Iterative Joint Equalization and Decoding based on Soft Cholesky Equalization for General Complex Valued Modulation Symbols

Jochem Egle and Jürgen Lindner
Department of Information Technology
University of Ulm
Albert-Einstein-Allee 43, D-89081 Ulm, Germany
e-mail: jochem.egle@e-technik.uni-ulm.de

Abstract

A new iterative equalizer based on the classical block decision feedback equalizer is derived. The performance of this equalizer benefits from the usage of a soft decision function, which is based on noise whitening and consecutive matched filtering. The derivation of the decision function is very general and therefore applicable to arbitrary complex valued modulation symbols.

After adding the possibility of processing extrinsic information to this decision function the equalizer is embedded in an iterative equalization/decoding scheme.

To demonstrate the performance of the scheme simulation results for frequency selective channels using BPSK and 16 QAM are presented.

1 Introduction

In recent years mobile communication has seen a tremendous development. The growth rate of number of subscribers for mobile communication systems is enormous and mobile multimedia communication demanding very high data rate are on the horizon. Since the available spectrum allocation for mobile communication systems and the transmit power is limited, suitable methods have to be found, which use the available bandwidth efficiently and require as little transmit power as possible.

To meet these requirements many modern commutation systems use non-binary modulation alphabets, e. g. M-ary phase shift keying (PSK) or M-ary quadrature amplitude modulation (QAM) schemes in combination with forward error correcting codes.

In the following we will introduce an algorithm which is suitable to detect a coded signal which uses arbitrary complex valued modulation alphabet and is subject to a high level of inter symbol interference (ISI) originating from a frequency selective channel.

A well-known detection scheme with moderate complexity and often near optimum performance is based on recurrent neural networks (RNN). It was introduced first for systems without coding in [14] and [7] for the case of BPSK and extended to general complex valued modulation alphabets in [12]. In [10] and [11] it was extended to systems with coding.

Other approaches leading to equalizer structures which are closely related to the one derived here are explained in [1] and [5]. Both publications use linear filtering in combination with iterative soft decision feedback. The difference between them is the filter design criterion. The first one uses the unbiased minimum variance criterion whereas the second one the minimum mean squared criterion.

First [3] applied the ideas of multiple iterations and soft feedback to the classical cholesky based block decision feedback equalizer (CBDFE), see e. g. [15] in a multi-user context. Using this as a basis we will drop some simplifying assumptions and generalize the results previously obtained, resulting in a class of equalizers which differ in complexity and performance.

Furthermore, we will test the suitability of this new equalizer to be used in a “Turbo” equalization scheme similar to the structure introduced first in [2].

After explaining a vector valued transmission model of a coded transmission over a frequency selective channel based on [6], we describe the classical CBDFE. Starting from this basis we derive a new iterative equalizer, i. e. , the iterative soft Cholesky block decision feedback equalizer (SCE), for an uncoded transmission. Subsequently we add the possibility of processing extrinsic information to the SCE and use this modified version in an iterative coding and equalization scheme.

In the simulation part we compare the performance of different versions of the SCE against each other, to the matched filter bound (MFB) and to an RNN type equalizer, see e. g. [12] for coded and uncoded transmission.

For mathematical notation we use double underlined letters to denote matrices $\underline{\underline{M}}$, whereas vectors $\underline{v}$ will be denoted single underlined letters. Furthermore, $(\cdot)^T$ denotes the transposed, $(\cdot)^H$ the hermitian, i. e. the transposed conjugate complex of a matrix, and $(\cdot)^{-1}$ the inverse of a matrix.

2 System Model

Figure 2 depicts a continuous-time model of a coded single carrier transmission in the equivalent low-pass domain. A sequence of statistical independent bits $\underline{q}$ is encoded using an optionally punctured terminated convolutional code. The code is deduced from a rate $1/n$

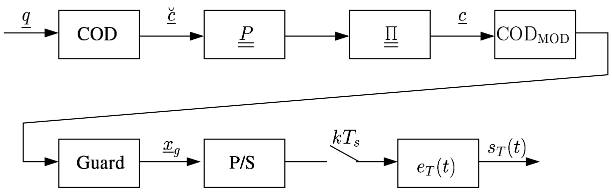

Figure 1: Transmit side of a coded single carrier transmission scheme

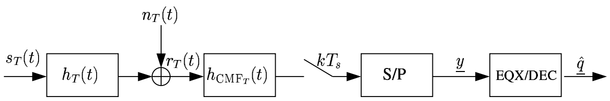

Figure 2: Transmit channel together with a receiver

mother code using a puncturing matrix $\underline{\underline{P}}$. After interleaving the encoded bit sequence using a pseudo random interleaver matrix $\underline{\underline{\Pi}}$ the modulation specific coding $\mathrm{COD_{MOD}}$ groups the interleaved binary codeword sequence $\underline{c}$ into blocks of $\log_2(M)$ bits. Each of those blocks is determining one complex transmit symbol $x(k) \in \mathcal{A} = \{a_1, a_2, \ldots, a_M\}$. In "Guard", periodically after N transmit symbols, a guard period of known symbols is inserted such that the receive stream may be separated into independent blocks in the receiver. For simplicity, we assume in the following the insertion of zeros . After parallel/serial (P/S) conversion the transmit symbols $x_g(k)$ are modulated using the basic waveform $e_T(t)$ resulting in the transmit signal $s_T(t)$.

Subsequently the modulated signal $s_T(t)$ is transmitted over a frequency selective additive white Gaussian noise having impulse response $h_T(t)$. In the receiver the symbol time sampled output sequence of the channel matched filter $h_{\mathrm{CMF}_T}(t)$

$$h_{\mathrm{CMF}_T}(t) = \frac{1}{2} e_T^*(-t) * h_T^*(-t) \tag{1}$$

is fed to an equalization/decoding algorithm.

