LUND UNIVERSITY

Rate Maximizing Channel Shortening Detector With Soft Feedback Side Information.

Further contributions and results.

Medhat Mohamad

Abstract

In this thesis, the detection problem of signals transmitted throughout multiple input multiple output (MIMO) and/or inter-symbol interference (ISI) channels is studied. Channel shortening of trellis based detection is examined taking into account the maximum achievable rate that can be obtained. Two cases are checked: static channel shortening where there is no feedback side information and joint channel shortening with iterative detection.

A rate maximizing channel shortening detector with soft feedback side information for MIMO channel implemented using the Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm is studied. The information rate is maximized throughout each iteration step, where the detector is formed of a front end filter.

This thesis discusses a concave optimization procedure that provides the feedback filter coefficients as well as the trellis branch labels. This optimization procedure for MIMO channel follows different approach than the one studied for ISI channel and can be considered as a general but suboptimal approach.

For the static channel shortening technique, this thesis provides a modified closed form formula for the maximum achievable information rate, as well as a comparison between minimum mean squared error (MMSE) channel shortening techniques and rate maximizing channel shortening techniques. A closed form formula of maximum information rate of MIMO-ISI channel is derived.

Table of contents:

Introduction:

Wireless communication is one of the fastest growing domains throughout the whole developing technologies in communications industry. The great fusion between data communication and wireless systems, such as cellular phones, and local area networks (LANs) lead to the evolution of new technologies that witness exponentially increasing demands.

Today more than two billion human beings use cellular systems. Smart phones are now everywhere. These days, LANs are necessity [1].This great demand on wireless technology accompanies very challenging issues that trigger competition between companies and provide the base of thousands of researches.

In his famous paper “Mathematical theory of communication”, that represents the building block of the information theory; Shannon mentioned that the task of a communication system is to convey a certain message from transmitter to receiver. The receiver’s responsibility is to extract the message from the transmitted signal. In his paper Shannon determined the maximum rate that can be conveyed over some types of channels (BSC, AWGN channels, etc.). Since that time, researchers work on and modify the systems that are able to approach the capacity of the channel that Shannon determined. Ultimately, the best exploitation of the communication channel is the main goal for all communication engineers.

Maximum transmission rate under the lowest signal to noise ratio (SNR) possible and the lowest bit error rate (BER) under a specific bandwidth (BW) is a tradeoff subject to optimization throughout all communication systems, beside the lowest possible complexity for those systems. Receiver design is a topic that attracted and still attracts scientists and engineers throughout the world.

The design of receivers becomes a challenging subject especially with the obstacles that can face the signaling and signal processing techniques. One of the biggest obstacles is the ISI that dramatically affects the performance of any receiver. As well as the other users’ interference that are sharing the same channel for communication. These obstacles, beside the AWGN, need very sophisticated and highly intelligent detectors that should be able to extract the message from the received noisy signal.

The situation becomes more challenging when “fairly” new technologies, such as MIMO, are introduced. On the other hand, MIMO opens new dimension for multiplexing and/or diversity exploitation using space- time coding. This technology appeared since only 20 years and now is used in almost all modern communication protocols like long term evolution (LTE), worldwide interoperability for microwave access (WIMAX) and LANs and is still developing.

Thousands of researches in different directions are carried like: increasing the number of antenna arrays, decreasing multiuser (MU) and co-channel interference, researches on space time

(ST) coding and, as in our case, finding reliable receivers for MIMO systems. Finding reliable receivers becomes a crucial need, especially when a new phenomenon beside ISI and AWGN is affecting our system, which is the interference of different space channels.

This thesis investigates the channel shortening concept as an important technique that suppresses the increase in the receiver complexity. In trellis based receivers, the complexity is exponential in the memory length of the channel. Decreasing this complexity is carried by decreasing the channel memory length. This trellis coding uses different detection algorithms like Viterbi algorithm , soft output Viterbi algorithm (SOVA) or the optimum BCJR.

Separating the detection and decoding problem is one of the solutions for the exhaustive inefficient joint detection solution but this separation causes a tremendous degeneration of the system’s efficiency. Turbo equalization seems to be a prosperous solution. An insightful look at its principle will be recommended. The performance of turbo equalization with channel shortening concepts for systems with ISI interference is checked and a novel work on the performance of channel shortening rate maximization with soft feedback for MIMO systems is provided.

Chapter1: Communication systems and MIMO Channel

1.1- Communication model

According to Shannon, the rule of any communication system is to transmit a message and successfully extract the message from the received signal [2]. In digital communication systems this can be done through the steps illustrated in figure 1.1:

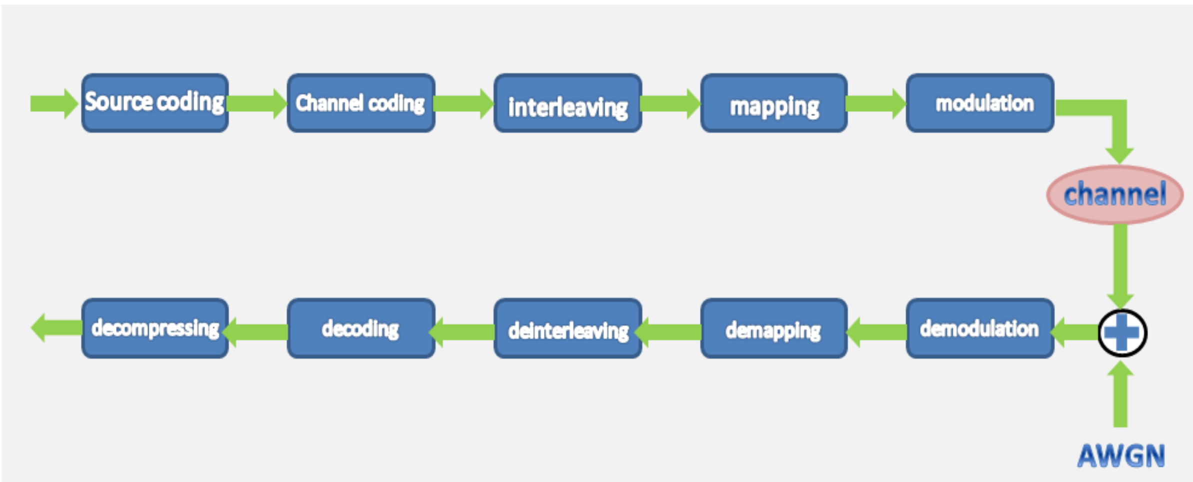


Fig. 1.1: communication model

The building blocks of figure 1.1 are briefly described below.

1- <u>Source coding:</u>

Transmission and storage of digital data require the information to be presented in digital form. Beside this main goal, source coding achieves two main points:

   a- Increase the bit rate for transmission throughout data compression

   b- Provide equally likely proportion between source encoded bits [3, 4].

2- <u>Channel coding:</u>

It is a type of redundancy added to the system so that to mitigate the effect of channel behavior such as noise and fading. According to Shannon, any system that transmits with rate higher than channel capacity will fail. On the other hand, systems that transmit with rates less than the capacity can be designed with low error probability. Decreasing error rate at the expense of decreasing transmission rate is done by coding techniques.

Coding techniques are classified into two main categories: error detection techniques and error correction techniques. Error detection techniques, such as cyclic redundancy check codes, allow the receiver to detect the error without the ability to correct it. While error correction techniques, such as convolutional codes, allow the receiver to find the error bits and correct them.

It is noteworthy to mention a class of coding techniques that approaches channel limits: Turbo codes and Low Density Parity Check (LDPC) codes, which are type of codes that attract the interest of researchers due to their high performance and reliability [5]. In this thesis, LDPC codes are used for simulations and comparisons.

3- Interleaving:

A way of shuffling adjacent bits in order to protect the message from burst errors. It also converts burst errors into single errors. There are a lot of interleaver models like S-Random interleavers [6]. S-Random interleavers randomly spread the bits with the restriction that each pair in group of S bits must be at least S indices apart after interleaving. Figure 1.2 depicts an S-random interleaver with S=3, for 18 code bits.

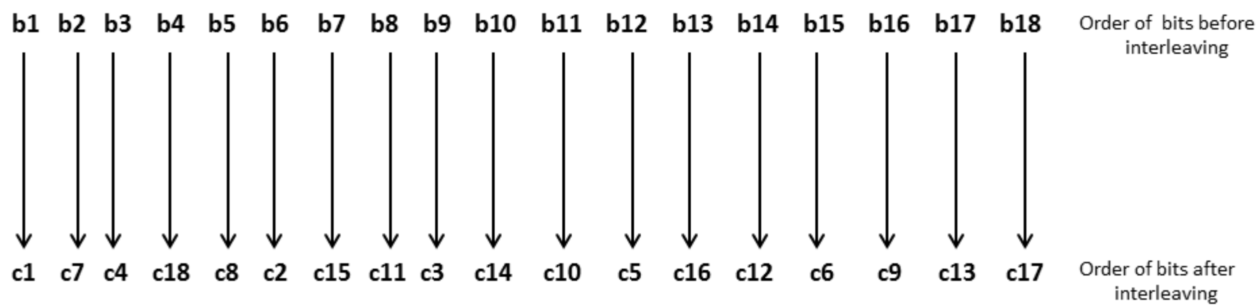


Fig. 1.2: 3-random interleaver for 18 code bits

4- Mapping:

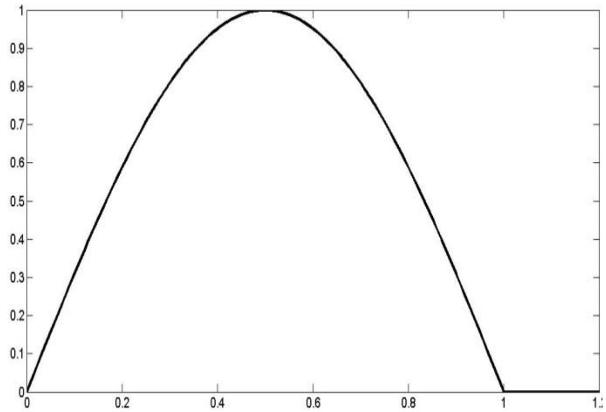
Depending on the bandwidth (BW) efficiency and required bit rate needed, different modulations schemes (BPSK, QPSK, QAM, M-QAM…) have to be chosen. This collects the encoded bits into symbols.

5- Modulation:

The final step is to modulate the symbols we have with the specific modulation pulse which is chosen in a way that provide the best spectral efficiency. Then the pulse is shifted to the carrier frequency that is used for transmission. Figure 1.3 illustrates some of the used modulation pulses in communication systems.

The signal is transmitted as an electromagnetic wave throughout a transmission antenna which could be a single or a multiple system. The signal passes through the channel that deforms the signal and adds AWGN before receiving and processing of the signal.

The challenge at the receiver lies on extracting the message from the deformed signal which is going to be done by reciprocating the transmission process plus equalizing the effect of the channel so that transmitted message 'hopefully' extracted.



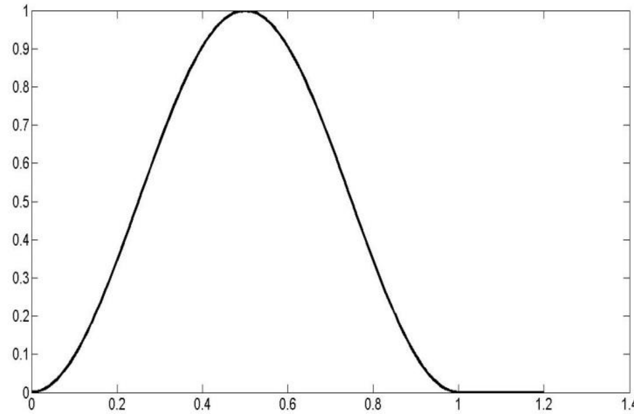


Fig.1.3: half cycle sine vs. raised cosine pulses: used for modulation of symbols

1.2- Channel Model

In this thesis, linearly modulated transmissions affected by AWGN are considered with the discrimination among three different cases:

1- **<u>ISI channel model</u>** is represented as:

$$\boldsymbol{y} = \boldsymbol{h} * \boldsymbol{x} + \boldsymbol{n}, \qquad (1.1)$$

where $\boldsymbol{h} = [h_0,\ h_1,\ h_2 \dots, h_L\ ]$ is the complex-valued channel impulse response (CIR) vector with L+1 coefficients, $\boldsymbol{x}$ denotes the input symbols' vector, $*$ is the convolution operator and $\boldsymbol{n}$ is the noise vector where the noise samples are assumed to be independent identically distributed (i.i.d) zero mean circularly symmetric complex Gaussian (ZMCSCG) random variables.

ISI channel is represented in matrix notation as follows:

$$\boldsymbol{y} = H\boldsymbol{x} + \boldsymbol{n}, \qquad (1.2)$$

where H is the convolutional matrix that is represented as:

$$\begin{bmatrix} h_0 & 0 & 0 & 0 & 0 & 0 & 0 \\ h_1 & h_0 & 0 & 0 & 0 & 0 & 0 \\ h_2 & h_1 & h_0 & 0 & 0 & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 & 0 & 0 \\ h_L & \dots & h_2 & h_1 & h_0 & 0 & 0 \\ 0 & h_L & \dots & h_2 & h_1 & h_0 & 0 \\ 0 & 0 & h_L & \dots & h_2 & h_1 & h_0 \end{bmatrix},$$

figure 1.4 shows the plot of a convolutional ISI channel.

2- **MIMO channel model:** is represented as in equation (1.2):

$\boldsymbol{y} = H\boldsymbol{x} + \boldsymbol{n},$

where H ,which is modeled as:

$$\begin{bmatrix} h_{1,1} & \cdots & h_{1,N} \\ \vdots & \ddots & \vdots \\ h_{N,1} & \cdots & h_{M,N} \end{bmatrix},$$

represents the inter-antenna interference channel matrix. M and N is the number of transmitting and receiving antennas respectively.

3- **MIMO-ISI channel model:** is given by channel matrix H:

$$\begin{bmatrix} H_0 & 0 & 0 & 0 & 0 & 0 & 0 \\ H_1 & H_0 & 0 & 0 & 0 & 0 & 0 \\ H_2 & H_1 & H_0 & 0 & 0 & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 & 0 & 0 \\ H_L & \cdots & H_2 & H_1 & H_0 & 0 & 0 \\ 0 & H_L & \cdots & H_2 & H_1 & H_0 & 0 \\ 0 & 0 & H_L & \cdots & H_2 & H_1 & H_0 \end{bmatrix},$$

where H_i is the i^{th} MIMO channel.

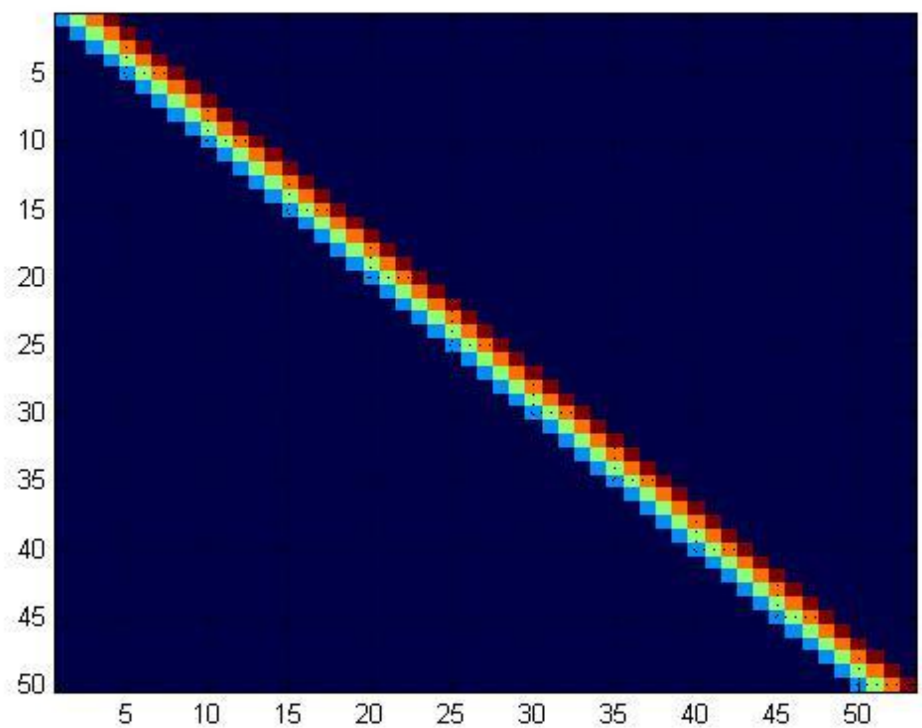


Fig. 1.4: ISI convolutional channel

1.3- MIMO channel capacity

The urgent need for higher transmission rates pushes communication engineers to continuously search for systems that can handle this tremendous growth. One of the developing systems that become very popular in recent years is the MIMO systems (or what is also known as space-time wireless systems). The benefit of those systems is that they opened a new important resource in communication systems beside time and bandwidth, namely space. This section gives a brief outlook on the MIMO channel capacity.

