Master's Thesis

Investigation of Interference Rejection Suppression in MIMO Systems

By

Abigaba JohnPaul and Al Ali Hayder

Department of Electrical and Information Technology
Faculty of Engineering, LTH, Lund University
SE-221 00 Lund, Sweden

Abstract.

In wireless communication networks, the cell edge users' throughput is limited by inter-cell interference. The Interference Rejection Combining (IRC) receiver can effectively suppress the inter-cell interference. The main working principle of the IRC receiver is the Minimum Mean Square Error (MMSE), and requires both channel estimation and the received signal covariance matrix estimation, whose quality determines the performance of the receiver.

Channel estimation is achieved using pilot boosting, where by more power is allocated to the pilot symbols than the data symbols. The signal covariance matrix can be estimated using the reference symbols, or the data symbols, but in this piece of work we aim to enhance the accuracy of the covariance matrix by optimizing both estimation schemes (data and reference symbols), secondly using the Rao Blackwell Ledoit Wolf (RBLW) method, and lastly by applying the RBLW method to the optimized covariance matrix.

The results indicate that the RBLW method (for coherence time, T = 10) dominates the rest of the covariance matrix estimation schemes for the case of the channel estimate, while the reference symbols estimation method outperforms the other schemes for the true (ideal) channel, due to more accurate channel and covariance matrices. We also investigate the optimal number of pilots versus channel conditions using different signal constellations to achieve maximum capacity.

Acknowledgments

First and foremost we would like to extend our sincere gratitude to Professor Fredrik Rusek for his unlimited guidance and support. Thank you for sharing your knowledge with us for the last two years.

Individually as Paul, I would like to thank the Swedish Institute for all the financial support rendered to me throughout my study period, this would still be a dream for me if it wasn't for the Swedish Institute.

Special thanks also to the wireless class of 2011/2013, and the entire EIT staff, thank you for making our stay in Lund worthwhile.

And finally we would like to thank our families for bearing with us all this time we have been away from them, and for their undying love, encouragement and moral support.

John Paul Abigaba & Al Ali Hayder

Table of Contents

CHAPTER 1

1 Introduction

The roots of interference rejection is in military applications, some of which aimed at protecting the desired signals from being hindered (signal jamming), while some needed to achieve signal jamming. However the rapid growth in cellular communication systems, coupled with increased user demand for more throughput have pushed for interference rejection suppression techniques to be used even for non-military applications.

Several interference rejection methods have evolved over time, with a lot of receivers combining interference rejections schemes alongside diversity to achieve robustness against interference. Diversity is the transmission of the same information via more than a single channel, so that in the possible event that data from one channel is corrupted, or irretrievable for some reason, a copy of the data can be retrieved from another channel. Interference rejection on the other hand involves estimating the residual covariance matrix of the interference, in the frequency domain, and combating /minimizing the effect of the interference. The channel and the interference covariance matrices could also be jointly estimated in time domain.

Channel estimation requires the use of reference symbols, which cause extra system overhead, and waste of bandwidth, as they do not carry any data. Interference rejection can also be implemented by using more than a single receive antenna, in which case, one of the antennas is used to cancel the interference while the other is in principle used for receiving the desired information, this however adds to the system complexity.

In this thesis, we model the interference and the desired user data in a similar manner, using quadrature amplitude modulation (QAM) and matlab for our simulations. The performance evaluation of the various interference rejection schemes is based on the bit error rate, and capacity.

1.1 Background

Over the last decade, the world has witnessed tremendous improvement in communication technologies as a result of increased data rates made available/possible by cutting edge research and innovation. Modern systems such as LTE-A, also use Orthogonal Frequency Division Multiplexing (OFDM) algorithms in combination with Multiple Input and Multiple Output antennas (MIMO) to achieve these high data rates and hence improve general system capacity. However, with the ever increasing demand in user traffic, more has to be done to cope up with the demand.

Increasing the system bandwidth clearly would go a long way to positively impact this cause, if only bandwidth was not such an expensive resource. A better approach is more efficient utilization of the available radio spectrum by re-using the existing bandwidth, this though comes at the expense of increased interference between parties using the same frequency band, leading to lower Signal to Noise plus Interference (SINR) levels, and therefore less capacity and increased error rates.

Building receivers and transmitters with interference rejection algorithms plays an important role to increase the signal to interference plus noise levels, by minimizing and possibly mitigating the generation and effect of interference. In this thesis however we focus more on the receiver side, if the receiver has knowledge about the structure of the interference, it is possible to partly, get rid of this interference (1).

1.2 Motivation

The investigation of interference rejection suppression in MIMO systems is a very important topic for quite a number of reasons. Cellular capacity is hugely interference limited especially by co channel interference (CCI) and adjacent channel interference (ACI) (1) (2). One remedy for this kind of cellular interference is cell splitting and reduction in transmission power (2). Cell splitting however leads to creation of more cells, and each cell would require separate equipment (base station) rendering this process rather expensive. A less expensive alternative would in fact be to use more sophisticated signal processing, and in particular IRC is a promising technique (1) (2).

Also with improvements and innovations in technology, we have seen an upspring of better user equipment, in form of smart phones, Tablet PCs etc, these however raise the issue of compatibility with the older technologies. Interference rejection algorithms are useful in facilitating a transition between old and new technologies. The usefulness of compatibility can be highlighted by a number of examples, such as the co-utilization of the same frequency band between narrow band Code Division Multiple Access (CDMA) systems and the older Time Division Multiple Access (TMA) systems (2).

Hence the motivation of this thesis evolves from the need to evaluate some of the existing interference rejection algorithms, such as estimation of the covariance matrix of the interference, using reference symbols, among others, and to combine these algorithms with the hope of coming up with an even better algorithm that minimizes or mitigates interference, and as a result achieve increased data rates and lower bit/symbol error rates.

1.3 Objectives of the Thesis

This thesis concerns the evaluation of various interference rejection suppression algorithms in MIMO systems, and investigating the factors which influence their performance.

The focus through the research has been put on methods using minimum mean square error, and how we can obtain better results by applying algorithms such as the Rao-Blackwell Ledoit Wolf method, and optimization in combination with the reference symbols and the data symbols covariance matrix estimation methods.

1.4 Structure of the Thesis

This thesis is made up of 6 chapters, and organized as follows. Chapter 2 provides a theoretical background of MIMO systems, QAM, MMSE and LDPC codes, all of which are used in the thesis. Chapter 3 gives a description of the various covariance matrix estimation methods that we investigate, explaining what we actually do. It also includes the system model used and the methods for performance evaluation. Chapter 4 covers the results obtained from the various covariance matrix estimation methods in terms of BER and capacity, and an analysis of these results. We also present results about the optimal amount of training needed, and how it varies with SNR, SIR, signal (16-QAM and 4-QAM) constellation and coherence interval (in symbols). Chapter 5 covers a brief conclusion of our findings and chapter 6 contains areas for possible future research in connection with interference rejection suppression.