By introducing the discrete-time auto correlation r(k) and the sampled colored noise $\tilde{n}(k)$

$$\begin{aligned} r(k) &= \frac{1}{4} h_{\mathrm{CMF}_T}(t) * h_T(t) * e_T(t)|_{t=kT_s} \\ \tilde{n}(k) &= \frac{1}{2} h_{\mathrm{CMF}_T}(t) * n_T(t)|_{t=kT_s} \end{aligned} \tag{2}$$

it is possible to reformulate the relation between the vector of transmit symbols $\underline{x} = (x(1), \ldots, x(N))^T$ and the sampled output of the matched filter (without guard time) $\underline{y} = (y(1), \ldots, y(N))^T$ in compact matrix vector notation

$$\begin{aligned} &\underline{y} = \underline{\underline{R}}\,\underline{x} + \underline{\tilde{n}} \text{ with} \\ &\underline{\underline{R}} = \underline{\underline{[r_{ij}]}} \text{ and } r_{ij} = r(i-j)\ . \end{aligned} \tag{3}$$

3 Equalization Scheme

Commencing with the conventional Cholesky based block decision feedback equalizer (CBDFE), which may be viewed as the block processing counterpart of the ordinary decision feedback equalizer (DFE) based on finite impulse response (FIR)-filtering, we will derive an iterative block decision feedback equalizer. Retaining the filter matrices of the CBDFE, the new scheme overcomes the unavoidable SNR-loss of a conventional CBDFE using a novel soft decision function and multiple iterations. Therefore, we will term this new scheme iterative soft Cholesky block decision feedback equalizer (SCE).

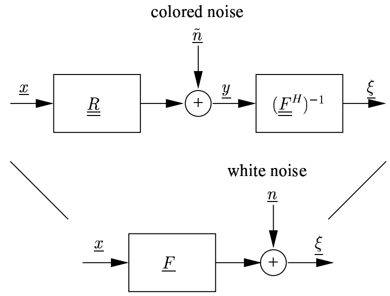

Figure 3: Equivalent vector valued transmission scheme

3.1 Cholesky Block Decision Feedback Equalizer

The correlation matrix $\underline{\underline{R}}$ is hermitian and positive semidefinite. Since a positive semi-definite matrix becomes positive definite by adding a small value greater zero, e. g. , the noise variance, on the main diagonal it possible to treat $\underline{\underline{R}}$ as positive definite. Under this assumption $\underline{\underline{R}}$ can be factorized in a product of an upper triangular matrix $\underline{\underline{F}}$ and the hermitian of this matrix using cholesky decomposition:

$$\underline{\underline{R}} = \underline{\underline{F}}^H \underline{\underline{F}} \tag{4}$$

The CBDFE employs $(\underline{\underline{F}}^H)^{-1}$ as feed-forward filter. The combination of the channel matrix $\underline{\underline{R}}$ and the feed-forward filter $(\underline{\underline{F}}^H)^{-1}$ results in the simplified model shown in Figure 3 and described by

$$\begin{aligned} \underline{\xi} &= (\underline{\underline{F}}^H)^{-1}(\underline{\underline{R}}\,\underline{x} + \underline{\tilde{n}}) \\ &= (\underline{\underline{F}}^H)^{-1}(\underline{\underline{F}}^H \underline{\underline{F}}\,\underline{x} + \underline{\underline{F}}^H \underline{n}) = \underline{\underline{F}}\,\underline{x} + \underline{n} \\ &= \begin{pmatrix} f_{11} & f_{12} & \cdots & f_{1N} \\ & f_{22} & \cdots & f_{2N} \\ & & \ddots & \vdots \\ \underline{0} & & & f_{NN} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} + \underline{n}\ . \end{aligned} \tag{5}$$

Introducing $\mathrm{dec}(\cdot)$ as a two-dimensional decision device matched to the modulation alphabet $\mathcal{A}$ and exploiting the upper triangular matrix property of $\underline{\underline{F}}$ the CBDFE sequentially detects the block of transmitted symbols starting from the end of the block

$$\begin{aligned}
\hat{x}_N &= \mathrm{dec}\left(\frac{1}{f_{NN}}\xi_N\right) \\
\hat{x}_{N-1} &= \mathrm{dec}\left(\frac{1}{f_{N-1N-1}}(\xi_{N-1} - f_{N-1N}\hat{x}_N)\right) \\
\vdots & \qquad\qquad \vdots \\
\hat{x}_1 &= \mathrm{dec}\left(\frac{1}{f_{11}}(\xi_1 - \sum_{i=2}^{N} f_{1i}\hat{x}_i)\right) \quad .
\end{aligned} \tag{6}$$

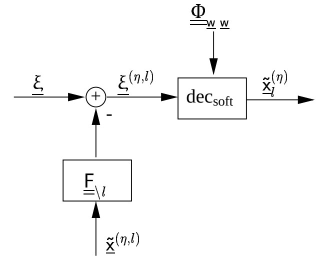

Figure 4: SCE updating estimate of symbol $\underline{x}_l$ in iteration η

3.2 Iterative Soft Block Decision Feedback Equalizer

The soft decision function of the SCE is based on the optimum detector for colored Gaussian noise. In each stage of the equalizer the covariance matrix of the interference as well as the transmit symbol vector is estimated.

3.2.1 Real valued matrix notation

Below we will introduce a notation which describes the quadrature components of a baseband symbol not by means of complex numbers but by a vector of length two. This representation enables us to handle general complex Gaussian random vectors like they are required for the derivation of the SCE and which may not be described correctly by means of a complex Gaussian distribution, e. g. , a random vector consisting of complex elements with non-zero covariance between the quadrature components. A deeper insight into the properties of such random vectors is given by Neeser and Massey in [8].

