

Long-Term Krylov-Prefilter for Reduced-Dimension Receive Processing in DS-CDMA Systems

Guido Dietl, Michael Joham, Wolfgang Utschick, and Josef A. Nossek

Abstract— Currently, long-term channel properties are only used for reduced-dimension processing in eigen subspaces. In this paper, we introduce a prerake at the transmitter in the uplink of a Direct Sequence (DS) CDMA system with non-orthogonal codes in order to exploit long-term channel properties for reduced-dimension Wiener filtering in Krylov subspaces at the receiver. Simulation results show that Krylov outperforms eigen prefiltering with and without previous correlation. Moreover, long-term is better than or equal to short-term processing except for optimal filtering in Krylov subspaces assuming multipath environments.

Keywords— Adaptive Filtering, DS-CDMA, Long-Term Channel Properties, Multi-Stage Nested Wiener Filter, Non-Orthogonal Codes, Reduced-Rank Equalization, Space-Time Processing, Wireless Communications.

I. INTRODUCTION

TIME dispersive channels and multi-user access in mobile communication systems cause *Intersymbol* (ISI) and *Multiple Access Interference* (MAI). Since equalizers process the received signal in space and time to restore the desired signal, the dimension of the observation space has to be reduced to avoid high computational complexity. Additionally, long-term channel properties like path delays and angles of arrival can be exploited to decrease the number of subspace updates as shown by Brunner et al. [1] who designed reduced-dimension filters in *eigen subspaces*. Long-term processing employs second order statistics which include the expectation with respect to the fast changing channel weights (short-term channel property), where long-term channel properties are assumed to be slow changing or approximately constant.

The *Multi-Stage Nested Wiener Filter* (MSNWF) introduced by Goldstein et al. [2] can be seen as *Wiener Filter* (WF) operating in a *Krylov subspace* [3, 4]. Unfortunately, the Krylov prefilter depends upon the crosscorrelation vector between the observation vector and the desired signal which is zero in the long-term case assuming zero mean complex normal distributed channel weights.

Our contribution is to combine the channel with a short-term matched filter either at the transmitter or the receiver of the communication system. Thus, the crosscorrelation vector is no longer zero and a long-term Krylov prefilter can be used to reduce dimension of the observation space.

In the next section, we define the system model and the receiver structures for the uplink of a DS-CDMA system with non-orthogonal codes. Before we show simulation results in Section IV, we derive short-term and long-term

The authors are with the Institute for Circuit Theory and Signal Processing, Munich University of Technology, 80290 Munich, Germany (e-mail: {gudi,mijo,wout,jono}@nws.ei.tum.de).

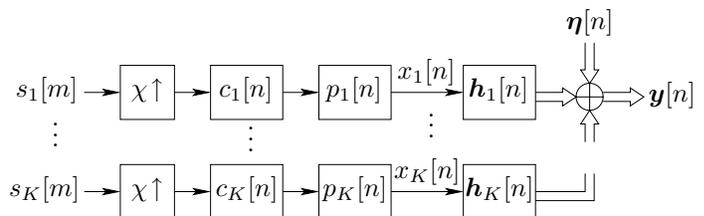


Fig. 1. System Model of DS-CDMA Uplink

second order statistics for time dispersive and non time dispersive channels in Section III.

II. SYSTEM MODEL AND RECEIVER STRUCTURES

We consider the uplink of a DS-CDMA system shown in Figure 1. The i.i.d. symbols $s_k[m]$ of user $k \in \{1, 2, \dots, K\}$ are spread with the chip sequence $c_k[n]$ of length χ chosen from a set of non-orthogonal codes. Thus, we have MAI even for flat Rayleigh fading channels. We employ a time processing short-term matched filter at the transmitter of user k , the so-called prerake [5]

$$p_k[n] = \sqrt{\frac{1}{\sum_{q=1}^{Q_k} |h_{k,q}|^2}} \sum_{q=1}^{Q_k} h_{k,q}^* \delta[\nu_{k,Q_k} - n - \nu_{k,q}],$$

where Q_k is the number of paths of user k , $h_{k,q}$ is the zero mean complex normal distributed channel weight of the q -th path with delay $\nu_{k,q}$, and ν_{k,Q_k} is the maximum delay of user k . The signal $x_k[n]$ of user k is transmitted over the channel

$$\mathbf{h}_k[n] = \sum_{q=1}^{Q_k} h_{k,q} \mathbf{a}_{k,q} \delta[n - \nu_{k,q}] \in \mathbb{C}^{N_a},$$

where $\mathbf{a}_{k,q}$ is the steering vector of the q -th path. N_a antennae receive the signal perturbed by *Additive White Gaussian Noise* (AWGN) $\boldsymbol{\eta}[n]$.