CHAPTER 2

2 Theoretical Background

2.1 The Quadrature Amplitude Modulation (QAM)

For the simulations, as mentioned earlier, Quadrature Amplitude Modulation (QAM) is used. In this information modulation technique, the data is carried on the amplitude values of two signals which are 90^0 out of phase with each other (orthogonal). At a given time, t, the QAM signal (with respect to band-pass transmission) can be expressed as

$$z(t) = I(t)\cos(2\pi f_c t) - Q(t) sin(2\pi f_c t) \quad (2.0)$$

where *I(t)* represents in the inphase signal component, consisting of the amplitude and pulse shape, *Q(t)* represents the quadrature component and f_c is the carrier frequency. The information is represented by changing the amplitude and the phase of the two components (I and Q components) (3). Figure 1 below is a diagrammatic representation of 16-QAM signal constellation, where each dot represents one signal made up of 4-bits.

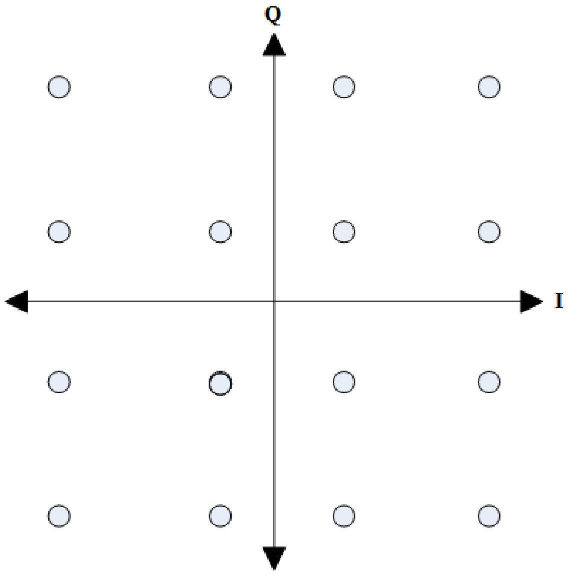

Figure 1. Signal Constellation for 16-QAM

2.2 Multiple Input Multiple Output (MIMO)

As stated earlier, MIMO stands for multiple input and multiple output antennas, which if used can achieve higher traffic data rates in communication systems. MIMO techniques also open up a whole new spatial dimension for exploitation, keeping in mind the limitations surrounding frequency and time dimensions. By using multiple transmit and multiple receive antennas, it is possible to transmit multiple data streams at the same time, and within the same frequency band, hence increasing the data rate with the number of transmit antennas and thereby improving system performance through diversity. Figure 2, below depicts a wireless transmission system using MIMO. By exploiting the structure of the channel matrix, independent paths can be obtained via which independent data streams can be transmitted, that is multiplexing.

Despite the positive side of MIMO deployment, using multiple antennas comes at a cost of space and extra power requirement (especially for hand held devices), and not to mention the increased complexity needed for multi-domain signal processing (4).

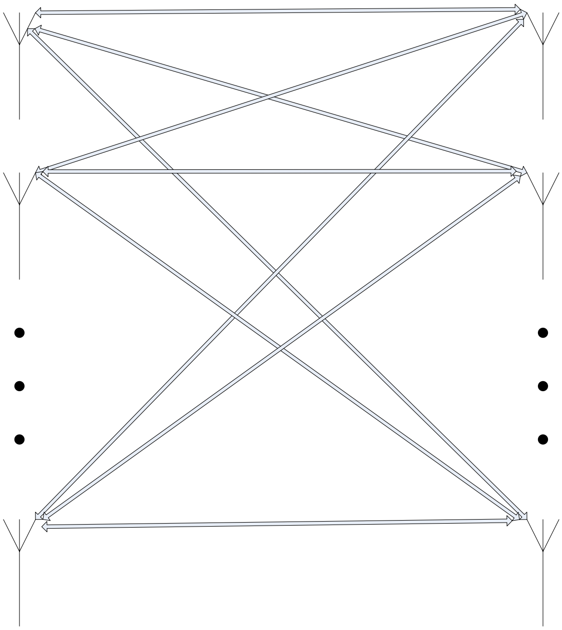

Figure 2. **Wireless Transmission MIMO system**

Using the above diagram, the MIMO system model can be derived as below.

$$\begin{bmatrix} y_1 \\ \vdots \\ y_{Mr} \end{bmatrix} = \begin{bmatrix} h_{11} & \cdots & h_{1Mt} \\ \vdots & \ddots & \vdots \\ h_{1Mr} & \cdots & h_{MrMt} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_{Mt} \end{bmatrix} + \begin{bmatrix} n_1 \\ \vdots \\ n_{Mr} \end{bmatrix} \tag{2.1}$$

Equation (2.1) above can alternatively be written as

$$\boldsymbol{y} = \boldsymbol{Hx} + \boldsymbol{n} \tag{2.2}$$

where, $\boldsymbol{y}$ is the received, M_r x 1, symbol column vector, $\boldsymbol{x}$ is the M_t x 1, column vector of the transmitted symbols, $\boldsymbol{H}$ is the channel matrix of dimension M_r x M_t, and $\boldsymbol{n}$ the M_t x 1 noise vector. M_t and M_r are the total number of transmit and receive antennas respectively.

2.2.1 Parallel Decomposition of the MIMO Channel (CSI available at Tx)

Having multiple transmit and receive antennas can further be exploited by sending independent data streams across the independent channels. The MIMO channel can be decomposed into independent channels, there by achieving spatial multiplexing. The performance of such a system can thus be improved by a factor, denoted *R*, where *R* is the rank of the channel. The rank is the number of independent streams which can be obtained from decomposing the MIMO channel, also called the multiplexing gain, and cannot exceed the number of rows and columns of the MIMO channel matrix, $R \leq min\ (M_t, M_r)$, for a channel matrix, $\boldsymbol{H}$ of dimension $M_r\ x\ M_t$.
Applying matrix theory, the singular value decomposition of the matrix can be obtained as;

$$\boldsymbol{H} = \boldsymbol{U\Sigma V}^{H} \tag{2.3}$$

Where $\boldsymbol{U}$ and $\boldsymbol{V}$ are singular-square matrices of dimensions M_r and M_t respectively, while $\boldsymbol{\Sigma}$ is an $M_r\ x\ M_t$, diagonal matrix with singular values of $\boldsymbol{H}$. Decomposition of the MIMO channel is illustrated in Figure 3, before transmission, the signal transmission vector, $\boldsymbol{x}$ is multiplied by the matrix $\boldsymbol{V}$ and at the receiver, the received signal vector, $\boldsymbol{y}$ is multiplied by the matrix $\boldsymbol{U^H}$ (5).

