Iterative Multi-User Detection Using Reduced-Complexity Equalization

Michael Botsch, Guido Dietl, and Wolfgang Utschick

Institute for Circuit Theory and Signal Processing, Munich University of Technology, 80333 Munich, Germany
E-mail: {botsch,dietl,utschick}@tum.de

Abstract

In this paper, we consider a reduced-complexity equalization method for turbo *Multi-User Detection* (MUD) in a frequency-selective *Code Division Multiple Access* (CDMA) uplink scenario. The non-linear trellis-based detector which is normally used in turbo receivers, is replaced by the *Matrix Wiener Filter* (MWF), the optimal linear filter based on the *Mean Square Error* (MSE) criterion. To further reduce the complexity, the *Block Conjugate Gradient* (BCG) algorithm is used to approximate the MWF in a low-dimensional Krylov subspace. Additionally, second order statistics of non-stationary random processes are approximated by their time-invariant averages. Simulation results show that embedding the suboptimal MWF in a turbo receiver does not lead to significant performance loss, while reducing the computational complexity enormously.

1 Introduction

In an uplink scenario, MUD denotes the detection of data from different transmitters when the data is observed at the receiver in a non-orthogonal multiplex. Such observation may occur in a CDMA scenario if the spreading codes are non-orthogonal or if the channel between transmitters and receiver is frequency-selective. Thus, one of the main tasks at the receiver is to compensate multiple-access and intersymbol interference.

Assuming that the transmitters use channel coding, the optimal *Maximum A Posteriori* (MAP) receiver which performs jointly symbol detection and decoding to combat interference, is computationally not feasible. Therefore, Douillard et al. [1] introduced the turbo equalizer consisting of a MAP detector and a MAP decoder exchanging iteratively soft information about the coded data bits. Here, the two sources of diversity which are required in every turbo system, are on the one hand the channel encoders at the transmitters and on the other hand the multipath channel. Thus, the system can be interpreted as an iterative decoding scheme for serial concatenated codes [2] where the inner encoder is the channel and the inner decoder is the detector. Since the complexity of implementing a MAP detector is still very high, Wang and Poor [3] replaced the non-linear detector by the optimal linear detector, the *Wiener Filter* (WF). To further reduce the complexity, Dietl [4] approximated the WF in a low-dimensional Krylov subspace and showed that after a few turbo iterations, the performance loss due to the reduced-rank WF is negligible in a single-user uplink scenario.

In this paper, we consider a multi-user *Direct Sequence CDMA* (DS-CDMA) uplink scenario for mobile communications with multiple antennas at the receiver. Despite orthogonal spreading codes, multiple-access and intersymbol interference occur due to the frequency-selective *Single-Input Multiple-Output* (SIMO) channel. To combat interference, the observation vector at the receiver is high-dimensional leading to a high computational complexity even if linear techniques for joint multi-user detection and equalization are used. Conventional methods aiming to reduce the dimensionality of the observation vector, are the *Principal Component* (PC) [5] or the *Cross-Spectral* (CS) [6] method which approximate the MWF in an eigen subspace of the auto-covariance matrix of the observation. An approximation of the MWF in the Krylov subspace of the auto-covariance matrix of the observation and the cross-covariance vector between the observation and the desired signal, has been shown to outperform the PC and CS method. The *Multi-Stage Wiener Filter* (MSWF) developed by Goldstein et al. [7] is a possible implementation of the WF approximation in Krylov subspaces. In this paper another implementation is used, namely the BCG algorithm [8].

The main contribution of this paper is to present how computational complexity can be reduced by using the BCG algorithm for multi-user detection in combination with the turbo principle. Another important result is that although the receiver detects all users simultaneously by using the BCG algorithm, the performance loss compared to separate detection of all transmitters is negligible. Simulation results in Section 4 show the benefit of using the BCG algorithm.

In order to reduce complexity not only the Krylov subspace approximation of the MWF will be considered but also the approximation of second order statistics of non-stationary random processes by their time-invariant averages. This idea was developed for full-rank WFs [9] but holds also for the reduced-rank MWF presented in this paper.

Throughout the paper, vectors and matrices are denoted by lower and upper case bold letters, and random variables are written using *sans serif* fonts. The matrix $\boldsymbol{I}_n$ is the $n \times n$ identity matrix, $\boldsymbol{e}_i$ its i-th column and $\mathbf{0}_{m\times n}$ the $m\times n$ zero matrix. The operation '$\otimes$' denotes the Kronecker product, '$*$' convolution, $\operatorname{tr}\{\cdot\}$ the trace of a matrix, $\mathrm{E}\{\cdot\}$ expectation, $(\cdot)^{\mathrm{T}}$ transpose, $(\cdot)^{\mathrm{H}}$ Hermitian, $\lceil\cdot\rceil$ ceil, $\lfloor\cdot\rfloor$ floor, and $\|\cdot\|_2$ the Euclidean norm. The matrix $\boldsymbol{S}_{(\ell,M,N)} = [\mathbf{0}_{M\times\ell}, \boldsymbol{I}_M, \mathbf{0}_{M\times(N-\ell)}] \in \{0,1\}^{M\times(M+N)}$ is used for selection. The soft information of a binary random variable $\mathsf{d} \in \{0,1\}$ is represented by the *Log-Likelihood Ratio* (LLR) $l = \ln(\mathrm{P}(\mathsf{d}=0)/\mathrm{P}(\mathsf{d}=1))$ [10]. The auto-correlation matrix of the random vector $\boldsymbol{\mathsf{u}}$ is denoted as $\boldsymbol{R}_{\boldsymbol{\mathsf{u}}} = \mathrm{E}\{\boldsymbol{\mathsf{u}}\boldsymbol{\mathsf{u}}^{\mathrm{H}}\}$ and the cross-correlation matrix between the vectors $\boldsymbol{\mathsf{u}}$ and $\boldsymbol{\mathsf{v}}$ is $\boldsymbol{R}_{\boldsymbol{\mathsf{u}},\boldsymbol{\mathsf{v}}} = \mathrm{E}\{\boldsymbol{\mathsf{u}}\boldsymbol{\mathsf{v}}^{\mathrm{H}}\}$.

