$, the output of the blocking filter $\boldsymbol{B}_1$ in Fig. 1. Since K stages are needed to obtain the full-rank MMSE filter, the output SINR of $\boldsymbol{c}_{1:K}$ is the full-rank SINR β^∞, which from (40) must satisfy

$$\beta_c = \beta^\infty = \frac{P}{\sigma^2 + \alpha \frac{P}{1+\beta^\infty}}. \tag{108}$$

Comparing (108) with (96) with $\boldsymbol{c}_1^\perp = \boldsymbol{c}_{2:K}$ shows that the output SINR of $\boldsymbol{c}_{2:K}$ must be $\beta_{2:K} \to \beta^\infty$ as $K = \alpha N \to \infty$. This completes the first step of the proof.

For the induction step we partition the MSWF into the first $i-1$ stages, consisting of $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_{i-1}$, and the rest of the filter, consisting of $\boldsymbol{c}_{i:K}$. By assumption, the large system output SINR of $\boldsymbol{c}_{i:K}$ is $\beta_{i:K}^\infty = \beta^\infty$. We also need to define the filter $\boldsymbol{c}_{i:L}$, which consists of stages i through L of a rank-L MSWF. Clearly, $\boldsymbol{c}_{i:L}$ is a linear combination of $\boldsymbol{c}_i, \ldots, \boldsymbol{c}_L$,

$$\boldsymbol{c}_{i:L} = \sum_{l=i}^{L} w_{l;(i:L)} \boldsymbol{c}_l \tag{109}$$

where $L \le K$, and $w_{l;(i:L)}$, $l = i, \ldots, L$, are the combining coefficients. We decompose $\boldsymbol{c}_{i:L}$ as

$$\boldsymbol{c}_{i:L} = \boldsymbol{c}_i + w_{i+1}\boldsymbol{c}_i^\perp \tag{110}$$

where $\boldsymbol{c}_i^\perp$ is orthogonal to $\boldsymbol{c}_i$, which appears in the MSWF. In what follows, we will choose $\boldsymbol{c}_i^\perp = \boldsymbol{c}_{i+1:L}$. That is, for $L = i+1$, $\boldsymbol{c}_i^\perp = \boldsymbol{c}_{i+1}$ and for $L = K$, $\boldsymbol{c}_i^\perp = \boldsymbol{c}_{i+1:K}$ is the "bottom part" of the full-rank MSWF below stage i. Note that $\boldsymbol{c}_{i+1:K}$ is the MMSE filter for estimating d_i from $\boldsymbol{r}_i$.

Let

$$d_i = \boldsymbol{c}_i^\dagger \boldsymbol{r} = S_i + I_i + N_i \tag{111}$$

$$d_i^\perp = (\boldsymbol{c}_i^\perp)^\dagger \boldsymbol{r} = S_i^\perp + I_i^\perp + N_i^\perp \tag{112}$$

$$d_{i:L} = \boldsymbol{c}_{i:L}^\dagger \boldsymbol{r} = d_i + w_{i+1} d_i^\perp \tag{113}$$

where $N_i = \boldsymbol{c}_i^\dagger \boldsymbol{n}$, $N_i^\perp = (\boldsymbol{c}_i^\perp)^\dagger \boldsymbol{n}$, the desired signal S_i is the projection of d_i onto I_{i-1}, and $S_i^\perp$ is the projection of $d_i^\perp$ onto I_i

$$S_i^\perp = a_i I_i, \qquad a_i = \frac{E[(d_i^\perp)^* I_i]}{E[|I_i|^2]}. \tag{114}$$

The variables needed for the induction step are illustrated in Fig. 9.

Lemma 2:

$$E[(d_i^\perp)^* N_i] = 0 \tag{115}$$

$$E[(d_i^\perp)^* S_i] = 0 \tag{116}$$

$$E[(d_i^\perp)^* I_i] = \boldsymbol{c}_i^\dagger \boldsymbol{R} \boldsymbol{c}_i^\perp. \tag{117}$$

Proof: First,

$$E[(d_i^\perp)^* N_i] = E[\boldsymbol{c}_i^\dagger \boldsymbol{n} \boldsymbol{r}^\dagger \boldsymbol{c}_i^\perp] = E(\boldsymbol{c}_i^\dagger \boldsymbol{n}\boldsymbol{n}^\dagger \boldsymbol{c}_i^\perp)$$
$$= \sigma^2 E[(\boldsymbol{c}_i^\perp)^\dagger \boldsymbol{c}_i] = 0.$$

To show (116), we write

$$E[(d_i^\perp)^* S_i] = a_{i-1} E[(d_i^\perp)^* I_{i-1}]$$
$$= a_{i-1} E[(d_i^\perp)^* (d_{i-1} - S_{i-1} - N_{i-1}).$$

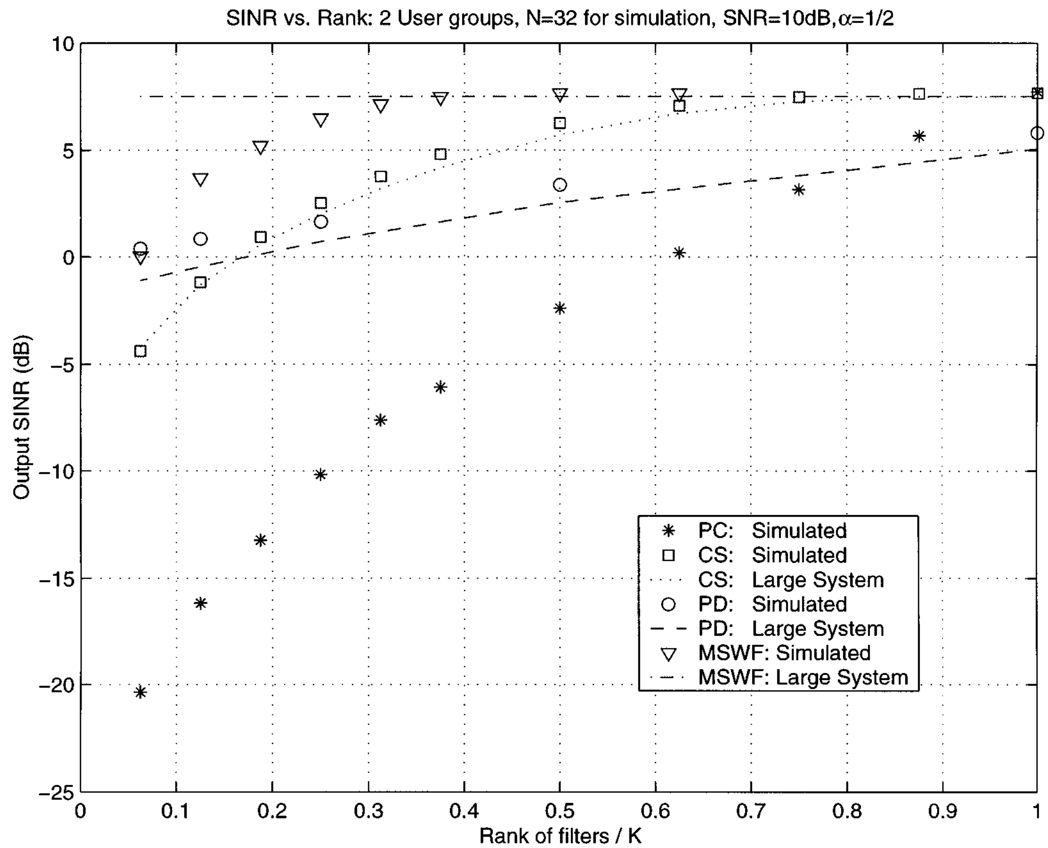

Fig. 8. Output SINR versus normalized rank D/K for reduced-rank filters with two groups of high- and low-power users.