Starting by the definition of the channel capacity as the highest information rate that can be transmitted over the channel, the information rate (in bits) is given by:

$$I(X;Y) = H(Y) - H(Y|X) = H(Y) - H(n)\,, \tag{1.3}$$

where $I(X;Y)$ is the mutual information operator, and $H(.)$ is the differential entropy operator.

Assuming AWGN "n" and a Gaussian distribution of the transmitted symbols:

$$H(n) = \log(\det(\pi e R_{nn})),$$

where R_{nn} is the covariance matrix of the noise, and equals to $N_o I$, and:

$$H(Y) = \log(\det(\pi e R_{YY})),$$

where R_{YY} is the covariance matrix of the received symbols and is given by:

$$R_{YY} = E\{yy^T\} = E\{(Hy+n)(Hy+n)^T\} = E\{(Hx+n)(x^T H^T + n^T)\}$$

$$= E\{Hxx^T + nx^T H^T + Hxn^T + nn^T\}.$$

This can be expressed as:

$$R_{YY} = \ E\{Hxx^T\} + E\{nx^T H^T\} + E\{Hxn^T\} + E\{nn^T\},$$

where $E\{\,.\}$ is the expectation operator.

Since noise and transmitted symbols are uncorrelated this gives:

$$R_{YY} = \ E\{Hxx^T\} + E\{nn^T\} = HR_{xx}H^T + N_o I,$$

which wraps up expression of the information rate as:

$$I(X;Y) = \log\left(\det\left(\frac{R_{YY}}{R_{nn}}\right)\right) = \log\left(\det\left(\frac{HR_{xx}H^T + N_o I}{N_o I}\right)\right) = \log\left(det\left(I + \frac{HR_{xx}H^T}{N_o}\right)\right),$$

therefore the capacity of the channel is given by:

$$C = \max_{Tr\{R_{xx}\}} I(X;Y) = \max_{Tr\{R_{xx}\}} \log\left(det\left(I + \frac{HR_{xx}H^T}{N_o}\right)\right). \tag{1.4}$$

Equation (1.4) shows that the capacity of the MIMO system is controlled by the distribution of the transmit symbols. This depends on the amount of channel knowledge available at the transmitter and receiver. This thesis assumes perfect knowledge of the channel at the receiver side. Channel knowledge at the transmitter side is difficult to obtain and increases the complexity of the communication system.

Channel knowledge at the transmitter is mainly achieved by reciprocity principle, which assumes that the behavior of the channel in the downlink is similar to its behavior on the uplink, or by using the feedback information of the channel from the receiver to the transmitter. If the channel is unknown to the transmitter, the assumption that the symbols are independent and possess equal power is valid, therefore $R_{xx} = I$ and then the capacity of the system will be

$$C = \log\left(det\left(I + \frac{HH^T}{N_o}\right)\right), \tag{1.5}$$

although this is not the exact Shannon capacity since R_{xx} is assumed to be equal to I where I is identity matrix.

Taking the Eigen decomposition of HH^T, the capacity relation becomes:

$$C = \log\left(det\left(I + \frac{Q\Lambda Q^T}{N_o}\right)\right) = \log\left(det\left(I + \frac{\Lambda}{N_o}\right)\right) = \sum_{i=1}^{r} \log\left(1 + \frac{\lambda_i}{N_o}\right), \tag{1.6}$$

The most favorable case would be when all the Eigen values are equal i.e. the channel is an orthogonal matrix, then $C = r\log\left(I + \frac{\lambda}{N_o}\right)$, where r is the rank of the channel matrix and is equal to the number of receiving antennas. Figure 1.5 shows the capacity of an orthogonal channel.

For a channel known at the transmitter, pre-filtering of the symbols can be made before transmission. The idea falls in the control of the R_{xx} matrix. This can be done by using the water filling algorithm which is achieved by putting more power in the channel branches that possess better behavior. The capacity after applying the algorithm will be as following:

$$C = \sum_{i=1}^{r} \log(1 + \frac{\lambda_i \gamma_i}{N_o}), \tag{1.7}$$

where γ_i represents the transmitted energy in i^{th} sub-channel. Figure 1.6 shows a schematic diagram of the water filling algorithm and figure 1.7 shows the performance of the known and unknown MIMO system channels [7].

<u>Ergodic VS outage capacities:</u>

Since the channel of a communication system is a stochastic process then the capacity behavior of the channel is assumed to be random as it depends on the channel itself. I.e. the capacity value is going to change with each realization of channel. From here, the overall capacity of a system is measured by taking the ensemble average of the information rate calculated for each realization of the channel. This capacity is called the Ergodic capacity (or mean capacity) and is equal to:

$$\bar{C} = E\left\{\log\left(det\left(I + \frac{HH^T}{N_o}\right)\right)\right\}, \tag{1.8}$$

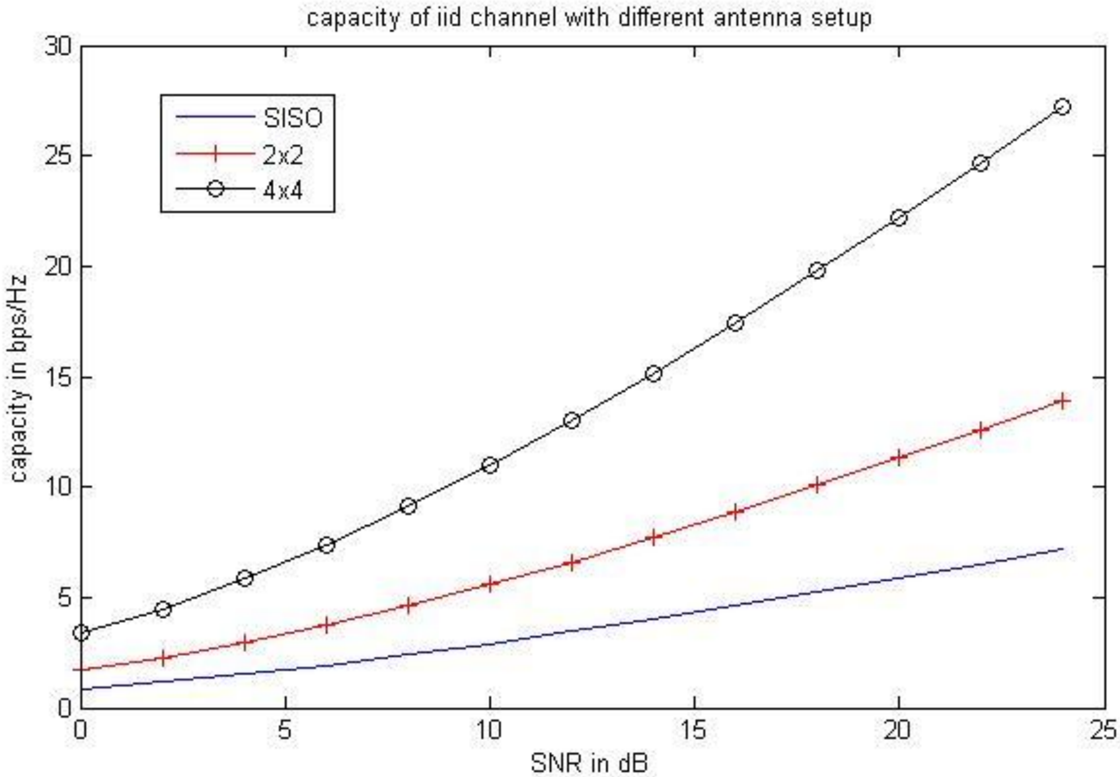


Fig. 1.5: Capacity of orthogonal MIMO channel

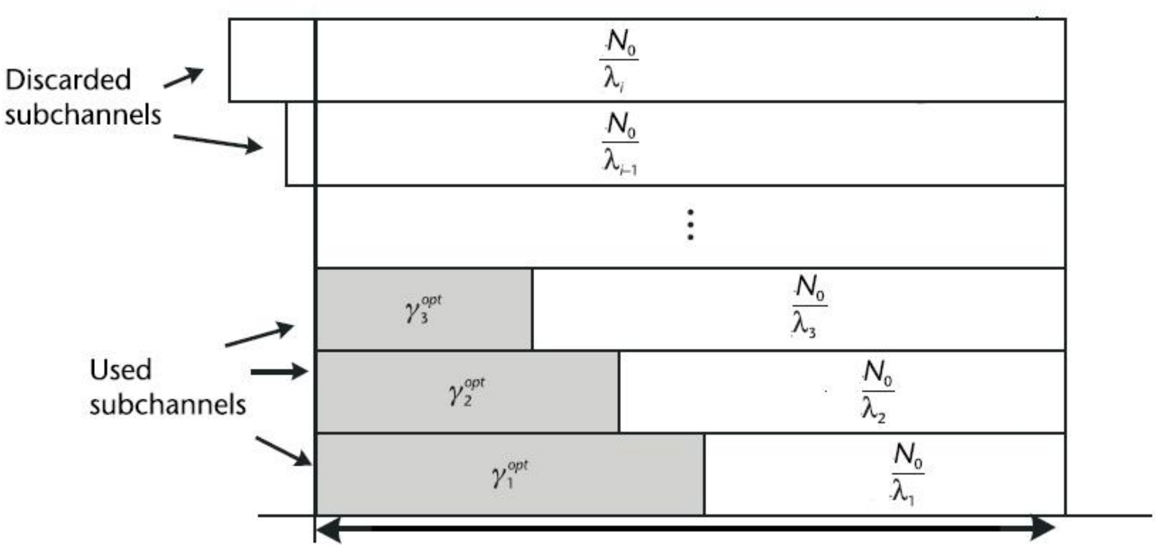


Fig. 1.6: water filling algorithm

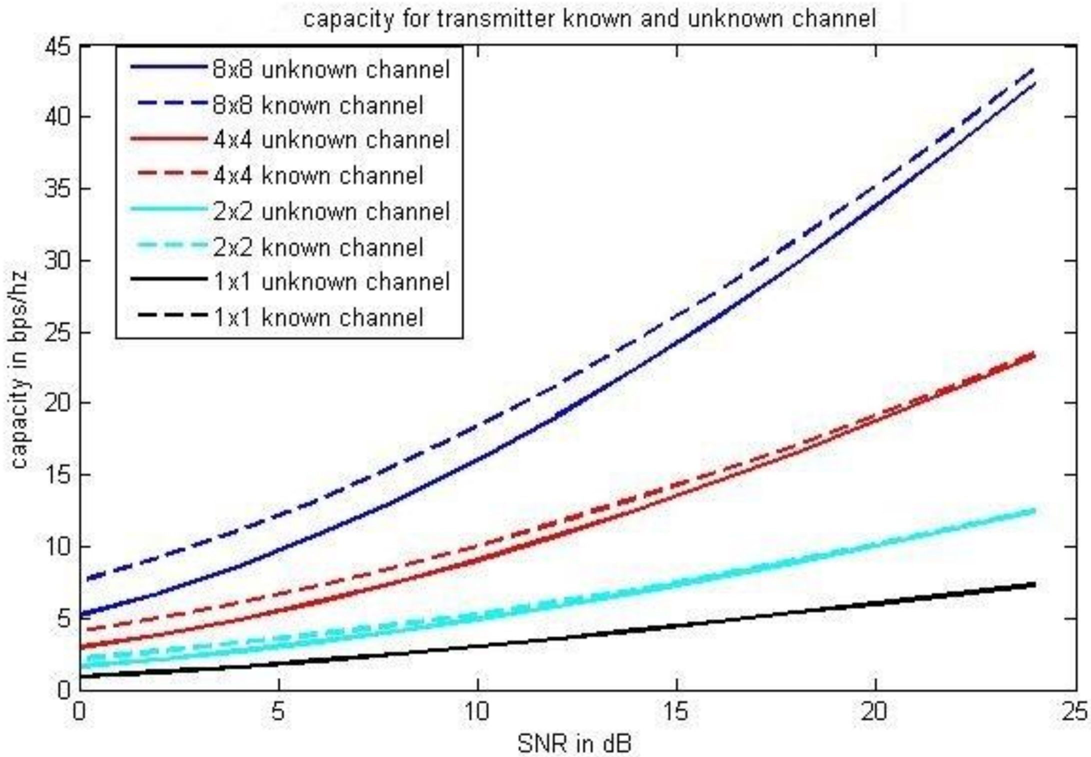


Fig. 1.7: capacity of MIMO systems with and without channel knowledge at the transmitter

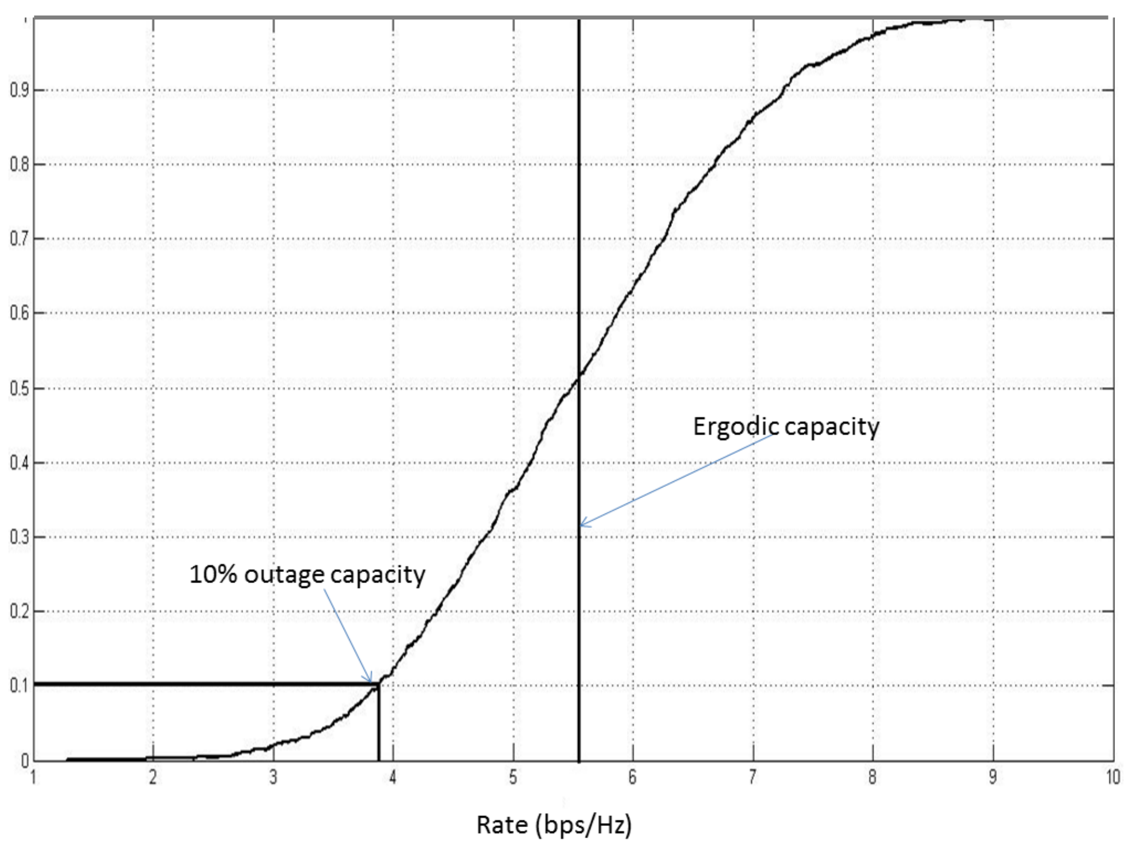


Fig.1.8: Ergodic and 10% outage capacity of a 2x2 MIMO system

The outage capacity is another estimation of the capacity performance. The q% channel capacity tells us how much capacity (100-q) % of realizations can achieve while the other q% of realizations cannot. For example, 10% outage capacity means that 90% of the realizations would achieve this information rate while 10% will fail. Outage capacity is a quantification of performance level under certain conditions of reliability. Figure 1.8 shows the cumulative distribution function (CDF) of I.I.D channel information rate with Ergodic capacity and 10% outage capacity [7]. This thesis sticks with the Ergodic measure of the channel capacity.

Chap 2: Equalization:

ISI, as well as Inter antenna interference, are serious problems that affect the performance of a communication system. Dealing with ISI and inter antenna interference is a necessity for a communication system to be reliable. For successful receiving of the message, detectors should have the ability to mitigate the effect of interference.

In CDMA technology rake receivers are used; a set of receivers detect the arriving multipath signals and collect (rake up) the energy from those different paths. Rather than taking the energy of the best peak arriving and discarding the other peaks.

Rake receivers consist of banks of correlators. Each is sampled at different delay time. The samples out of the correlator are then weighted and combined [8]. Rake receivers suffer from a crucial problem which is cost. As the receiver supports more multipath components the correlators' bank will increase and thus the cost of the receiver.

Another solution for interference problem is equalization which aims to cancel the effect of the channel by assuming that the channel is a type of filter. So, concatenation of this filter with a filter that opposes its action will hopefully get rid of the interference. Figure 3.1 shows schematics of rake receivers and equalizers.