2 System Model

2.1 Transmitter

Fig. 1 depicts the block diagram of the communication system from the receiver's point of view. Because the signals at the transmitter are not known at the receiver they are modelled as random processes. There are K users in the uplink scenario, each aiming to transmit an information bit block $\boldsymbol{\mathsf{d}}_i \in \{0,1\}^B$, $i = 1,2,\ldots,K$, to the base station. After encoding the information blocks for user i with a rate r convolutional code, the resulting coded bit block $\boldsymbol{\mathsf{b}}_i \in \{0,1\}^{\frac{B}{r}}$ has the length $SQ = B/r$. The coded block is then interleaved using the same permutation matrix $\boldsymbol{\Pi} \in \{0,1\}^{SQ\times SQ}$ for all users leading to $\boldsymbol{\mathsf{b}}'_i \in \{0,1\}^{SQ}$, and mapped to the complex symbol blocks $\boldsymbol{\mathsf{s}}_i \in \mathbb{M}^S$ using the modulation alphabet $\mathbb{M}$ whose cardinality is 2^Q. The mapper can be described by the bijective function $M : \{0,1\}^Q \to \mathbb{M}$, $\boldsymbol{\mathsf{b}}'_{i,m+1} \mapsto \mathsf{s}_i[m] = M(\boldsymbol{\mathsf{b}}'_{i,m+1})$, where $\boldsymbol{\mathsf{b}}'_{i,m+1} = \boldsymbol{S}_{(mQ,Q,(S-1)Q)}\boldsymbol{\mathsf{b}}'_i$ and $\mathsf{s}_i[m] = \boldsymbol{e}_{m+1}^{\mathrm{T}}\boldsymbol{\mathsf{s}}_i$, $m = 0,1,\ldots,S-1$. The symbol block $\boldsymbol{\mathsf{s}}_i$ is then transformed into a chip block $\boldsymbol{\mathsf{x}}_i$ of length χS using a *Orthogonal Variable Spreading Factor* (OVSF) code of length χ, and transmitted over the channel.

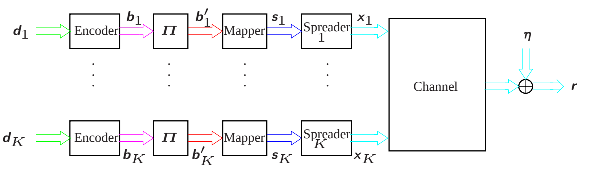

Figure 1. Block diagram of transmitters and channel

Considering the block $\boldsymbol{\mathsf{s}}_i$, $i = 1,2,\ldots,K$, as a vector containing elements of the sequence $\mathsf{s}_i[m]$, the spreading of $\mathsf{s}_i[m]$ using the OVSF codes $c_i[n]$ of length χ which leads to the chip sequence $\mathsf{x}_i[n]$, can be written using zero padding:

$$\begin{aligned}\mathsf{x}_i[n] &= c_i[n] * \sum_{m=-\infty}^{\infty} \mathsf{s}_i[m]\delta[n-\chi m] \\ &= c_i[n] * \mathsf{z}_i[n], \end{aligned} \tag{1}$$

where $\delta[n]$ denotes the unit impulse function and $\mathsf{z}_i[n]$ is defined implicitely by Eq. (1). Then the vector sequence $\boldsymbol{\mathsf{x}}'[n] = [\mathsf{x}_1[n], \mathsf{x}_2[n], \ldots, \mathsf{x}_K[n]]^{\mathrm{T}} \in \mathbb{C}^K$ is expressed using the matrix-vector notation

$$\boldsymbol{\mathsf{x}}'[n] = \boldsymbol{C}'\boldsymbol{\mathsf{z}}'[n], \tag{2}$$

with $\boldsymbol{C}' = [\boldsymbol{C}_0, \boldsymbol{C}_1, \ldots, \boldsymbol{C}_{\chi-1}] \in \mathbb{C}^{K\times K\chi}$ and $\boldsymbol{\mathsf{z}}'[n] = [\bar{\boldsymbol{\mathsf{z}}}^{\mathrm{T}}[n], \bar{\boldsymbol{\mathsf{z}}}^{\mathrm{T}}[n-1], \ldots, \bar{\boldsymbol{\mathsf{z}}}^{\mathrm{T}}[n-\chi+1]]^{\mathrm{T}} \in \mathbb{C}^{K\chi}$. Hereby, $\boldsymbol{C}_p = \operatorname{diag}\{c_1[p], c_2[p], \ldots, c_K[p]\} \in \{-1,+1\}^{K\times K}$ and $\bar{\boldsymbol{\mathsf{z}}}[n] = [\mathsf{z}_1[n], \mathsf{z}_2[n], \ldots, \mathsf{z}_K[n]]^{\mathrm{T}} \in \mathbb{C}^K$.

$\boldsymbol{\mathsf{x}}'[n]$ is transmitted via a frequency-selective MU-SIMO channel of order L with impulse response

$$\boldsymbol{H}'[n] = \sum_{\ell=0}^{L} \boldsymbol{H}_\ell \delta[n-\ell] \quad \in \mathbb{C}^{R\times K}, \tag{3}$$

where R is the number of receive antennas. The received signal vector

$$\boldsymbol{\mathsf{r}}'[n] = \boldsymbol{H}'[n] * \boldsymbol{\mathsf{x}}'[n] + \boldsymbol{\eta}'[n] \quad \in \mathbb{C}^R, \tag{4}$$

is perturbed by stationary *Additive White Gaussian Noise* (AWGN) $\boldsymbol{\eta}'[n] = [\eta_1[n], \eta_2[n], \ldots, \eta_R[n],]^{\mathrm{T}} \in \mathbb{C}^R$ with the circular complex normal distribution $\mathcal{N}_{\mathrm{c}}(\mathbf{0}_R, \sigma_{\boldsymbol{\eta}}^2 \boldsymbol{I}_R)$.

In order to compute the linear equalizer filter of order G in Section 3, an alternative matrix-vector model of the time-dispersive MU-SIMO channel is derived in the following. The vector $\boldsymbol{\mathsf{r}}[n] = [\boldsymbol{\mathsf{r}}'^{,\mathrm{T}}[n], \boldsymbol{\mathsf{r}}'^{,\mathrm{T}}[n-1], \ldots, \boldsymbol{\mathsf{r}}'^{,\mathrm{T}}[n-G]]^{\mathrm{T}} \in \mathbb{C}^{R(G+1)}$ is composed of $G+1$ adjacent received signal vectors $\boldsymbol{\mathsf{r}}'[n]$. Using the block Toeplitz matrix

$$\boldsymbol{H} = \sum_{\ell=0}^{L} \boldsymbol{S}_{(\ell,G+1,L)} \otimes \boldsymbol{H}_\ell \quad \in \mathbb{C}^{R(G+1)\times K(L+G+1)}, \tag{5}$$

Eq. (4) may be rewritten as

$$\boldsymbol{\mathsf{r}}[n] = \boldsymbol{H}\boldsymbol{\mathsf{x}}[n] + \boldsymbol{\eta}[n] \quad \in \mathbb{C}^{R(G+1)}, \tag{6}$$

where, analogous to $\boldsymbol{\mathsf{r}}[n]$, the vector $\boldsymbol{\mathsf{x}}[n] \in \mathbb{C}^{K(L+G+1)}$ is composed of $L+G+1$ adjacent transmit signal vectors $\boldsymbol{\mathsf{x}}'[n]$, and $\boldsymbol{\eta}[n]$ of $G+1$ adjacent noise vectors $\boldsymbol{\eta}'[n]$.