An alternative system might be obtained by moving the time processing short-term matched filter to the receiver. Nevertheless, we consider only preprocessing at the transmitter since matched filtering at the receiver increases computational complexity because of additional space processing due to multiple antennae and the estimation of short-term channel properties in the full observation space. In our approach, the receiver estimates only long-term channel properties in the full observation space whereas short-term channel properties may be determined in the reduced-dimension subspace. The main disadvantage of the chosen system is that short-term channel properties have to be

known at the transmitter which is no problem in *Time Division Duplex* (TDD) systems if we assume a channel coherence time greater than the time between two successive uplink and downlink slots (see e. g. [6]).

In the following, we deal with two different receiver structures. The first receiver (cf. Figure 2) prefilters directly the received signal whereas the second structure (cf. Figure 3) correlates the received signal with the chip sequence of the user of interest before applying the reduced-dimension WF.

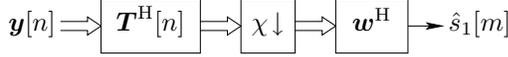


Fig. 2. Receiver without Previous Correlation

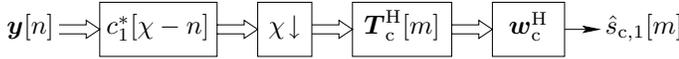


Fig. 3. Receiver with Previous Correlation

In order to compute the prefilter impulse responses $\mathbf{T}[n] \in \mathbb{C}^{N_a \times D}$ and $\mathbf{T}_c[m] \in \mathbb{C}^{N_a \times D}$ of length L and L_c (observation length), and the D -dimensional filter vectors \mathbf{w} and \mathbf{w}_c , respectively, we derive a matrix-vector model of the channel. We define the $N_a L$ -dimensional space-time observation vector $\bar{\mathbf{y}}[n] = [\mathbf{y}^T[n], \mathbf{y}^T[n-1], \dots, \mathbf{y}^T[n-L+1]]^T$, $n = \chi(m+1) - 1$, and the symbol vector $\mathbf{s}_k[m] = [s_k[m], s_k[m-1], \dots, s_k[m-L+1]]^T \in \mathbb{C}^{\mathcal{L}}$, $\mathcal{L} = \lceil (L + 2\nu_{k,Q_k})/\chi \rceil$, of user k , where $\lceil x \rceil$ denotes the next integer greater than or equal to x . Furthermore, we define the selection matrix

$$\mathbf{S}_{(\ell, M, N)} = \begin{bmatrix} \mathbf{0}_{M \times \ell} & \mathbf{1}_M & \mathbf{0}_{M \times (N-\ell)} \end{bmatrix} \in \{0, 1\}^{M \times (M+N)},$$

where $\mathbf{0}_{M \times \ell}$ denotes the $M \times \ell$ zero matrix and $\mathbf{1}_M$ the $M \times M$ identity matrix. The selection matrix selects the ℓ -th main diagonal of the $M \times (M+N)$ convolutional matrix.

Finally, we obtain the following representation of the channel:

$$\bar{\mathbf{y}}[n] = \sum_{k=1}^K \mathcal{H}_k \mathbf{s}_k[m] + \bar{\boldsymbol{\eta}}[n], \quad (1)$$

with the complete channel matrix

$$\mathcal{H}_k = \mathbf{H}_k \mathbf{P}_k \mathbf{C}_k^{(L+2\nu_{k,Q_k})},$$

where

$$\mathbf{H}_k = \sum_{\ell=0}^{\nu_{k,Q_k}} \mathbf{S}_{(\ell, L, \nu_{k,Q_k})} \otimes \mathbf{h}_k[\ell]$$

is the $N_a L \times (L + \nu_{k,Q_k})$ convolutional matrix of the physical channel,