Equalizers can be classified into linear equalizers, like zero forcing and MMSE equalizers, and nonlinear equalizers, like decision feedback equalizers (DFE) and maximum likelihood sequence estimators (MLSE). Moreover, a powerful tool is to use iterative equalizers that perform the equalization (and decoding) process iteratively.

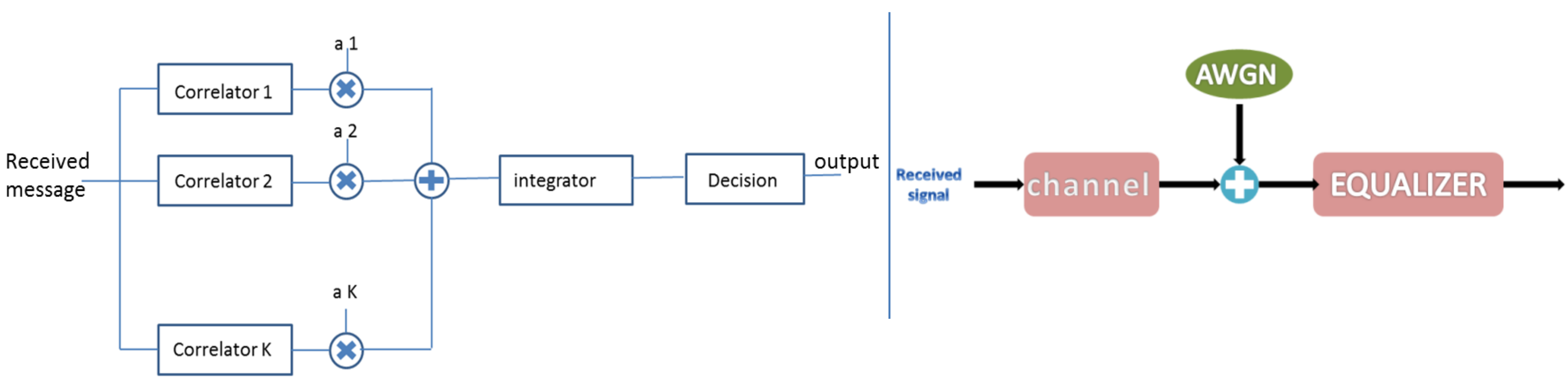


Fig. 2.1: rake receivers vs. equalizer schematic

2.1- Linear equalizers

The main goal of linear equalizers is to simply invert the effect of the channel. Linear equalizers are either zero forcing equalizers (ZF) or MMSE equalizers.

a- <u>Zero forcing equalizers:</u>

ZF equalizer is a linear equalizer that completely cancels the effect of the channel, so that the transfer function resulting from the product of the channel and equalizer transfer function is a completely flat function. This is done by simply multiplying the channel matrix by its inverse. A transmitted symbol vector **x** passing in a MIMO ISI channel **H** which adds AWGN vector **n** would be received as a symbol vector **y.** Filtering **y** with ZF filter $\boldsymbol{g} = \boldsymbol{H}^{-1}$ gives:

$$\mathbf{z} = \boldsymbol{g}\,\mathbf{y} = \boldsymbol{H}^{-1}\mathbf{y} = \boldsymbol{H}^{-1}(\boldsymbol{H}\mathbf{x}+\mathbf{n}) = \mathbf{x} + \boldsymbol{H}^{-1}\mathbf{n}. \qquad (2.1)$$

As it is obvious form the equation a serious problem appears from zero forcing equalization which is noise coloring and enhancement. This problem makes the efficiency of a ZF equalizer to degrade significantly. Figure 2.2 shows an illustration of the noise enhancement problem [8].

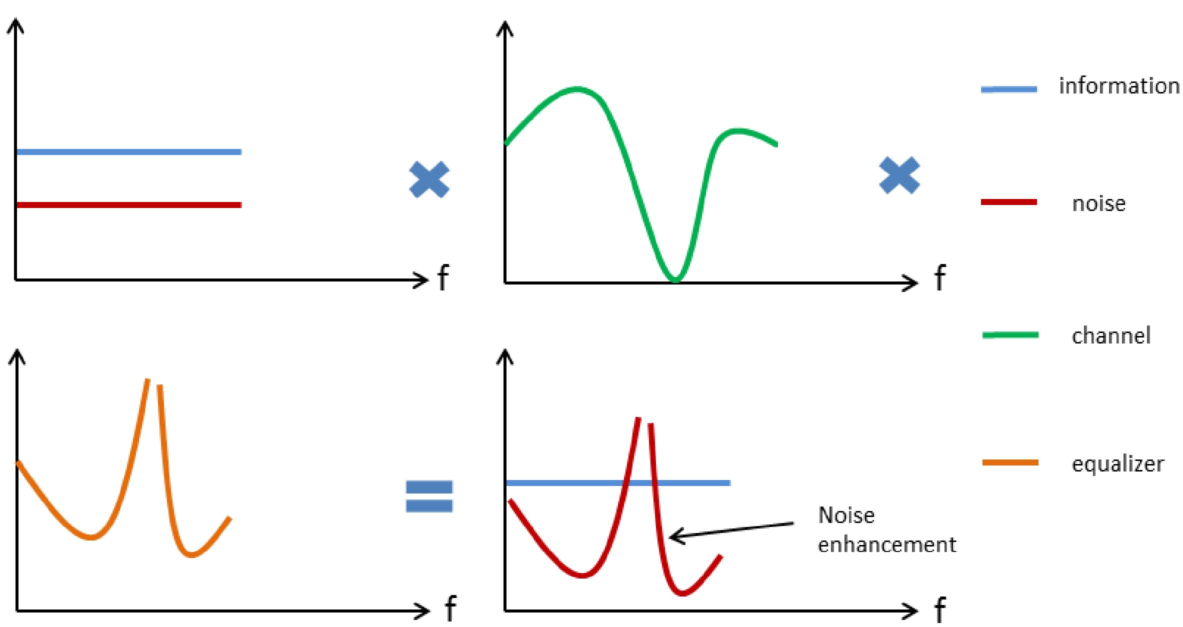


Fig 2.2: noise enhancement with ZF equalizers

b- <u>MMSE equalization:</u>

Changing the main goal of the equalizer from completely suppressing the effect of ISI (like in the ZF equalizer) to minimization of the BER, a better choice will be is to use the MMSE equalization. The main goal of the mean square equalization is to minimize the mean square error between the transmitted symbols and the equalizer output:

$\mathbf{g}_{MMSE} = \text{argmin}E\{|e^2|\}=\text{argmin}E\{ee^*\}$

where e is the error between transmitted and filtered signal, but e=**gy**-**x**. the orthogonality principle states that the error should be orthogonal to the received sequence **y** therefore E{e**y***}=**0**.

Thus E{(**gy-x**)**y***} =**0** implies E{**gyy***}-E{**xy***}=**0** implies: $\mathbf{g}\mathbf{R}_{yy} - \mathbf{R}_{xy} = \mathbf{0}$,

where: $\mathbf{R}_{yy}$ and $\mathbf{R}_{xy}$ are the autocorrelation matrix of the received symbol and the cross correlation matrix between received and transmitted symbols respectively. Therefore:

$$\mathbf{g} = \mathbf{R}_{xy}\mathbf{R}_{yy}^{-1}. \tag{2.2}$$

Figure 3.3 shows a schematic where MMSE mitigates the noise enhancement problem found in ZF equalizers. Figure 3.4 shows the BER performance of an MMSE equalizer compared with ZF equalizer. Performance of MMSE is better than that of ZF equalizers but still far away from the optimum solution [8].

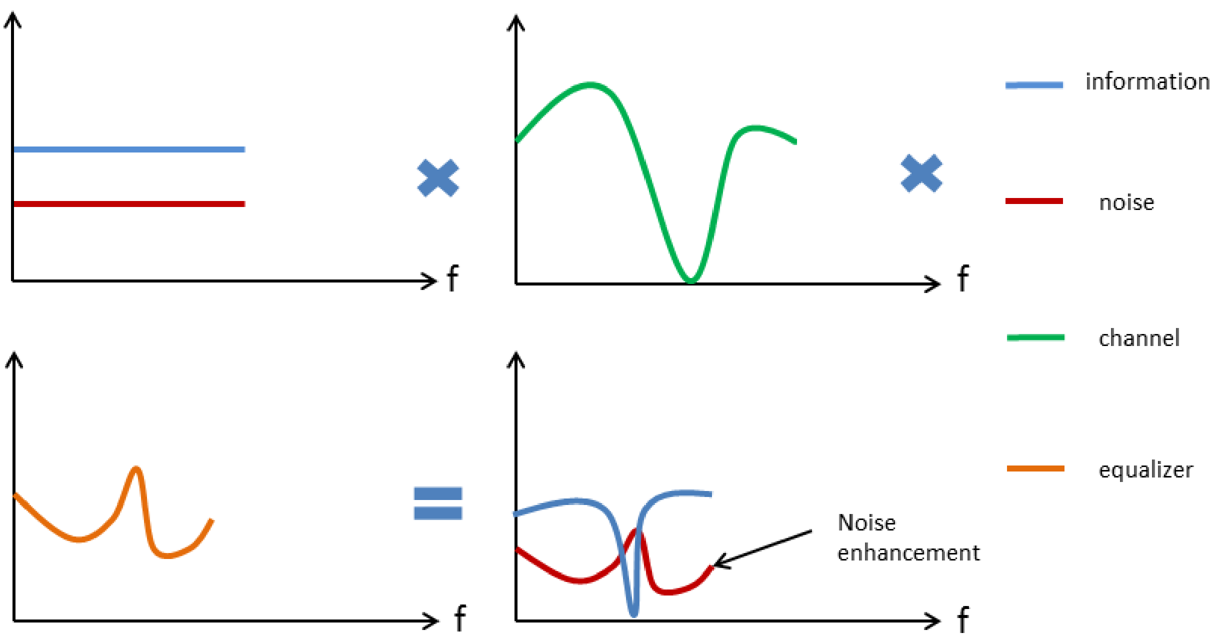


Fig.2.3: MMSE effect on noise enhancement problem

2.2- Nonlinear equalizers

a- Decision feedback equalizers (DFE):

ISI (or spatial channel interference) can be represented as the effect of one symbol (or more) on the other symbols. So as soon as one symbol (or more) is already known, its effect on the incoming symbols is known and mitigated. This is the main goal of DFE. Figure 2.5 shows a schematic of a DFE. The equalizer is formed of two filters: forward filter which is a conventional equalizer (can be ZF or MMSE) and a feedback filter. The main problem that can be found in DFE is error propagation. Since the current symbol decision depends on the previous symbols decisions occurrence of the error in previous symbols, will propagate the error on the current symbols [8].

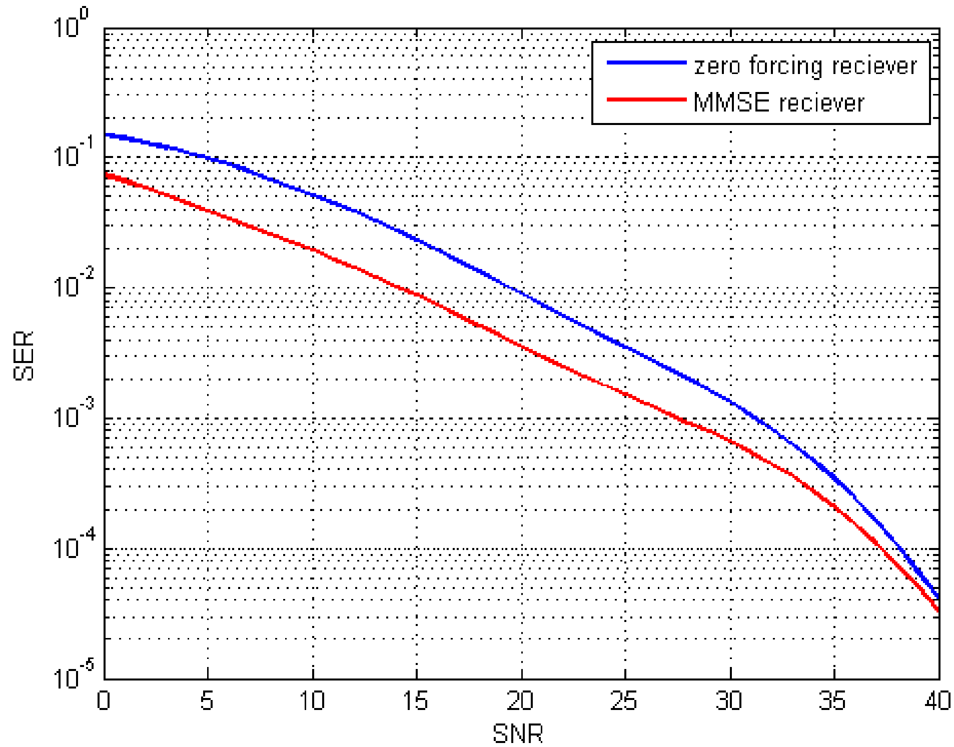


Fig. 2.4: ZF VS MMSE SER performance of 2x2 MIMO system with BPSK modulation scheme

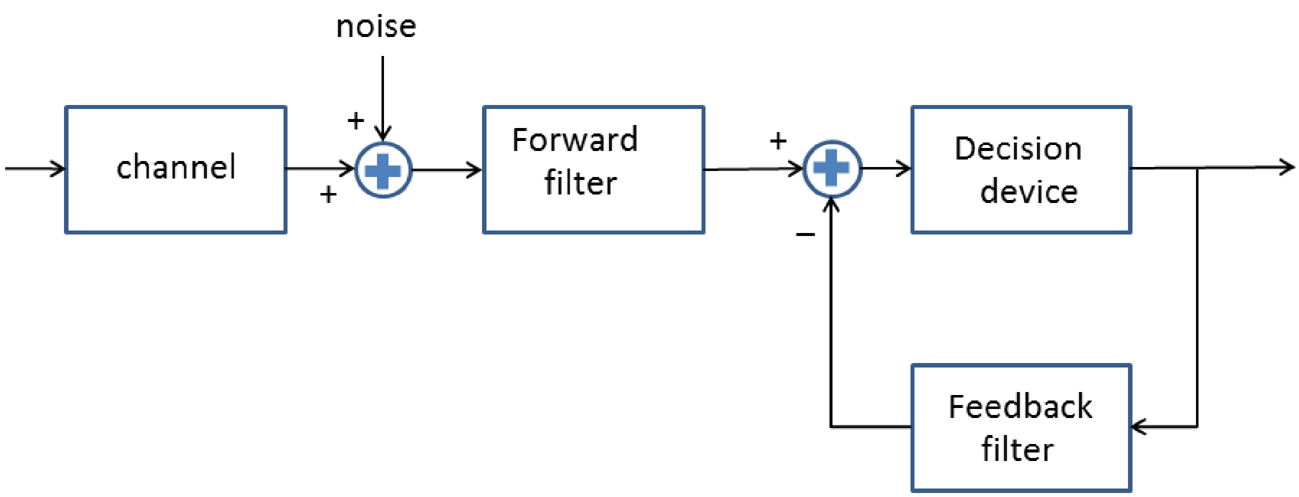


Fig.2.5: DFE scheme

b- <u>maximum likelihood sequence estimators (MLSE):</u>

In ISI, the channel is expressed as a memory system similar to a convolutional encoder with rate 1/1. The equalization then can be obtained using the trellis based approach. This trellis based approach can be also used for equalization of spatial interference channels. Therefore, sequence algorithms such as Viterbi or BCJR algorithms can be used for equalization purpose. The main difference between MLSE and previous equalizers is that the previous systems make symbol by symbol operation while MLSE equalizers work on sequence of symbols. MLSE have the best performance among all equalizers.

Although MLSE have the best performance, they have the highest complexity. The complexity of the MLSE increases exponentially with the channel memory. On the other hand the complexity of the linear equalizers increases (linearly or quadraticaly in case of MMSE).

The exponential growth of MLSE with memory taps push researchers to find ways to decrease this complexity. One path is shortening the channel memory which is the main topic of the thesis and will be discussed in chapter 4.

2.3- Turbo equalization

In optimal detection the receiver considers equalization, demapping and decoding as a joint process and then judge on the sequence of bits that are most probably transmitted given some received sequence. So in this system we are going to compare the received sequence with all the probable sequences, which is a very exhaustive complex process. It is obvious that computational complexity of the system grows exponentially with the block size. Because of this high level of complexity, practical detectors work on equalization and decoding as separated processes.

The receiving process can be decomposed (and then become suboptimum but less complex) into equalization process and decoding process. Figure 3.6 shows the detection process. First equalization is held to mitigate the effect of interference and then demapping and decoding as separated processes.

For equalization, as discussed previously, either the linear equalizers or the efficient trellis approach is followed. This sequence estimation can take place in two ways: either hard decision like Viterbi algorithm and in this case the hard decision will destroy the soft information, or the soft decision which is the case with the BCJR algorithm.