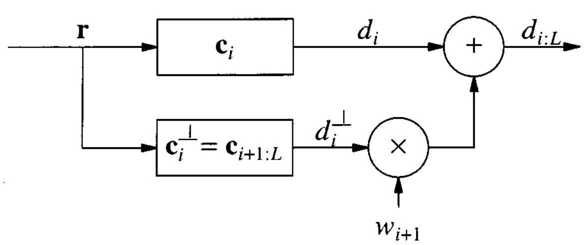

Fig. 9. Illustration of variables used in the proof of Theorem 1.

Now

$$E[(d_i^\perp)^* N_{i-1}] = \sigma^2 \boldsymbol{c}_{i-1}^\dagger \boldsymbol{c}_i^\perp = 0 \tag{118}$$

and

$$E[(d_i^\perp)^* d_{i-1}] = \boldsymbol{c}_{i-1}^\dagger \boldsymbol{R} \boldsymbol{c}_i^\perp = 0 \tag{119}$$

follows from (32) and the fact that $\boldsymbol{c}_i^\perp$ is a linear combination of $\boldsymbol{c}_{i+1}, \ldots, \boldsymbol{c}_K$. Consequently,

$$\begin{aligned} E[(d_i^\perp)^* S_i] &= -a_{i-1} E[(d_i^\perp)^* S_{i-1}] \\ &= -a_{i-1} a_{i-2} E[(d_i^\perp)^* I_{i-2})] \\ &= -a_{i-1} a_{i-2} E[(d_i^\perp)^* (d_{i-2} - S_{i-2} - N_{i-2})] \\ &= a_{i-1} a_{i-2} E[(d_i^\perp)^* S_{i-2}] \\ &= \text{constant} \cdot E[(d_i^\perp)^* S_1] \\ &= \text{constant} \cdot E[\boldsymbol{c}_1^\dagger \boldsymbol{R} \boldsymbol{c}_i^\perp] = 0. \end{aligned}$$

Finally,

$$\begin{aligned} E[(d_i^\perp)^* I_i] &= E[(d_i^\perp)^* (d_i - S_i - N_i)] \\ &= E[(d_i^\perp)^* d_i] = \boldsymbol{c}_i^\dagger \boldsymbol{R} \boldsymbol{c}_i^\perp \end{aligned} \tag{120}$$

which completes the proof of Lemma 2.

We now compute the large system limit of

$$\beta_{i:L} = \frac{E[|S_i|^2]}{E[|d_{i:L}|^2] - E[|S_i|^2]} \tag{121}$$

in terms of the output SINR for $\boldsymbol{c}_i^\perp$, which is

$$\begin{aligned} \beta_i^\perp &= \frac{E[|S_i^\perp|^2]}{E[|d_i^\perp|^2] - E[|S_i^\perp|^2]} \\ &= \frac{\frac{|\boldsymbol{c}_i^\dagger \boldsymbol{R} \boldsymbol{c}_i^\perp|^2}{E[|I_i|^2]}}{(\boldsymbol{c}_i^\perp)^\dagger \boldsymbol{R} \boldsymbol{c}_i^\perp - \frac{|\boldsymbol{c}_i^\dagger \boldsymbol{R} \boldsymbol{c}_i^\perp|^2}{E[|I_i|^2]}} \end{aligned} \tag{122}$$

and which satisfies

$$\frac{\beta_i^\perp}{1+\beta_i^\perp} = \frac{E[|S_i^\perp|^2]}{E[|d_i^\perp|^2]} = \frac{1}{E[|I_i|^2]} \frac{|\boldsymbol{c}_i^\dagger \boldsymbol{R} \boldsymbol{c}_i^\perp|^2}{(\boldsymbol{c}_i^\perp)^\dagger \boldsymbol{R} \boldsymbol{c}_i^\perp}. \tag{123}$$

Selecting w_{i+1} to minimize $E[|d_{i:L}|^2]$ in (113) gives

$$w_{i+1} = -\frac{E[(d_i^\perp)^* d_i]}{E[|d_i^\perp|^2]} = -\frac{E[(d_i^\perp)^* I_i]}{(\boldsymbol{c}_i^\perp)^\dagger \boldsymbol{R} \boldsymbol{c}_i^\perp} \tag{124}$$

$$= -\frac{\boldsymbol{c}_i^\dagger \boldsymbol{R} \boldsymbol{c}_i^\perp}{(\boldsymbol{c}_i^\perp)^\dagger \boldsymbol{R} \boldsymbol{c}_i^\perp} \tag{125}$$

and

$$\begin{aligned} E[|d_{i:L}|^2] &= E[|d_i|^2] - \frac{|E[(d_i^\perp)^* d_i]|^2}{E[|d_i^\perp|^2]} \\ &= E[|d_i|^2] - \frac{|\boldsymbol{c}_i^\dagger \boldsymbol{R} \boldsymbol{c}_i^\perp|^2}{(\boldsymbol{c}_i^\perp)^\dagger \boldsymbol{R} \boldsymbol{c}_i^\perp} \\ &= E[|S_i|^2] + E[|I_i|^2] + \sigma^2 - E[|I_i|^2] \frac{\beta_i^\perp}{1+\beta_i^\perp} \\ &= E[|S_i|^2] + \sigma^2 + E[|I_i|^2] \frac{1}{1+\beta_i^\perp} \end{aligned} \tag{126}$$

where (123) has been used. Combining (121)–(126) gives

$$\beta_{i:L} = \frac{E[|S_i|^2]}{\sigma^2 + \frac{E[|I_i|^2]}{1+\beta_i^\perp}}. \tag{127}$$

For $\beta_i^\perp = \beta_{i+1:K}$, corresponding to $\boldsymbol{c}_i^\perp = \boldsymbol{c}_{i+1:K}$, we have $\beta_{i:K}^\infty = \beta^\infty$ (by assumption), so that (127) and (40) imply that

$$\beta_{i:K} = \frac{E[|S_i|^2]}{\sigma^2 + \frac{E[|I_i|^2]}{1+\beta_{i+1:K}}} \xrightarrow[K\to\infty]{} \frac{P}{\sigma^2 + \frac{\alpha P}{1+\beta^\infty}}. \tag{128}$$

We can rewrite (128) as

$$\sigma^2 \left[\frac{1}{E^\infty[|S_i|^2]} - \frac{1}{P}\right] = \frac{\alpha}{1+\beta^\infty} - \frac{E^\infty[|I_i|^2]}{E^\infty[|S_i|^2]} \frac{1}{1+\beta_{i+1:K}^\infty} \tag{129}$$

where the superscript "∞" denotes the large system limit of the associated variable.