$$\mathbf{P}_k = \sum_{\ell=0}^{\nu_{k,Q_k}} p_k[\ell] \mathbf{S}_{(\ell, L + \nu_{k,Q_k}, \nu_{k,Q_k})}$$

is the $(L + \nu_{k,Q_k}) \times (L + 2\nu_{k,Q_k})$ prerake matrix, and

$$\mathbf{C}_k^{(N)} = \mathbf{S}_{(0, N, N\chi - N)} (\mathbf{1}_N \otimes \mathbf{c}_k), \quad \mathcal{N} = \lceil N/\chi \rceil,$$

is the $N \times \mathcal{N}$ code matrix of user k with the code vector $\mathbf{c}_k = [c_k[0], c_k[1], \dots, c_k[\chi-1]]^T \in \mathbb{C}^{\mathcal{X}}$ of user k . Note that $\mathcal{N} = \mathcal{L}$ for $N = L + 2\nu_{k,Q_k}$. This choice of the code matrix implies already the upsampling of the symbol sequence. The operator ‘ \otimes ’ denotes the Kronecker multiplication.

Without loss of generality, we consider only the detection of the first user. We set $\mathcal{T} = [\mathbf{T}^H[0], \mathbf{T}^H[1], \dots, \mathbf{T}^H[L-1]]^H \in \mathbb{C}^{N_a L \times D}$, $D < N_a L$, $\mathcal{T}_c \in \mathbb{C}^{N_a L_c \times D}$, $D < N_a L_c$, in the same way, and get the estimate $\hat{s}_1[m]$ by applying the reduced-dimension WF $\mathbf{w} \in \mathbb{C}^D$, i. e.

$$\hat{s}_1[m] = \mathbf{w}^H \mathcal{T}^H \bar{\mathbf{y}}[n], \quad n = \chi(m+1) - 1.$$

In the second receiver case, we decorrelate the signal before reduced-rank Wiener processing and obtain

$$\hat{s}_1[m] = \mathbf{w}_c^H \mathcal{T}_c^H \mathbf{D}_1^H \bar{\mathbf{y}}[n], \quad n = \chi(m+1) - 1,$$

where the decorrelation matrix $\mathbf{D}_1 = \mathbf{C}_1^{(L)} \otimes \mathbf{1}_{N_a} \in \mathbb{C}^{N_a L \times N_a L_c}$ and $\mathbf{w}_c \in \mathbb{C}^D$. Note that $L_c = \lceil L/\chi \rceil$. It remains to determine the prefilter matrices \mathcal{T} and \mathcal{T}_c , and the reduced-dimension WF vectors \mathbf{w} and \mathbf{w}_c .

The known eigen based approach [7] chooses the D columns of the prefilter matrices \mathcal{T} and \mathcal{T}_c to be the D eigenvectors of the covariance matrices $\mathbf{R}_{\bar{\mathbf{y}}}$ and $\mathbf{D}_1^H \mathbf{R}_{\bar{\mathbf{y}}} \mathbf{D}_1$, respectively, corresponding to the largest eigenvalues. Thereby, the second order statistics defined in Section III include either short- or long-term information which is denoted in the following with the superscript ‘‘short’’ and ‘‘long’’.

In comparison to that, we choose the D columns of \mathcal{T} and \mathcal{T}_c to be the base vectors of the D -dimensional Krylov subspaces of the covariance matrices $\mathbf{R}_{\bar{\mathbf{y}}}$ and $\mathbf{D}_1^H \mathbf{R}_{\bar{\mathbf{y}}} \mathbf{D}_1$, and the crosscorrelation vectors $\mathbf{r}_{\bar{\mathbf{y}}, s_1}$ and $\mathbf{D}_1^H \mathbf{r}_{\bar{\mathbf{y}}, s_1}$, respectively (cf. e. g. [3]). The base vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_D$ of a D -dimensional Krylov subspace of a matrix \mathbf{A} and a vector \mathbf{b} , i. e.