The BCJR algorithm can be represented as a system that takes two inputs: A priori information and some observations and produce a one output which is the a posteriori information. Figure 2.6 shows the BCJR algorithm block.

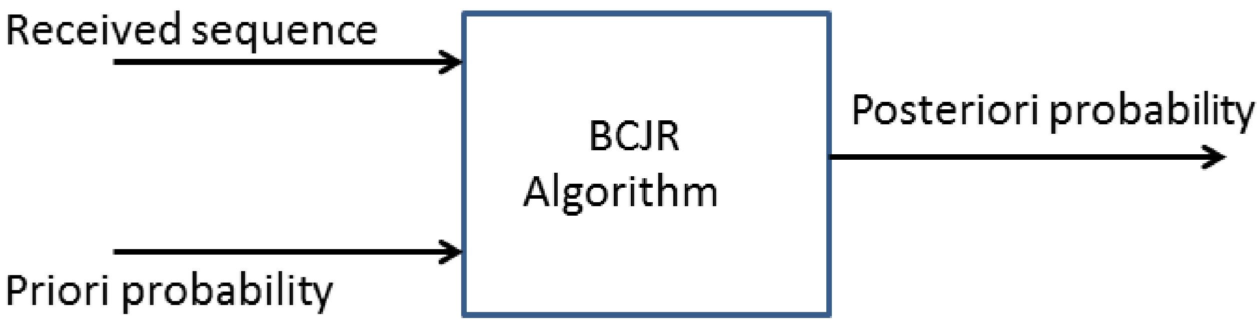


Fig. 2.6: BCJR algorithm with priori prob and received seq as input and post. Prob. as output

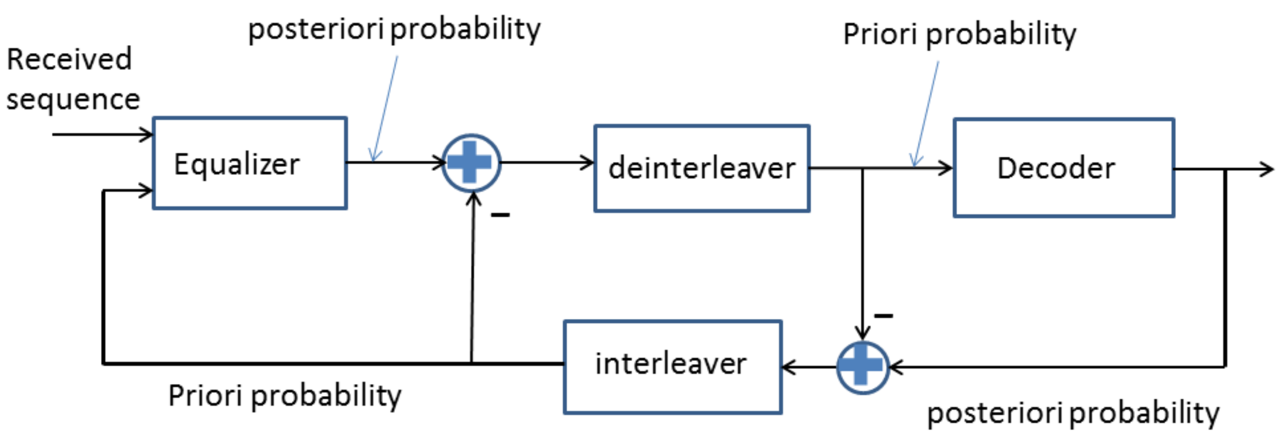


Fig. 2.7: turbo equalizer with two BCJR blocks equalizer and decoder

In the separated equalization decoding strategy, the equalizer doesn't have any prior knowledge available and solely depends on the local observations coming from the channel. On the other hand the decoder builds its decisions on the priori information provided from the equalizer without any local observations. Therefore the performance of the BCJR algorithm can be improved by feeding back the output of the decoder (APP) into the input of the equalizer as a priori information. This is the core of the turbo equalization principle. Fig. 2.7 shows the diagram of a turbo equalizer [9].

One point to mention here: the only feedback information needed is the extrinsic information which represents the new information to the equalizer (or the decoder) otherwise, feeding the intrinsic information as well, means that the equalizer (or the decoder) is told what it already knows. And this causes a fast convergence. Moreover, due to the high correlation between neighboring symbols, shuffling those symbols away from each other will further prohibit the fast convergence effect. This shuffling is done using interleaver/deinterleaver.

Chap 3 channel shortening concept:

MLSE and MAP detectors are optimum compared to the other types of equalizers. But as mentioned earlier, the main problem is the complexity of those detectors that increases exponentially with the length of channel memory. Assume, for example, a constellation set ♀ with cardinality Q and an ISI channel with memory length L. Using a MIMO system of MxM antennas, the number of paths in the trellis algorithm will be $Q^{M(L-1)}$. For terrestrial terrain channel, for example, L can extend for more than 50 taps which is a severe exhaustive calculation situation. Due to this fact, decreasing the complexity becomes a crucial demand on such systems.

Decreasing the complexity of the system can be carried by one of the two strategies:

1- Processing part of the full original trellis, so that a fraction of the trellis paths is checked.
2- Decreasing the number of states which the trellis has, so that a reduced trellis is obtained. This can be achieved using channel shortening techniques.

There are many examples regarding the first approach such as sphere detection and M algorithms.

3.1- Trellis pruning

Sphere detection:

The MLSE is a way of searching among a set of lattice points in order to find a lattice point that is closest to the received sequence. Graphically, the received sequence is going to be a point that falls in the lattice space. The nearest lattice point to the received point is assumed as the transmitted sequence. Figure 3.1 shows a schematic for lattice structure with the received sequence point.

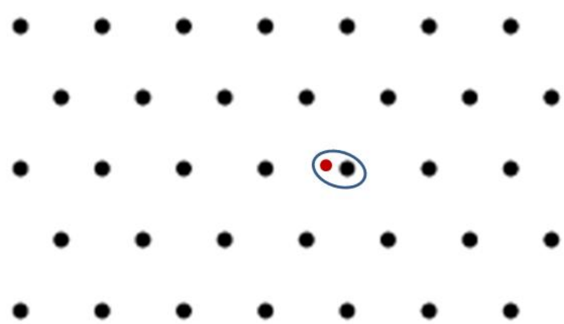


Fig. 3.1: lattice concept (black pts are the valid lattice pts and the red dot is the received sequence)

The major idea of sphere decoding is to search over a set of lattice points that lie in a hyper sphere instead of searching the entire lattice. The sphere is of radius “R”, while the center is represented by the received sequence point. Thus, the searching complexity is reduced. Figure 3.2 shows the sphere detection scheme. The question that should be asked is how the radius R is chosen and how the optimal lattice point that falls inside this radius is determined? Fincke and Pohst proposed an algorithm in 1985 that effectively does this [10].

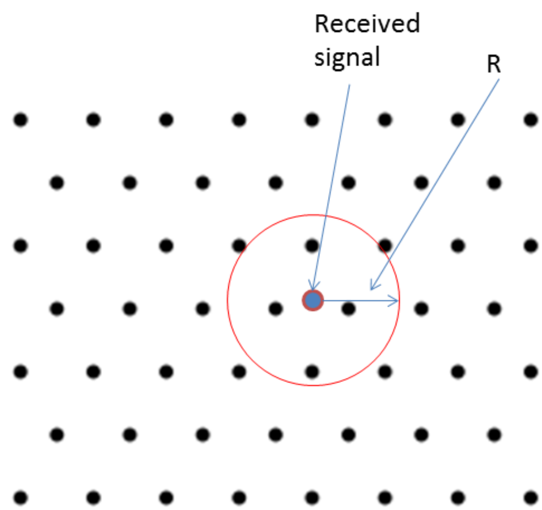


Fig. 3.2: sphere detection: lattice inside the circle are the only points decoded.

M-algorithm:

In the M- algorithm the trellis is pruned in a way that only the M states, which have the highest cumulative metrics, are kept. The other states are terminated and no branches from the terminated states are extended. So as each new input is received, the algorithm extends the M surviving states to the next trellis depth, checks the cumulative metrics of the new states, keeps the best M states and terminates the others.

The process is repeated till the end of trellis is reached. The M-algorithm can possess a complexity of $QM \times \log(QM)$ instead of Q^L, where Q is the number of overall states, M is the number of survival states, and L is the memory length of the channel [11].

3.2- Reduction of channel complexity by channel shortening:

The other strategy to decrease the complexity of trellis detection is by shortening the trellis itself. This is done by decreasing the number of states in the trellis. To decrease the memory length of the channel, filtering is carried. After filtering, the desired impulse response (DIR) output will be shorter than the exact channel impulse response.

Since Forney's novel work in 1972, where he proved that ML detection for ISI channels (and for MIMO channels later) can be applied by using the Viterbi algorithm, and researches are done to find a way of shortening the channel length. These researches vary from simple techniques, like minimum phase filtering of the channel, till the wide area of researches following the MMSE approach suggested by Magee and Falconer in 1973.

Channel shortening using Minimum phase filtering:

A trivial way of channel shortening (used sometimes in GSM technologies) is the minimum phase filtering of the channel. The received symbols are filtered by a minimum phase filter. The minimum phase filter pushes the channel taps that have the highest energy level to be the first received taps and those with lower energy level follow i.e. sorting of the tapes according to descending power level. Figure 3.3 shows the effect of phase filtering. After phase filtering, taps of low power under a certain threshold of energy constrain are truncated. Although this method of channel shortening is very simple, yet it is far from optimality. Normally, truncation causes loss of performance.

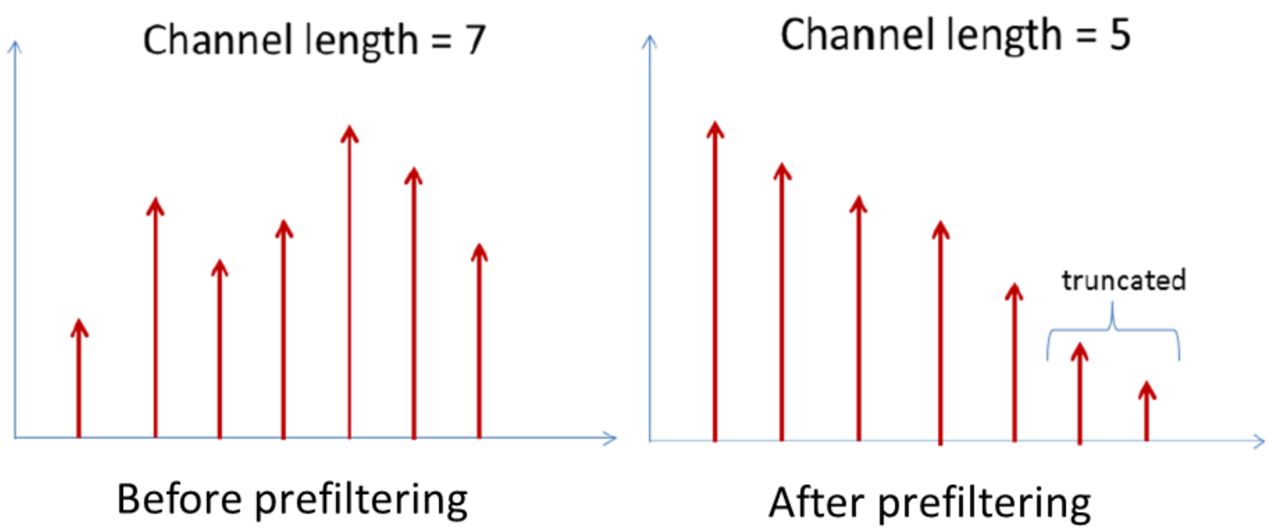


Fig. 3.3: minimum phase filtering

Channel shortening using MMSE criterion:

The model proposed by Magee and Falconer is shown in figure 3.4. As illustrated by the figure, the received sequence is passed to a pre-filter before the Viterbi equalization. The outputs of the equalizer, which are the sequence estimates, are passed to DIR filter. The output of the filter is compared with the received sequence and the MSE of the two sequences is optimized as follows:

$$\min_{p,q} E\{e^2\}, \tag{3.1}$$

where e is the error between received sequence and DIR outputs. Using equation (1.2), the filtered output will be $\boldsymbol{py}$. On the other hand the DIR output will be $\boldsymbol{qI}$. Therefore: $e = \boldsymbol{py} - \boldsymbol{qI}$, and thus

$$E\{e^2\} = E\{ee^*\} = E\{(\boldsymbol{py}\text{-}\ \boldsymbol{qI})(\boldsymbol{y}^*\boldsymbol{p}^* - \boldsymbol{I}^*\boldsymbol{q}^*)\} = E\{\boldsymbol{p}^*A\boldsymbol{p} + \boldsymbol{q}^*\boldsymbol{q} - 2\boldsymbol{p}^*H\boldsymbol{q}\}, \tag{3.2}$$

where $A = R_{yy}$.

The error is minimized with respect to the pre-filter $\boldsymbol{p}$ by differentiating $E\{e^2\}$ w.r.t $\boldsymbol{p}$ and setting it equal to zero

$$2A\boldsymbol{p} - 2H\boldsymbol{q} = 0, \tag{3.3}$$

and thus

$$\boldsymbol{p}_{opt} = A^{-1}H\boldsymbol{q}, \tag{3.4}$$

and

$$E\{e^2\} = \boldsymbol{q}^*[I - H^*A^{-1}H]\boldsymbol{q}. \tag{3.5}$$

Next, optimization with respect to $\boldsymbol{q}$ is carried out, where minimum mean square of error is the minimum Eigen value of the matrix $I - H^*A^{-1}H$ and thus $\boldsymbol{q}$ represents the minimum Eigen vector of the matrix [12].

The Magee-Falconer model assumes equal distribution of ISI, although in reality this is not the case. So a modification of this proposal is provided by Ödling et al. where weighting for the ISI residuals is carried out.

Figure 3.5 shows the modified model, where D is the delay parameter and W is the weighting filter [13]. Note that D is identity matrix for Magee-Falconer model.

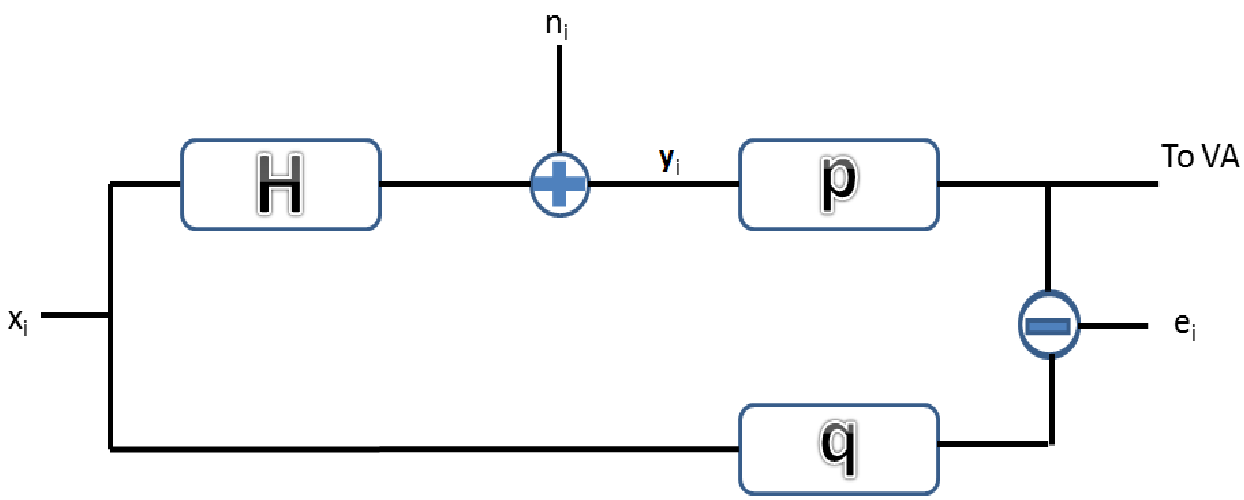


Fig. 3.4: MMSE channel shortener proposed by Magee and Falconer

Therefore we get:

$$E\{e^2\} = |W(H\boldsymbol{p} - D\boldsymbol{q})|^2 + |\boldsymbol{p}|^2 \times {\sigma_n}^2. \tag{3.6}$$

If $W = I$ (i.e. no weighting) the model ends up in Magee-Falconer model.

A similar approach for this work is held for MIMO-ISI channels in [14]. The scheme of a MIMO-ISI system is shown in figure 3.5 where W represents the received symbols filter,D^{Δ} represents the delay parameter, B represents the target impulse response(TIR), ni, no and L represent the number of Tx and Rx antennas, and the memory length respectively. It looks very similar to the model proposed by Magee and Falconer. Figure 3.6 shows the plot of a block toeplitz matrix which represents a MIMO ISI channel.