In this notation a complex vectors $\underline{v}$ of length a is represented by real vectors $\underline{\mathsf{v}}$ of length $2a$ and a complex matrix $\underline{\underline{M}}$ of dimension $a \times b$ by real matrices $\underline{\underline{\mathsf{M}}}$ of dimension $2a \times 2b$. The corresponding transformation rule is shown in Equation (7). $\otimes$ denotes the Kronecker product

$$\begin{aligned}
\underline{\mathsf{v}} &= \begin{pmatrix}\Re(v_1) & \Im(v_1) & \dots & \Re(v_a) & \Im(v_a)\end{pmatrix}^T \\
&= \Re(\underline{v}) \otimes \begin{pmatrix}1\\0\end{pmatrix} + \Im(\underline{v}) \otimes \begin{pmatrix}0\\1\end{pmatrix} \\
\underline{\underline{\mathsf{M}}} &= \begin{pmatrix}
\Re(m_{11}) & -\Im(m_{11}) & \dots & -\Im(m_{1b}) \\
\Im(m_{11}) & \Re(m_{11}) & \dots & \Re(m_{1b}) \\
\vdots & \vdots & & \vdots \\
\Im(m_{a1}) & \Re(m_{a1}) & \dots & \Re(m_{ab})
\end{pmatrix} \\
&= \Re(\underline{\underline{M}}) \otimes \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix} + \Im(\underline{\underline{M}}) \otimes \begin{pmatrix}0 & -1\\1 & 0\end{pmatrix}
\end{aligned} \tag{7}$$

In this notation a complex symbol corresponds to a real vector of length 2, e.g. , the modulation alphabet is given by $\mathcal{A} = \{\underline{\mathsf{a}}_1, \dots, \underline{\mathsf{a}}_M\}$ with $\underline{\mathsf{a}}_j = (\Re(a_j), \Im(a_j))^T$.

3.2.2 SCE – iteration scheme

The SCE replaces the hard decision function of dec$(\cdot)$ by a soft decision function dec$_{soft}(\cdot)$ which is designed under the constraint of minimizing the mean squared error (MSE) of the decision. Furthermore, the SCE uses multiple iterations, each iteration starting from the end of the block, during which previous estimates of interfering symbols are fed back.

A model of the SCE when calculating an estimate $\underline{\tilde{\mathsf{x}}}_l^{(\eta)}$ of the l-th symbol $\underline{\mathsf{x}}_l$ in the η-th iteration is shown in Figure 4 and described by

$$\begin{aligned}
\underline{\tilde{\mathsf{x}}}_l^{(\eta)} &= \mathrm{dec_{soft}}(\underline{\xi}^{(\eta,l)}\underline{\underline{\Phi}}^{(\eta,l)}_{\underline{\mathsf{w}}\,\underline{\mathsf{w}}}) \quad \text{with} \\
\underline{\xi}^{(\eta,l)} &= \underline{\xi} - \underline{\underline{\mathsf{F}}}_{\setminus l}\underline{\tilde{\mathsf{x}}}^{(\eta,l)} \quad \text{and} \\
\underline{\tilde{\mathsf{x}}}^{(\eta,l)} &= [\underline{\tilde{\mathsf{x}}}_1^{(\eta-1)^T}, \dots, \underline{\tilde{\mathsf{x}}}_l^{(\eta-1)^T}, \underline{\tilde{\mathsf{x}}}_{l+1}^{(\eta)\,T}, \dots, \underline{\tilde{\mathsf{x}}}_N^{(\eta)^T}]^T \quad .
\end{aligned} \tag{8}$$

$\underline{\underline{\mathsf{F}}}_{\setminus l}$ denotes a matrix derived from $\underline{\underline{\mathsf{F}}}$ by setting the $(2l-1)$-th and the $2l$-th column to the zero vector. $\underline{\tilde{\mathsf{x}}}^{(\eta,l)}$ contains previous soft decisions of interfering symbols. The matrix $\underline{\underline{\Phi}}^{(\eta,l)}_{\underline{\mathsf{w}}\,\underline{\mathsf{w}}}$ is the covariance matrix of the noise plus interference signal.

Initially, no estimates of the transmit symbols are available. Therefore, we will initialize the feedback vector $\underline{\tilde{\mathsf{x}}}^{(1,N)}$ by the unconditioned mean of the transmit vector, $E\{\underline{\mathsf{x}}\}$, which corresponds under the assumption of an unbiased and equiprobable modulation alphabet to the all-zero vector.

The desired end state is reached for $\underline{\tilde{\mathsf{x}}}^{(\eta,l)} = \underline{\mathsf{x}}$, i. e. all interference has been subtracted and therefore $\underline{\xi}^{(\eta,l)}$ equals

$$\underline{\xi}^{(\eta,l)} = \underline{\underline{\mathsf{f}}}_l \underline{\mathsf{x}}_l + \underline{\mathsf{n}} \quad . \tag{9}$$

$\underline{\underline{\mathsf{f}}}_l$ is a sub-matrix of $\underline{\underline{\mathsf{F}}}$ containing the $(2l-1)$-th and $2l$-th column of $\underline{\underline{\mathsf{F}}}$, reflecting the impact of the l-th transmit symbol $\underline{\mathsf{x}}_l$ on the receive vector $\underline{\xi}$. In case Equation (9) is valid $\underline{\xi}^{(\eta,l)}$ is only dependent on transmit symbol $\underline{\mathsf{x}}_l$.

3.2.3 Decision function

First in [13] a decision rule aiming to minimize the mean squared error (MSE) of the decision was proposed and further investigated in [4]. In [12] a similar idea was applied to determine the decision function of a RNN.

Then in [3] the ideas mentioned above were applied to the CBDFE to overcome the DFE inherent SNR-loss. In the following we generalize the results gained in [3] resulting in a class of decision functions.