$$\mathcal{K}^{(D)}(\mathbf{A}, \mathbf{b}) = \text{span} \{ \mathbf{b}, \mathbf{A}\mathbf{b}, \dots, \mathbf{A}^{D-1}\mathbf{b} \},$$

may be computed by applying the *Lanczos algorithm* [8]

$$\mathbf{x}_i = \frac{\mathbf{A}\mathbf{x}_{i-1} - \mathbf{x}_{i-2}^H \mathbf{A}\mathbf{x}_{i-1} \mathbf{x}_{i-2} - \mathbf{x}_{i-1}^H \mathbf{A}\mathbf{x}_{i-1} \mathbf{x}_{i-1}}{\| \mathbf{A}\mathbf{x}_{i-1} - \mathbf{x}_{i-2}^H \mathbf{A}\mathbf{x}_{i-1} \mathbf{x}_{i-2} - \mathbf{x}_{i-1}^H \mathbf{A}\mathbf{x}_{i-1} \mathbf{x}_{i-1} \|_2},$$

if the matrix \mathbf{A} is assumed to be Hermitian.

Finally, we define the *mean square errors*

$$\xi(\mathbf{w}) = \mathbb{E} \left\{ |\hat{s}_1[m] - s_1[m - \mu]|^2 \right\},$$

$$\xi_c(\mathbf{w}_c) = \mathbb{E} \left\{ |\hat{s}_{c,1}[m] - s_1[m - \mu]|^2 \right\},$$

where μ is the latency, and obtain the D -dimensional WFs \mathbf{w} and \mathbf{w}_c by solving the following optimization criteria:

$$\mathbf{w} = \arg \min_{\mathbf{w}'} \xi(\mathbf{w}') \quad \text{and} \quad \mathbf{w}_c = \arg \min_{\mathbf{w}_c'} \xi(\mathbf{w}_c').$$

The solutions are

$$\mathbf{w} = \left(\mathcal{T}^H \mathbf{R}_{\bar{\mathbf{y}}}^{\text{short}} \mathcal{T} \right)^{-1} \mathcal{T}^H \mathbf{r}_{\bar{\mathbf{y}}, s_1}^{\text{short}},$$

$$\mathbf{w}_c = \left(\mathcal{T}_c^H \mathbf{D}_1^H \mathbf{R}_{\bar{\mathbf{y}}}^{\text{short}} \mathbf{D}_1 \mathcal{T}_c \right)^{-1} \mathcal{T}_c^H \mathbf{D}_1^H \mathbf{r}_{\bar{\mathbf{y}}, s_1}^{\text{short}},$$

where the short-term covariance matrix $\mathbf{R}_{\bar{\mathbf{y}}}^{\text{short}} = \text{E}\{\bar{\mathbf{y}}[n]\bar{\mathbf{y}}[n]^H\}$ and the short-term crosscorrelation vector $\mathbf{r}_{\bar{\mathbf{y}},s_1}^{\text{short}} = \text{E}\{\bar{\mathbf{y}}[n]s_1^*[m-\mu]\}$ are derived and explained in the next section. Although we compare prefiltering based on short-term channel properties to prefiltering based on long-term channel properties, the reduced-dimension WFs depend always on short-term knowledge.

III. SHORT-TERM AND LONG-TERM CHANNEL PROPERTIES

The short-term crosscorrelation vector between the observation vector $\bar{\mathbf{y}}[n]$ and the desired signal $s_1[m-\mu]$, and the short-term covariance matrix of the observation $\bar{\mathbf{y}}[n]$ can be written as

$$\begin{aligned} \mathbf{r}_{\bar{\mathbf{y}},s_1}^{\text{short}} &= \sigma_{s_1}^2 \mathcal{H}_1 \mathbf{e}_{\mu+1} \quad \text{and} \\ \mathbf{R}_{\bar{\mathbf{y}}}^{\text{short}} &= \sum_{k=1}^K \sigma_{s_k}^2 \mathcal{H}_k \mathcal{H}_k^H + \sigma_{\eta}^2 \mathbf{1}, \end{aligned}$$

respectively, where short-term means instantaneously known or estimated channel weights $h_{k,q}$. Here, $\mathbf{e}_{\mu+1}$ denotes a unit vector with a one at the $(\mu+1)$ -th position and $\sigma_{s_k}^2 = \text{E}\{|s_k[m]|^2\}$ is the power of user k .