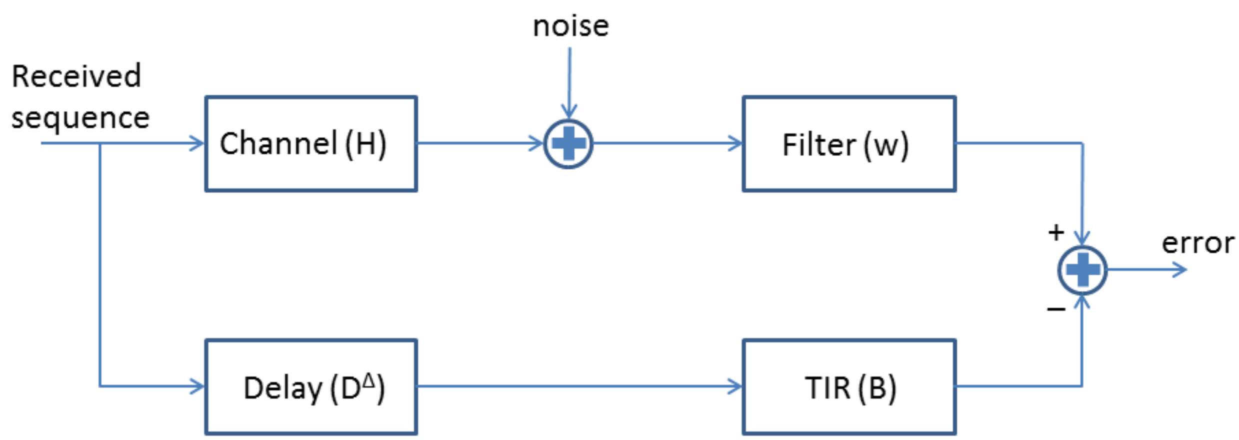


Fig. 3.5: MIMO-ISI channel shortening criterion

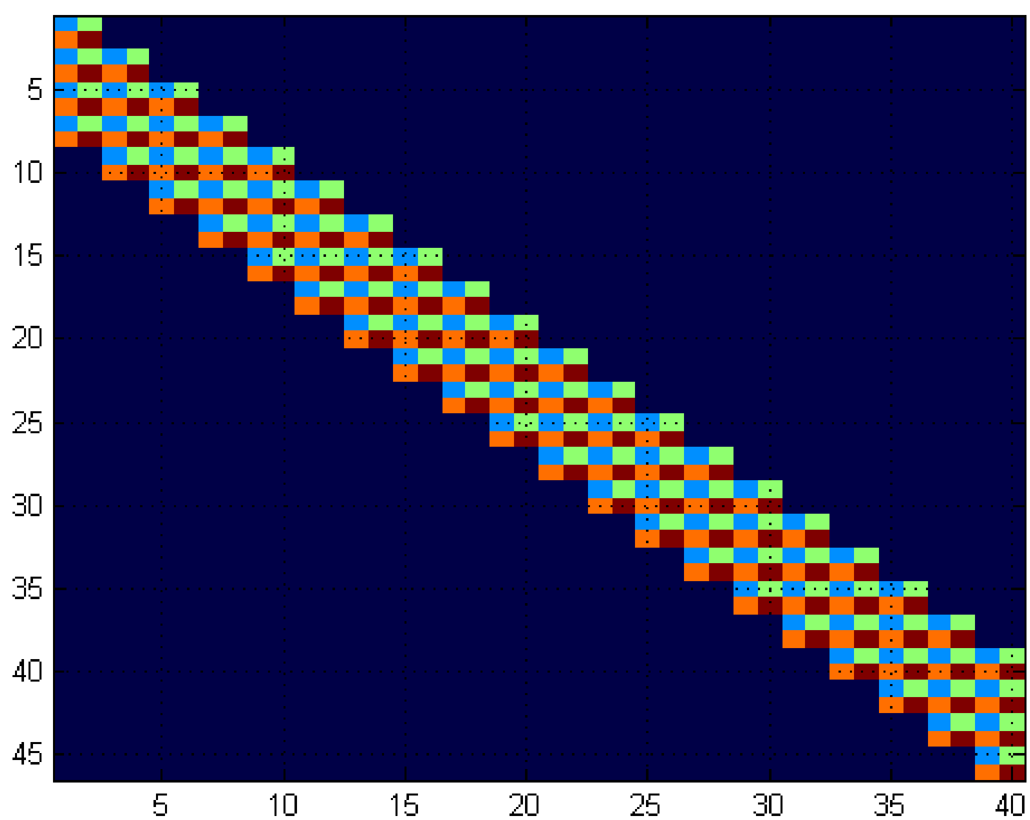


Fig. 3.6: MIMO-ISI block convolutional matrix

The optimum prefilter is represented as [14]

$$W^*{}_{opt} = B^*{}_{opt}\left(R_{xx}{}^{-1} + H^* R_{nn}{}^{-1} H\right)^{-1} H^* R_{nn}{}^{-1} . \tag{3.7}$$

Under the assumption that $R_{xx} = I$ and $R_{nn} = I\sigma_n{}^2$,then

$$W^*{}_{opt} = B^*{}_{opt}(I\sigma_n{}^2 + H^* H)^{-1} H^* \tag{3.8}$$

On the other hand B_{opt} is chosen under two criterions:

Criterion 1: Orthonormal constrain (ONC) is exactly the same as Magee-Falconer model except it is for MIMO-ISI channel where $B^* B = I_{ni}$.

Criterion 2: Identity tap constrain (ITC), where the matrix B is chosen to be an Identity matrix such that $B_{opt} = \min_B trace(R_{ee})$ subject to $B^* \varphi = I_{ni}$, where $\varphi^* = [\mathbf{0}_{ni\times nim} \quad I \quad \mathbf{0}_{ni\times ni(Nb-m)}]$ and $0 \leq m < Nb$, Nb is the memory length and R_{ee} is the autocorrelation matrix of error.

3.3- Rate maximizing channel shortening

All the previous channel shortening detectors have been optimized from MMSE point of view. However, MSE is a secondary metric that does not fulfill the maximum information rate according to Shannon. This section talks about optimal channel shortening for MIMO and ISI channels. It shows that the optimal shortening filter is somehow related to the MMSE filter.

For a non-shortened channel, the received symbols are first filtered with a match filter before being processed by the trellis as follows [15]:

$$\boldsymbol{z} \triangleq H^T \boldsymbol{y} = H^T H \boldsymbol{x} + H^T \boldsymbol{n} = G\boldsymbol{x} + \boldsymbol{\eta} , \tag{3.9}$$

where $\boldsymbol{y}$, H, $\boldsymbol{x}$, and $\boldsymbol{n}$ are as defined in equation (1.2). $G \triangleq H^T H$ and $\boldsymbol{\eta}$ is the colored noise. This model is known as the Ungerboeck model. Another model, proposed by Forney, whitens the noise by further filtering $\boldsymbol{z}$ with a linear filter.

By taking Forney model into account then the input output relationship of the channel is described by:

$$p(\boldsymbol{y}|\boldsymbol{x}) = \frac{1}{(\pi N_o)^N} \exp\left(-\frac{||\boldsymbol{y}-H\boldsymbol{x}||^2}{N_o}\right) \tag{3.10}$$

Where: $||\boldsymbol{y} - H\boldsymbol{x}||^2$ can be expanded as: $\boldsymbol{y}^*\boldsymbol{y} - 2Re\{\boldsymbol{y}^* H\boldsymbol{x}\} + \boldsymbol{x}^* G\boldsymbol{x}$, therefore:

$$p(\boldsymbol{y}|\boldsymbol{x}) = \frac{1}{(\pi N_o)^N} \exp\left(-\frac{\boldsymbol{y}^*\boldsymbol{y} - 2Re\{\,\boldsymbol{y}^* H\boldsymbol{x}\} + \boldsymbol{x}^* G\boldsymbol{x}}{N_o}\right), \tag{3.11}$$

Thus, the non-shortened system is completely described by the filter H^* and the coupling matrix G.

Channel shortening systems introduce a mismatched filter H_r and a coupling matrix G_r. Therefore, the input output relationship of the system will be defined as:

$$\tilde{p}(\boldsymbol{y}|\boldsymbol{x}) = \frac{1}{(\pi N_o)^N} \exp\left(-\frac{\boldsymbol{y}^*\boldsymbol{y} - 2Re\{\,\boldsymbol{y}^* H_r\boldsymbol{x}\} + \boldsymbol{x}^* G_r\boldsymbol{x}}{N_r}\right), \tag{3.12}$$

where N_r is the mismatched filtered noise, H_r and G_r will be optimized to get the maximum achievable rate.

Since $\boldsymbol{y}^*\boldsymbol{y}$ is not effective in the detection process, as well as the term $\frac{1}{(\pi)^N}$, they can be neglected. Moreover, with the absorption of N_r in G_r and H_r respectively (i.e. set $N_r = 1$) then

$$\tilde{p}(\boldsymbol{y}|\boldsymbol{x}) = \exp(2Re\{\,\boldsymbol{y}^* H_r\boldsymbol{x}\} - \boldsymbol{x}^* G_r\boldsymbol{x}). \tag{3.13}$$

As defined in equation (1.3), the achievable rate is:

$$I(X;Y) = H(Y) - H(Y|X) = H(Y) - H(n).$$

In this case: $H(Y) = E\{-\tilde{p}(\boldsymbol{y})\log(\check{p}(\boldsymbol{y}))\}$, and $H(Y|X) = E\{-\tilde{p}(\boldsymbol{y}|\boldsymbol{x})\log(\tilde{p}(\boldsymbol{y}|\boldsymbol{x}))\}$.

H_r and G_r have to be optimized to get the maximum achievable rate. An optimization constrain due to the shortened channel is given by: $(G_r\,)_{mn} = 0$ if $|m - n| > \nu$, where $(G_r\,)_{mn}$ denotes the element of the coupling matrix at row m and column n, and ν is the shortened length of the channel. After optimization (check [15] for optimization procedure), the optimum filter will be

$$H_r = [HH^T + N_o I]^{-1} H[G_r + I], \tag{3.14}$$

on the other hand, the maximum achievable rate will be:

$$I_{max} = \log(\det(G_r + I)) + tr\{(G_r + I)H^T[HH^T + N_o I]^{-1}H - tr\{G_r\}, \tag{3.15}$$

where $G_r = UU^T - I$ and U is defined in appendix (2) of [21]. It is important to observe that $G_r + I$ is a positive semi definite (PSD) matrix since $G_r + I = UU^T$. This means that minimum Eigen value of G_r should be > -1 and $\neq 0$. The optimum filter relation shows that the trellis depends on $G_r + I$ rather than G_r. Moreover, the optimized filter H_r is formed of the typical MMSE filter compensated by the trellis processor.

It is worth to mention that with full complexity detection (i.e. $\nu = L$), the achievable rate is going to be the normal full capacity defined by equation (1.5). On the other hand, when only the main diagonal of G_r matrix is conserved (i.e. $\nu = 0$), the performance of the MMSE detector proposed by Magee and Falconer is obtained.

The maximum achievable rate derived, consists of the conventional mutual information term ($\log(\det(I+\frac{HH^T}{N_o}))$) and a penalty term $tr\{(G_r+I)H^T[HH^T+N_oI]^{-1}H-tr\{G_r\}$. In this thesis, a proof that the penalty term diminishes to zero for the optimal case is set. Thus, the penalty term has no effect on the achievable rate of the shortened channel.

Starting from equation (3.15):

$$
\begin{aligned}
I_{max} &= \log(\det(G_r+I)) + tr\{(G_r+I)H^T[HH^T+N_oI]^{-1}H - tr\{G_r\} \\
&= \log(\det(U^\dagger U)) + tr\{U^\dagger U)H^T[HH^T+N_oI]^{-1}H - tr\{U^\dagger U - I\} \\
&= \log(\det(U^\dagger U)) + tr\{U^\dagger U)H^T[HH^T+N_oI]^{-1}H - tr\{U^\dagger U\} + tr\{I\} \\
&= \log(\det(U^\dagger U)) + tr\{U[H^\dagger[HH^\dagger+N_0I]^{-1}H - I]U^\dagger\} + tr\{I\} \\
&= 2\textstyle\sum_{n=1}^{N}\log(u_{nn}) - tr\{UBU^\dagger\} + N,
\end{aligned}
\tag{3.16}
$$

where B is defined as: $B \triangleq -H^T[HH^T+N_oI]^{-1}H+I$ and represents the MMS error.

Taking the term $tr\{UBU^\dagger\}$

$$tr\{UBU^\dagger\} = \sum[u_{nn}\ \boldsymbol{u}_n^\nu]\begin{bmatrix} B_{nn} & \boldsymbol{b}_n^\nu \\ \boldsymbol{b}_n^{\nu\dagger} & \tilde{B}_n^\nu \end{bmatrix}\begin{bmatrix} u_{nn} \\ \boldsymbol{u}_n^{\nu\dagger} \end{bmatrix}, \tag{3.17}$$

where $\boldsymbol{u}_n^\nu = [u_{n,n+1}\ u_{n,n+2}\ \dots\ u_{\mathrm{n,min}(L,n+\nu)}\ ]$, $\tilde{B}_n^\nu = \begin{bmatrix} B_{n+1,n+1} & \cdots & B_{n+1,\min(N,n+\nu)} \\ \vdots & \ddots & \vdots \\ B_{\min(N,n+\nu),n+1} & \cdots & B_{\min(N,n+\nu),\min(N,n+\nu)} \end{bmatrix}$,

and $\boldsymbol{b}_n^\nu = [b_{n,n+1}\ b_{n,n+2}\ \dots\ b_{\mathrm{n,min}(L,n+\nu)}\ ]$. According to [15]

$$\boldsymbol{u}_n^\nu = -u_{nn}\boldsymbol{b}_n^\nu\tilde{B}_n^{\nu^{-\mathbf{1}}}. \tag{3.18}$$

Inserting equation (3.18) into equation (3.17) gives:

$$tr\{UBU^\dagger\} = \sum\left[u_{nn}\quad -u_{nn}\boldsymbol{b}_n^\nu\tilde{B}_n^{\nu^{-\mathbf{1}}}\right]\begin{bmatrix} B_{nn} & \boldsymbol{b}_n^\nu \\ \boldsymbol{b}_n^{\nu\dagger} & \tilde{B}_n^\nu \end{bmatrix}\begin{bmatrix} u_{nn} \\ (-u_{nn}\boldsymbol{b}_n^\nu\tilde{B}_n^{\nu^{-\mathbf{1}}})^\dagger \end{bmatrix}$$

$$= \sum\left[u_{nn}B_{nn} - u_{nn}\boldsymbol{b}_n^\nu\tilde{B}_n^{\nu^{-\mathbf{1}}}\ \boldsymbol{b}_n^{\nu\dagger}\quad u_{nn}\boldsymbol{b}_n^\nu - u_{nn}\boldsymbol{b}_n^\nu\right]\begin{bmatrix} u_{nn} \\ (-u_{nn}\boldsymbol{b}_n^\nu\tilde{B}_n^{\nu^{-\mathbf{1}}})^\dagger \end{bmatrix}$$

$$= \sum\left[u_{nn}B_{nn} - u_{nn}\boldsymbol{b}_n^\nu\tilde{B}_n^{\nu^{-\mathbf{1}}}\ \boldsymbol{b}_n^{\nu\dagger}\quad 0\right]\begin{bmatrix} u_{nn} \\ (-u_{nn}\boldsymbol{b}_n^\nu\tilde{B}_n^{\nu^{-\mathbf{1}}})^\dagger \end{bmatrix}$$

$$= \sum u_{nn}{}^2B_{nn} - u_{nn}{}^2\boldsymbol{b}_n^\nu\tilde{B}_n^{\nu^{-\mathbf{1}}}\ \boldsymbol{b}_n^{\nu\dagger}$$

$$= \sum {u_{nn}}^2 [\, B_{nn} - \boldsymbol{b}_n^{\nu} \tilde{B}_n^{\nu^{-\mathbf{1}}} \;\; \boldsymbol{b}_n^{\nu^{\dagger}}]. \tag{3.19}$$

Equation (12) in [15] states that $c_{nn} = B_{nn} - \boldsymbol{b}_n^{\nu} \tilde{B}_n^{\nu^{-\mathbf{1}}} \;\; \boldsymbol{b}_n^{\nu^{\dagger}}$, thus inserting equation (12) in [15] into equation (3.19) we get

$$tr\{UBU^{\dagger}\} = \sum {u_{nn}}^2 c_{nn} \,, \tag{3.20}$$

but according to equation (12) in [15], $c_{nn} = \frac{1}{{u_{nn}}^2}$ which implies

$$tr\{UBU^{\dagger}\} = \sum_{i=1}^{N} 1 = N \,. \tag{3.21}$$

Inserting equation (3.21) into equation (3.15) gives

$$\begin{aligned} I_{max} &= \log(\det(U^{\dagger}U)) + N - N \\ &= \log(\det(U^{\dagger}U)) \\ &= \log(\det(Gr + I)) \,. \end{aligned} \tag{3.22}$$

This means that the maximum achievable rate is depending only on the Gram matrix and is very similar to the full memory capacity formula.