Lemma 3: $\boldsymbol{c}_i$, $E[|S_i|^2]$, and $E[|I_i|^2]$ are independent of σ^2 $\forall\, i$.

The proof is given in Appendix C.

As $\sigma^2 \to \infty$, $\beta_{i+1:K} \to 0$ and $\beta^\infty \to 0$, so that (129) can only be true if $E^\infty[|S_i|^2] = P$. Consequently, from (128) as $K = \alpha N \to \infty$

$$E[|S_i|^2] \to P, \qquad E[|I_i|^2] \to \alpha P, \qquad \beta_{i+1:K}^\infty \to \beta^\infty \tag{130}$$

and the SINR associated with the output of $\boldsymbol{c}_i$ is

$$\beta_{i:i}^\infty = \frac{E^\infty[|S_i|^2]}{\sigma^2 + E^\infty[|I_i|^2]} = \frac{P}{\sigma^2 + \alpha P} \tag{131}$$

for $i = 1, \ldots, K$, where convergence in probability follows from the fact that the variables are continuous and bounded functions of the moments $\boldsymbol{s}_1^\dagger \boldsymbol{R}_I^k \boldsymbol{s}_1$, $k = 1, \ldots, 2D-1$.

From (127)–(130), we can write

$$\beta_{i-1:i}^\infty = \frac{P}{\sigma^2 + \alpha \frac{P}{1+\beta_{i:i}^\infty}} \tag{132}$$

where $\beta_{i:i}^\infty$ is given by (131). Similarly, (127) can be used again to express $\beta_{i-2:i}^\infty$ in terms of $\beta_{i-1:i}^\infty$, which is given by (132) and (131). Iterating in this manner gives the theorem.

B. Theorem 2: Basis for the MSWF

In what follows, it will be convenient to replace the normalized MSWF filters $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_D$, given by (30), by the unnormalized filters $\check{\boldsymbol{c}}_i$, which satisfy

$$\check{\boldsymbol{c}}_{i+1} = \left(\boldsymbol{I} - \sum_{l=1}^{i} \check{\boldsymbol{c}}_l \check{\boldsymbol{c}}_l^\dagger\right) \boldsymbol{R}_I \check{\boldsymbol{c}}_i, \qquad i = 1, \ldots, D-1. \tag{133}$$

Clearly, $\check{\boldsymbol{c}}_1, \ldots, \check{\boldsymbol{c}}_D$ span the same space $\mathcal{S}_D$. The theorem is obviously true for $D = 1$ since $\boldsymbol{c}_1 = \check{\boldsymbol{c}}_1 = \boldsymbol{s}_1$. Since

$$\check{\boldsymbol{c}}_2 = (\boldsymbol{I} - \boldsymbol{c}_1 \boldsymbol{c}_1^\dagger) \boldsymbol{R}_I \boldsymbol{c}_1 = \boldsymbol{R}_I \boldsymbol{s}_1 - (\boldsymbol{s}_1^\dagger \boldsymbol{R}_I \boldsymbol{s}_1) \boldsymbol{s}_1 \tag{134}$$

$[\boldsymbol{c}_1 \check{\boldsymbol{c}}_2] = [\boldsymbol{c}_1 \quad \boldsymbol{R}_I \boldsymbol{c}_1] \boldsymbol{A}_2$ where

$$\boldsymbol{A}_2 = \begin{bmatrix} 1 & -(\boldsymbol{s}_1^\dagger \boldsymbol{R}_I \boldsymbol{s}_1) \\ 0 & 1 \end{bmatrix}$$

is nonsingular, so the theorem is true for $D = 2$.

Assume that the theorem is true for $D = i$, and that

$$[\,\boldsymbol{c}_1 \quad \check{\boldsymbol{c}}_2 \quad \cdots \quad \check{\boldsymbol{c}}_i\,] = [\,\boldsymbol{s}_1 \quad \boldsymbol{R}_I \boldsymbol{s}_1 \quad \cdots \quad \boldsymbol{R}_I^{i-1} \boldsymbol{s}_1\,] \boldsymbol{A}_i \tag{135}$$

where $\boldsymbol{A}_i$ is a nonsingular upper triangular $i \times i$ matrix with diagonal elements equal to one. From (30)

$$\begin{aligned} \check{\boldsymbol{c}}_i &= \left(\boldsymbol{I} - \sum_{l=1}^{i-1} \check{\boldsymbol{c}}_l \check{\boldsymbol{c}}_l^\dagger\right) \boldsymbol{R}_I \check{\boldsymbol{c}}_{i-1} \\ &= \boldsymbol{R}_I \check{\boldsymbol{c}}_{i-1} - [\,\check{\boldsymbol{c}}_1 \quad \check{\boldsymbol{c}}_2 \quad \cdots \quad \check{\boldsymbol{c}}_{i-1}\,] \boldsymbol{u}, \qquad i = 1, \ldots, D \end{aligned}$$

where

$$\boldsymbol{u} = [\,\check{\boldsymbol{c}}_1^\dagger \boldsymbol{R} \check{\boldsymbol{c}}_{i-1} \quad \check{\boldsymbol{c}}_2^\dagger \boldsymbol{R} \check{\boldsymbol{c}}_{i-1} \quad \cdots \quad \check{\boldsymbol{c}}_{i-1}^\dagger \boldsymbol{R} \check{\boldsymbol{c}}_{i-1}\,]^T.$$

Denoting the jth column of $\boldsymbol{A}_i$ as $\boldsymbol{A}_i^{(j)}$, we have

$$\begin{aligned} \check{\boldsymbol{c}}_i &= \boldsymbol{R}_I [\,\boldsymbol{s}_1 \quad \boldsymbol{R}_I \boldsymbol{s}_1 \quad \cdots \quad \boldsymbol{R}_I^{i-1} \boldsymbol{s}_1\,] \boldsymbol{A}_i^{(i)} \\ &\quad - [\,\boldsymbol{s}_1 \quad \boldsymbol{R}_I \boldsymbol{s}_1 \quad \cdots \quad \boldsymbol{R}_I^{i-2} \boldsymbol{s}_1\,] \boldsymbol{A}_{i-1} \boldsymbol{u} \\ &= [\,\boldsymbol{s}_1 \quad \boldsymbol{R}_I \boldsymbol{s}_1 \quad \cdots \quad \boldsymbol{R}_I^{i} \boldsymbol{s}_1\,] \boldsymbol{A}_{i+1}^{(i+1)} \end{aligned}$$

where

$$\boldsymbol{A}_{i+1}^{(i+1)} = \begin{bmatrix} -\boldsymbol{A}_{i-1} \boldsymbol{u} \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \boldsymbol{A}_i^{(i)} \end{bmatrix}.$$

Hence, $\boldsymbol{A}_{i+1}$ is also a nonsingular upper triangular matrix with diagonal elements equal to one, which establishes the theorem.