At the transmitter

$$\tilde{y} = U\Sigma V^{H}V\tilde{x} + \tilde{n} \tag{2.4}$$

$$\tilde{y} = U\Sigma\tilde{x} + \tilde{n} \tag{2.5}$$

At the receiver

$$\tilde{y} = U^H U \Sigma \tilde{x} + U^H \tilde{n} \tag{2.6}$$
$$\tilde{y} = \Sigma \tilde{x} + n \tag{2.7}$$

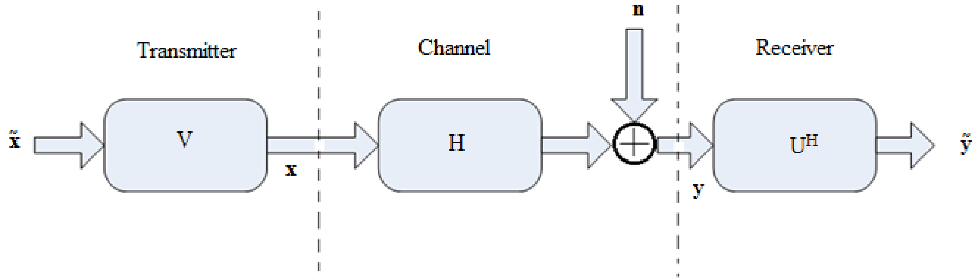

Figure 3. Decomposition of the MIMO channel

Since U^H is a unitary matrix, then $U^H\tilde{n}$=***n*** (4).

2.2.2 Capacity of the MIMO Channel

The Shannon capacity refers to the maximum amount of information rate that can be sent across the channel with minimal error rate. The channel capacity very much depends on the availability of Channel State Information (CSI) at the receiver/transmitter or both, however we assume that CSI is available at the receiver.

The capacity, *C* of the channel is given by the relationship between the input data vector, ***x*** and the output vector ***y***.

$$C = max_{p(x)} I\,(X;Y) \tag{2.8}$$

Where, *p(**x**)* is the probability distribution of the input vector ***x***, while *I(x;y)* represents the mutual information between ***x***, and the receiver vector ***y***.

for assuming a memory-less channel, the mutual information, is defined as;

$$I(x;y) = \sum_{x\in X, y\in Y} p(x,y)\, log_2\left(\frac{p(x,y)}{p(x)p(y)}\right) \;, \tag{2.9}$$

Taking the sum over all possible input and output pairs $x \in \mathcal{X}, y \in \mathcal{Y}$ respectively, in which case Equation 2.8 becomes (4),

$$C \;= max_{p(x)} \sum_{x\in X, y\in Y} p(x,y)\, log_2\left(\frac{p(x,y)}{p(x)p(y)}\right) \tag{2.10}$$

Mutual information can also be written in terms of the channel differential entropy of the channel output and the conditional differential entropy of the channel output vector, $\boldsymbol{y}$, given knowledge of the input vector $\boldsymbol{x}$, as below,

$$I(X,Y) = H(Y) - H(Y|X) \tag{2.11}$$

Where,

$$H(Y) = -\sum_{y\in Y} p(y)\, log_2\, p(y) \tag{2.12}$$

and,

$$H(Y|X) = -\sum_{x\in X, y\in Y} p(x,y)\, log_2\, p(y|x) \tag{2.13}$$

However since $\boldsymbol{x}$ and $\boldsymbol{n}$ are independent, $H(Y|X) = H(\boldsymbol{n})$, and Equation 2.11 simplifies to

$$I(X,Y) = H(y) - H(n) \tag{2.14}$$

From Equation 2.14, it is clear that to maximize the mutual information, $\boldsymbol{y}$ has to be maximized, and the mutual information of $\boldsymbol{y}$ depends on its covariance matrix which is given by,

$$\boldsymbol{R_{yy}} = E\{\boldsymbol{yy}^H\} = \boldsymbol{HR_{xx}H}^H + N_o\boldsymbol{I}_{Mr} \tag{2.15}$$

Where $\boldsymbol{R}_{xx} = E\{\boldsymbol{y}\boldsymbol{y}^H\}$, the covariance matrix of the input vector, and $E\{.\}$ is the expectation operator. It so happens that for all random output vectors with covariance matrix $\boldsymbol{R}_{yy}$, the entropy ***H(y)*** is maximized when **y** is a zero mean circularly symmetric complex- Gaussian (ZMCSCG) random vector. But for this to happen, the input vector, $\boldsymbol{x}$ must also be a ZMCSCG random vector, hence this is the ideal distribution on $\boldsymbol{x}$ (5). The differential entropies of $\boldsymbol{y}$ and $\boldsymbol{n}$ are given in the equations below

$$H(y) = \ log_2 \, det\ [\pi e R_{yy}] \tag{2.16}$$

$$H(n) = \ log_2 \, det\ [\pi e N_o I_{Mr}] \tag{2.17}$$

This implies that the mutual information in Equation (2.14) translates into,

$$I(X;Y) = \ log_2 \, det\ \left(I_{Mr} + \frac{E_S}{M_T N_o}\ H R_{xx} H^H\right) \tag{2.18}$$

And hence the capacity in Equation 2.8 reduces to

$$C = max_{Tr(R_{xx} = M_T)} \, log_2 \, det\ \left(I_{Mr} + \frac{E_S}{M_T N_o}\ H R_{xx} H^H\right) \tag{2.19}$$

This is the capacity for unit bandwidth, therefore for a bandwidth, W Hz, the maximum achievable data rate would be W.*C* bits/s (4) (5) (6).

2.2.3 Capacity when channel is unknown to Transmitter

Without channel state information the transmitter is unable to optimize its power, the best strategy would be to equally distribute the available power across all the M_t transmit antennas. The resulting input covariance, $\boldsymbol{R}_{xx}$ = I_{Mt}, which implies that the input signal vector $\boldsymbol{x}$ is chosen to be statistically independent, and the capacity follows from Equation 2.19**.**

$$C = \ log_2 \, det \, \left(\boldsymbol{I}_{Mr} + \frac{E_s}{M_T N_o} \, \boldsymbol{HH^H} \right) \tag{2.20}$$

Using eigen decomposition, $\boldsymbol{HH^H}$ can be expressed as, $\boldsymbol{Q\Lambda Q^H}$, where $\boldsymbol{Q}$ is an M_r x M_r unitary matrix while $\boldsymbol{\Lambda}$ is a diagonal matrix with the eigen values of the matrix $\boldsymbol{HH^H}$ on its main diagonal.

$$C = \ log_2 \, det \, \left(\boldsymbol{I}_{Mr} + \frac{E_s}{M_T N_o} \boldsymbol{Q\Lambda Q^H} \right) \tag{2.21}$$

Using the fact that $\boldsymbol{\Lambda} = diag\{\lambda_1 \, \lambda_2 \, \lambda_3 \dots \lambda_{Mr}\}$, with $\lambda_i \geq 0$, in connection with basic matrix theory, the capacity further reduces to

$$C = \ log_2 \, det \, \left(\boldsymbol{I}_{Mr} + \frac{E_s}{M_T N_o} \, \boldsymbol{\Lambda} \right) \tag{2.22}$$

$$C = \sum_{i=1}^{r} log_2 \, \left(1 + \frac{E_s}{M_T N_o} \, \lambda_i \right) \tag{2.23}$$

where r is the rank of the channel and λ_i are the eigen values of $\boldsymbol{HH^H}$.
From Equation 2.23, the MIMO capacity can be seen as a combination of r SISO links, each with a power gain, λ_i, and transmit power, E_s/M_t.