A special case can be derived for the ISI channel. Since the block length extends over 1000 symbols, which is considered infinitely large, the matrix representing the ISI channel will be large. This increases the computation complexity. Although, simplification is possible, since the ISI channel matrix is a toeplitz matrix formed from the repeated channel sequence **h**.

As the block length grows large, the toeplitz matrix can be represented as a circulant matrix. Thus, circular convolution is approximated to linear convolution. Therefore, the ISI-Matrix can be diagonlaized with the discrete Fourier transmit (DFT) matrix such that $Gr = Q\Lambda Q^{-1}$, where Q is the unitary DFT matrix and

$$I_{max} = \log(\det(Gr + I)) \;= \log(\det(I + \Lambda)). \tag{3.23}$$

Normalizing the capacity over the block size and using equation (3.23), the achievable rate is given by

$$I_{max} = \frac{1}{N} \sum_{k=1}^{N} 1 + \lambda_k \,, \tag{3.24}$$

where λ_k is the k[th] Eigen value of Gr.

Szegö's theorem states that as $N \to \infty$ the Eigen values of the toeplitz matrix converge to the Fourier transform of the Hermitian sequence which forms the matrix. Using Szegö's theorem, the achievable rate is expressed as

$$I_{max} = \frac{1}{N}\sum_{k=1}^{N} 1+\lambda_k = \frac{1}{\pi}\int_{-\pi}^{\pi} \log(1+Gr(w))dw\ , \tag{3.25}$$

where $Gr(w)$ is the discrete time Fourier transform of the sequence **g** given by

$$Gr(w) = \sum_n g(n)e^{-jwn}. \tag{3.26}$$

In the same manner, a closed form expression regarding the capacity of MIMO-ISI channel can be established. MIMO-ISI channel is represented as a block toeplitz matrix, which has some special mathematical properties.

By setting $Gr = Q\Lambda Q^{-1}$ in equation (3.21), then I_{LB} can be expressed as in equation (3.24).

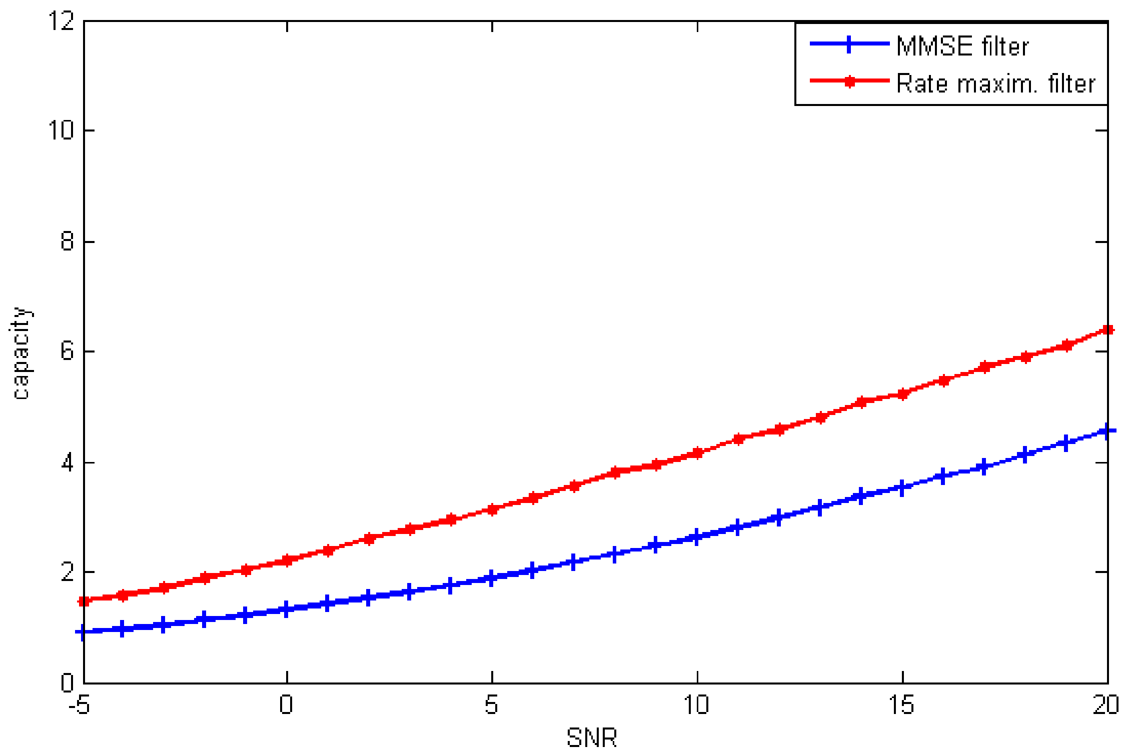


Fig. 3.7: Capacity performance of MMSE filter and rate maximization filter for a 2x2 MIMO-ISI channel

For MIMO-ISI channel, the coupling Matrix Gr is repeated as a block Toeplitz matrix and according to equation (5) in [16] which represents the modified Szegö's relation

$$\lim_{N\to\infty} \frac{1}{NK}\sum_{k\in\lambda\ of\ block\ toeplitz}^{NK} F(\lambda) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{K}\sum_{k\in\lambda\ of\ block\ matrix}^{K} F(\ Gr(w))dw, \tag{3.27}$$

where $Gr(w)$ is the Fourier transform of gr_k. While gr_k is the kth repeated sequence of the block forming the toeplitz matrix. Then

$$I_{LB} = \log(1+\lambda_k) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\sum_{k=1}^{K}\log(1+\ Gr(w)_k)dw. \tag{3.28}$$

In words, to get the achievable rate of MIMO-ISI channel, the Fourier transform is taken for each channel branch of the MIMO channel, and then the former relation is used.

Figure 3.7 shows a comparison of 2x2 MIMO-ISI capacity performances between optimal maximizing rate filter and MMSE filter proposed in [14]. It is obvious that the performance of the rate maximizing filter is significantly better than that of the MMSE filter.

N.B. for the MMSE filter, the closed form maximum achievable rate formula of the optimum maximum rate filter is used although; this capacity may not be achieved by the MMSE filter. Yet, it is used only for the sake of comparison of performance.

3.4- Rate maximizing channel shortening detector with soft feedback information:

As the importance of channel shortening on the performance and complexity of the receiver is studied, the performance of channel shortening combined with iterative detection is going to be checked. The channel shortening detector proposed earlier does not consider any priori information while the receiver discussed here is going to provide a channel shortening detector with soft side information.

We start by introducing some previous work considering the ISI case as discussed in [17].

The schematic diagram of the receiver is shown in figure 3.8.The main difference between this scheme and the normal scheme for full complexity iterative detector is the addition of two filters: one is the received sequence channel shortening filter and the other is the feedback side information filter. The input-output relation of the shortened channel discussed earlier will be modified to [17]

$$\tilde{p}(\boldsymbol{y}|\boldsymbol{x}) = \exp(2Re\{\,\boldsymbol{x}^*[Hr^*\boldsymbol{y} - F\boldsymbol{c}]\} - \boldsymbol{x}^* G_r \boldsymbol{x})\,, \tag{3.29}$$

where Hr is the mismatched shortening filter, F is the feedback filter matrix such that $F_{k,l} = 0$ where $|k - l| >$ the memory length(ν) , G_r is the coupling matrix , $\boldsymbol{x}$ is the transmitted sequence, $\boldsymbol{y}$ is the received sequence and $\boldsymbol{c}$ is the feedback sequence.

Two quantities are defined: $\beta \ni \beta I = E\{\boldsymbol{c}\boldsymbol{c}^*\}$ represents the energy of the estimated symbols and $\sigma \ni \sigma I = E\{\boldsymbol{c}\boldsymbol{x}^*\}$ represents the correlation between the estimated and transmitted symbols (i.e. detection accuracy). For perfect feedback estimates i.e. $\boldsymbol{c} = \boldsymbol{x}$, we get $\beta = \sigma = 1$ otherwise $0 < \beta, \sigma < 1$.

N.B. for simplicity, in simulations it is assumed that β and σ are always equal.

Two special cases can be mentioned, for $\beta = 0$, then there is no feedback information and therefore we end up with a normal channel shortening problem. On the other hand, when $\sigma = 1$ then it is optimal that F perfectly suppresses all the interference. So that

$$F = H^T H - diag(H^T H)\,, \tag{3.30}$$

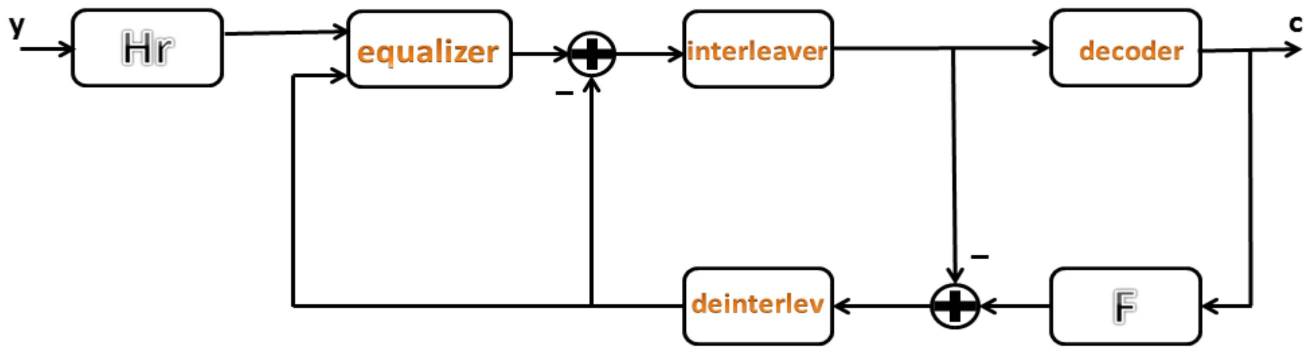


Fig. 3.8: channel shortening iterative receiver: two filter (Hr and F) are added to the normal iterative scheme

which means that the front end filter will be a match filter H^T and the constricted memory is $\nu = 0$. So the coupling matrix is

$$G_r = \ diag(H^T H). \tag{3.31}$$

The achievable information rate is determined as $\max_{G_r,F,H_r} I_{LB}$. In contrast to the static channel shortening, optimization over the feedback matrix F is carried out.

Starting by optimizing I_{LB} over the mismatched filter H_r gives according to appendix (I) in [15]

$$Hr_{opt} = [HH^T + N_o I]^{-1} H[G_r + I + \sigma F^T], \tag{3.32}$$

and

$$I_{LB} = \log(\det(G_r + I) + Re\{tr\{(A - I)(G_r + I + 2\sigma F)\}\}$$

$$+tr\{(A\sigma^2 - \beta I)F^T (G_r + I)^{-1} F\} + N, \tag{3.33}$$

where $A = H^T[HH^T + N_o I]^{-1}H$. The optimum mismatching frontend filter $Hr_{opt}$is given by a Wiener filter with a feedback cancellation part and trellis processing part. The optimization over F and G_r is the next step.

Optimization over F is held by vectorizing the capacity relation which gives a quadratic expression, which means that a closed form optimization of F can be reached. Yet, the optimization of G_r is complex. Thus, two cases are discriminated: ISI case and MIMO case.

For ISI channel the situation simplifies since all the matrices are toeplitz matrices and can be approximated into circulant matrices which means that they can be diagonalized by unitary DFT matrix; using Szegö's theorem

$$I_{LB} = \ \frac{1}{\pi}\int_{-\pi}^{\pi} \log\big(1 + Gr(w)\big) - D_1(w)\big(1 + Gr(w)\big) - \frac{D_2(w)|F(w)|^2}{1+Gr(w)} + 2\sigma Re\{D_1(w)F(w)\}d(w) \tag{3.34}$$

where: $D_1(w)$ and $D_2(w)$ are the Fourier transforms of the induced sequences for the matrices $(I-A)$ and $(\beta I-\sigma^2 A)$ respectively. Moreover, $|F(w)|^2 = F(w)^2$.

Since $F(w)^2$ is real according to

$$F(w) \triangleq F(w)_R + jF(w)_{Im}\ , \tag{3.35}$$

this implies:

$$\frac{1}{\pi}\int_{-\pi}^{\pi} 2\sigma D_1(w)\, F(w)_R - \frac{D_2(w)}{1+Gr(w)}\left[F(w)_R{}^2 + F(w)_{Im}{}^2\right]. \tag{3.36}$$

Since $D_2(w) \geq 0$ and $1+Gr(w) \geq 0$, it follows that $F(w)_{Im} = 0$ so that equation (3.36) is maximized. Thus, $|F(w)|^2 = F(w)^2$ and $F(w)$ is real. Then the achievable rate can be written as

$$I_{LB} = \frac{1}{\pi}\int_{-\pi}^{\pi} \log\big(1+Gr(w)\big) - D_1(w)\big(1+Gr(w)\big) - \frac{D_2(w)F(w)^2}{1+Gr(w)} + 2\sigma D_1(w)F(w)d(w)\ . \tag{3.37}$$

Now, optimization over $F(w)$ and $Gr(w)$ can be checked by taking the Hessian matrix of I_{LB} with respect to the two quantities. The Hessian matrix

$$\text{Ḩ}(w) = -\frac{2D_2(w)}{\big(1+Gr(w)\big)^2}\begin{bmatrix} \frac{1}{2D_2(w)} + \frac{F(w)^2}{1+Gr(w)} & -F(w) \\ -F(w) & 1+Gr(w) \end{bmatrix}, \tag{3.38}$$

is always negative definite (ND) which implies that the optimization over $F(w)$ and $Gr(w)$ is concave optimization.

Figure 3.9 shows the variation of the information rate achieved w.r.t β. Channel **h**= [0.5 0.5 -0.5 -0.5] and BPSK system are used. The output of the decoder is a soft (log likelihood ratio) LLR's. To get the soft bit sequence, $p(a_k=1)$ and $p(a_k=-1)$ are calculated such that $p(a_k=1) = \frac{\exp(LLR)}{1+\exp(LLR)}$ and $p(a_k=-1) = 1-\frac{\exp(LLR)}{1+\exp(LLR)}$, then c_k is calculated as

$$c_k = (+1)p(a_k=1) \quad p(a_k=-1)\text{, and } \beta = E\{\boldsymbol{cc}^*\}.$$

The figure shows that as the memory length increase, the change of capacity due to the quality of side information decrease. Figure 3.10 shows the BER performance of the system. It is worth to mention that LDPC code from digital video broadcasting (DVB) (32400, 64800) is used.

The BER achieved with L=1 and 2 are plotted . The heavy black line is the BER performance of the full memory MAP detector. The two red lines are the BER in case of L=2 with soft feedback (circles) and without soft feedback. The two blue lines are the BER in case of L=1 with soft feedback (stars) and without soft feedback. It is clear that there is more than 1 dB gain in case of L=0 when feedback exists.