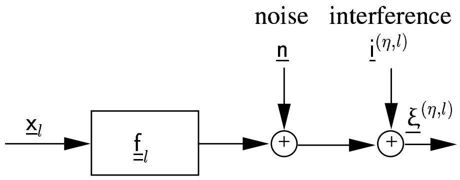

Figure 5: Model of the decision of the SCE

The new decision function allows to take decision based on all entries of $\underline{\underline{f}}_l$, whereas the decision function of a conventional CBDFE benefits only from energy on the main diagonal of $\underline{\underline{F}}$. A model for decision of one symbol $\underline{x}_l$ is shown in Figure 5. Multiplying the transmit symbol $\underline{x}_l$ with $\underline{\underline{f}}_l$ spreads its energy over the receive vector $\underline{\xi}^{(\eta,l)}$.

$$\underline{\xi}^{(\eta,l)} = \underline{\underline{f}}_l \underline{x}_l + \underline{n} + \underline{i}^{(\eta,l)} \tag{10}$$

Two parts disturb the reception. The first one, $\underline{n}$, is caused by additive white Gaussian noise which is assumed to be independent of the transmit signal, whereas the second part, $\underline{i}^{(\eta,l)}$, comprises the remaining interference of other symbols.

For the sake of computational tractability we treat the remaining interference $\underline{i}^{(\eta,l)}$

$$\underline{i}^{(\eta,l)} = \underline{\underline{F}}_{\backslash l}(\underline{x} - \underline{\tilde{x}}^{(\eta,l)}) \tag{11}$$

as a random vector having a multivariate Gaussian distribution with correlation matrix $\underline{\underline{\Phi}}_{\underline{i}\,\underline{i}}^{(\eta,l)}$.

As already stated, the derivation of the decision function is based on minimizing the MSE of decision $\underline{\tilde{x}}_l^{(\eta)}$.

$$\underline{\tilde{x}}_l^{(\eta)} = \arg\left(\min_{\underline{\tilde{x}}_l} \mathrm{E}\{(\underline{x}_l - \underline{\tilde{x}}_l)^T(\underline{x}_l - \underline{\tilde{x}}_l) | \underline{\xi}^{(\eta,l)}\}\right) \tag{12}$$

The solution of this minimization problem is given by:

$$\begin{aligned}\underline{\tilde{x}}_l^{(\eta)} &= \mathrm{E}\{\underline{x}_l | \underline{\xi}^{\eta,l}\} \\ &= \sum_{j=1}^{M} \underline{a}_j P(\underline{x}_l = \underline{a}_j | \underline{\xi}^{(\eta,l)})\end{aligned} \tag{13}$$

To evaluate Equation (13) we have to determine $P(\underline{x}_l = \underline{a}_j | \underline{\xi}^{(\eta,l)})$. Applying Bayes' rule to $P(\underline{x}_j = \underline{a}_m | \underline{\xi}^{(\eta,l)})$ gives us:

$$P(\underline{x}_l = \underline{a}_j | \underline{\xi}^{(\eta,l)}) = \frac{p(\underline{\xi}^{(\eta,l)} | \underline{x}_l = \underline{a}_j) P(\underline{x}_l = \underline{a}_j)}{p(\underline{\xi}^{(\eta,l)})} \tag{14}$$

The additional assumption of equiprobable transmit symbols, i. e. $P(\underline{x}_l = \underline{a}_m) = 1/M \,\forall\, \underline{a}_m \in \mathcal{A}$, combined with Equation (14) leads to

$$P(\underline{x}_l = \underline{a}_j | \underline{\xi}^{(\eta,l)}) = \frac{p(\underline{\xi}^{(\eta,l)} | \underline{x}_l = \underline{a}_j)}{\sum_{m=1}^{M} p(\underline{\xi}^{(\eta,l)} | \underline{x}_l = \underline{a}_m)} \,. \tag{15}$$

For the next step we will introduce the abbreviation $\underline{w}^{(\eta,l)}$

$$\underline{w}^{(\eta,l)} = \underline{i}^{(\eta,l)} + \underline{n} \tag{16}$$

as the sum of residual interference $\underline{i}^{(\eta,l)}$ plus noise $\underline{n}$. The covariance matrix of $\underline{w}^{(\eta,l)}$ is given by:

$$\underline{\underline{\Phi}}_{\underline{w}\,\underline{w}}^{(\eta,l)} = \underline{\underline{\Phi}}_{\underline{i}\,\underline{i}}^{(\eta,l)} + \sigma_n^2 \underline{\underline{I}} \tag{17}$$

Exploiting additionally the assumption of having interference with Gaussian distribution $p(\underline{\xi}^{(\eta,l)} | \underline{x}_l = a_j)$ is described by a $2N$ dimensional Gaussian distribution with variance $\underline{\underline{\Phi}}_{\underline{w}\,\underline{w}}^{(\eta,l)}$ and mean $\underline{\underline{f}}_l \underline{a}_j$

$$\begin{aligned}p(\underline{\xi}^{(\eta,l)} | \underline{x}_l = \underline{a}_j) &= \frac{1}{(2\pi)^N \det(\underline{\underline{\Phi}}_{\underline{w}\,\underline{w}}^{(\eta,l)})^{1/2}} \\ &\cdot \exp\left(-\frac{1}{2}(\underline{\xi}^{(\eta,l)} - \underline{\underline{f}}_l \underline{a}_j)^T \underline{\underline{\Phi}}_{\underline{w}\,\underline{w}}^{(\eta,l)-1} (\underline{\xi}^{(\eta,l)} - \underline{\underline{f}}_l \underline{a}_j)\right) \,.\end{aligned} \tag{18}$$

To reduce computational complexity it is quite common to use a set of sufficient statistics $\underline{\tilde{\xi}}^{(\eta,l)}$ of the random process, which still contains all information of $\underline{\xi}^{(\eta,l)}$, but requires less dimensions. The condition for a sufficient statistic is summarized in Equation (19).

$$p(\underline{\xi}^{(\eta,l)} | \underline{x}_l = \underline{a}_j) = p(\underline{\tilde{\xi}}^{(\eta,l)} | \underline{x}_l = \underline{a}_j) f(\underline{\xi}^{(\eta,l)}) \tag{19}$$

In this case as it can be easily seen from Equation (15)

$$P(\underline{x}_l = \underline{a}_j | \underline{\xi}^{(\eta,l)}) = P(\underline{x}_l = \underline{a}_j | \underline{\tilde{\xi}}^{(\eta,l)}) \tag{20}$$

is valid.