Taking into account Eq. (2) where the vector $\boldsymbol{\mathsf{z}}'[n]$ is built up by K values $\mathsf{s}_i[m]$ and $K(\chi-1)$ zeros, $\boldsymbol{\mathsf{x}}[n]$ can be rewritten as

$$\boldsymbol{\mathsf{x}}[n] = \boldsymbol{\mathcal{C}}\boldsymbol{\mathsf{s}}[m], \tag{7}$$

with $\boldsymbol{\mathsf{s}}[m] = [\boldsymbol{\mathsf{s}}'^{,\mathrm{T}}[m], \boldsymbol{\mathsf{s}}'^{,\mathrm{T}}[m-1], \ldots, \boldsymbol{\mathsf{s}}'^{,\mathrm{T}}[m-e_{\mathrm{s}}+1]]^{\mathrm{T}} \in \mathbb{C}^{Ke_{\mathrm{s}}}$, $\boldsymbol{\mathsf{s}}'[m] = [\mathsf{s}_1[m], \mathsf{s}_2[m], \ldots, \mathsf{s}_K[m]]^{\mathrm{T}}$, $e_{\mathrm{s}} = \lceil (L+G+\chi)/\chi \rceil$, $\boldsymbol{\mathcal{C}} \in \{-1,+1\}^{K(L+G+1)\times Ke_{\mathrm{s}}}$, and $n = \chi m$. Thus, Eq. (6) can be rewritten as

$$\boldsymbol{\mathsf{r}}[n] = \boldsymbol{H}\boldsymbol{\mathcal{C}}\boldsymbol{\mathsf{s}}[m] + \boldsymbol{\eta}[n], \tag{8}$$

where $\boldsymbol{\mathcal{C}}$ is defined implicitly.

2.2 Receiver

At the receiver side, the *Soft-Input Soft-Output* (SISO) detector calculates the extrinsic information $\boldsymbol{l}_{\mathrm{ext},i}^{(\mathrm{Det})} \in \mathbb{R}^{SQ}$ about the interleaved and coded bit block $\boldsymbol{b}_i'$ for each user i. Roughly speaking, extrinsic information refers to the incremental information about the current bit obtained through the decoding process from all other bits. The extrinsic information $\boldsymbol{l}_{\mathrm{ext},i}^{(\mathrm{Det})}$ is computed using the observation signal block $\boldsymbol{r}$ and the *a priori* information $\boldsymbol{l}_{\mathrm{apr},i}^{(\mathrm{Det})} \in \mathbb{R}^{SQ}$ about $\boldsymbol{b}_i'$ which is calculated by interleaving the extrinsic information $\boldsymbol{l}_{\mathrm{ext},i}^{(\mathrm{Dec})} \in \mathbb{R}^{SQ}$ about the coded bit block $\boldsymbol{b}_i$ computed by the decoder at the previous iteration step. A detailed description of the detector is given in the next section.

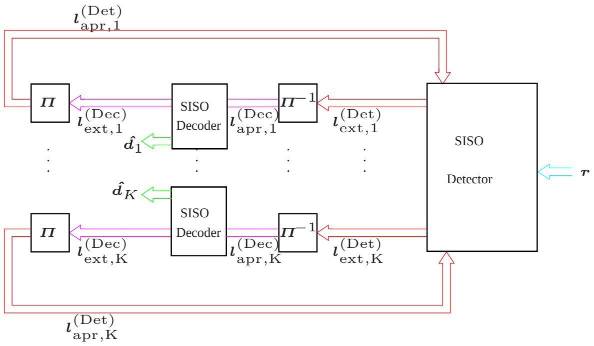

Figure 2. Block diagram of receiver

The SISO decoders fulfil two tasks. During the iterative process, the i-th decoder computes the extrinsic information $\boldsymbol{l}_{\mathrm{ext},i}^{(\mathrm{Dec})}$ about the coded bit block $\boldsymbol{b}_i$ from its *a posteriori* LLR vector $\boldsymbol{l}_{\mathrm{apo},i}^{(\mathrm{Dec})} \in \mathbb{R}^{SQ}$ using its *a priori* information $\boldsymbol{l}_{\mathrm{apr},i}^{(\mathrm{Dec})} \in \mathbb{R}^{SQ}$. The *a priori* information $\boldsymbol{l}_{\mathrm{apr},i}^{(\mathrm{Dec})}$ is the deinterleaved extrinsic information $\boldsymbol{l}_{\mathrm{ext},i}^{(\mathrm{Det})}$ about the interleaved coded bit block $\boldsymbol{b}_i'$ at the output of the detector at the previous iteration step. The second task is to compute $\hat{\boldsymbol{d}}_i$, the detected information bit block $\boldsymbol{d}_i$ after the last turbo iteration. Being a MAP decoder, the k-th bit, $\hat{d}_{i,k}$, $k = 1, 2, \ldots, B$, in the block $\hat{\boldsymbol{d}}_i$ is obtained from the optimization problem

$$\hat{d}_{i,k} = \underset{d \in \{0,1\}}{\operatorname{argmax}} \, \mathrm{P}(\boldsymbol{e}_k^{\mathrm{T}} \boldsymbol{d}_i = d | \boldsymbol{I}_{\mathrm{apr},i}^{(\mathrm{Dec})} = \boldsymbol{l}_{\mathrm{apr},i}^{(\mathrm{Dec})}), \qquad (9)$$

where $\boldsymbol{l}_{\mathrm{apr},i}^{(\mathrm{Dec})}$ is a realization of the vector random variable $\boldsymbol{I}_{\mathrm{apr},i}^{(\mathrm{Dec})}$. For more details see e. g. [11].