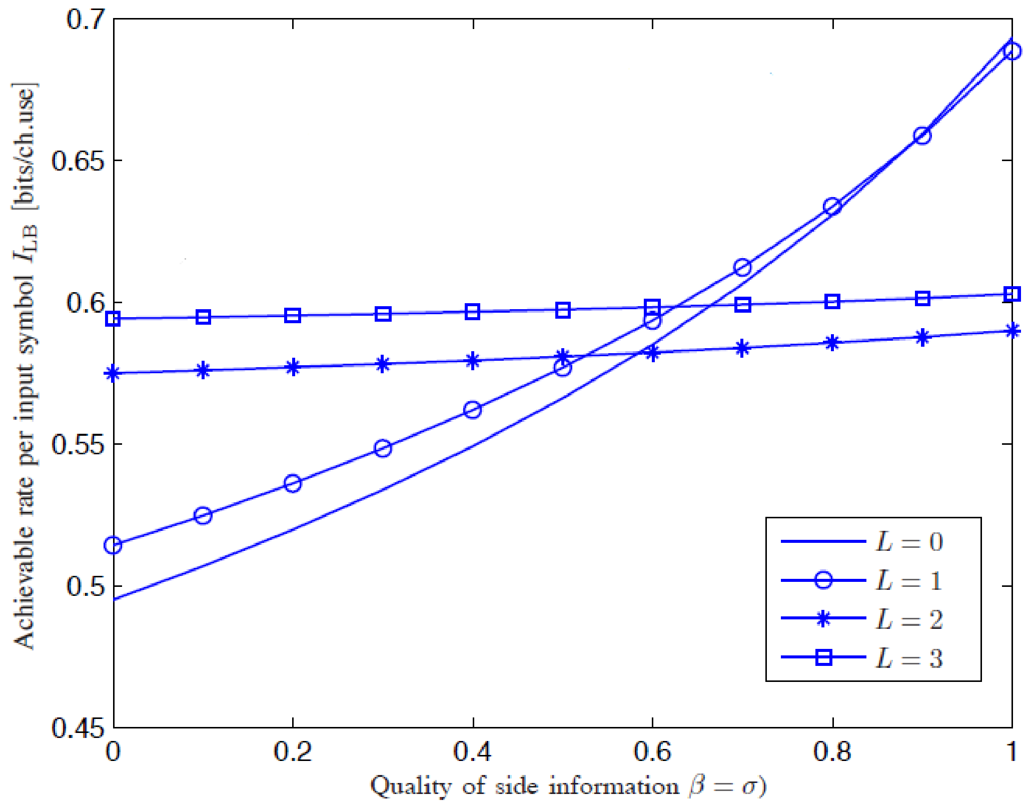


Fig. 3.9: variation of the information rate with respect to side information quality

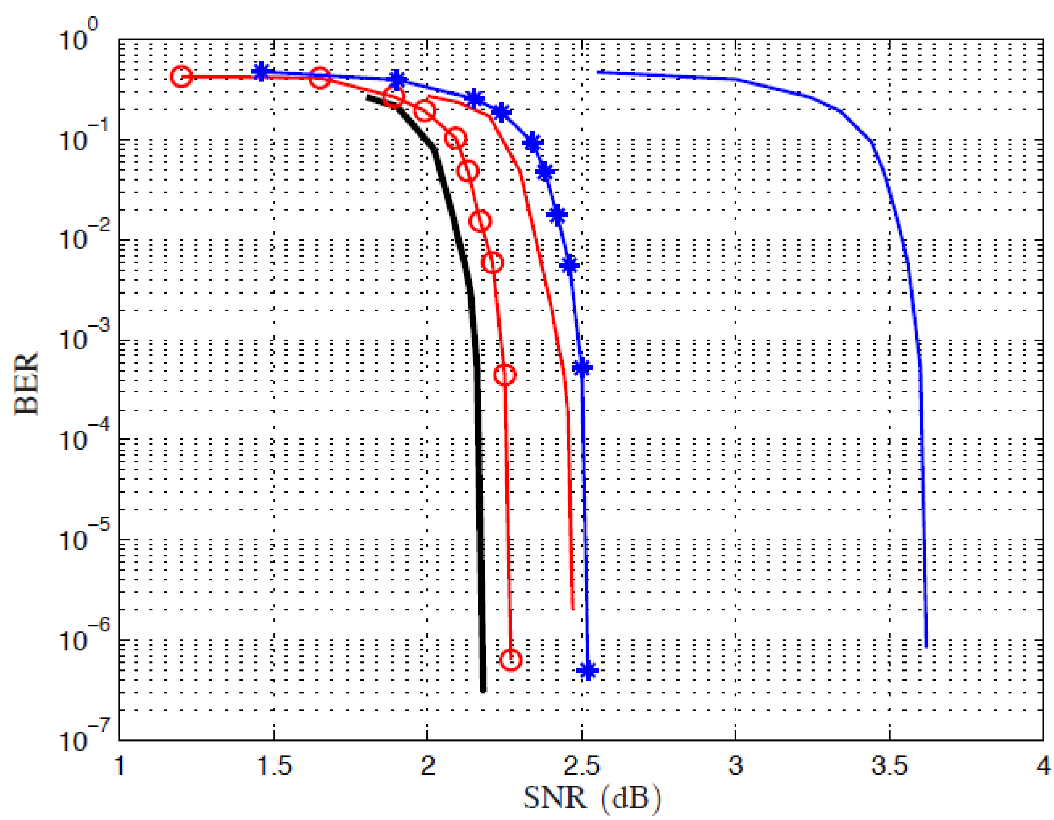


Fig. 3.10: BER performance of ISI channel shortening detector with soft feedback side information

For MIMO channel, as discussed in this thesis, the channel matrix does not possess any special characteristics (like the toeplitz matrix in ISI case). Therefore, optimization will be tricky. For optimization information rate relation is approximated.

It is proposed that the coupling matrix (*Gr*) at any value of σ is a linear combination between the coupling matrices Gr_0 and Gr_1 . At Gr_0: $\sigma = 0$ (i.e. the case of static channel shortening), where the capacity is minimum; at Gr_1: $\sigma = 1$ (i.e. in case of perfect interference cancellation), where the capacity is maximum. Therefore, Gr is going to be

$$Gr = (x)Gr_0 + (1-x)\, Gr_1 \,, \tag{3.39}$$

where x is a factor between 0 and 1. On the other hand, due to strict relation between the interference cancellation matrix F and the coupling matrix Gr (at $\sigma = 1$: $F_1 = H^T H - Gr_1$), we can represent F at any σ as linear combination with x factor. When $\sigma = 0$, $F = 0$ since there is no feedback information. F_1 represents F when $\sigma = 1$. F is expressed at any σ as follows:

$$F = (1-x)\, F_1 \,. \tag{3.40}$$

The achievable rate is expressed as a function of x (which is a scalar value) instead of Gr and F. This is given as

$$\begin{gathered} I_{LB} = \log(\det\big((x)Gr_0 + (1-x)Gr_1 + I\big) + \\ Re\left\{tr\{(A-I)\big((x)Gr_0 + (1-x)Gr_1 + I + 2\sigma(1-x)\, F_1\big)\}\right\} + \\ tr\left\{(A\sigma^2 - \beta I)(1-x)\, {F_1}^T\big((x)Gr_0 + (1-x)Gr_1 + I\big)^{-1}(1-x)\, F_1\right\}, \end{gathered} \tag{3.41}$$

where: Gr_0, $Gr_1$and F_1 are constant valued matrices and therefore the optimization is carried over x.

The optimization is held over three terms:

<u>First term is</u>

$$\log(\det\big((x)Gr_0 + (1-x)Gr_1 + I\big) = \log(\det\big((Gr_0 - Gr_1)x + Gr_1 + I\big), \tag{3.42}$$

let $Gr_0 - Gr_1 \triangleq V$, and $Gr_1 + I \triangleq Z \succ 0$, then

$$\log(\det\big((Gr_0 - Gr_1)x + Gr_1 + I\big) = \log(\det(Vx + Z)), \tag{3.43}$$

where V is any symmetric matrix. Using the fact that a function is convex /concave when it is constricted to a line (i.e. if g(t)=f(x+vt) is convex/concave implies that f(x) is convex/concave) [18]. Applying this theorem to equation (3.42):

$f(Z) = \log(\det(Z + Vx))$, this implies that

$$g(x) = \log(det(Z)) + \log(\det(I + xZ^{-\frac{1}{2}}VZ^{-\frac{1}{2}} = \log(\det(Z)) + \sum(1 + x\lambda_i), \quad (3.44)$$

where λ is the Eigen value of matrix $Z^{-\frac{1}{2}}VZ^{-\frac{1}{2}}$. Since $\log(det(Z))$ is constant and $\sum(1 + x\lambda_i)$ is a concave function, then g(x) is concave. Therefore, equation (3.43) is concave which means that equation (3.42) is concave.

<u>Second term is</u>

$$tr\{(A - I)\big((x)Gr_0 + (1 - x)Gr_1 + I + 2\sigma(1 - x)\, F_1\big)\}$$

$$= -tr\{(I - A)\big((x)Gr_0 + (1 - x)Gr_1 + I + 2\sigma(1 - x)\, F_1\big)\}$$

$$= -tr\{(I - A)\big((Gr_0 - Gr_1 - 2\sigma\, F_1)x + Gr_1 + I + 2\sigma\, F_1\big)\}\,. \quad (3.45)$$

Defining $\Lambda \triangleq\ I - A$ as a PSD matrix (check the proof at the end of the chapter), $\Delta \triangleq Gr_1 + I + 2\sigma\, F_1$ to be a PSD matrix (since $Gr_1, I\ and\ F_1$ are all PSD matrices) and $\Gamma \triangleq Gr_0 - Gr_1 - 2\sigma\, F_1$ to be a symmetric matrix, then the second part can be written as

$$-tr\{\Lambda[\Gamma x + \Delta]\}\,, \quad (3.46)$$

and since $\Gamma x + \Delta$ is linear function then equation (3.46) is linear function too which means that equation (3.45) is linear function.

<u>Third term is</u>

$$-tr\left\{(\beta I - A\sigma^2)(1 - x)\, {F_1}^T\big((x)Gr_0 + (1 - x)Gr_1 + I\big)^{-1}(1 - x)\, F_1\right\}$$

$$= -tr\left\{(\beta I - A\sigma^2)(1 - x)^2\left[\, {F_1}^T\big((x)Gr_0 + (1 - x)Gr_1 + I\big)^{-1}\, F_1\right]\right\}. \quad (3.47)$$

Defining $\beta I - A\sigma^2 \triangleq \Lambda$ as a PSD matrix (check the end of the chapter for proof).

Therefore the third part can be represented as:

$$-tr\left\{\left(\, F_1\Lambda\, {F_1}^T\right)^{-1}(1 - x)^2\left[\big((Gr_0 - Gr_1)x + Gr_1 + I\big)^{-1}\right]\right\}. \quad (3.48)$$

Defining $\Gamma \triangleq \left(\, F_1\Lambda\, {F_1}^T\right)^{-1}(Gr_0 - Gr_1)$ a PSD matrix (since F_1, Λ are PSD, and since x is less than one then $Gr_0 - Gr_1$ is PD) and $\Delta \triangleq \left(\, F_1\Lambda\, {F_1}^T\right)^{-1}(Gr_1 + I)$ is PSD matrix. Then the third part can be written as:

$$-tr\{(1 - x)^2[\Gamma x + \Delta]^{-1}\}\,, \quad (3.49)$$

differentiating twice to check the convexity of the function

$$-2tr\{[\Gamma x + \Delta]^{-1}\} + (2x - 2)tr\{-[\Gamma x + \Delta]^{-1}\Gamma[\Gamma x + \Delta]^{-1}\}$$

$$= -2tr\{[\Gamma x + \Delta]^{-1}[I + (1-x)[\Gamma x + \Delta]^{-1}][I + (1-x)[\Gamma x + \Delta]^{-1}\Gamma]^T\}\,, \tag{3.50}$$

but this term is ND. Therefore, the third part is concave.

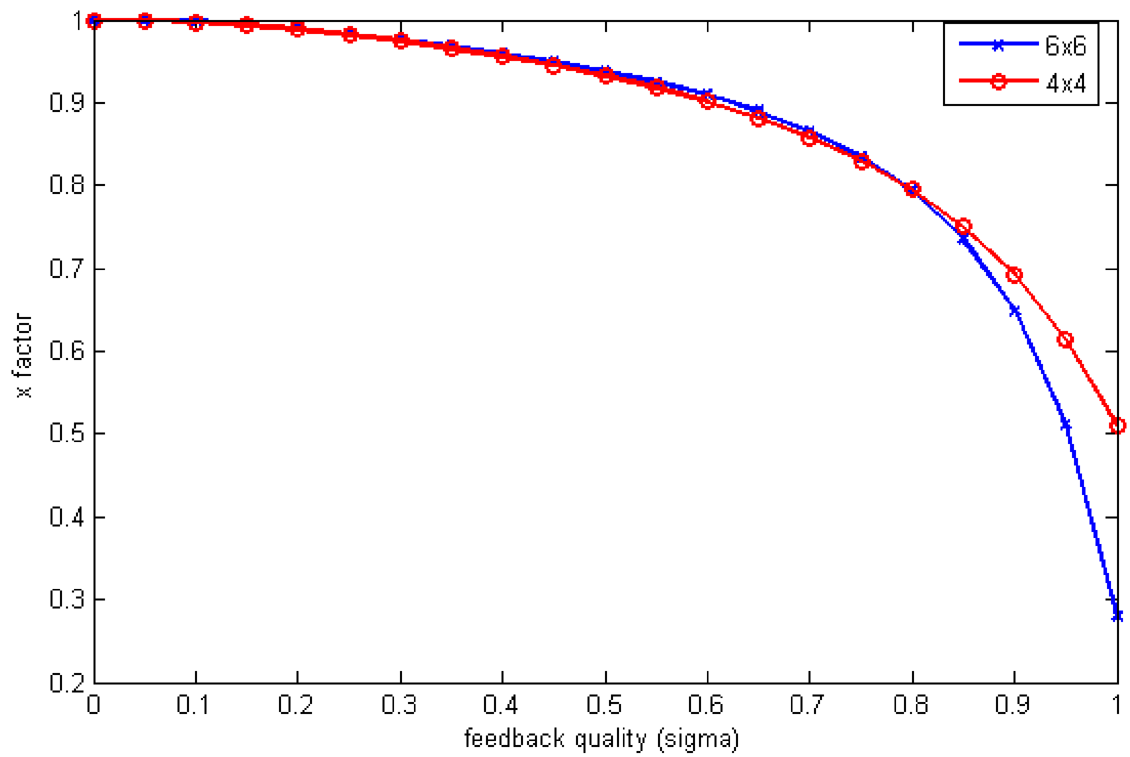


Fig. 3.11: variation of the linear factor x with respect to σ for 4x4 and 6x6 MIMO channels.

Since the first part is concave and the third part is concave then their sum is also concave. This is due to the fact that concave + concave is concave and according to the fact that concave + affine is concave, the maximum achievable rate is concave over the factor x.

Figure 3.11 shows the variation of optimal x with respect to σ for 6x6 as well as 4x4 MIMO systems. It is noteworthy to mention that the variation of x with respect to σ is constant as the memory length L and the signal to noise ratio vary.

In figure 3.12 and 3.13, the variation of the achievable rate with σ for 4x4 MIMO and 6x6 MIMO systems is illustrated. Again, as in ISI case, it can be seen from the two plots that the curves of capacity with different memory lengths intercept each other. Moreover, as the memory length increases the impact provided by the feedback information decreases, since the cancellation of the interference decreases.