2.2.4 Capacity when channel is known to Transmitter

The capacity obtained in Equation 2.21 above cannot be considered as the Shannon capacity, since with availability of channel state information at the transmitter, better capacity, and improved system performance can be registered. This is achieved by assigning different energy levels to the various links that make up the MIMO channel, with more energy being allocated to links with better quality and vice versa, a technique referred to as water filling, refer to Equation 2.34 below (6).

$$C = \max_{\sum_i^r \gamma_i} \sum_{i=1}^{r} log_2 \left(1 + \frac{E_s\,\gamma_i}{M_T N_o}\,\boldsymbol{\lambda_i}\right) \tag{2.34}$$

Where γ_i is the transmit energy in the i-th sub channel so that $\sum_i^r \gamma_i = M_t$.

2.3 Minimum Mean Square Error (MMSE) Detection

The MMSE equalizer offers a tradeoff between minimizing the interference level and noise enhancement, unlike the zero forcing equalizer which is ignorant to noise enhancement. The basic principle of the MMSE is to minimize the Mean Square Error (MSE) between the correct data symbols, **x**, and the filter output. The filter output, $\boldsymbol{z} = w\boldsymbol{y}$ and the minimum square error for the k-th symbol is evaluated as in Equation 2.35 and $\boldsymbol{w}$ is the MMSE filter. The ideal MMSE filter will be dealt with later and is given in Equation 3.7.

$$\sigma^2{}_{MMSE} = min_{w_k}\, \mathcal{E}[|x_k - z_k|^2] \tag{2.35}$$

2.4 Low Density Parity-Check (LDPC) Codes

In communication systems, errors are bound to happen for a number of reasons, such as channel dispersion, which in turn leads to inter-symbol interference, noise in the channel, which causes erroneous decisions at the receiver, power surge in the electronic circuits, to mention but a few. One way to detect and correct these errors is through coding, this introduces redundancy in the system, that is, extra bits added to the original number of bits. There are number of coding schemes employed in communication systems, with turbo codes being used in LTE, others include reed Solomon codes, convolutional codes, etc.

In this thesis however, LDPC codes are used because their performance is closer to the Shannon limit for various channels, for large block lengths, and they also experience less error floor as compared to turbo codes. They derive their name from the nature of their parity-check matrix, which is comprised of few 1's and a lot of 0's. LDPC codes are divided into regular and irregular LDPC codes. Regular LDPC codes have a constant number of 1's across the rows and columns of the parity-check matrix, while the opposite is true for irregular LDPC codes. The irregular LDPC codes however are the ones whose performance is close to the Shannon capacity limit.

$$\boldsymbol{M} = \begin{bmatrix} 1\,0\,0\,1\,0\,1\,0\,0\,0\,1\,0 \\ 0\,0\,1\,1\,1\,0\,0\,1\,0\,0\,0 \\ 0\,0\,1\,0\,1\,0\,1\,0\,1\,0\,0 \\ 1\,0\,0\,0\,0\,1\,0\,1\,0\,1\,0 \end{bmatrix} \tag{2.36}$$

If the $\boldsymbol{M}$-above is a parity-check matrix, the code word, $\boldsymbol{v}$ is constructed such that

$$\boldsymbol{Mv} = 0\ (mod\ 2) \tag{2.37}$$

However to construct the code, a generator matrix, is used, the generator matrix can be calculated from the parity check matrix, $\boldsymbol{M}$. The parity check matrix in Equation 2.36 can also be represented graphically by Tanner graphs. (7) Decoding can be done by either hard decoding, where each bit is considered either a 0 or 1, for binary codes, or soft decoding which offers additional information about how probable each of the possible channel symbols might have been (8).

However the problem with LDPC codes is that the code lengths have to be quite long in order to obtain capacity performance close to the Shannon limit, more over long block lengths lead to large parity-check and generator matrices which require a lot of computational resources and thereby making the encoders for the LDPC codes somewhat complex and expensive. However similar problems exist for most near-capacity codes and LDPC are usually better than Turbo-like codes from this perspective.

2.5 Coherence Time

One of the important channel measurements is the coherence time, which characterizes the channel changes. The Channel coherence time defines the time interval in which the channel is assumed to be constant (9).

According to (10) the values of coherence time T for a fading channel can be calculated using

$$T = \frac{1}{2 f_{max} T_s} \tag{2.38}$$

where f_{max} is the maximum Doppler frequency, which can be obtained using

$$f_{max} = \left(\frac{v}{c}\right) f_c \tag{2.39}$$

and T_s is the symbol time , whereas f_c is the carrier frequency. Finally v and c denote the velocity of the receiver and the speed of light respectively.

2.6 Covariance Matrix

The covariance matrix is an important statistic that describes the pair-wise correlation between random variables. These variables are considered to have a Gaussian distribution and zero mean (11).
For an n x 1 random vector ***P***, below with mean μ,

$$P = \begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_n \end{bmatrix} \tag{2.40}$$

the covariance matrix of ***P***, denoted as *cov[**P**]* is defined as;

$$cov[P] = E[(P - \mu)(P - \mu)^T] \tag{2.41}$$

where $\mu = E[P]$, and $E[.]$ denotes expectation. Equation 2.14 can also be expressed as below.

$$cov[P] = \begin{bmatrix} E[(P_1 - \mu_1)(P_1 - \mu_1)] & \cdots & E[(P_1 - \mu_1)(P_n - \mu_n)] \\ \vdots & \ddots & \vdots \\ E[(P_n - \mu_n)(P_1 - \mu_1)] & \cdots & E[(P_n - \mu_n)(P_n - \mu_n)] \end{bmatrix} \tag{2.42}$$

2.7 Channel Estimation and Reference Symbols

In wireless communications it is important to know the channel properties of the communication link because with this information, it is possible to adapt the signal transmissions to the current channel conditions. This in turn leads to reliable data transmissions and increased information rate.

A popular approach used to estimate the channel is using reference symbols, also called pilot symbols or training sequence, where symbols known to the receiver are transmitted, and the channel estimate obtained by combining the knowledge of the transmitted reference symbol and the received copy of the signal. This is achieved at the cost of reduced number of symbols for data transmission (12). If C denotes the pilots sent, we can simply estimate the channel using the Least Squares (LS) approach as

$$H^{LS} = \frac{R}{C} \tag{2.43}$$

where R is the received noisy-distorted version of the pilot symbols (9).