3 Linear Detection and Rank Reduction

3.1 A Priori Based Linear Equalization

In order to compute $\boldsymbol{l}_{\mathrm{ext},i}^{(\mathrm{Det})}$, the optimal detector based on the MAP criterion has an extremely high computational complexity and therefore, in the following, a reduced-rank linear equalizer based on the MSE-criterion is presented. The detector's inputs are the received signal and the *a priori* information about the interleaved coded bit block. Due to the *a priori* information, the random processes $s_i[m]$, $i = 1, 2, \ldots, K$, must be assumed to be non-zero mean and non-stationary. Taking into account the latency time ν on chip level, the optimal linear detector $(\boldsymbol{W}[m], \boldsymbol{a}[m])$ must compute an estimate of $\boldsymbol{s}'[m]$ according to

$$\begin{aligned} \hat{\boldsymbol{s}}'[m] &= [\hat{s}_1[m], \ldots, \hat{s}_K[m]]^{\mathrm{T}} \\ &= \boldsymbol{W}[m]\boldsymbol{r}[\chi m + \nu] + \boldsymbol{a}[m], \end{aligned} \qquad (10)$$

where

$$\boldsymbol{r}[\chi m + \nu] = \boldsymbol{H}\boldsymbol{\mathcal{C}}_\nu \boldsymbol{s}_\nu[m] + \boldsymbol{\eta}[\chi m + \nu]. \qquad (11)$$

The matrix $\boldsymbol{\mathcal{C}}_\nu \in \{-1, +1\}^{K(L+G+1) \times K e_{\mathrm{s}}^\nu}$ and the vector $\boldsymbol{s}_\nu[m] \in \mathbb{C}^{K e_{\mathrm{s}}^\nu}$ are the modified versions of $\boldsymbol{\mathcal{C}}$ and $\boldsymbol{s}[m]$ from Eq. (8) in order to consider the latency time ν, i. e., e_{s} must be replaced by e_{s}^ν which is defined by $L, G, \chi,$ and ν. With $e_{\mathrm{z}} = L + G + 2\chi - \chi\lceil (L + G + \chi)/\chi \rceil - 1$, e_{s}^ν can be computed by $e_{\mathrm{s}}^\nu = \lceil (L + G + \chi)/\chi \rceil$ for $\nu \leq e_{\mathrm{z}}$ and $e_{\mathrm{s}}^\nu = \lceil (L + G + \chi)/\chi \rceil + \lfloor \nu/\chi \rfloor - \lceil (\nu - e_{\mathrm{z}})/\chi \rceil$ for $\nu > e_{\mathrm{z}}$.

The vector $\boldsymbol{a}[m]$ must be used in order to take into account the mean $\mathrm{E}\{\boldsymbol{s}'[m]\}$ of $\boldsymbol{s}'[m] = \boldsymbol{S}_\nu \boldsymbol{s}_\nu[m]$, with $\boldsymbol{S}_\nu = \boldsymbol{S}_{(K\lfloor \nu/\chi \rfloor, K, K(e_{\mathrm{s}}^\nu - 1))}$. With

$$\xi_m(\boldsymbol{\mathcal{W}}, \boldsymbol{a}) = \mathrm{E}\left\{ \| \boldsymbol{s}'[m] - (\boldsymbol{\mathcal{W}}\boldsymbol{r}[\chi m + \nu] + \boldsymbol{a}) \|_2^2 \right\} \qquad (12)$$

being the MSE produced by the detector $(\boldsymbol{\mathcal{W}}, \boldsymbol{a})$ when estimating $\boldsymbol{s}'[m]$, we obtain the *Minimum MSE* (MMSE) based optimal linear equalizer by solving the optimization problem

$$(\boldsymbol{W}[m], \boldsymbol{a}[m]) = \underset{(\boldsymbol{\mathcal{W}}, \boldsymbol{a})}{\operatorname{argmin}} \, \mathrm{E}\{ \| \boldsymbol{s}'[m] - \boldsymbol{\mathcal{W}}\boldsymbol{r}[\chi m + \nu] - \boldsymbol{a} \|_2^2 \}. \qquad (13)$$

Using Eq. (11) and the relation $\boldsymbol{s}'[m] = \boldsymbol{S}_\nu \boldsymbol{s}_\nu[m]$, $\xi_m(\boldsymbol{\mathcal{W}}, \boldsymbol{a})$ can be rewritten as

$$\begin{aligned} \xi_m(\boldsymbol{\mathcal{W}}, \boldsymbol{a}) = \mathrm{tr}\{ \mathrm{E}\{ & (\boldsymbol{\mathcal{W}}\boldsymbol{H}\boldsymbol{\mathcal{C}}_\nu \boldsymbol{s}_\nu[m] + \boldsymbol{\mathcal{W}}\boldsymbol{\eta}[\chi m + \nu] \\ & + \boldsymbol{a} - \boldsymbol{S}_\nu \boldsymbol{s}_\nu[m])(\boldsymbol{\mathcal{W}}\boldsymbol{H}\boldsymbol{\mathcal{C}}_\nu \boldsymbol{s}_\nu[m] \\ & + \boldsymbol{\mathcal{W}}\boldsymbol{\eta}[\chi m + \nu] + \boldsymbol{a} - \boldsymbol{S}_\nu \boldsymbol{s}_\nu[m])^{\mathrm{H}} \} \}. \end{aligned}$$

The minimum of $\xi_m(\boldsymbol{\mathcal{W}}, \boldsymbol{a})$ can be found by jointly solving the equations resulting from setting the derivatives of $\xi_m(\boldsymbol{\mathcal{W}}, \boldsymbol{a})$ with respect to $\boldsymbol{\mathcal{W}}$ and $\boldsymbol{a}$, respectively to zero. This leads to

$$\boldsymbol{a}[m] = -\boldsymbol{W}[m]\boldsymbol{H}\boldsymbol{\mathcal{C}}_\nu \, \mathrm{E}\{\boldsymbol{s}_\nu[m]\} + \boldsymbol{S}_\nu \, \mathrm{E}\{\boldsymbol{s}_\nu[m]\}, \qquad (14)$$

and

$$\boldsymbol{W}[m] = \boldsymbol{S}_\nu \boldsymbol{R}_{\boldsymbol{s}_\nu}[m] \boldsymbol{\mathcal{C}}_\nu^{\mathrm{T}} \boldsymbol{H}^{\mathrm{H}} (\boldsymbol{H}\boldsymbol{\mathcal{C}}_\nu \boldsymbol{R}_{\boldsymbol{s}_\nu}[m] \boldsymbol{\mathcal{C}}_\nu^{\mathrm{T}} \boldsymbol{H}^{\mathrm{H}} + \boldsymbol{R}_{\boldsymbol{\eta}})^{-1}, \qquad (15)$$

where $\boldsymbol{R}_{\boldsymbol{s}_\nu}[m]$ is the covariance matrix of $\boldsymbol{s}_\nu[m]$.