C. Principal Components

We assume all users are received with power P. From (63) we have

$$E(v_D) = P \sum_{n=1}^{D} E\left(\frac{|\boldsymbol{v}_n^\dagger \boldsymbol{s}_1|^2}{\lambda_n + \sigma^2}\right) \tag{136}$$

where the expectation is with respect to the random signature matrix $\boldsymbol{S}$. Since the elements of $\boldsymbol{S}$ are i.i.d., we can replace $\boldsymbol{s}_1$ by $\boldsymbol{s}_k$, so that

$$\begin{aligned} E\left(\frac{|\boldsymbol{v}_n^\dagger \boldsymbol{s}_1|^2}{\lambda_n + \sigma^2}\right) &= \frac{1}{K} E\left(\sum_{k=1}^{K} \frac{|\boldsymbol{v}_n^\dagger \boldsymbol{s}_k|^2}{\lambda_n + \sigma^2}\right) \\ &= \frac{1}{K} E\left(\frac{\boldsymbol{v}_n^\dagger \boldsymbol{S} \boldsymbol{S}^\dagger \boldsymbol{v}_n}{\lambda_n + \sigma^2}\right) \\ &= \frac{1}{K} E\left(\frac{\lambda_n}{\lambda_n + \sigma^2}\right). \end{aligned} \tag{137}$$

Combining (136) and (137) gives

$$E(v_D) = \frac{P}{K}\sum_{n=1}^{D} E\left(\frac{\lambda_n}{\lambda_n+\sigma^2}\right) = \frac{P}{\alpha N}\sum_{n=1}^{D} E\left(\frac{\lambda_n}{\lambda_n+\sigma^2}\right). \tag{138}$$

As $K = \alpha N \to \infty$, the distribution of $\{\lambda_n\}$ converges to a deterministic limit $G(\cdot)$ with associated density [27, Theorem 2.1]

$$g(x) = \begin{cases} \frac{\sqrt{[x-a(\alpha)][b(\alpha)-x]}}{2\pi x}, & \text{for } a(\alpha) < x < b(\alpha) \\ 0, & \text{otherwise} \end{cases} \tag{139}$$

where $a(\alpha)$ and $b(\alpha)$ are given by (65), and for $0 < \alpha < 1$ there is an additional mass point at $x = 0$

$$\Pr\{\lambda_n = 0\} = 1 - \alpha. \tag{140}$$

Combining (138) with the asymptotic eigenvalue distribution gives

$$\lim_{\substack{K=\alpha N\to\infty \\ D=\delta N\to\infty}} E[v_D] = \frac{1}{\alpha}\int_c^{b(\alpha)} \frac{\lambda}{\lambda+\sigma^2}\, dG(\lambda) = \int_{c+\sigma^2}^{b(\alpha)+\sigma^2} \frac{\sqrt{[x-a(\alpha)-\sigma^2][b(\alpha)+\sigma^2-x]}}{2\pi x}\, dx \tag{141}$$

where c satisfies

$$G(c) = \int_{a(\alpha)}^{c} \frac{\sqrt{[x-a(\alpha)][b(\alpha)-x]}}{2\pi x}\, dx + (1-\alpha)\mathbf{1}\{\alpha<1\} = 1 - \delta. \tag{142}$$

The preceding integrals can be evaluated in closed form giving (64) and (66). The conjecture that v_D converges to a deterministic limit depends on showing that the large system limit of the variance of the random variable $|\theta_n|^2$ is zero where $\theta_n = \boldsymbol{v}_n^\dagger \boldsymbol{s}_1$. We remark that θ_n and θ_l are uncorrelated for $n \neq l$, i.e.,

$$\sum_{k=1}^{K} \boldsymbol{v}_n^\dagger \boldsymbol{s}_k \boldsymbol{s}_k^\dagger \boldsymbol{v}_l = \boldsymbol{v}_n^\dagger \boldsymbol{S}\boldsymbol{S}^\dagger \boldsymbol{v}_l = 0 \tag{143}$$

so that

$$E\left[\sum_{k=1}^{K} \boldsymbol{v}_n^\dagger \boldsymbol{s}_k \boldsymbol{s}_k^\dagger \boldsymbol{v}_l\right] = \sum_{k=1}^{K} E[\boldsymbol{v}_n^\dagger \boldsymbol{s}_k \boldsymbol{s}_k^\dagger \boldsymbol{v}_l] = KE[\theta_n\theta_l^*] = 0.$$

In the case of an arbitrary power distribution F, (137) no longer holds, since the projection of the desired user's signal onto the eigenvectors depends on F and P_1. It, therefore, appears to be difficult to compute the corresponding large system limit for arbitrary F.

D. Cross-Spectral Method

Here we compute the variance of ξ_n defined by (68), where n is chosen randomly. We have

$$\sigma_\xi^2 = \sum_{n=1}^{K} \frac{|\boldsymbol{v}_n^\dagger \boldsymbol{s}_1|^2}{\lambda_n+\sigma^2} = \boldsymbol{s}_1^\dagger \boldsymbol{R}^{-1}\boldsymbol{s}_1 \xrightarrow[K\to\infty]{} \frac{\beta^\infty}{1+\beta^\infty} \tag{144}$$

where β^∞ is the asymptotic SINR for the full-rank MMSE filter. Alternatively, from (137) and (138) we have that

$$\sigma_\xi^2 = E\left(\frac{\lambda_n}{\lambda_n+\sigma^2}\right) = \frac{1}{\alpha}\int_{a(\alpha)}^{b(\alpha)} \frac{\lambda}{\lambda+\sigma^2}\, dG(\lambda) \tag{145}$$

which for $G(\cdot)$ given by (139) can be evaluated as (74).

E. Theorem 4: Partial Despreading (PD)

The projection matrix is

$$\boldsymbol{M}_D = \begin{bmatrix} \boldsymbol{s}_{1,1} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{s}_{1,2} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{s}_{1,D} \end{bmatrix} \tag{146}$$

where $\boldsymbol{s}_{1,d} = [\,\boldsymbol{s}_{1,d(M-1)+1} \;\cdots\; \boldsymbol{s}_{1,dM}\,]'$ is a segment of $\boldsymbol{s}_1$ containing M chips, and the columns of $\boldsymbol{M}_D$ contain nonoverlapping segments. In analogy with (14) and (15) we define

$$\tilde{\boldsymbol{R}} = \boldsymbol{M}_D^\dagger \boldsymbol{R}\boldsymbol{M}_D = \tilde{\boldsymbol{S}}_I \tilde{\boldsymbol{P}}_I \tilde{\boldsymbol{S}}_I^\dagger + \tilde{P}_1 \tilde{\boldsymbol{s}}_1 \tilde{\boldsymbol{s}}_1^\dagger + \tilde{\boldsymbol{N}} \tag{147}$$

where $\tilde{\boldsymbol{S}}_I = \sqrt{N}\boldsymbol{M}_D^\dagger \boldsymbol{S}_I$ is $D \times K$ with (d, k) element

$$\tilde{\boldsymbol{S}}_{I;(d,k)} = \sqrt{N}\boldsymbol{s}_{1,d}^\dagger \boldsymbol{s}_{k,d} \tag{148}$$