However the LS channel estimate in Equation 2.34 can be improved by taking into account the correlation between different frequencies and using the linear MMSE (LMMSE) method. For more details about the LMMSE channel estimation method, the reader is referred to (9). It should also be noted that the LMMSE approach has a better performance but high computational complexity in comparison with the LS method. In this thesis, the LS approach for channel estimation is used, refer to Section 3.1.1 for more details.

CHAPTER 3

3 Covariance Matrix Estimation Methods

In order to build an MMSE receiver that can combat inter-cell interference and improve system capacity and BER, the interference covariance matrix has to be estimated, the more accurate the estimate of the covariance matrix the better the receiver performance. For this matrix estimate, we focus on two different schemes, using the received data symbols (data-samples-based interference covariance matrix, $\widehat{\boldsymbol{R}}_{yy}$), and secondly using the reference symbols (reference symbols based interference covariance matrix, $\boldsymbol{R}_{RS}$). Earlier research indicates that algorithms based on the reference symbols outperform data-samples-based algorithms (1).

In this thesis, the aim is to improve the accuracy of these two estimated covariance matrices by employing different techniques. First by optimizing over the estimated interference covariance matrices (data samples and reference symbols) according to the criteria that will be presented in Section 3.5, and secondly by applying the Rao-Blackwell Ledoit-Wolf (RBLW) (13) method, which is presented in Section 3.6 to both the data samples based covariance matrix estimation method, and the optimized interference covariance matrix.

Also presented in this chapter is the criteria used to obtain the channel matrix estimate, and the system performance evaluation. However before getting into the details of the covariance matrix estimation, let us look at the system model employed in this thesis work.

3.1 System Model

Depending on the multiplexing schemes applied at the transmitter side, various transmit-receive configurations can be achieved. In this thesis we assume spatial multiplexing at the transmitter and multiple receive antennas with a focus on the downlink, for a single user experiencing interference from another cell. The received signal vector $\boldsymbol{y}_t$ at the t-th time instant is given by, $\boldsymbol{y}_t = k\boldsymbol{H}\boldsymbol{x}_t + l\boldsymbol{S}\boldsymbol{z}_t + \boldsymbol{n}_t$, which can further be expressed as

$$\begin{bmatrix} y_1 \\ \vdots \\ y_{Mr} \end{bmatrix} = k \begin{bmatrix} h_{11} & \cdots & h_{1Mt} \\ \vdots & \ddots & \vdots \\ h_{1Mr} & \cdots & h_{MrMt} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_{Mt} \end{bmatrix} + l \begin{bmatrix} s_{11} & \cdots & s_{1mt} \\ \vdots & \ddots & \vdots \\ s_{1Mr} & \cdots & s_{MrMt} \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_{Mt} \end{bmatrix} + \begin{bmatrix} n1 \\ \vdots \\ n_{Mr} \end{bmatrix} \tag{3.0}$$

where $\boldsymbol{y}_t$ = [y_1 y_2 y_3 … y_{Mr} $]^T$, M_r being the number of antennas at the receiver, $\boldsymbol{x}_t$ = [x_1 x_2 x_3 … x_{Mt} $]^T$ is a column vector of the transmitted symbols from M_t transmit antennas while $\boldsymbol{z}_t$ = [z_1 z_2 z_3 … z_{Mr} $]^T$ is a column vector of the received symbols at M_r receive antennas, $\boldsymbol{n}_t$= [n_1 n_2 n_3 … n_{Mt} $]^T$ is the white Gaussian noise with zero mean and unit variance. Finally $\boldsymbol{H}$ and $\boldsymbol{S}$ are the corresponding desired channel matrix and the interference channel matrix respectively.

The desired channel matrix, $\boldsymbol{H}$ and the interference channel matrix, $\boldsymbol{S}$ are both normalized to have unit mean energy, while $\boldsymbol{x}_t$ and $\boldsymbol{z}_t$ are the mapped data and the interference symbols respectively which are also normalized to have unit mean energy. *k* and *l* are scaling factors for the SNR and the SIR respectively and their derivation is shown in appendix A. SNR is a ratio of the desired signal power to the background noise power, while the SIR is a ratio of the desired signal power to the interference power.

Initially simulations were done without coding, however to further improve our results, LPDC coding was introduced, see Section 2.4

3.1.1 Channel Estimation by Pilot Boosting

To achieve better channel estimates and system performance, the reference symbols can be assigned more power than the data carrying symbols.
This is achieved as follows, assume a randomly generated matrix, **R** and make the singular value decomposition.

$$\boldsymbol{R} = \boldsymbol{U\Sigma}\mathbf{V}^{\mathbf{H}} \tag{3.1}$$

Where $\boldsymbol{U}$ and $\boldsymbol{V}$ are unitary matrices, while $\boldsymbol{\Sigma}$ is a diagonal matrix with the singular amplitudes of $\boldsymbol{R}$ on its main diagonal.
We then create a new matrix, call it $\boldsymbol{W}$**,** by taking M_t rows and T_p columns of the unitary matrix **U**, where T_p is the number of pilots, and the boosted pilots are achieved in the following equation.

$$P = W \times \sqrt{T_p} \tag{3.2}$$

where $T_p \geq M_t$, this is optimal when we using pilot boosting
The channel estimate is obtained as below using the least mean squares criterion.

$$\hat{\mathrm{H}} = YP^{\dagger} \tag{3.3}$$

$\mathrm{P}^{\dagger}$ is the pseudo inverse of the boosted pilots, (10) $\hat{\mathrm{H}}$ is the channel estimate used in the simulations.

However its also of importance to keep in mind that channel estimation using reference symbols has its drawbacks. Obviously the reference symbols consume bandwidth yet they carry no useful information, for instance LTE uses 12 refence symbols for each resource block pair (12 x 16), if these were instead used for data symbols, it could provide better spectral efficiency. Therefore in order to have a sensible spectral efficiency, the training sequence should be kept short, but short training sequencies are more sensitive to noise than longer sequencies which can average out the noise. Also if the channel changes after pilot transmission, the receiver is unable to detect these changes and would use an out-dated channel estimate, resulting into decision error (9).

3.1.2 System Performance Evaluation

The system performance is evaluated based on two factors, the first being the BER, and the other is the system throughput. The BER for MIMO systems with M-QAM modulation is calculated using matlab, as the number of erroneous bits detected by the interference aware receiver, divided by the total number of bits in the transmitted signal.
System throughput is obtained by finding the receiver filter w, refer to Section 2.3, for the different covariance matrix estimation schemes under investigation, which maximizes the signal to interference plus noise ratio (SINR). The SINR per receive symbol per receive antenna is calculated by:

$$SINR_i = \frac{|w_{ij}\, h_{ij}|^2{}_{i=j}}{\sum_{i,i\neq j}^{M} |w_{ij}\, h_{ij}|^2 + \sum_{j}^{M} |w_{ij}\, s_{ij}|^2 + N_o \sum_{j}^{M} |w_{ij}|^2} \tag{3.4}$$

where i represents the transmit antenna while j represents the receive antenna and $i = j$, refers to the main diagonal elements which represent the desired symbol or channel for specific antenna while the case of $i \neq j$ refers to the off-diagonal elements, also considered as interference. For each receive antenna, the SINR is calculated by finding the energy in the desired data symbol with respect to the interference power and the noise power.