In order to apply the *turbo principle* [12], the detector from Eqs. (14) and (15) must be modified to take into account the fact that there is no *a priori* information in the extrinsic information [4]. Applied to the design of the equalizer, this means that $\mathrm{E}\{\boldsymbol{s}'[m]\} = \boldsymbol{0}_{K \times 1}$ and $\sigma_{\boldsymbol{s},i}^2[m] = r_{\boldsymbol{s},i}[m] = 2^{-Q} \sum_{s \in \mathbb{M}} |s|^2 =: \varrho_{\boldsymbol{s},i}$,

$i = 1, \ldots, K$, must be assumed. Therefore, in Eq. (10), $\boldsymbol{W}[m]$ is replaced by

$$\boldsymbol{\Omega}[m]=\boldsymbol{S}_\nu \boldsymbol{\Gamma}_{\boldsymbol{s}_\nu}[m]\boldsymbol{\mathcal{C}}_\nu^{\mathrm{T}}\boldsymbol{H}^{\mathrm{H}}(\boldsymbol{H}\boldsymbol{\mathcal{C}}_\nu \boldsymbol{\Gamma}_{\boldsymbol{s}_\nu}[m]\boldsymbol{\mathcal{C}}_\nu^{\mathrm{T}}\boldsymbol{H}^{\mathrm{H}}+\boldsymbol{R}_{\boldsymbol{\eta}})^{-1}, \tag{16}$$

and $\boldsymbol{a}[m]$ by

$$\boldsymbol{\alpha}[m] = -\boldsymbol{\Omega}[m]\boldsymbol{H}\boldsymbol{\mathcal{C}}_\nu \boldsymbol{\mu}_{\boldsymbol{s}_\nu}[m], \tag{17}$$

where $\boldsymbol{\Gamma}_{\boldsymbol{s}_\nu}[m]$ is the adjusted auto-covariance matrix and $\boldsymbol{\mu}_{\boldsymbol{s}_\nu}[m]$ the adjusted mean of the vector $\boldsymbol{s}_\nu[m]$. The difference between $\boldsymbol{R}_{\boldsymbol{s}_\nu}[m]$ and $\boldsymbol{\Gamma}_{\boldsymbol{s}_\nu}[m]$ is that in the adjusted auto-covariance matrix the variances $\sigma^2_{\boldsymbol{s},i}[m]$ of the $(\lfloor \nu/\chi \rfloor + 1)$-th diagonal $K \times K$-block are replaced by the second moments $\varrho_{\boldsymbol{s},i}$ where $i = 1, 2, \ldots, K$. The non-diagonal elements in $\boldsymbol{R}_{\boldsymbol{s}_\nu}[m]$ and $\boldsymbol{\Gamma}_{\boldsymbol{s}_\nu}[m]$ vanish because the symbols $\boldsymbol{s}_\nu[m]$ are temporally uncorrelated due to the interleaver and because we assume that the sent signals of the different users are uncorrelated. In $\boldsymbol{\mu}_{\boldsymbol{s}_\nu}[m]$, the $(K\lfloor \nu/\chi \rfloor + 1)$-th to the $(K\lfloor \nu/\chi \rfloor + K)$-th entry are zero due to the requirement that $\mathrm{E}\{\boldsymbol{s}'[m]\} = \boldsymbol{0}_{K\times 1}$.

For a given modulation scheme, it is easy to calculate $r_{\boldsymbol{s},i}[m]$ and therefore, the only missing part in order to compute $\boldsymbol{\Gamma}_{\boldsymbol{s}_\nu}[m]$ and $\boldsymbol{\mu}_{\boldsymbol{s}_\nu}[m]$ are the expectations $\mathrm{E}\{\boldsymbol{s}_i[m]\}$ with $i = 1, 2, \ldots, K$ and $m = 1, 2, \ldots, S-1$. In the following, a way to express $\mathrm{E}\{\boldsymbol{s}_i[m]\}$ is presented using the *a priori* LLR $\boldsymbol{l}^{(\mathrm{Det})}_{\mathrm{apr},i}$ which is delivered by the decoder. It holds that

$$\mathrm{E}\left\{\boldsymbol{s}_i[m]\right\} = \sum_{s\in\mathbb{M}} s \mathrm{P}(\boldsymbol{s}_i[m] = s). \tag{18}$$

With $\boldsymbol{D} : \mathbb{M} \to \{0,1\}^Q$, $s_i[m] \mapsto \boldsymbol{b}'_{i,m+1} = \boldsymbol{D}(s_i[m])$ being the inverse function of the mapper M, for $\mathrm{P}(\boldsymbol{s}_i[m] = s) = \mathrm{P}(\boldsymbol{b}'_{i,m+1} = \boldsymbol{D}(s))$ it can be written

$$\begin{aligned}\mathrm{P}(\boldsymbol{s}_i[m] = s) &= \mathrm{P}(\boldsymbol{b}'_{i,m+1} = \boldsymbol{D}(s)) \\ &= \prod_{q=1}^{Q} \mathrm{P}(b'_{i,mQ+q} = \boldsymbol{e}_q^{\mathrm{T}}\boldsymbol{D}(s)).\end{aligned} \tag{19}$$

With

$$\mathrm{P}(b'_{i,mQ+q} = \pm 1) = \frac{\exp(\pm l^{(\mathrm{Det})}_{\mathrm{apr},i,mQ+q})}{(1+\exp(\pm l^{(\mathrm{Det})}_{\mathrm{apr},i,mQ+q}))}, \tag{20}$$

one obtains

$$\begin{aligned}\mathrm{P}(b'_{i,mQ+q} = \boldsymbol{e}_q^{\mathrm{T}}\boldsymbol{D}(s)) =& \frac{1}{2}\Big(1 + (1 - 2\boldsymbol{e}_q^{\mathrm{T}}\boldsymbol{D}(s)) \\ &\cdot \tanh \frac{l^{(\mathrm{Det})}_{\mathrm{apr},i,mQ+q}}{2}\Big).\end{aligned} \tag{21}$$

Now, this expression and Eq. (19) can be plugged into Eq. (18) and depending on the modulation alphabet which is used, the expectation $\mathrm{E}\{\boldsymbol{s}_i[m]\}$ can be computed. For QPSK, one obtains

$$\begin{aligned}\mathrm{E}\{\boldsymbol{s}_i[m]\} = \frac{1}{\sqrt{2}}\Bigg(&\tanh \frac{l^{(\mathrm{Det})}_{\mathrm{apr},i,mQ+1}}{2} \\ &+ \mathrm{j} \tanh \frac{l^{(\mathrm{Det})}_{\mathrm{apr},i,mQ+2}}{2}\Bigg).\end{aligned} \tag{22}$$

With this result and knowing that $r_{\boldsymbol{s},i}[m] = 1$ for QPSK, one can firstly compute $\boldsymbol{\Gamma}_{\boldsymbol{s}_\nu}[m]$ and $\boldsymbol{\mu}_{\boldsymbol{s}_\nu}[m]$, and then $\boldsymbol{\Omega}[m]$ and $\boldsymbol{\alpha}[m]$.