A quite well-known method in the presence of colored noise, which fulfills Equation (19) is the whitening approach. It is composed of a whitening filter and a consecutive matched filter.

$$\underline{\tilde{\xi}}^{(\eta,l)} = (\underline{\underline{f}}_l^T \underline{\underline{\Phi}}_{\underline{w}\,\underline{w}}^{(\eta,l)-1} \underline{\underline{f}}_l)^{-1} \underline{\underline{f}}_l^T \underline{\underline{\Phi}}_{\underline{w}\,\underline{w}}^{(\eta,l)-1} \underline{\xi}^{(\eta,l)} \tag{21}$$

Usage of this approach delivers us a sufficient statistic $\underline{\tilde{\xi}}^{(\eta,l)}$ of $\underline{\xi}^{(\eta,l)}$ of length 2 with mean $\underline{a}_j$ and variance $(\underline{\underline{f}}_l^T \underline{\underline{\Phi}}_{\underline{w}\,\underline{w}}^{(\eta,l)-1} \underline{\underline{f}}_l)^{-1}$

$$\begin{aligned}p(\underline{\tilde{\xi}}^{(\eta,l)} | \underline{x}_l = \underline{a}_j) &= \frac{\det(\underline{\underline{f}}_l^T \underline{\underline{\Phi}}_{\underline{w}\,\underline{w}}^{(\eta,l)-1} \underline{\underline{f}}_l)^{\frac{1}{2}}}{2\pi} \\ &\cdot \exp(-\frac{1}{2}(\underline{\tilde{\xi}}^{(\eta,l)} - \underline{a}_j)^T \underline{\underline{f}}_l^T \underline{\underline{\Phi}}_{\underline{w}\,\underline{w}}^{(\eta,l)-1} \underline{\underline{f}}_l (\underline{\tilde{\xi}}^{(\eta,l)} - \underline{a}_j))\end{aligned} \tag{22}$$

which can be used in Equation (15) instead of $p(\underline{\xi}^{(\eta,l)} | \underline{x}_l = \underline{a}_j)$

3.2.4 Estimation of the Covariance matrix

The covariance matrix of the interference may be expressed by means of a conditional expectation value as shown in Equation (23).

$$\begin{aligned}\underline{\underline{\Phi}}_{\underline{i}\,\underline{i}}^{(\eta,l)} &= \mathrm{E}\{\underline{i}^{(\eta,l)}\, \underline{i}^{(\eta,l)T} | \underline{\xi}\} \\ &= \underline{\underline{F}}_{\backslash l} \mathrm{E}\{(\underline{x} - \underline{\tilde{x}}^{(\eta,l)})(\underline{x} - \underline{\tilde{x}}^{(\eta,l)})^T | \underline{\xi}\} \underline{\underline{F}}_{\backslash l}^T \\ &= \underline{\underline{F}}_{\backslash l} \underline{\underline{\Phi}}_{\underline{x}\,\underline{x}}^{(\eta,l)} \underline{\underline{F}}_{\backslash l}^T\end{aligned} \tag{23}$$

It depends on the matrix $\underline{\underline{F}}_{\backslash l}$ and the covariance matrix of the remaining decision error $\underline{\underline{\Phi}}_{\underline{x}\,\underline{x}}^{(\eta,l)}$. Under the assumption of statistical independent decision errors of the symbols in one block $\underline{\underline{\Phi}}^{(\eta,l)}_{\underline{x}\,\underline{x}}$ is a block diagonal matrix

$$\underline{\underline{\Phi}}^{(\eta,l)}_{\underline{x}\,\underline{x}} = \underline{\underline{\mathrm{diag}}}([\underline{\underline{\Phi}}^{(\eta-1)}_{\underline{x}_1\underline{x}_1}, \dots, \underline{\underline{\Phi}}^{(\eta-1)}_{\underline{x}_l\underline{x}_l}, \underline{\underline{\Phi}}^{(\eta)}_{\underline{x}_{l+1}\underline{x}_{l+1}}, \dots, \underline{\underline{\Phi}}^{(\eta)}_{\underline{x}_N\underline{x}_N}]) \quad (24)$$

with 2×2 sub-matrices $\underline{\underline{\Phi}}^{(\eta)}_{\underline{x}_l\underline{x}_l}$.

$$\underline{\underline{\Phi}}^{(\eta)}_{\underline{x}_l\underline{x}_l} = \mathrm{E}\{\underline{x}_l\underline{x}_l^T|\underline{\xi}^{(\eta,l)}\} - \tilde{\underline{x}}_l^{(\eta)}\tilde{\underline{x}}_l^{(\eta)T} \quad (25)$$

These sub-matrices describe the decision error of the l-th symbol which can be calculated using Equation (25) together with Equation (15)

$$\mathrm{E}\{\underline{x}_l\underline{x}_l^T|\underline{\xi}^{(\eta,l)}\} = \sum_{j=1}^{M} \underline{a}_j\underline{a}_j^T P(\underline{x}_l = \underline{a}_j|\underline{\xi}^{(\eta,l)}) \quad (26)$$

3.2.5 Complexity Reduction

The task requiring most computational complexity in the equalization scheme is the matrix inversion of the covariance matrix $\underline{\underline{\Phi}}^{(\eta,l)}_{\underline{w}\,\underline{w}}$.