Then the system capacity is calculated according to Shannon capacity (3). Keep in mind that the system capacity has been covered in Section 2.2.2.

$$C = \sum_{i=1}^{M_t} BW\ log_2\,(1 + SINR_i) \tag{3.5}$$

Where, BW is the physical bandwidth (3) , in our simulations we assume the unit bandwidth. This equation above holds for Gaussian inputs symbols, and not for QAM.

3.2 Interference Rejection Combining

Interference rejection combining (IRC) algorithms are effective in improving the cell edge user throughput and can suppress the inter-cell interference when the information on the interference covariance matrix assumed to be known, which is an ideal assumption and can be considered as the upper bound for any receiver based on MMSE algorithms and in this thesis as well (1) (14). With the fore mentioned assumptions in consideration, the true covariance matrix of the received signal would be given as;

$$\boldsymbol{R}_{\boldsymbol{yy\_IRC}} = [k^2\boldsymbol{H}\boldsymbol{H}^H + l^2\boldsymbol{S}\boldsymbol{S}^H + N_o\boldsymbol{I}]^{-1} \tag{3.6}$$

The ideal MMSE filter is obtained as

$$w_{IRC} = k\boldsymbol{H}^H\left[R_{yy\_IRC}\right]^{-1} \tag{3.7}$$

where, N_o is the noise power.

3.3 Reference Symbols based Covariance Matrix Estimation

This is one of the most common methods of estimating the desired channel in communication systems. In our simulations, it is assumed that the receiver has information about T_p reference symbols of each transmitted block, in a coherence time T, where $T_P \leq T$. We further investigate the optimal number of pilots, based on different signal constellations (Section 3.3.1) and various channel conditions (SIR and SNR). Based on the received block, it is possible to estimate the effective interference plus noise vector, $\hat{z}_k$ as below,

$$\hat{z}_k = \boldsymbol{y}_k - k\boldsymbol{H}\boldsymbol{x}_k \tag{3.8}$$

where $\boldsymbol{x}_k$ is a 1 x Mr column vector of containing the known pilot symbols for the k-th time instant. Using the estimate in Equation 3.8 above, the interference covariance matrix is computed below.

$$\widehat{\boldsymbol{R}}_{zz} = \frac{1}{T_p}\sum_{k=1}^{T_p} \hat{z}_k \hat{z}_k{}^H = \frac{1}{T_p}\sum_{k=1}^{T_p}(\boldsymbol{y}_k - k\boldsymbol{H}\boldsymbol{x}_k)(\boldsymbol{y}_k - k\boldsymbol{H}\boldsymbol{x}_k)^H \tag{3.9}$$

$$1 \leq k \leq T_p$$

The overall covariance matrix for the reference symbol based method, assuming the channel is known (an assumption that can be made due to the structure of LTE) is then calculated as

$$\widehat{\boldsymbol{R}}_{RS} = \widehat{\boldsymbol{R}}_{zz} + k^2\boldsymbol{H}\boldsymbol{H}^{\boldsymbol{H}} \tag{3.10}$$

while the filter, w_{RS} is formulated as,

$$w_{RS} = k\boldsymbol{H}^{\boldsymbol{H}}[\widehat{\boldsymbol{R}}_{RS}]^{-\boldsymbol{1}} \tag{3.11}$$

The equalized block (filter output) follows as,

$$\boldsymbol{r_k} = w_{RS}\boldsymbol{y_k} \tag{3.12}$$

3.3.1 Optimal Training Interval

It is possible to achieve high data rates in multiple antenna systems especially when the receiver has knowledge about the channel, which can be achieved using reference symbols (15). As we are study the case of the user throughput on the cell edge, the interference covariance matrix has to be estimated, but in order to achieve maximum capacity, the number of reference symbols has to be chosen properly and appropriately. However, the optimal number of reference symbols depends on the SNR, SIR, and coherence time of the channel, (16) see section 4.4.

Increasing the number of pilots, T_p improves the estimate of the covariance matrix ($\hat{\boldsymbol{R}}_{RS}$), but if T_p is too large, the slots left for data transmission (T-T_p), could be too small. Therefore we compute the optimal value of T_p versus the total block length through the scale factor (T-T_p) /T, for different SNR, SIR, and block lengths. According to Equation 3.13, the maximum capacity is dependent on the optimal value of T_p. Thus, the capacity decreases linearly when the time interval of T_p increases beyond the optimal value through the coefficient (T-T_p) /T.

$$C = \sum_{i=1}^{M_r} \arg\max_{T_p} \frac{T - T_p}{T} \ log_2 \left(1 + \ SINR_i\right) \tag{3.13}$$

where C is the capacity, arg max denotes the argument maximization function, and the *SINR* is defined in Section 3.1.2. The channel of the serving cell is assumed to be constant for the entire block length.

3.4 Data Samples Based Covariance Matrix Estimation

Sometimes using reference symbols to estimate the covariance matrix is a suboptimal choice, in such an instance, the received data symbols could be used to estimate the covariance matrix. Since the received data symbols contain both the channel effect and that of the interfering channel, it is not possible to estimate the covariance matrix of the undesired channel and that of the desired channel independently, like was the case for the method using reference symbols. Instead the estimated covariance matrix in this case is a combination of both the desired channel and the interfering channel effect (1). From the system model in Equation (3.0), the received signal covariance matrix is estimated as,

$$\widehat{\boldsymbol{R}}_{yy} = \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{y}_t \boldsymbol{y}_t{}^H \tag{3.14}$$

where T is the number of symbols considered. The covariance matrix estimate in the equation above approaches the true covariance matrix of the received signal when the observation window grows large. Since the covariance matrix has been estimated, we can easily formulate the receiver filter, w_d, for this case as,

$$w_d = kH^H[\widehat{\boldsymbol{R}}_{yy}]^{-1} \tag{3.15}$$

and the filter output, $r_d = w_d \times y$. However the above estimate of the receiver filter is not realistic because the ideal channel is unknown at the receiver side, hence we build the filter using the estimate channel ($\hat{\mathrm{H}}$) presented in Section 3.1.1. Considering the estimation channel, the receiver filter, $\widetilde{w}_d$ can be expressed as

$$\widetilde{w}_d = k\hat{\mathrm{H}}^H[\widehat{\boldsymbol{R}}_{yy}]^{-1} \tag{3.16}$$

3.5 Covariance Matrix Optimization

We also investigate the case of optimizing the interference covariance matrix based on the data samples method with the covariance matrix based on reference symbols method. Optimization generally leads to improved system performance, though at a cost of increased system complexity. However after optimization of the covariance matrices, we also apply the RBLW method presented in Section 3.6.

$$\widehat{\boldsymbol{R}}_{yy\_optz}(Q) = \widehat{\boldsymbol{R}}_{yy}Q + (1-Q)\widehat{\boldsymbol{R}}_{RS} \tag{3.17}$$

The optimized interference covariance matrix should be one that minimizes the minimum square error, in comparison with the IRC and maximizes the capacity of the system. The optimized interference covariance matrix $\widehat{\boldsymbol{R}}_{yy-optz}$ is a function of Q, which is chosen to vary between 0 and 1.