3.2 Soft Demapping

The second step to obtain the SISO detector is to softly demap of the estimated symbols which were calculated using $(\boldsymbol{\Omega}[m], \boldsymbol{\alpha}[m])$, to the extrinsic LLRs of the coded and interleaved bits. It can be written

$$l^{(\mathrm{Det})}_{\mathrm{ext},i,mQ+q} = \ln \frac{\displaystyle\sum_{\substack{\boldsymbol{b}_i \in \{0,1\}^Q \\ \boldsymbol{e}_q^{\mathrm{T}}\boldsymbol{b}_i = 0}} g_1(\hat{\boldsymbol{s}}_i[m], \boldsymbol{b}_i)}{\displaystyle\sum_{\substack{\boldsymbol{b}_i \in \{0,1\}^Q \\ \boldsymbol{e}_q^{\mathrm{T}}\boldsymbol{b}_i = 1}} g_1(\hat{\boldsymbol{s}}_i[m], \boldsymbol{b}_i)}, \tag{23}$$

with

$$\begin{aligned}g_1(\hat{\boldsymbol{s}}_i[m], \boldsymbol{b}_i) =& \mathrm{p}_{\hat{\boldsymbol{s}}_i[m]}(\hat{s}_i[m]|\boldsymbol{b}'_{i,m+1} = \boldsymbol{b}_i) \\ &\cdot \prod_{\substack{\ell=1 \\ \ell \neq q}}^{Q} \mathrm{P}(\boldsymbol{e}_\ell^{\mathrm{T}}\boldsymbol{b}'_{i,m+1} = \boldsymbol{e}_\ell^{\mathrm{T}}\boldsymbol{b}_i),\end{aligned} \tag{24}$$

where $q = 1, 2, \ldots, Q$ and $m = 0, 1, \ldots, S-1$. $\mathrm{P}(\boldsymbol{e}_\ell^{\mathrm{T}}\boldsymbol{b}'_{i,m+1} = \boldsymbol{e}_\ell^{\mathrm{T}}\boldsymbol{b}_i)$ can be expressed by replacing $\boldsymbol{D}(s)$ by $\boldsymbol{b}_i$ in Eqs. (19) and (21). The probability density function $\mathrm{p}_{\hat{\boldsymbol{s}}_i[m]}(\hat{s}_i[m]|\boldsymbol{b}'_{i,m+1} = \boldsymbol{b}_i)$ can be approximated by the complex Gaussian distribution [9], [13]

$$\mathrm{p}_{\hat{\boldsymbol{s}}_i[m]}(\hat{s}_i[m]|\boldsymbol{b}'_{i,m+1} = \boldsymbol{b}_i) = \frac{\exp\left(-\frac{|\hat{s}_i[m]-\mu_{(\boldsymbol{b}_i)}[m]|^2}{\sigma^2_{(\boldsymbol{b}_i)}[m]}\right)}{\pi\sigma^2_{(\boldsymbol{b}_i)}[m]}, \tag{25}$$

where

$$\begin{aligned}\mu_{(\boldsymbol{b}_i)}[m] &= \mathrm{E}\left\{\hat{\boldsymbol{s}}_i[m]|\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{S}_\nu \boldsymbol{s}_\nu[m] = M(\boldsymbol{b}_i)\right\} \\ &= M(\boldsymbol{b}_i)\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{\Omega}[m]\boldsymbol{H}\boldsymbol{\mathcal{C}}_\nu \boldsymbol{S}_\nu^{\mathrm{T}}\boldsymbol{e}_i,\end{aligned} \tag{26}$$

and

$$\begin{aligned}\sigma^2_{(\boldsymbol{b}_i)}[m] &= \mathrm{E}\{|\hat{\boldsymbol{s}}_i[m]|^2|\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{S}_\nu \boldsymbol{s}_\nu[m] = M(\boldsymbol{b}_i)\} \\ &\quad - |\mu_{(\boldsymbol{b}_i)}[m]|^2 \\ &= \frac{\mu_{(\boldsymbol{b}_i)}[m]}{M(\boldsymbol{b}_i)} - |\mu_{(\boldsymbol{b}_i)}[m]|^2.\end{aligned} \tag{27}$$

Hereby, it has been assumed that $\varrho_{\boldsymbol{s},i} = |M(\boldsymbol{b}_i)|^2$ and it has been used that

$$\hat{\boldsymbol{s}}_i[m] = \boldsymbol{e}_i^{\mathrm{T}}(\boldsymbol{\Omega}[m]\boldsymbol{r}[\chi m + \nu] + \boldsymbol{\alpha}[m]). \tag{28}$$

Plugging these results into Eq. (23), one obtains for QPSK modulation

$$l^{(\mathrm{Det})}_{\mathrm{ext},i,mQ+1} + \mathrm{j}\, l^{(\mathrm{Det})}_{\mathrm{ext},i,mQ+2} = \frac{\sqrt{8}\ \hat{s}_i[m]}{1 - \boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{\Omega}[m] \boldsymbol{H} \boldsymbol{\mathcal{C}}_\nu \boldsymbol{S}_\nu^{\mathrm{T}} \boldsymbol{e}_i}. \tag{29}$$

At this point, all required relations have been derived to determine $\boldsymbol{l}^{(\mathrm{Det})}_{\mathrm{ext},i}$ given the received signal block $\boldsymbol{r}$ and $\boldsymbol{l}^{(\mathrm{Det})}_{\mathrm{apr},i}$ with $i = 1, 2, \ldots, K$, and thereby all components which built up the detector.

3.3 Complexity Reduction

The computation of $\boldsymbol{\Omega}[m]$ and $\boldsymbol{\alpha}[m]$ using Eqs. (16) and (17) must be carried out for every vector $\hat{\boldsymbol{s}}'[m] \in \mathbb{C}^K$, $m = 0, 1, \ldots, S-1$, at each turbo iteration. Although, using an MMSE based linear equalizer requires much less resources than a MAP based equalizer, the complexity is still large. Therefore, two attempts are presented to reduce the receiver's complexity, knowing that the matrix inversion in Eq. (16) causes the major computational work.

The first attempt is to replace the second order matrix $\boldsymbol{\Gamma}_{\boldsymbol{s}_\nu}[m]$ by its time average $\bar{\boldsymbol{\Gamma}}_{\boldsymbol{s}_\nu}$ which is the mean over all $\boldsymbol{\Gamma}_{\boldsymbol{s}_\nu}[m]$ with $m = 0, 1, \ldots, S-1$ at the current turbo iteration. This reduces the complexity enormously, because the matrix inversion in Eq. (16) must be calculated only once per iteration for the entire block leading to

$$\bar{\boldsymbol{\Omega}} = \boldsymbol{S}_\nu \bar{\boldsymbol{\Gamma}}_{\boldsymbol{s}_\nu} \boldsymbol{\mathcal{C}}_\nu^{\mathrm{T}} \boldsymbol{H}^{\mathrm{H}} (\boldsymbol{H} \boldsymbol{\mathcal{C}}_\nu \bar{\boldsymbol{\Gamma}}_{\boldsymbol{s}_\nu} \boldsymbol{\mathcal{C}}_\nu^{\mathrm{T}} \boldsymbol{H}^{\mathrm{H}} + \boldsymbol{R}_{\boldsymbol{\eta}})^{-1}. \tag{30}$$