If we consider only long-term channel information, we have to compute the expectation value over all channel realizations where long-term properties, i.e. path delays, angles of arrival and codes, don't change. Thus, we take the expectation with respect to the time varying channel weights $h_{k,q}$. We get

$$\begin{aligned} \mathbf{r}_{\bar{\mathbf{y}},s_1}^{\text{long}} &= \sigma_{s_1}^2 \text{E}_{h_{1,1},\dots,h_{1,Q_1}} \{\mathcal{H}_1\} \mathbf{e}_{\mu+1} \quad \text{and} \\ \mathbf{R}_{\bar{\mathbf{y}}}^{\text{long}} &= \sum_{k=1}^K \sigma_{s_k}^2 \text{E}_{h_{k,1},\dots,h_{k,Q_k}} \{\mathcal{H}_k \mathcal{H}_k^H\} + \sigma_{\eta}^2 \mathbf{1}. \end{aligned}$$

Note that in general, $\mathbf{R}_{\bar{\mathbf{y}}}^{\text{long}}$ and $\mathbf{r}_{\bar{\mathbf{y}},s_1}^{\text{long}}$ are not analytically computable except for some special cases because the complete channel matrix \mathcal{H}_k already includes the convolution of the physical channel and the prerake. Thus, we have to build the expectation over the combination of higher order powers of the channel weights. Consequently, in our simulations presented in Section IV, we estimate the long-term second order statistics by averaging over several samples of short-term channel realizations in the general multipath channel case.

If we assume flat Rayleigh fading channels for all K users, i.e. $Q_k = 1$ and $\nu_{k,1} = 0 \forall k$, we may derive the long-term statistics analytically. Note that $\mu = 0$ in this case. The matrix-vector model of the channel (cf. Equation 1) can be simplified to

$$\bar{\mathbf{y}}[n] = \sum_{k=1}^K |h_{k,1}| (\mathbf{c}_k \otimes \mathbf{a}_{k,1}) s_k[m] + \bar{\boldsymbol{\eta}}[n],$$

where again $n = \chi(m+1) - 1$ but this time $L = \chi$.

It can easily be seen that the short-term statistics may

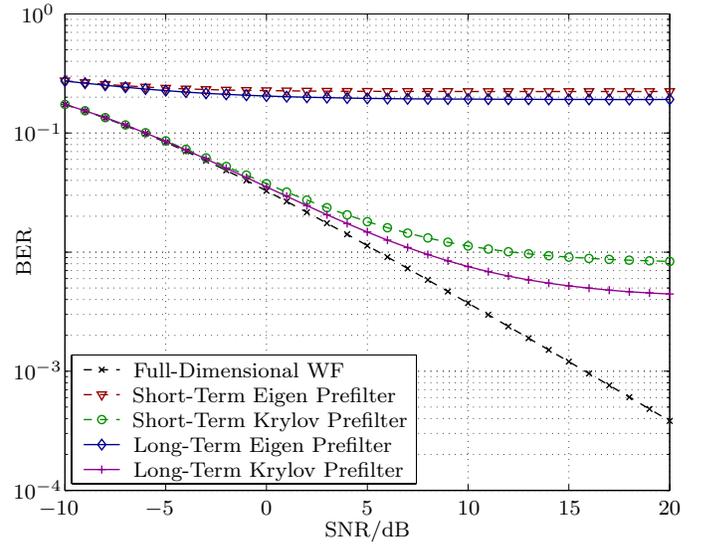


Fig. 4. BER without Previous Correlation (Flat Rayleigh Fading)

be written as

$$\begin{aligned} \mathbf{r}_{\bar{\mathbf{y}},s_1}^{\text{short}} &= \sigma_{s_1}^2 |h_{1,1}| (\mathbf{c}_1 \otimes \mathbf{a}_{1,1}) \quad \text{and} \\ \mathbf{R}_{\bar{\mathbf{y}}}^{\text{short}} &= \sum_{k=1}^K \sigma_{s_k}^2 |h_{k,1}|^2 (\mathbf{c}_k \otimes \mathbf{a}_{k,1}) (\mathbf{c}_k \otimes \mathbf{a}_{k,1})^H + \sigma_{\eta}^2 \mathbf{1}. \end{aligned}$$