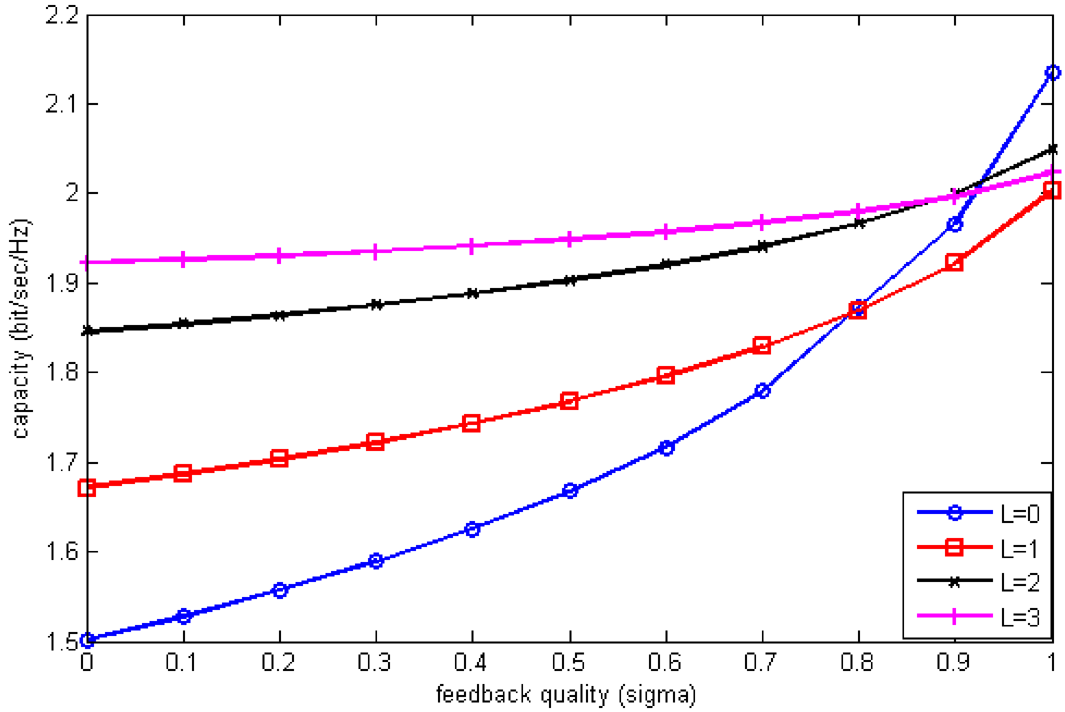


Fig. 3.12: variation of capacity w.r.t feedback quality for a 6x6 MIMO system

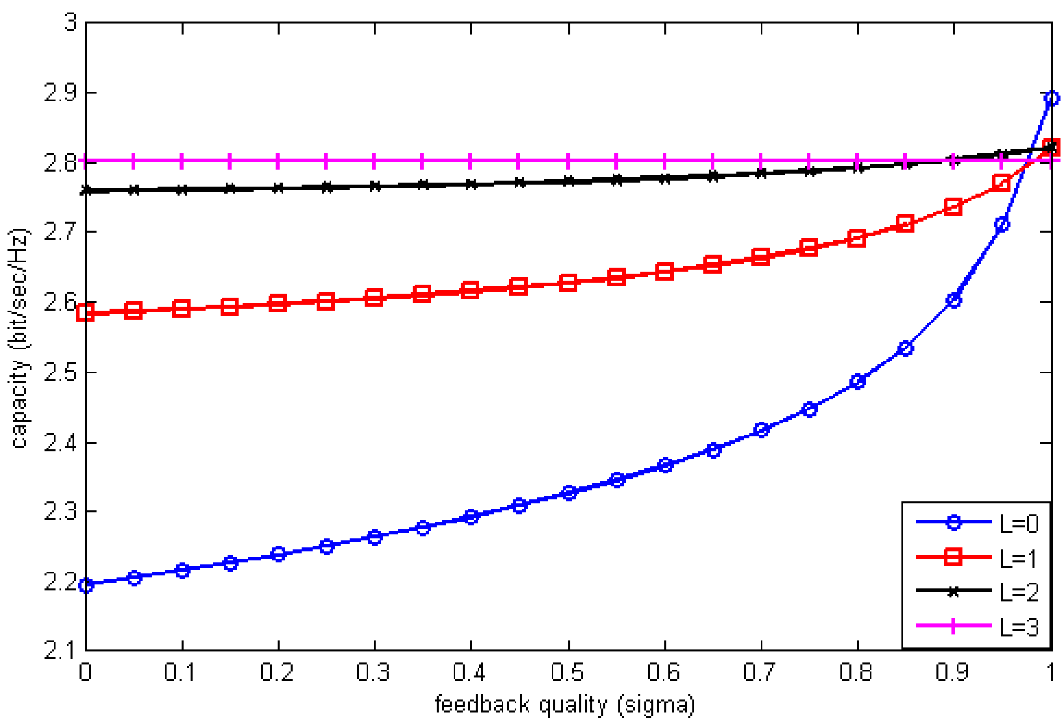


Fig. 3.13: variation of capacity w.r.t feedback quality for a 4x4 MIMO system

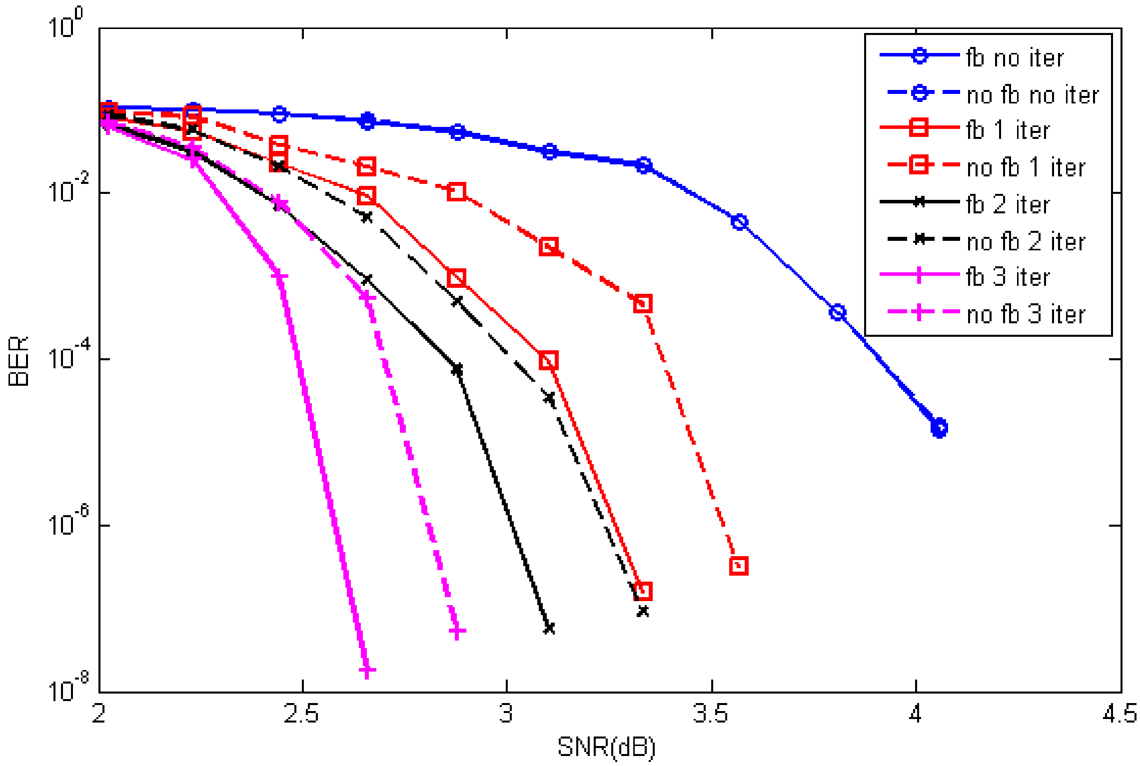


Fig. 3.14: BER performance of a 6x6 MIMO detector with and without feedback

The BER performance of the shortening detector with soft feedback information for 6x6 and 4x4 MIMO systems is shown in Figures 3.14 and 3.15 respectively. The setup of the detector is the same as that for the ISI case. The BER performance is studied for both feedback case and static channel shortening case where no feedback exists (dashed lines). The figures illustrate 1, 2 and 3 iterations cases.

It is important to notice that with soft feedback information, a complete iteration can be saved (like in the case with 6x6 MIMO: second iteration with soft feedback is very similar to the third iteration with no feedback while 3^{rd} iteration with feedback exceeds the second iteration with no feedback). Moreover, for the case of no iteration, no feedback plot and feedback plot are the same.

General suboptimal optimization for information rate relation:

As we discriminated between ISI and MIMO case for the optimization of the information rate expression we can return to the fundamental expression and optimize for a general case. The optimization approach is suboptimal since no joint optimization of F and Gr is carried.

For the information rate relation described by equation (3.33), we need to optimize over F and Gr: The optimization over F is quadratic and closed form can be reached.

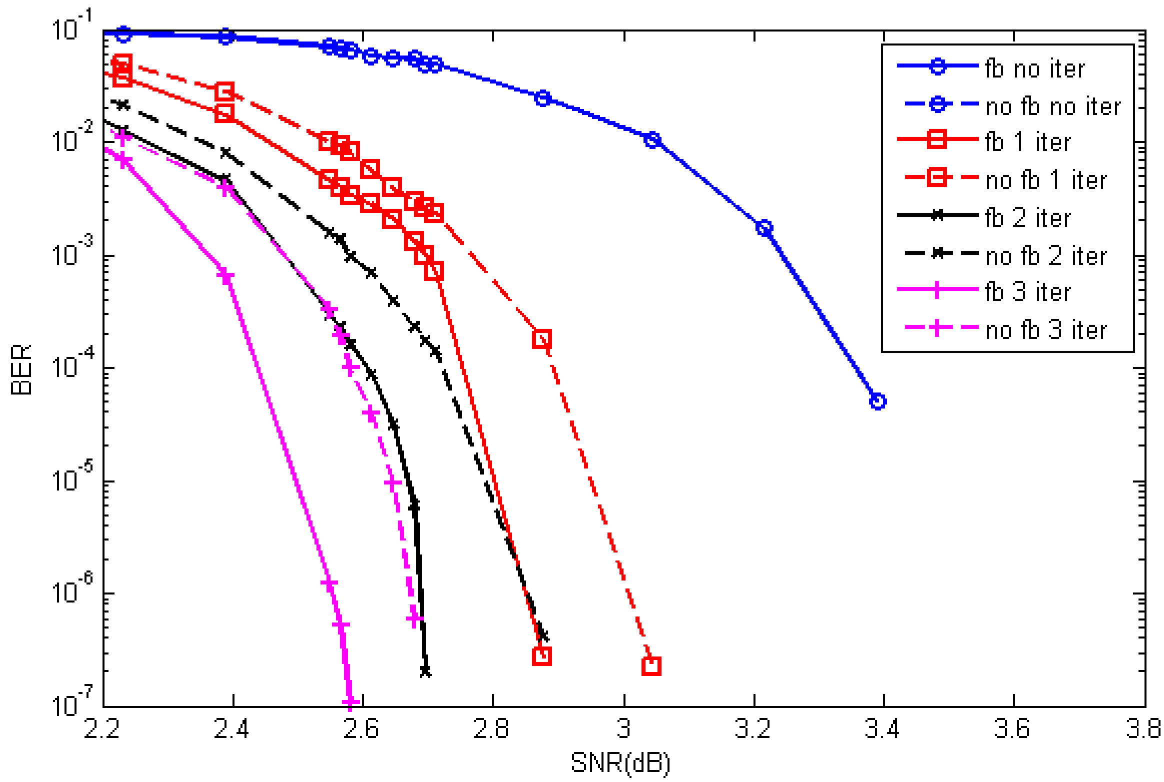


Fig. 3.15: BER performance of a 4x4 MIMO detector with and without feedback

While for the optimization over Gr, N is a constant and is not going to affect the optimization, therefore it is neglected. Again we partition the optimization over three terms:

<u>First term:</u> following the same procedure as in equation (3.43) and equation (3.44) we prove that

$\log(\det(G_r + I)$ is concave function.

<u>Second term:</u>

$$tr\{(A - I)(G_r + I + 2\sigma F)\}\,. \tag{3.51}$$

Define $\Gamma \triangleq -(A - I)$ is a PSD (proof is made below), and $\Delta \triangleq 2\sigma F + I$ is a PSD since F is PSD

then (3.51) becomes

$$-tr\{\,\Gamma Gr + \Gamma\Delta\}\,, \tag{3.52}$$

which is linear because $\Gamma Gr + \Gamma\Delta$ is linear and trace of linear is linear.

<u>Third term:</u>

$$tr\{(A\sigma^2 - \beta I)F^T(G_r + I)^{-1}F\} = -tr\{(\beta I - A\sigma^2)F^T(G_r + I)^{-1}F\}\,. \tag{3.53}$$

Define $\beta I - A\sigma^2 \triangleq \Gamma$ is a PSD matrix and $(G_r + I)^{-1} \triangleq Z$ where Z is linear with G_r, then equation (3.53) becomes

$$-tr\{\Gamma F^T Z^{-1} F\}\,. \tag{3.54}$$

Γ, F and Z^{-1} are PSD matrices, and let $\Delta \triangleq \Gamma^{-1}$ and $\Lambda \triangleq F^{-1}$ is PSD matrices then equation (3.54) becomes

$$-tr\left\{\Delta^{-1}\left(\Lambda^{-1}\right)^T Z^{-1}\,\Lambda^{-1}\right\} = -tr\{\left[\Lambda Z\,\Lambda^T \Delta\right]^{-1}\}\,. \tag{3.55}$$

Define $X \triangleq \Lambda Z\,\Lambda^T$ which is linear w.r.t Z. Therefore equation (3.55) becomes $-tr\{\,X^{-1}\}$ but $tr\{\,X^{-1}\}$ is convex, thus equation (3.55) is concave which proves that equation (3.53) is concave. This proves that the achievable rate relation defined in equation (3.33) is concave since sum of concave is concave and sum of affine to concave is concave.

Proof of some relations:

1- Proof of $tr\{\,X^{-1}\}$ is convex: is similar to the proof found in equations (3.42) and (3.43).

2- proof that $(A - I)$ is ND: Defined $A = H^T[HH^T + N_o I]^{-1}H$, then

$$A = \left[H^{-1}[HH^T + N_o I]H^{T^{-1}}\right]^{-1}, \tag{3.56}$$

Taking the SVD of H we can write A as

$$A = [V^{T^{-1}}\Sigma^{-1}U^{-1}[U\Sigma V^T H^T + N_o I]U^{T^{-1}}\Sigma^{T^{-1}}V^{-1}]^{-1}$$

$$= [[V^{T^{-1}}\Sigma^{-1}U^{-1}U\Sigma V^T H^T + N_o V^{T^{-1}}\Sigma^{-1}U^{-1}]U^{T^{-1}}\Sigma^{T^{-1}}V^{-1}]^{-1}$$

$$= \left[\left[V\Sigma^T U^T + N_o V^{T^{-1}}\Sigma^{-1}U^{-1}\right]U^{T^{-1}}\Sigma^{T^{-1}}V^{-1}\right]^{-1}$$

$$= \left[I + N_o V^{T^{-1}}\Sigma^{-1}U^{-1}U^{T^{-1}}\Sigma^{T^{-1}}V^{-1}\right]^{-1}$$

$$= \left[I + N_o V^{T^{-1}}\Sigma^{-1}\Sigma^{T^{-1}}V^{-1}\right]^{-1}$$

$$= \left[V^{T^{-1}}V^{-1} + N_o V^{T^{-1}}\Sigma^{-1}\Sigma^{T^{-1}}V^{-1}\right]^{-1}$$

$$= V\left[I + N_o\Sigma^{-1}\Sigma^{T^{-1}}\right]^{-1}V^T$$

$$= V \sum \frac{1}{1+\frac{N_O}{\lambda_i}} V^T \text{ ,} \tag{3.57}$$

where λ_i are the Eigen values of H.

Taking the Eigen decomposition of $A = Q\Gamma Q^{-1}$, we show that $\gamma_i = \frac{1}{1+\frac{N_O}{\lambda_i}}$ is always less than 1.This proves that $A - I$ is a ND matrix.

Similarly and since σ is always less than or equal 1 and as we assume $\sigma = \beta$ we can say that:

$\sigma A - I$ is a ND matrix as well as $\sigma^2 A - \beta I$ is also a ND matrix.

Chap 4: conclusion and future work:

This thesis discusses the interference (ISI as well is spatial) effect as a crucial problem on the performance of communication system. We discuss the techniques to mitigate this phenomenon using equalization. We find that the equalization techniques can vary between low complexities receivers with low efficiency (like linear equalizers) and high complexity receivers with high performance (like sequence estimators). As well as, we discuss the iterative receivers and the way they approach joint equalization decoding processing.

In chapter 3, we discuss the techniques that can decrease the complexity of sequence equalization and discriminate between trellis pruning techniques and channel shortening techniques. We find that most of the previous techniques used MMSE shortening methodology which is a secondary metric and is not optimal from the information rate point of view.

We examine the recently established rate maximizing channel shortening techniques and check their superiority over MMSE techniques.

We also study the rate maximizing channel shortening detector with side information for the ISI channel and introduce a similar detector for the MIMO channel where we prove that the performance of the detector with feedback information is superior to detectors with no feedback information.

A closed form expression for MIMO-ISI shortened channel maximum achievable rate is also established.

Future works include development of MIMO-ISI channel detectors with soft side information which would be a special (but simplified) case of the general achievable rate expression.

Our study focuses on linear transmission systems that are found in mobile communications and WLANS. On the other hand, satellite communications are non-linear systems where eligibility of those techniques would be a matter of study too.

Mitigating of interference in OFDM systems is another promising field. Although ISI is mitigated in OFDM systems using the cyclic prefix (CP) technique, CP can be a big wastage of resources where sometimes we need to tolerate less CP on the expense of ISI and thus we need to equalize. Moreover spatial interference still a matter of study in OFDM systems and techniques to mitigate this interference are still developing.

Complexity challenge in trellis sequence equalization can hold further researches such as finding ways of combining trellis shortening techniques with trellis pruning techniques such as sphere detection techniques. Such joint simplification for the trellis would provide a significant modification in the complexity of the receivers.

References:

1- A. Goldsmith," Wireless communications" PP 1.
2- C. Shannon," Mathematical theory of communication".
3- R. Togneri and C. Desilva," Fundamentals of information theory and coding design" PP 105.
4- T. Wieg and H. Schwarz," Source and Video coding" PP 6.
5- S. Lin an D. Castello," Error control coding" second edition.
6- C. Heegard and S. Wicker," Turbo coding", Boston, MA: Kluwer 1999.
7- A. Paulraj, R. Nabar and D. Gore," Introduction to space time wireless communications". PP 63-76
8- A. F. Molisch," wireless communications".
9- R. Koetter, A. Singer and M. Tuchler," Turbo equalization" IEEE signal processing magazine January 2004.
10- A. Paulraj, R. Nabar and D. Gore," Introduction to space time wireless communications". PP 149-150
11- K. Kar and Y. wong," The soft output M-algorithm and its applications" PHD thesis Queen's university-Ontario-Canada. PP 21-22.
12- D. D. Falconer and F.R Magee," Adaptive channel memory truncation for maximum likelihood sequence estimation". The bell system technical J. vol. 52 no. 9 PP 1541-1562 November 1973.
13- O. Edfors, P. Ödling, H. Erikson, T. Koski and P. Borjesson," Combined linear viterbi equalization- a comparative study and minimax design". LuLeå university of technology-Sweden.
14- N. Al-Dhaher,"FIR channel shortening equalizers for MIMO-ISI channels" IEEE transactions on communication vol. 49 no. 2, February 2001.
15- F. Rusek and A. Pralja," Optimal channel shortening for MIMO-ISI channels". IEEE transactions on wireless communications vol. 11 no. 2 February 2012.
16- M. Miranda and P. Tilli," Asymptotic spectra of Hermitian block toeplitz matrices an preconditioning result". SIAM J MATRIX ANAL. APPL. Vol. 21 no. 3 PP 867-881.
17- F. Rusek, N. Al-Dhaher and A. Gomaa," A rate maximizing channel shortening detector with soft feedback side information". Globecom. 2012 communication theory symposium.
18- S. Boyd," Convex optimization" PP 74.