A simplification which leads for many scenarios to a negligible loss in performance is the assumption of having mutually uncorrelated interference terms but taking the correlation of the quadrature components of each interfering term into account. In this case $\underline{\underline{\Phi}}^{(\eta,l)}_{\underline{i}\,\underline{i}}$ and as a consequence of this $\underline{\underline{\Phi}}^{(\eta,l)}_{\underline{w}\,\underline{w}}$ is approximated by a block diagonal matrix with blocks of 2×2 sub-matrices. This simplification reduces the complexity from inverting a $2N \times 2N$ matrix to the complexity of inverting N 2×2 matrices.

When additionally neglecting the correlation between the quadrature components of the interference vector $\underline{i}^{(\eta,l)}$, $\underline{\underline{\Phi}}^{(\eta,l)}_{\underline{i}\,\underline{i}}$ is given by a diagonal matrix according to:

$$\underline{\underline{\Phi}}^{(\eta,l)}_{\underline{i}\,\underline{i}} \approx \underline{\underline{\mathrm{diag}}}(\underline{\mathrm{diag}}(\underline{\underline{\Phi}}^{(\eta,l)}_{\underline{i}\,\underline{i}})) \quad (27)$$

This simplifying approach was pursued in [3].

4 Joint Equalization and Decoding

In the previous sections we have described an iterative equalizer based on the optimum receive filter in the presence of colored Gaussian noise. In this section we will explain how to combine equalization and decoding such that a combined equalization and decoding block with an unrealizable high complexity is approximated by a structure using two separated devices passing in several iterations information between each other. In the literature this receiver structure is known as “Turbo” equalizer. It was first introduced in [2].

A common probability measure in iterative schemes for binary random variables, e. g. code bits, are the log likelihood ratio (LLR) L:

$$L(x) = \log \frac{P(x=0)}{P(x=1)} \quad (28)$$

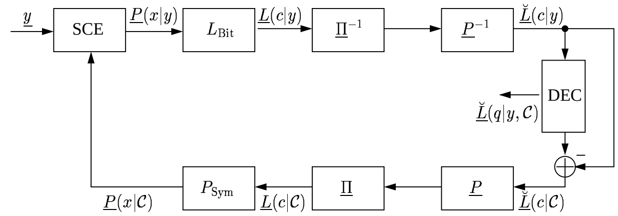

Figure 6: Combined Equalization and Decoding

A receiver structure using the ideas mentioned above is shown in Figure 6. For better readability the indices in the L values and probability arguments have been omitted.

First we will describe the exchange of information between decoder and equalizer. To take the constraint of the code into account we modify the probability $P(\underline{x}_l = \underline{a}_j|\underline{\xi}^{(\eta,l)})$ in Equation (13) and (26) by adding an additional condition $\mathcal{C}$ which represents the constraint of the code. Application of Bayes' rule results in:

$$P(\underline{x}_l = \underline{a}_j|\tilde{\underline{\xi}}^{(\eta,l)}, \mathcal{C}) = \frac{p(\tilde{\underline{\xi}}^{(\eta,l)}|\underline{x}_l = \underline{a}_j, \mathcal{C})P(\underline{x}_l = \underline{a}_j|\mathcal{C})}{p(\tilde{\underline{\xi}}^{(\eta,l)}|\mathcal{C})} \quad (29)$$

The additional assumption of statistical independent constraints $\underline{\xi}^{(\eta,l)}$ and $\mathcal{C}$, which is at least justified in the first iteration of the combined equalization decoding scheme by the usage of the random interleaver $\underline{\underline{\Pi}}$, leads to:

$$P(\underline{x}_l = \underline{a}_j|\tilde{\underline{\xi}}^{(\eta,l)}, \mathcal{C}) = \frac{p(\tilde{\underline{\xi}}^{(\eta,l)}|\underline{x}_l = \underline{a}_j)P(\underline{x}_l = \underline{a}_j|\mathcal{C})}{p(\tilde{\underline{\xi}}^{(\eta,l)})} \quad (30)$$

With $\mathcal{B}$ as the mapping rule from symbols to bits, e. g. , $\mathcal{B}(\underline{a}_1, 1)$ is the bit value of the first bit of symbol $\underline{a}_1$ and assuming that the additional information $P(c_{li}|\mathcal{C})$ of the code bits determining the l-th symbol are mutually statistical independent $P(\underline{x}_l = \underline{a}_j|\mathcal{C})$ is given by:

$$P(\underline{x}_l = \underline{a}_j|\mathcal{C}) = \prod_{i=1}^{\log_2(M)} P(c_{li} = \mathcal{B}(\underline{a}_j, i)|\mathcal{C}) \quad (31)$$

Using LLR's Equation (31) can be reformulated as

$$P(\underline{x}_l = \underline{a}_j|\mathcal{C}) = \prod_{i=1}^{\log_2(M)} \frac{\exp(1 - \mathcal{B}(\underline{a}_j, i)L(c_{li}|\mathcal{C}))}{(1 + \exp(L(c_{li}|\mathcal{C})))}. \quad (32)$$

Up to now we have described the exchange of information between decoder and equalizer. The exchange between equalizer and decoder is accomplished by a block L_{Bit} which calculates Bit LLR's $L(c_{li}|\underline{y})$ based on the symbol probabilities $P(\underline{x}_l|\underline{\xi}^{(\eta,l)})$ of the equalizer.

$$L(c_{li}|\underline{y}) = L(c_{li}|\tilde{\underline{\xi}}^{(\eta,l)}) = \log \frac{\sum_{\underline{a}_j \in \mathcal{A}_i^{[0]}} P(\underline{x}_l = \underline{a}_j | \tilde{\underline{\xi}}^{(\eta,l)})}{\sum_{\underline{a}_j \in \mathcal{A}_i^{[1]}} P(\underline{x}_l = \underline{a}_j | \tilde{\underline{\xi}}^{(\eta,l)})} \quad (33)$$

Herein $\mathcal{A}_i^{[0]}$ defines the set of modulation symbols which are defined by a bit sequence having a zero at position i whereas $\mathcal{A}_i^{[1]}$ corresponds to the set of modulation symbols having a one at position i.