The second attempt focuses on the reduced-rank implementation of the equalizers from Eq. (16) or Eq. (30), respectively. In order to obtain $\boldsymbol{\Omega}$ or $\bar{\boldsymbol{\Omega}}$, the systems of equations can be solved iteratively using the BCG algorithm. The BCG algorithm is a method to solve the system of equations $\boldsymbol{A}\boldsymbol{X} = \boldsymbol{B}$ iteratively, with $\boldsymbol{A} \in \mathbb{C}^{M\times M}$ being non-singular, $\boldsymbol{X} \in \mathbb{C}^{M\times N}$, and $\boldsymbol{B} \in \mathbb{C}^{M\times N}$. At the k-th iteration the algorithm approximates $\boldsymbol{X}$ by $\boldsymbol{X}_k$ which lies in the block Krylov subspace span$\{\boldsymbol{B}, \boldsymbol{A}\boldsymbol{B}, \boldsymbol{A}^2\boldsymbol{B}, \ldots, \boldsymbol{A}^{k-1}\boldsymbol{B}\}$. The pseudocode for the BCG algorithm is shown in Alg. 1. Stopping the BCG iterations before the algorithm converges, reduces the computational complexity as well as the performance loss due to channel estimation errors which is especially a problem in low-sample support scenarios. The latter property of the BCG algorithm is due to its inherent regularizing effect described in [14].

Algorithm 1 (BCG algorithm):

$\boldsymbol{X}_{(0)} = \boldsymbol{0}_{M\times N}$

$\boldsymbol{D}_{(0)} = \boldsymbol{R}_{(0)} = \boldsymbol{B} - \boldsymbol{A}\boldsymbol{X}_{(0)}$

for $k \in \{0, 1, \ldots, \lceil m/n \rceil - 1\}$ **do**

  $\boldsymbol{\Phi}_{(k)} = (\boldsymbol{D}_{(k)}^{\mathrm{H}} \boldsymbol{A} \boldsymbol{D}_{(k)})^{-1} \boldsymbol{R}_{(k)}^{\mathrm{H}} \boldsymbol{R}_{(k)}$

  $\boldsymbol{X}_{(k+1)} = \boldsymbol{X}_{(k)} + \boldsymbol{D}_{(k)} \boldsymbol{\Phi}_{(k)}$

  $\boldsymbol{R}_{(k+1)} = \boldsymbol{R}_{(k)} - \boldsymbol{A}\boldsymbol{D}_{(k)} \boldsymbol{\Phi}_{(k)}$

  $\boldsymbol{\Psi}_{(k+1)} = (\boldsymbol{R}_{(k)}^{\mathrm{H}} \boldsymbol{R}_{(k)})^{-1} \boldsymbol{R}_{(k+1)}^{\mathrm{H}} \boldsymbol{R}_{(k+1)}$

  $\boldsymbol{D}_{(k+1)} = \boldsymbol{R}_{(k+1)} + \boldsymbol{D}_{(k)} \boldsymbol{\Psi}_{(k+1)}$

end for

With $\boldsymbol{A}$ corresponding to $\boldsymbol{H}\boldsymbol{\mathcal{C}}_\nu \bar{\boldsymbol{\Gamma}}_{\boldsymbol{s}_\nu} \boldsymbol{\mathcal{C}}_\nu^{\mathrm{T}} \boldsymbol{H}^{\mathrm{H}} + \boldsymbol{R}_{\boldsymbol{\eta}}$, $\boldsymbol{B}$ to $\boldsymbol{S}_\nu \bar{\boldsymbol{\Gamma}}_{\boldsymbol{s}_\nu} \boldsymbol{\mathcal{C}}_\nu^{\mathrm{T}} \boldsymbol{H}^{\mathrm{H}}$, and $\boldsymbol{X}$ to $\bar{\boldsymbol{\Omega}}$, Alg. 1 can be used to determine the detector. Taking advantage of the iterative way in which the BCG algorithm solves the system of equations one can stop the computation before convergence to the exact solution and use an approximation of $\bar{\boldsymbol{\Omega}}$. By doing this, the computational complexity can be reduced enormously while the performance of the receiver remains almost unchanged as the simulations in Section 4 show.

4 Simulation Results

In the sequel, we present results from a Monte Carlo simulation with 1000 trials for a representative channel realization showing the *Bit Error Rate* (BER) as a function of the *Signal-to-Noise Ratio* (SNR) $10\,\mathrm{lg}(E_{\mathrm{b}}/N_0)$ in dB with E_{b} being the energy per information bit and N_0 the noise power density for the CDMA-system with the following specifications: QPSK modulation, $K = 8$ number of users, $R = 2$ receive antennas at the base station ordered in an uniform linear array with spacing $\lambda/2$ where λ is the carrier wavelength, spreading factor $\chi = 8$, 13 pilot symbols per user, 512 information bits per user, order of FIR filter in the equalizer $G = 13$, $L + 1 = 4$ number of propagation paths for each user, and latency time $\nu = 9\,T_{\mathrm{c}}$ where T_{c} is the chip duration. The channels between the K transmitters are uncorrelated, the angles of arrival at the receiver are Laplacian distributed with an angular spread of $10°$ and the power delay profile is exponential with rate of decay T_{c}. It is assumed that all users encode their information bits using a rate-$1/2$ non-recursive convolutional encoder with the generator polynomial $\mathcal{G}(D) = (1 + D^2, 1 + D + D^2)$ where D represents the delay operator. Moreover, a random interleaver is used where the permutation matrix $\boldsymbol{\Pi}$ is randomly occupied. At the receiver, the BCJR algorithmis used for MAP decoding.

Figure 3 shows the BER curves obtained when calculating the detector using Eqs. (16) and (17), and Fig. 4 the BER curves when using Eqs. (17) and (30). The

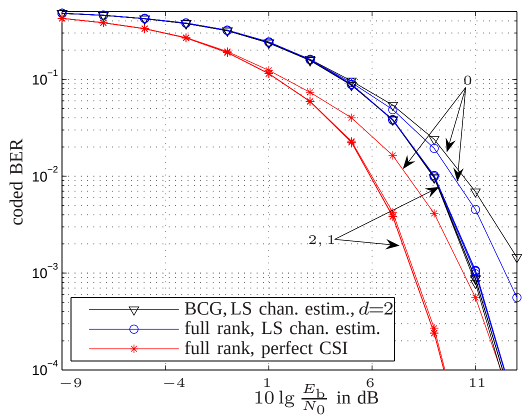

Figure 3. Iterative decoding and detection using $\boldsymbol{\Omega}[m]$, BCG curves for the turbo iterations $0, 1$, and 2 (marked with

arrows) are plotted for the three cases: perfect Channel State Information (CSI) and full-rank MWF solution, Least Squares (LS) channel estimation and full-rank MWF solution, and LS channel state estimation and MWF approximation in the Krylov subspace of order d. In Fig. 3 as well as in Fig. 4, it can be seen that after the first turbo iteration, there is no significant difference between the MWF solution and the Krylov subspace approximation of the MWF. Moreover, it can be observed that the main gain is reached when the detector receives a priori information from the decoder for the first time, i. e., at the first turbo iteration. It is remarkable and one of the main results presented in this paper that maximum rank-reduction leads to approximately the same BER as the full-rank equalizer when performing iterative detection. A comparison of Fig. 3 with Fig. 4 justifies the use of $\bar{\boldsymbol{\Omega}}$ instead of $\boldsymbol{\Omega}$.