The long-term second order statistics may be obtained by using the variances $\sigma_{h_{k,1}}^2 = \text{E}_{h_{k,1}}\{|h_{k,1}|^2\}$ of the channel weights. It follows [9]

$$\begin{aligned} \mathbf{r}_{\bar{\mathbf{y}},s_1}^{\text{long}} &= \sqrt{\frac{\pi}{4}} \sigma_{s_1}^2 \sigma_{h_{1,1}} (\mathbf{c}_1 \otimes \mathbf{a}_{1,1}) \quad \text{and} \\ \mathbf{R}_{\bar{\mathbf{y}}}^{\text{long}} &= \sum_{k=1}^K \sigma_{s_k}^2 \sigma_{h_{k,1}}^2 (\mathbf{c}_k \otimes \mathbf{a}_{k,1}) (\mathbf{c}_k \otimes \mathbf{a}_{k,1})^H + \sigma_{\eta}^2 \mathbf{1}. \end{aligned}$$

IV. SIMULATION RESULTS

In this section, we present simulation results where we compare eigen with Krylov based prefiltering and where we use either short-term or long-term information to compute the prefilter matrices. Furthermore, we give a comparison to the optimal WF without prefiltering.

First, we consider flat Rayleigh fading. We transmit QPSK symbols of $K = 4$ users with the same power over Rayleigh fading channels with unit variances. A spreading factor of $\chi = 4$ and $N_a = 4$ antennae establish a 16-dimensional observation space which is reduced to $D = 2$ dimensions. The channel is assumed to be known. Thus, the second order statistics can be computed analytically. Figure 4 and Figure 5 show the *Bit Error Rate* (BER) without and with previous correlation, respectively.

It can be seen that Krylov outperforms eigen prefiltering for both receiver structures and results in similar BERs as the full-dimensional WF for SNRs below 0 dB although it leads to a tremendous complexity decrease. Besides, long-term is always better than or equal to short-term processing. The short-term eigen prefilter chooses the subspace

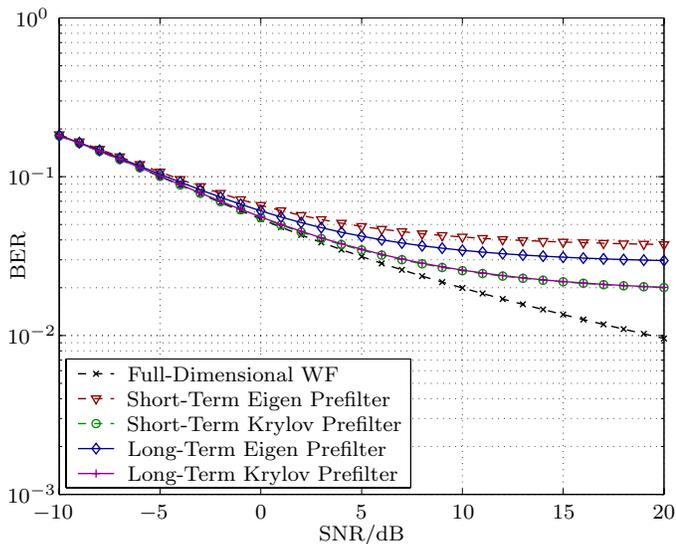


Fig. 5. BER with Previous Correlation (Flat Rayleigh Fading)

with the largest instantaneous signal power regardless if it is signal power of the user of interest or interference. In the long-term case the eigen prefilter projects the received signal in the subspace with the largest average signal power. Thus, it gains if we look at the BER averaged over several long-term and short-term channel realizations.

If we correlate the received signal before applying the reduced-dimension WF (cf. Figure 5), the BER increases since we lose degrees of freedom. Nevertheless, the eigen prefilter gains in this case because the correlation magnifies the desired signal compared to the interfering signals.

Second, we simulate a multipath fading channel with $Q_k = 2$ paths for all $K = 4$ users. The path delays $\nu_{k,1} = 0$ and $\nu_{k,2} = 1$. Note that we sample in multiples of a chip duration. We average over 1000 short-term channel realizations where path delays and angles of arrival are kept constant in order to estimate the long-term second order statistics. Figure 6 shows the results.