$\tilde{\boldsymbol{P}}_I = \boldsymbol{P}_I/N$ where $\boldsymbol{P}_I$ is the diagonal matrix of the powers of the interferers, $\tilde{P}_1 = P_1/D$

$$\tilde{\boldsymbol{s}}_1 = \sqrt{D}\boldsymbol{M}_D^\dagger \boldsymbol{s}_1 = \sqrt{D}[(\boldsymbol{s}_{1,1}^\dagger \boldsymbol{s}_{1,1})\cdots(\boldsymbol{s}_{1,D}^\dagger \boldsymbol{s}_{1,D})]^T \tag{149}$$

and

$$\tilde{\boldsymbol{N}} = \mathrm{diag}[\tilde{\sigma}_1^2 \cdots \tilde{\sigma}_D^2] \tag{150}$$

is the noise covariance matrix where $\tilde{\sigma}_k^2 = \sigma^2\|\boldsymbol{s}_{1,k}\|^2$. With these definitions, the output SINR $\beta_D = \tilde{P}_1 \tilde{\boldsymbol{s}}_1^\dagger \tilde{\boldsymbol{R}}_I^{-1}\tilde{\boldsymbol{s}}_1$, and we can apply the large system results for the full-rank MMSE filter presented in Section IV.

We consider the following two cases.

1) The elements of $\boldsymbol{S}$ are i.i.d. binary random variables chosen from $\{-\frac{1}{\sqrt{N}}, \frac{1}{\sqrt{N}}\}$ with equal probabilities.
2) The elements of $\boldsymbol{S}$ are i.i.d. Gaussian random variables with zero mean and variance $\frac{1}{N}$.

For case 1, $\tilde{\boldsymbol{N}} = \frac{\sigma^2}{D}\boldsymbol{I}_{D\times D}$, and $\tilde{\boldsymbol{s}}_1 = [\frac{1}{\sqrt{D}} \cdots \frac{1}{\sqrt{D}}]^T$. It is easy to verify that the elements of $\tilde{\boldsymbol{S}}_I$ are i.i.d. random variables with zero mean and variance $1/D$. Now the covariance matrix has the same form as for a full-rank MMSE filter with interferer powers given by $\boldsymbol{P}_I/N$, desired user power P_1/D, noise variance σ^2/D, and system load $\tilde{\alpha} = K/D = M\alpha$. According to (38) and (42), we have

$$\beta^{\text{PD}} = \mathcal{B}_F[M\alpha,\, MP_1,\, M\sigma^2,\, F(\cdot)]. \tag{151}$$

As $M \to \infty$, or $\delta \to 0$, we have

$$\begin{aligned}\beta^{\text{PD}} &= \lim_{M\to\infty} \frac{MP_1}{M\sigma^2 + M\alpha \int \frac{MP_1 P}{MP_1 + P\beta^{\infty}_{\text{PD},\,0}}\, dF(P)} \\ &= \frac{P_1}{\sigma^2 + \alpha \int P dF(P)}.\end{aligned} \tag{152}$$

Hence, for a fixed $D < \infty$, as $\delta \to 0$, the large system SINR for the PD filter converges to the matched-filter SINR.

For case 2, as $N, M \to \infty$ with $N/M = D$, we have $\tilde{\boldsymbol{N}} \to (\sigma^2/D)\boldsymbol{I}_{D\times D}$ and $\tilde{\boldsymbol{s}}_1 \to [\frac{1}{\sqrt{D}} \cdots \frac{1}{\sqrt{D}}]^T$, and the elements of $\tilde{\boldsymbol{S}}_I$ are i.i.d. Gaussian random variables with zero mean and variance $1/D$. The output SINR is again given by (151) where $M \to \infty$, which is the same as the matched-filter SINR (152).

For finite M, we must assume that the energy of each segment of $\boldsymbol{s}_{1,\,d}$ is the same, so that $\tilde{\boldsymbol{N}} = \frac{\sigma^2}{D} I_{D\times D}$ and $\tilde{\boldsymbol{s}}_1 = [\frac{1}{\sqrt{D}} \cdots \frac{1}{\sqrt{D}}]^T$. To prove that the elements in $\tilde{\boldsymbol{S}}_I$, given by (148), are i.i.d. Gaussian, we first note that conditioned on $\boldsymbol{s}_{1,\,d}$, element $(d,\, k)$ of $\tilde{\boldsymbol{S}}_I$ ($\tilde{\boldsymbol{S}}_{I;(d,\,k)}$) is Gaussian with zero mean and variance $\frac{1}{D}$. That is, the conditional probability density function (pdf) $f(\tilde{\boldsymbol{S}}_{I;(d,\,k)}|\boldsymbol{s}_{1,\,d})$ is independent of $\boldsymbol{s}_{1,\,d}$, so that

$$f(\tilde{\boldsymbol{S}}_{I;(d,\,k)}) = f(\tilde{\boldsymbol{S}}_{I;(d,\,k)}|\boldsymbol{s}_{1,\,d}).$$

Hence, the elements of $\tilde{\boldsymbol{S}}_I$ are uncorrelated Gaussian with zero mean and variance $\frac{1}{D}$, and are, therefore, i.i.d., so that (151) also applies to this case.

VIII. Conclusion

We have characterized the performance of the reduced-rank MSWF when used for multiple-access interference suppression. For the uniform power case the large system, SINR is easily computed as a continued fraction, and converges rapidly to the full-rank SINR as the rank D increases. Numerical results show that this large system analysis accurately predicts average performance for moderately sized systems. We do not have a similar type of simple expression for an arbitrary power distribution, although the performance can be computed numerically. An important conclusion, based on this analysis, is that the rank of the MSWF needed to achieve a target SINR in the neighborhood of the full-rank SINR does not scale with system size. Numerical results show that $D = 8$ is sufficient to achieve near full-rank performance for all cases considered.

The large system SINRs of PC, CS, and PD reduced-rank techniques have also been evaluated. Whether or not the expression for the CS technique is exact is an open question. In all cases, the large system analysis accurately predicts performance for moderately sized systems. Numerical results show that for very small D/K, the techniques based on eigendecomposition perform relatively poorly. The CS filter requires $D/K > 0.3$ to perform better than the matched filter, and $D/K > 0.33$ to perform better than PD. PC performs well only when $D/K \geq 1$. The MSWF performs significantly better than the other techniques considered for small D/K, and furthermore, does not require explicit tracking of the signal subspace.

Our results pertain to optimal (MMSE) filters. Adaptive algorithms based on the MSWF are presented in [9], [31], and are observed to converge significantly faster than a full-rank Least Squares algorithm. An analysis of the convergence of these adaptive reduced-rank filters with random data is presented in [32]. Investigation of tracking performance in the presence of time- and frequency-selective fading is being pursued.