Several strategies how to pass the information between equalizer and decoder are known in the literature. The one which will pursue in the simulations is a blockwise update per iteration, i. e. per iteration the information between equalizer and decoder and vice versa is exchanged once.

5 Simulation Results

To investigate the performance of the combined equalization and decoding scheme we have computed bit error rates (BER) based on Monte Carlo simulations. We have chosen time invariant channel models according to [9]. In the coded case we have used a constraint length

Channel b	$\underline{h} = [0.407\ 0.815\ 0.407]$
Channel c	$\underline{h} = [0.227\ 0.460\ 0.688\ 0.4600.227]$

Table 1: symbol time spaced channel models according to [9]

3 rate $\frac{1}{2}$ convolutional code with generator polynomial $[7\ 5]$. The blocklength of the terminated convolutional code was adjusted such that the corresponding bit sequence of K equalizer blocks is determined by one code word of length N_c. The corresponding parameter set is addressed by $K \times N$.

Figure 7 shows the results obtained for uncoded BPSK, an equalizer blocklength $N = 8$ and channel c. As an upper bound the classical CBDFE is depicted whereas maximum likelihood (ML) detection serves as a lower bound. In this case the full complexity and the diagonal approximation of the SCE perform nearly the same. Furthermore the improvement when using more than 2 iterations in negligible. The performance gain at a BER of 10^{-1} of 3.5 dB of the SCE in comparison to the CBDFE as well as the performance loss in comparison to ML of about 1.5 d is remarkable. Furthermore the RNN type equalizer does not converge at all. Nevertheless, at a BER 10^{-4} the gain decreases to 1 dB whereas the loss increases to 4 dB.

The results for a coded BPSK transmission with blocklength 100×16 over channel c are shown in figure 8. As a lower bound the code performance in an AWGN channel without ISI is depicted. In this case the SCE with full complexity nearly achieves the lower bound and outperforms the SCE with diagonal approximation.

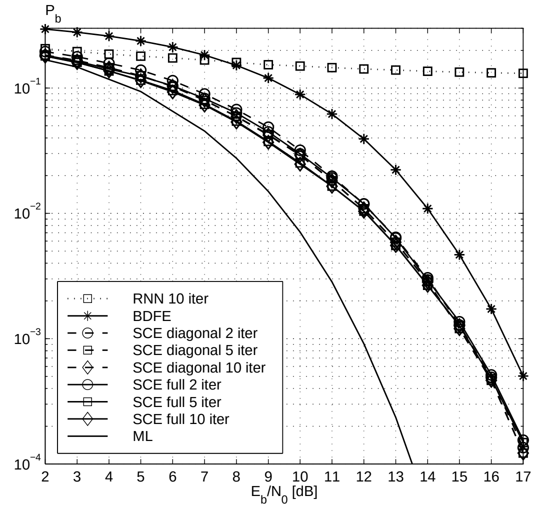

Figure 7: Comparison between SCE using diagonal and full covariance matrix, BDFE and RNN for BPSK, channel c, and blocklength $N = 8$

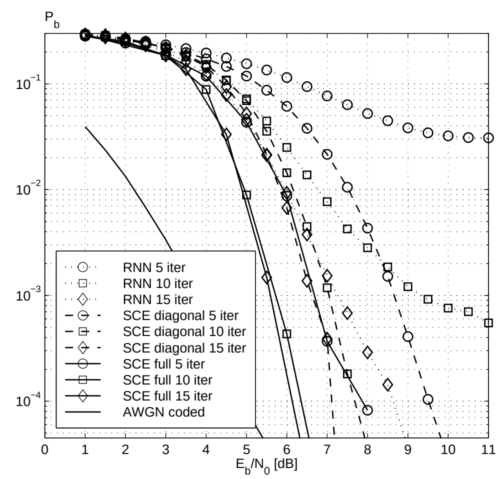

Figure 8: Comparison between SCE using diagonal and full covariance matrix and RNN for coded BPSK, channel c, and 100x16 blocks

The RNN which did fail in the uncoded case offers in the coded case moderate performance. Furthermore an increase of the number of iterations up to a number of 15 improves the performance significantly.

Figure 9 shows the results for coded 16 QAM, channel b, and 100x16 blocks. In this channel even the diagonal approximation nearly achieves the lower coded AWGN bound, whereas the RNN looses 2.5 dB. Additionally Figure 9 depicts a comparison between the combined equalization and decoding scheme as proposed in this paper and a conventional receiver scheme, which consists of separated equalization and decoding devices. The curves of the conventional scheme are labeled with sequential. When comparing the results of the conventional receiver with the corresponding ones of the combined receiver, one can see for both equalizer types a re-

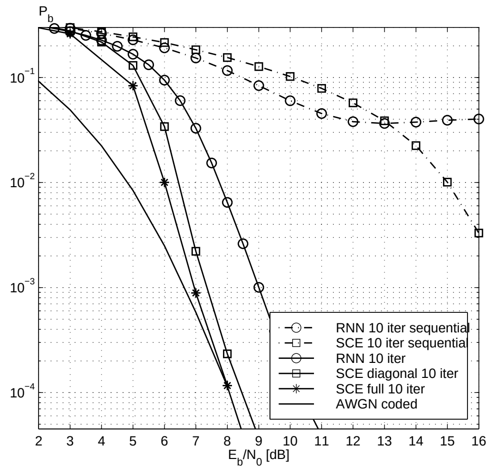

Figure 9: Comparison between SCE using diagonal and full covariance matrix and RNN using coded 16 QAM, channel b, 100x16 blocks, and 10 iterations

markable performance gain of at least 7 dB of the combined equalization and decoding scheme.