The Q which minimizes the error is chosen to be the optimal value, and the corresponding optimized covariance matrix considered as the best choice. We obtain the optimal value of Q according to the criteria below

$$\epsilon = \arg\min_{Q} E\left\{\left|\widehat{\boldsymbol{R}}_{yy\_optz}(Q) - R_{yy\_IRC}\right|^2\right\} \tag{3.18}$$

where E{.} denotes expectation. Once the optimal value is obtained, we can employ this covariance matrix with less error to determine the receiver filter, which, assuming the true channel can be expressed as,

$$\boldsymbol{w_{optz}} = k\boldsymbol{H^H}[\widehat{\boldsymbol{R}}_{yy\_optz}]^{-\mathbf{1}} \tag{3.19}$$

while the receiver filter $\widetilde{w}_{optz}$ with considering the estimate channel reads

$$\widetilde{w}_{optz} = k\hat{\mathrm{H}}^H[\widehat{\boldsymbol{R}}_{yy\_optz}]^{-1} \tag{3.20}$$

3.6 The Rao-Blackwell Ledoit-Wolf (RBLW) Method

The performance of the MMSE method for interference refection is greatly dependent on the covariance matrix estimate and how accurate it is in comparison to the true covariance matrix of the received signal. Different methods can be used to improve the estimate of the covariance matrix such as the Oracle Approximating Shrinkage (OAS) estimator, Ledoit-Wolf (LW) method, to mention but a few. The RBLW estimator as the name suggests, is a combination of the LW method and the Rao-Blackwell theorem, in an attempt to improve or reduce the estimation error of the LW method (13).

The Rao-Blackwell theorem states that if *g(X)* is an estimator of a parameter θ, then the conditional expectation of *g(X)* given *T(X)*, where *T* is a sufficient statistic is never worse than the original estimator *g(X)* under any convex loss criterion. (13) (17). The estimated sample covariance matrix in Equation 3.14, which is the most used estimator in a lot of applied problems, has its shortcomings, when its dimension, *p*, is bigger than the number of symbols under consideration, *T*, then it is not possible to invert the sample covariance matrix. On the other hand when $T \geq p$, the covariance matrix can be inverted, but inverting it leads to an even bigger estimation error, hence, it is 'ill conditioned' (18).

However the covariance matrix based on the data method does not necessarily achieve the MSE due to its high variance, (especially with few samples, a case we investigate) despite the fact that its unbiased for a large number of samples, therefore with the RBLW method, we aim to reduce this high variance of the covariance matrix by shrinking it towards the shrinkage target in Equation 3.21 below using the shrinkage coefficient given by Equation 3.23.

Thus, we can force the covariance matrix to be well conditioned, by imposing some ad hoc structure, for instance by diagonality or a factor model (18) (13)

$$\boldsymbol{F} = \frac{Tr(\widehat{\boldsymbol{R}}_{yy})}{p} \boldsymbol{I}_{mr} \tag{3.21}$$

where F is the shrinkage target.

For $\{ y_t \}_{t=1}^{T}$ vectors that have a Gaussian distribution, and of dimension, *p*, and a sample covariance matrix $\widehat{\boldsymbol{R}}_{yy}$, the RBLW estimator is obtained as a weighted average of the well conditioned covariance matrix, which is given by Equation 3.21, and the sample covariance matrix, $\widehat{\boldsymbol{R}}_{yy}$ (18) (13).

$$\widehat{\boldsymbol{R}}_{RBLW} = (1-\rho)\widehat{\boldsymbol{R}}_{yy} + \rho \boldsymbol{F} \tag{3.22}$$

Where, ρ is the shrinkage coefficient and is given by,

$$\rho = \frac{\frac{T-2}{T}.Tr\left(\widehat{\boldsymbol{R}}_{yy}{}^{2}\right) + Tr^{2}(\widehat{\boldsymbol{R}}_{yy})}{(T+2)\left[Tr\left(\widehat{\boldsymbol{R}}_{yy}{}^{2}\right) - \frac{Tr^{2}(\widehat{\boldsymbol{R}}_{yy})}{p}\right]}. \tag{3.23}$$

the receiver filter w_{RBLW} assuming the true channel reads

$$w_{RBLW} = kH^{H}[\widehat{\boldsymbol{R}}_{RBLW}]^{-1} \tag{3.24}$$

while the receiver filter $\widetilde{w}_{RBLW}$ with considering the estimate channel reads

$$\widetilde{w}_{RBLW} = k\hat{\mathrm{H}}^{H}[\widehat{\boldsymbol{R}}_{RBLW}]^{-1} \tag{3.25}$$

The RBLW method described above, is applied on the covariance matrix estimation method using data symbols, and the optimization method, so as to obtain more accurate covariance matrices and hence improved system performance.

CHAPTER 4

4 Results

As mentioned in the previous chapter, system performance evaluation is based on the BER and the system capacity, more so the 5% outage capacity and the Cumulative Distribution Function (CDF) of the capacity. The BER and capacity simulations are carried out for two assumptions about channel and interference knowledge at the receiver. The first being the availability of perfect channel state information at the receiver, in which case we refer to the desired channel as the true channel, secondly we simulate using the estimated channel from the reference symbols. Besides the difference in channel knowledge, different channel conditions (SNR, SIR and coherence time) are also considered, for a user of interest with 4 receive antennas and a transmitter also comprising 4 antennas.

Two data modulations schemes are employed, 4-QAM and16-QAM. The channel capacity for 4-QAM has been computed based on the Gaussian formula, Equation 3.5, which only holds for Gaussian input symbols, and not for 16-QAM or QPSK. Therefore the capacity results presented with respect to these schemes are rather "indicative" than hard facts.

4.1 BER for True Channel at the Receiver

Referring to Figure 4 below, the SIR is fixed to 10 dB, assuming a coherence time of 10 (in symbols), we evaluate the BER for various signals to noise ratios of the system using 4-QAM. As depicted in the figure, the IRC receiver performs best, which is as expected considering that the IRC receiver assumes complete knowledge of the true channel and the interference channel. The IRC is followed closely by the reference symbols based estimation method while the data symbols based method exhibits the poorest performance. The good performance of the reference symbols based method can be explained by the fact that since we assume knowledge of the true channel we are able to accurately estimate the interference and noise component in the received signal according to Equation 3.8, which directly translates to a good estimation of the covariance matrix and a rather efficient receiver filter as in Equation 3.11.