Appendix A
Equation (47) Is not Exact

Here we show that (47) does not hold with equality for arbitrary F. Consider a rank-two MSWF. According to (61), the SINR can be computed as

$$\beta_2^{MS} = \frac{1}{\gamma_1^{\infty} - \frac{[(\gamma_1^{\infty})^2 - \gamma_2^{\infty}]^2}{\gamma_3^{\infty} - 2\gamma_1^{\infty}\gamma_2^{\infty} + (\gamma_1^{\infty})^3}} \tag{153}$$

where $\gamma_n(\alpha,\, \sigma) = \boldsymbol{s}_1^{\dagger}\boldsymbol{R}_I^n\boldsymbol{s}_1$. Combining the expressions for $\gamma_n^{\infty}(\alpha,\, 0)$, $n = 1,\, 2,\, 3$, in Appendix B with (60) gives

$$\begin{aligned}\gamma_1^{\infty}(\alpha,\, \sigma) &= \gamma_1^{\infty}(\alpha,\, 0) + \sigma^2 \\ \gamma_2^{\infty}(\alpha,\, \sigma) &= \gamma_2^{\infty}(\alpha,\, 0) + 2\alpha\sigma^2 E(P) + \sigma^4 \\ \gamma_3^{\infty}(\alpha,\, \sigma) &= \gamma_3^{\infty}(\alpha,\, 0) + 3\alpha E(P^2)\sigma^2 + 3\alpha^2[E(P)]^2\sigma^2 \\ &\quad + 3\alpha E(P)\sigma^4 + \sigma^6.\end{aligned}$$

The asymptotic SINR of the two-stage MSWF is therefore a function of σ^2, α and $E[P^n]$, $n = 1,\, 2,\, 3$, whereas (47) cannot be expressed in terms of a finite number moments of P for an arbitrary power distribution. Nevertheless, the numerical results in Section VI show that (47) gives a reasonable approximation to (61) for the case examined.

Appendix B
Computation of $\gamma_n^{\infty}(\alpha,\, 0)$

Here we give two alternate methods to (57) for computing the large system limit $\gamma_n^{\infty}(\alpha,\, \sigma^2)$ for $\sigma^2 = 0$, where γ_n is defined by (49). In what follows we will abbreviate $\gamma_n(\alpha,\, 0)$ as $\gamma_n(\alpha)$. For the first method, we define the z-transform

$$\begin{aligned}\Phi(z) &= \sum_{m=0}^{\infty} \gamma_m z^{-m} \\ &= \boldsymbol{s}_1^{\dagger}\left(\sum_{m=0}^{\infty}(z^{-1}\boldsymbol{R}_I)^m\right)\boldsymbol{s}_1 \\ &= \boldsymbol{s}_1^{\dagger}(\boldsymbol{I} - z^{-1}\boldsymbol{R}_I)^{-1}\boldsymbol{s}_1 \\ &= z\boldsymbol{s}_1^{\dagger}(z\boldsymbol{I} - \boldsymbol{S}_I\boldsymbol{P}_I\boldsymbol{S}_I^{\dagger})^{-1}\boldsymbol{s}_1\end{aligned} \tag{154}$$

where the preceding sum converges for $|z^{-1}\lambda_{\max}(\boldsymbol{R}_I)| < 1$, where $\lambda_{\max}(\boldsymbol{R}_I)$ is the maximum eigenvalue of $\boldsymbol{R}_I$. We observe that $z^{-1}\Phi(-z)$ with $z = \sigma^2$ is the output SINR of the full-rank linear MMSE filter with signature sequence matrix $\boldsymbol{S}_I$ and received power matrix $\boldsymbol{P}_I$. Taking the large system limit gives [17]

$$\Phi^{\infty}(z) = -z\mathcal{B}_F[\alpha, 1, -z, F(\cdot)] = -zm_G(z) \tag{155}$$

where $\mathcal{B}_F$ is defined by (42) and $m_G(z)$ is defined by (36) and satisfies (37). For uniform power we have from (37) and (40)

$$z^{-1}\Phi^{\infty}(z) = \frac{1}{z - \frac{\alpha}{1-z^{-1}\Phi^{\infty}(z)}} \tag{156}$$

which is the z-transform of the sequence defined by (59) with $P = 1$ and $\gamma_0^{\infty} = 1$.

The second method for computing γ_n^{∞} is based on relating γ_n with $K + 1$ users to γ_n with K users. Specifically, let $\gamma_n(K, N) = \boldsymbol{s}_1^{\dagger}\boldsymbol{R}_I^n\boldsymbol{s}_1$ and

$$\gamma_n(K+1, N) = \boldsymbol{s}_1^{\dagger}(\boldsymbol{R}_I + P\boldsymbol{s}_{K+1}\boldsymbol{s}_{K+1}^{\dagger})\boldsymbol{s}_1 \tag{157}$$

where P is the power assigned to user $K+1$. We can write

$$\gamma_1(K+1, N) = \gamma_1(K, N) + P\boldsymbol{s}_1^{\dagger}(\boldsymbol{s}_{K+1}\boldsymbol{s}_{K+1}^{\dagger})\boldsymbol{s}_1 \tag{158}$$

and taking expectation with respect to $\boldsymbol{s}_{K+1}$ gives

$$E_{\boldsymbol{s}_{K+1}}\{\gamma_1(K+1, N)\} = \gamma_1(K, N) + \frac{P}{N}. \tag{159}$$

Taking the large system limit, and noting that $1/N$ becomes $d\alpha$ gives

$$\gamma_1^{\infty}(\alpha + d\alpha) - \gamma_1^{\infty}(\alpha) = P\, d\alpha \tag{160}$$

or

$$\frac{d\gamma_1^{\infty}}{d\alpha} = P. \tag{161}$$

Since $\gamma_1(0) = 0$, we have $\gamma_1^{\infty}(\alpha) = P\alpha$. If P is chosen from a distribution, then $E_P[\gamma_1^{\infty}(\alpha)] = E(P)\alpha$; however, since γ_1 converges to γ_1^{∞} in probability, we have

$$\gamma_1^{\infty}(\alpha) = E(P)\alpha. \tag{162}$$

Similarly, we can compute γ_2^{∞} and γ_3^{∞}, which are used in the proof of Theorem 1. We have

$$\begin{aligned} E_{\boldsymbol{s}_{K+1}}&[\boldsymbol{s}_1^{\dagger}(\boldsymbol{R}_I + P\boldsymbol{s}_{K+1}\boldsymbol{s}_{K+1}^{\dagger})^2\boldsymbol{s}_1] \\ &= \gamma_2(K, N) + 2\frac{P}{N}\gamma_1(K, N) + \frac{P^2}{N} \end{aligned} \tag{163}$$

which gives

$$\frac{d\gamma_2^{\infty}}{d\alpha} = 2P\gamma_1^{\infty}(\alpha) + P^2. \tag{164}$$