We see that optimal linear receive processing without any prefilter gains compared to the results of the flat Rayleigh fading channel due to diversity. For SNRs greater than -5 dB, long-term eigen prefiltering produces again smaller BERs than short-term eigen preprocessing. The reason for this fact is the same as mentioned above. Again, Krylov outperforms eigen prefiltering even in the long-term case, but in time dispersive channels, long-term Krylov preprocessing is worse than short-term Krylov preprocessing. This is due to the long-term crosscorrelation vector. Compared to the long-term covariance matrix and short-term second order statistics, the long-term crosscorrelation vector includes no information to combat ISI because it is the expectation of the combination of prerake and physical channel over all short-term channel realizations. Thus, it remains only the sum of the squared channel weights at one time position where the channel matches the prerake perfectly. The mixed products vanish due to expectation. In flat Rayleigh fading scenarios, long-term is better than

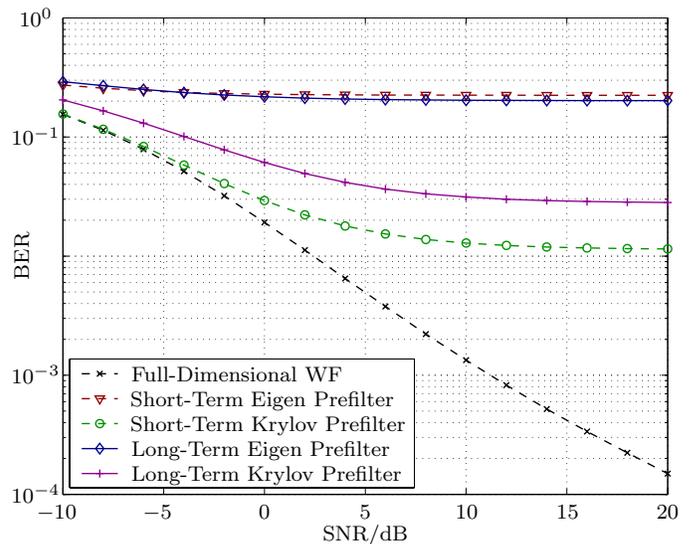


Fig. 6. BER without Previous Correlation (Multipath Channel)

or equal to Krylov prefiltering since there exists no ISI.

V. CONCLUSIONS

In this paper, we considered the uplink of a DS-CDMA system with non-orthogonal codes where we put a prerake at the transmitter in order to reduce dimension of the WF at the receiver with long-term Krylov prefiltering. Simulation results showed that Krylov increases the performance compared to eigen preprocessing. Moreover, long-term is better than short-term processing if we assume flat Rayleigh fading channels although it reduces computational complexity dramatically.

REFERENCES

- [1] C. Brunner, W. Utschick, and J. A. Nossek, "Exploiting the Short-Term and Long-Term Channel Properties in Space and Time: Eigenbeamforming Concepts for the BS in WCDMA," *European Transactions on Telecommunications*, vol. 12, no. 5, pp. 365–378, September/October 2001.
- [2] J. S. Goldstein, I. S. Reed, and L. L. Scharf, "A Multistage Representation of the Wiener Filter Based on Orthogonal Projections," *IEEE Transactions on Information Theory*, vol. 44, no. 7, pp. 2943–2959, November 1998.
- [3] M. L. Honig and W. Xiao, "Performance of Reduced-Rank Linear Interference Suppression," *IEEE Transactions on Information Theory*, vol. 47, no. 5, pp. 1928–1946, July 2001.
- [4] M. Joham, Y. Sun, M. D. Zoltowski, M. Honig, and J. S. Goldstein, "A New Backward Recursion for the Multi-Stage Nested Wiener Filter Employing Krylov Subspace Methods," in *Proc. Milcom 2001*, October 2001.
- [5] R. Esmailzadeh and M. Nakagawa, "Pre-RAKE Diversity Combination for Direct Sequence Spread Spectrum Mobile Communications Systems," *IEICE Transactions on Communications*, vol. E76-B, no. 8, pp. 1008–1015, August 1993.
- [6] F. Kowalewski and P. Mangold, "Joint Predistortion and Transmit Diversity," in *Proc. GLOBECOM*, November 2000, vol. 1, pp. 245–249.
- [7] H. Hotelling, "Analysis of a Complex of Statistical Variables into Principal Components," *Journal of Educational Psychology*, vol. 24, no. 6/7, pp. 417–441, 498–520, September/October 1933.
- [8] Y. Saad, *Iterative Methods for Sparse Linear Systems*, PWS – out of print, 1996.
- [9] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, 1991.