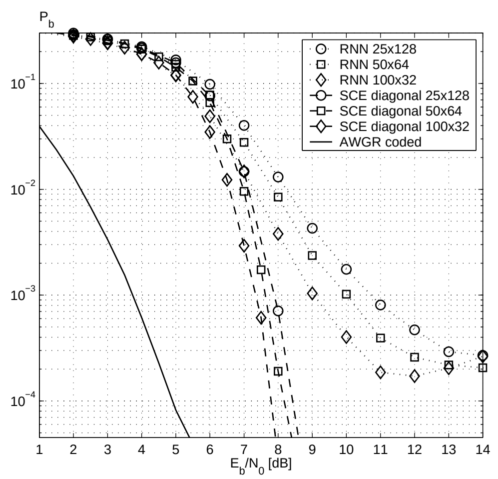

Figure 10: Comparison between SCE using diagonal RNN for varying blocklength using coded BPSK, channel c, and 10 iterations

Results for varying equalizer blocklength while keeping the codeword length constant are shown in Figure 10 for the BPSK case in channel c and in figure 11 in channel b. Both equalizers in both scenarios suffer from a performance degradation for increasing equalizer blocklength. Nevertheless, the decrease of the SCE is quite moderate in comparison compared with the RNN.

6 Conclusions

A new iterative equalization and decoding scheme based on the classical CBDFE has been introduced. It has been shown by simulation that this combination, even under some simplifying assumptions which reduce complexity, can achieve near optimum performance.

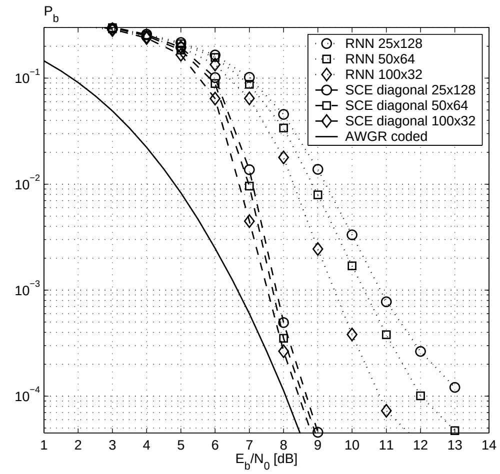

Figure 11: Comparison between SCE using diagonal RNN for varying blocklength using coded 16 QAM, channel b, and 10 iterations

References

[1] A. Bury and J. Lindner. Block transmission equalizers using constrained mininumum variance filters with application to mc-cdm. In *IEEE 6th International Symposium on Spread-Sprectrum Techniques and Applications, 6.9.-8.9.2000, New Jersey*, pages 159–163, Sept 1999.

[2] C. Douillard, M. Jezequel, C. Berrou, A. Picart, P. Didier, and A. Glavieux. Iterative Correction of Intersymbol Interference: Turbo-Equalization. *European Transactions on Telecommunications*, vol. 6(no. 5):pp. 507–511, Sept.-Oct. 1995.

[3] J. Egle, C. Sgraja, and J. Lindner. Iteratitive soft cholesky block decision feedback equalizer – a promising approach to combat interference. In *Vehicular Technology Conference 2001 (spring)*, 2001.

[4] T. Frey and M. Reinhardt. Signal Estimation for Interference Cancellation and Decision Feedback Equalization. In *Vehicular Technology Conference 1997*, pages 155–159, 1997.

[5] A. Lampe, R. Schober, and W. Gerstacker. A novel iterative multiuser detector for complex modulation schemes. In *Proceedings COST 262 Workhop Multiuser Detection in Spread Spectrum Communications*, pages 25–29, 2001.

[6] J. Lindner. MC-CDMA in the Context of General Multiuser / Multisubchannel Transmission methods. *European Transactions on Telecommunications*, no. 10:pp. 351–367, 1999.

[7] R. R. Müller and J. B. Huber. Iterated Soft-Decision Interference Cancellation for CDMA. In *Proceedings of 9th International Thyrrenian Workshop on Digital Communications*, 1997.

[8] F. D. Neeser and J. L. Massey. Proper complex random processes with application to information theory. *IEEE Transactions on Information Theory*, vol. 39(no. 4):pp. 1293–1302, 1994.

[9] J. G. Proakis. *Digital Communications*. McGraw-Hill, 1989.

[10] M. C. Reed. *Iterative Receiver Techniques for Code Multiple Access Communication Systems*. Ph.d. dissertation, University of South Australia, Faculty of Information Technology, 1998.

[11] C. Sgraja, A. Engelhart, W. G. Teich, and J. Lindner. Combined equalization and decoding for BFDM packet transmission schemes. In *Proceedings 1st international OFDM-Workshop, 22./23. Sept.*, 1999.

[12] C. Sgraja, A. Engelhart, W. G. Teich, and J. Lindner. Equalization with recurrent neural networks for complex-valued modulation schemes. In *Proc. 3rd Workshop Kommunikationstechnik, Schloß Reisensburg/Germany*, 1999.

[13] F. Tarköy. MMSE-optimal feedback and its applications. In *IEEE International Symposium on Information Theory, Whistler, B.C., Canada*, page 195, 1995.

[14] W. G. Teich and M. Seidl. Code Division Multiple Access Communications: Multiuser Detection based on a Recurrent Neural Network Structure. In *International Symposium on Spread Spectrum Techniques and Applications, 22.9.-25.9.1996, Mainz, Germany*, pages 979–984, 1996.

[15] S. Verdu. *Multiuser Detection*. Cambridge University Press, 1998.