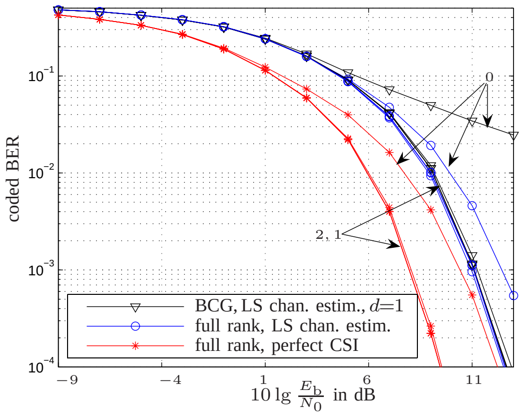

Figure 4. Iterative decoding and detection using $\bar{\boldsymbol{\Omega}}$; BCG

Using the BCG algorithm for equalization implies that at each BCG iteration the values $\hat{s}'_i[m]$ for all K users, i. e., $\hat{\boldsymbol{s}}'[m]$, are computed at once due to the block nature of the algorithm. This means less computational work than solving an equation for each user separately, e. g., using the *Conjugate Gradient* (CG) algorithm. On the other hand, when considering $\hat{s}'_i[m]$ of user i, it also means that the detector does not use $\mathrm{E}\{\boldsymbol{s}'_i[m]\}$ from the other users to combat multiple-access interference. Fig. 5 shows the BER curves when the detector computes an equalizer for every user. It can be seen that after three turbo iterations there is almost no difference to the block procedure. Thus, using the BCG algorithm for the detector is justified.

5 Conclusions

The main purpose of this work is to present two possibilities how complexity can be reduced in an iterative turbo multi-user CDMA system. The first measure to decrease the complexity is to approximate second order statistics of non-stationary random processes by their time-invariant averages. The second measure implies using the BCG algorithm such that the solution of the Wiener-Hopf equation lies in a Krylov subspace.

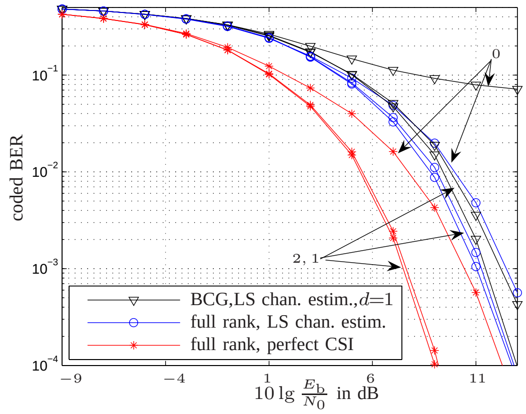

Figure 5. Iterative decoding and detection using $\bar{\boldsymbol{\Omega}}$, CG

Simulations show that despite the enormous complexity reduction both methods lead to excellent results.

References

[1] C. Douillard, M. Jézéquel, C. Berrou, A. Picart, P. Didier, and A. Glavieux, "Iterative Correction of Intersymbol Interference: Turbo-Equalization," in *European Transactions on Telecommunications*, September 1995, vol. 6, pp. 507–511.
[2] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon Limit Error-Correcting Coding and Decoding: Turbo-Codes," in *Proceedings of the* 1993 *IEEE International Conference on Communications*, May 1993, pp. 1064–1070.
[3] X. Wang and H. V. Poor, "Iterative (Turbo) Soft Interference Cancellation and Decoding for Coded CDMA," in *IEEE Transactions on Communications*, July 1999, vol. 47, pp. 1046–1061.
[4] G. Dietl, C. Mensing, and W. Utschick, "Iterative Detection Based on Reduced-Rank Equalization," in *Proceedings of the 2004 IEEE 60th Vehicular Technology Conference (VTC 2004-Fall)*, September 2004, vol. 3, pp. 1533–1537.
[5] H. Hotelling, "Analysis of a Complex of Statistical Variables into Principal Components," in *Journal of Educational Psychology*, September/October 1933, vol. 24, pp. 417–441.
[6] J. S. Goldstein and I. S. Reed, "Reduced-Rank Adaptive Filtering," in *IEEE Transactions on Signal Processing*, February 1997, vol. 45, pp. 492–496.
[7] J. S. Goldstein, I. S. Reed, and L. L. Scharf, "A Multistage Representation of the Wiener Filter Based on Orthogonal Projections," in *IEEE Transactions on Information Theory*, November 1998, vol. 44, pp. 2943–2959.
[8] D. P. O'Leary, "The Block Conjugate Gradient Algorithm and Related Methods," *Linear Algebra and its Applications*, vol. 29, pp. 293–322, 1980.
[9] M. Tüchler, A. C. Singer, and R. Koetter, "Minimum Mean Squared Error Equalization Using A-priori Information," in *IEEE Transactions on Signal Processing*, March 2002, vol. 50, pp. 673–683.
[10] J. Hagenauer, E. Offer, and L. Papke, "Iterative Decoding of Binary Block and Convolutional Codes," in *IEEE Transactions on Information Theory*, March 1996, vol. 42, pp. 429–445.
[11] M. Botsch, G. Dietl, and F. A. Dietrich, "Robust Reduced-Complexity Receivers for Multi-User Mobile Communication Systems," Tech. Rep. TUM-LNS-TR-05-04, Munich University of Technology, April 2005.
[12] M. Tüchler, R. Koetter, and A. C. Singer, "Turbo Equalization: Principles and New Results," in *IEEE Transactions on Communications*, May 2002, vol. 50, pp. 754–767.
[13] G. Dietl and W. Utschick, "Complexity Reduction of Iterative Receivers Using Low-Rank Equalization," in *IEEE Transactions on Signal Processing*, 2005, *To appear*.
[14] G. Dietl, M. Botsch, and W. Utschick, "Robust and Reduced-Rank Matrix Wiener Filter Based on the Block Conjugate Gradient Algorithm," in *SPAWC*, June 2005.