Solving and averaging over P gives

$$\gamma_2^{\infty}(\alpha) = [\alpha E(P)]^2 + \alpha E(P^2). \tag{165}$$

Finally, following this approach for $\gamma_3(K+1, N)$ gives

$$\frac{d\gamma_3^{\infty}}{d\alpha} = 3P\gamma_2^{\infty} + 3P^2\gamma_1^{\infty} + P^3 \tag{166}$$

so that

$$\gamma_3^{\infty}(\alpha) = \alpha E[P^3] + 3\alpha^2 E[P^2]E[P] + \alpha^3 E^3[P]. \tag{167}$$

This approach can be used to compute higher moments as follows. First, note that γ_n^{∞} is the large system limit of

$$\mathrm{trace}[(\boldsymbol{R}_I + P\boldsymbol{s}_{K+1}\boldsymbol{s}_{K+1})^n].$$

Since $\boldsymbol{s}_{K+1}\boldsymbol{s}_{K+1}^{\dagger}$ is idempotent, we can express γ_n as the sum of terms of the form

$$\begin{aligned} \mathrm{trace}\Big[P^{n-q}&\boldsymbol{R}_I^{n_1}(\boldsymbol{s}_{K+1}\boldsymbol{s}_{K+1}^{\dagger})\boldsymbol{R}_I^{n_2}(\boldsymbol{s}_{K+1}\boldsymbol{s}_{K+1}^{\dagger})\cdots \\ &\boldsymbol{R}_I^{n_{m-1}}(\boldsymbol{s}_{K+1}\boldsymbol{s}_{K+1}^{\dagger})\boldsymbol{R}_I^{n_m}\Big] \\ &= P^{n-q}\boldsymbol{s}_{K+1}^{\dagger}\boldsymbol{R}_I^{n_1+n_m}(\boldsymbol{s}_{K+1}\boldsymbol{s}_{K+1}^{\dagger})\boldsymbol{R}_I^{n_2} \\ &\cdot(\boldsymbol{s}_{K+1}\boldsymbol{s}_{K+1}^{\dagger})\cdots\boldsymbol{R}_I^{n_{m-1}}\boldsymbol{s}_{K+1} \\ &\to P^{n-q}\gamma_{n_1+n_m}^{\infty}\gamma_{n_2}^{\infty}\cdots\gamma_{n_{m-1}}^{\infty} \end{aligned}$$

where $q = \sum_{i=1}^{m} n_i \le n$ and the fact $\mathrm{trace}(\boldsymbol{AB}) = \mathrm{trace}(\boldsymbol{BA})$ has been used for matrices $\boldsymbol{A}$ and $\boldsymbol{B}$. Enumerating the terms in the expansion and averaging over P gives $d\gamma_n^{\infty}/d\alpha$ as a function of $\gamma_1^{\infty}, \ldots, \gamma_{n-1}^{\infty}$ and the moments of F. Since these are polynomials in α, the expression can be integrated to give γ_n^{∞}.

APPENDIX C
PROOF OF LEMMA 3

From (30)

$$\begin{aligned} \boldsymbol{c}_{i+1} &= \left(\boldsymbol{I} - \sum_{l=0}^{i}\boldsymbol{c}_l\boldsymbol{c}_l^{\dagger}\right)(\boldsymbol{S}_I\boldsymbol{S}_I^{\dagger} + \sigma^2\boldsymbol{I})\boldsymbol{c}_i \\ &= \left(\boldsymbol{I} - \sum_{l=0}^{i}\boldsymbol{c}_l\boldsymbol{c}_l^{\dagger}\right)\boldsymbol{S}_I\boldsymbol{S}_I^{\dagger}\boldsymbol{c}_i \end{aligned}$$

since $\boldsymbol{c}_i^{\dagger}\boldsymbol{c}_l$ for $l \neq i$. If $\boldsymbol{c}_l$, $l \le i$, is independent of σ^2, then $\boldsymbol{c}_{i+1}$ is independent of σ^2. Since $\boldsymbol{c}_1 = \boldsymbol{s}_1$ is independent of σ^2, by induction $\boldsymbol{c}_i$ is independent of $\sigma^2\ \forall i$.

To prove that $E[|S_i|^2]$ and $E[|I_i|^2]$ are independent of σ^2, we note that S_i is the projection of d_i onto I_{i-1}, and

$$E[|S_i|^2] = \frac{|\boldsymbol{c}_i^{\dagger}\boldsymbol{R}\boldsymbol{c}_i^{\perp}|^2}{E[|I_{i-1}|^2]}.$$

Now

$$\boldsymbol{c}_i^{\dagger}\boldsymbol{R}\boldsymbol{c}_i^{\perp} = \boldsymbol{c}_i^{\dagger}(\boldsymbol{S}_I\boldsymbol{S}_I^{\dagger} + \boldsymbol{s}_1\boldsymbol{s}_1^{\dagger} + \sigma^2\boldsymbol{I})\boldsymbol{c}_i^{\perp} = \boldsymbol{c}_i^{\dagger}(\boldsymbol{S}_I\boldsymbol{S}_I^{\dagger})\boldsymbol{c}_i^{\perp}$$

and

$$\begin{aligned} E[|I_i|^2] &= E[|d_i|^2] - E[|S_i|^2] - E[|N_i|^2] \\ &= \boldsymbol{c}_i^{\dagger}(\boldsymbol{S}_I\boldsymbol{S}_I^{\dagger} + \boldsymbol{s}_1\boldsymbol{s}_1^{\dagger} + \sigma^2\boldsymbol{I})\boldsymbol{c}_i - E[|S_i|^2] - \sigma^2 \\ &= \boldsymbol{c}_i^{\dagger}(\boldsymbol{S}_I\boldsymbol{S}_I^{\dagger})\boldsymbol{c}_i - E[|S_i|^2] \end{aligned}$$

for $i > 1$. Hence, if $E[|I_{i-1}|^2]$ is independent of σ^2, then $E[|S_i|^2]$ and $E[|I_i|^2]$ are independent of σ^2. Since $E[|I_1|^2] = \alpha P$, the Lemma follows by induction.

APPENDIX D
DERIVATION OF (104)

Since $c_2^\dagger c_1 = 0$, we have

$$c_2^\dagger R c_2 = c_2^\dagger (S_I S_I^\dagger) c_2 + \sigma^2 \tag{168}$$

and from (99)

$$\begin{aligned} c_2^\dagger (S_I S_I^\dagger) c_2 &= \kappa_2^2 (\mathcal{P}_{c_1}^\perp [(S_I S_I^\dagger) c_1])^\dagger (S_I S_I^\dagger)(\mathcal{P}_{c_1}^\perp [(S_I S_I^\dagger) c_1]) \\ &= \kappa_2^2 [(S_I S_I^\dagger) c_1 - \gamma_1 c_1]^\dagger (S_I S_I^\dagger)[(S_I S_I^\dagger) c_1 - \gamma_1 c_1] \\ &= \kappa_2^2 [\gamma_3 - 2\gamma_1\gamma_2 + (\gamma_1)^3] \end{aligned} \tag{169}$$

where $\gamma_n = s_1^\dagger (S_I S_I^\dagger)^n s_1$. (We abuse notation somewhat by reusing the γ_n originally defined in (49).) Also,

$$\kappa_2^2 = \|\mathcal{P}_{c_1}^\perp (R c_1)\|^2 = \gamma_2 - \gamma_1^2. \tag{170}$$

Combining (168)–(170) with the computation of the large system limits $\gamma_n^\infty(\alpha, 0)$ in Appendix B gives (104).

REFERENCES

[1] L. L. Scharf, “The SVD and reduced rank signal processing,” *Signal Processing*, vol. 25, no. 2, pp. 113–133, 1991.
[2] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
[3] P. A. Zulch, J. S. Goldstein, J. R. Guerci, and I. S. Reed, “Comparison of reduced-rank signal processing techniques,” in *Proc. 32nd Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, CA, Nov. 1998.
[4] L. L. Scharf and J. K. Thomas, “Wiener filters in canonical coordinates for transform coding, filtering, and quantizing,” *IEEE Trans. Signal Processing*, vol. 46, pp. 647–654, Mar. 1998.
[5] D. A. Pados and S. N. Batalama, “Low-complexity blind detection of DS/CDMA signals: Auxiliary-vector receivers,” *IEEE Trans. Commun.*, vol. 45, pp. 1586–1594, Dec. 1997.
[6] M. L. Honig, “A comparison of subspace adaptive filtering techniques for DS-CDMA interference suppression,” in *Proc. IEEE MILCOM*, Monterey, CA, Nov. 1997, pp. 836–840.
[7] H. V. Poor and G. W. Wornell, Eds., *Wireless Communications: Signal Processing Perspectives*. Englewood Cliffs, NJ: Prentice-Hall, 1998.
[8] X. Wang and H. V. Poor, “Blind multiuser detection: A subspace approach,” *IEEE Trans. Inform. Theory*, vol. 44, pp. 677–690, Mar. 1998.
[9] M. L. Honig and J. S. Goldstein, “Adaptive reduced-rank residual correlation algorithms for DS-CDMA interference suppression,” in *Proc. 32nd Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, CA, Nov. 1998.
[10] D. A. Pados and S. N. Batalama, “Joint space-time auxiliary-vector filtering for DS/CDMA systems with antenna arrays,” *IEEE Trans. Commun.*, vol. 47, pp. 1406–1415, Sept. 1999.
[11] M. L. Honig and H. V. Poor, “Adaptive interference suppression,” in *Wireless Communications: Signal Processing Perspectives*, H. V. Poor and G. W. Wornell, Eds. Englewood Cliffs, NJ: Prentice-Hall, 1998, ch. 2, pp. 64–128.
[12] M. L. Honig, “Adaptive linear interference suppression for packet DS-CDMA,” *Europ. Trans. Telecommun.*, vol. 9, pp. 173–182, Mar./Apr. 1998.
[13] J. S. Goldstein and I. S. Reed, “A new method of Wiener filtering and its application to interference mitigation for communications,” in *Proc. IEEE MILCOM*, vol. 3, Monterey, CA, Nov. 1997, pp. 1087–1091.
[14] J. S. Goldstein, I. S. Reed, and L. L. Scharf, “A multistage representation of the Wiener filter based on orthogonal projections,” *IEEE Trans. Inform. Theory*, vol. 44, pp. 2943–2959, Nov. 1998.
[15] A. J. Grant and P. D. Alexander, “Random sequence multisets for synchronous code-division multiple-access channels,” *IEEE Trans. Inform. Theory*, vol. 44, pp. 2832–2836, Nov. 1998.
[16] S. Verdú and S. Shamai (Shitz), “Spectral efficiency of CDMA with random spreading,” *IEEE Trans. Inform. Theory*, vol. 45, pp. 622–640, Mar. 1999.
[17] D. N. C. Tse and S. V. Hanly, “Linear multiuser receivers: Effective interference, effective bandwidth and user capacity,” *IEEE Trans. Inform. Theory*, vol. 45, pp. 641–657, Mar. 1999.
[18] J. W. Silverstein and Z. D. Bai, “On the empirical distribution of eigenvalues of a class of large dimensional random matrices,” *J. Multivariate Anal.*, vol. 54, no. 2, pp. 175–192, 1995.
[19] K. A. Byerly and R. A. Roberts, “Output power based partial adaptive array design,” in *Proc. 23rd Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, CA, Oct. 1989, pp. 576–580.
[20] J. S. Goldstein and I. S. Reed, “Reduced rank adaptive filtering,” *IEEE Trans. Signal Processing*, vol. 45, pp. 492–496, Feb. 1997.
[21] R. Singh and L. B. Milstein, “Interference suppression for DS/CDMA,” *IEEE Trans. Commun.*, vol. 47, pp. 446–453, Mar. 1999.
[22] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Reading, MA: Addison-Wesley, 1993.
[23] U. Madhow and M. L. Honig, “MMSE interference suppression for direct-sequence spread-spectrum CDMA,” *IEEE Trans. Commun.*, vol. 42, pp. 3178–3188, Dec. 1994.
[24] G. Woodward, M. Honig, and B. Vucetic, “Lattice-based subspace decomposition for DS-CDMA detection,” in *Proc. Int. Symp. Spread Spectrum Techniques and Applications (ISSSTA)*, Sun City, South Africa, Sept. 1998.
[25] L. L. Scharf and C. T. Mullis, “Canonical coordinates and the geometry of inference, rate, and capacity,” *IEEE Trans. Signal Processing*, vol. 48, pp. 824–831, Mar. 2000.
[26] E. G. Ström and S. L. Miller, “Properties of the single-bit single-user MMSE receiver for DS-CDMA system,” *IEEE Trans. Commun.*, vol. 47, pp. 416–425, Mar. 1999.
[27] D. Jonsson, “Some limit theorems for the eigenvalues of a sample covariance matrix,” *J. Multivariate Anal.*, vol. 12, pp. 1–38, 1982.
[28] S. Moshavi, E. G. Kanterakis, and D. L. Schilling, “Multistage linear receivers for DS-CDMA systems,” *Int. J. Wireless Inform. Networks*, vol. 3, pp. 1–17, Jan. 1996.
[29] R. R. Müller and S. Verdú, “Spectral efficiency of low-complexity multiuser detectors,” in *Proc. IEEE Int. Symp. Information Theory*, Sorrento, Italy, June 2000.
[30] D. N. C. Tse and O. Zeitouni, “Linear multiuser receivers in random environments,” *IEEE Trans. Inform. Theory*, vol. 46, pp. 171–188, Jan. 2000.
[31] M. L. Honig and J. S. Goldstein, “Adaptive reduced-rank interference suppression based on the multistage wiener filter,” *IEEE Trans. Commun.*, submitted for publication.
[32] W. Xiao and M. Honig, “Convergence analysis of adaptive full-rank and multi-stage redu ced-rank interference suppression,” in *Proc. CISS’2000*, Princeton, NJ, Mar. 2000.