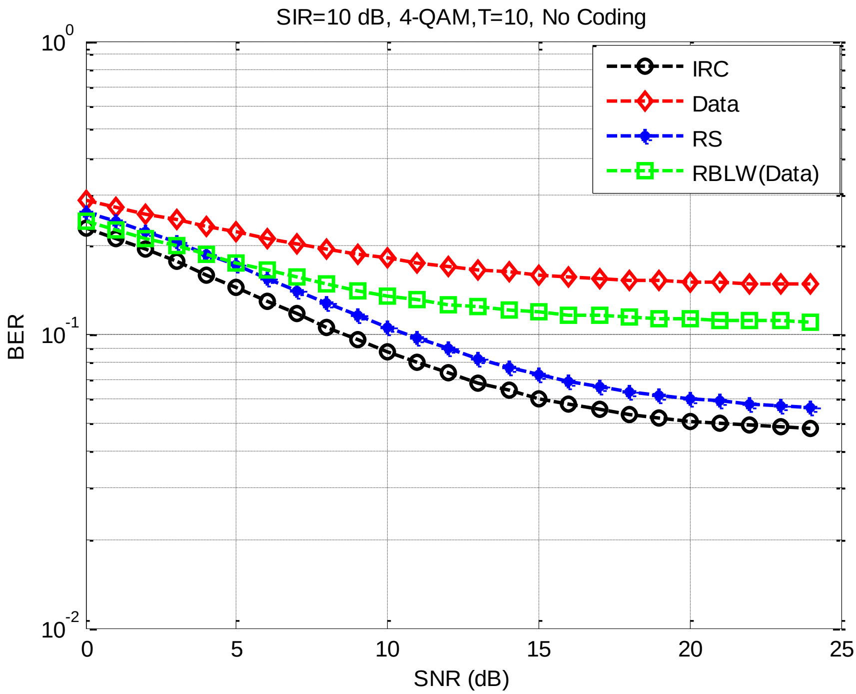

Figure 4. BER for true channel, T=10, SIR=10 dB, 4-QAM

The poor performance of the data symbols based covariance matrix estimation method is because few data samples were considered, 10, in this case. The RBLW method only offers a slight improvement in BER, since the data based method already exhibits a not so good BER performance.

However with increased number of data samples, the performance of the data samples based method can be improved. This can be explained by the fact that the noise is uncorrelated and so is the interference, so over a large observation window, the two can be averaged to zero, leaving only the data signals.

The results in Figure 4 show a rather high BER (because of the high interference levels and considering high order QAM, and 4 transmit/ receive antennas with spatial multiplexing), this can be reduced by coding. In this thesis, LDPC codes from the DVB-S2 standard with a code rate 1/2 and a block length of 32400 information bits have been used, the results of which are shown in Figures 5 and 6 for channel coherence intervals of 10 and 100 (in symbols) respectively.

The data symbols based method shows the same performance pattern, with BER decreasing with increasing number of samples taken into consideration even with coding. Proving further that the data based covariance estimation method performs well when a large number of data samples are taken into consideration. In attempt to further improve the performance of the data based method, the RBLW method is applied. The combination the data based method and the RBLW method offers a better performance, however the performance gain is not that much in comparison with just the data symbols based method, as we consider the user at the cell-edge.

The IRC receiver and the pilots based covariance estimation method on the other hand show a drastic drop in the BER, of course the IRC being the better performer, with BERs close to zero at SNR of about 10 dB and beyond, while the reference symbols based method achieves this after about 10.2 dB of SNR (Figure 5). By optimizing the data based covariance matrix and the reference symbols based covariance matrix, considerable gain is obtained comparing to the data symbols covariance estimation method alone. However applying the RBLW method on the optimized covariance matrix deteriorates its performance and the reason for this is discussed in the next section. The reference symbols method still out-performs the optimization method.

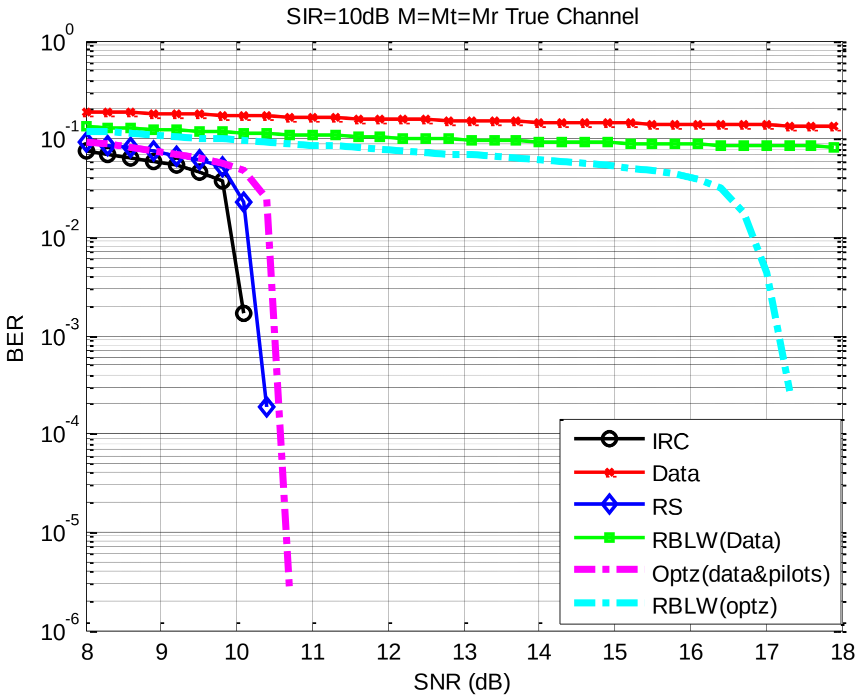

Figure 5. BER for true channel, T=10, SIR=10 dB, 4-QAM with coding

In Figure 6 below, the coherence time is set to be 100, which is rather a rare occurrence, so the main purpose of this particular simulation is to illustrate the conditions under which the data based covariance estimation method can obtain zero BER. According to the Equation 2.38 and Equation 2.39, *Ts* the symbol time chosen to be 100μs, and fc the carrier frequency, 2GHz. Thus $T \approx 10$ for high train velocities $v \approx 250\ Km/h$ and $T \approx 100$ for average human walking speed $v \approx 5\ km/h$.

However for large T there is no gain using the RBLW method as can be seen and this is because the RBLW is not well conditioned for covariance matrices with large number of samples (13). It is also observed that using optimization slightly decreases the system BER compared to the reference symbols based method, while there is a penalty applying the RBLW algorithm on the optimization result especially at high SNR.

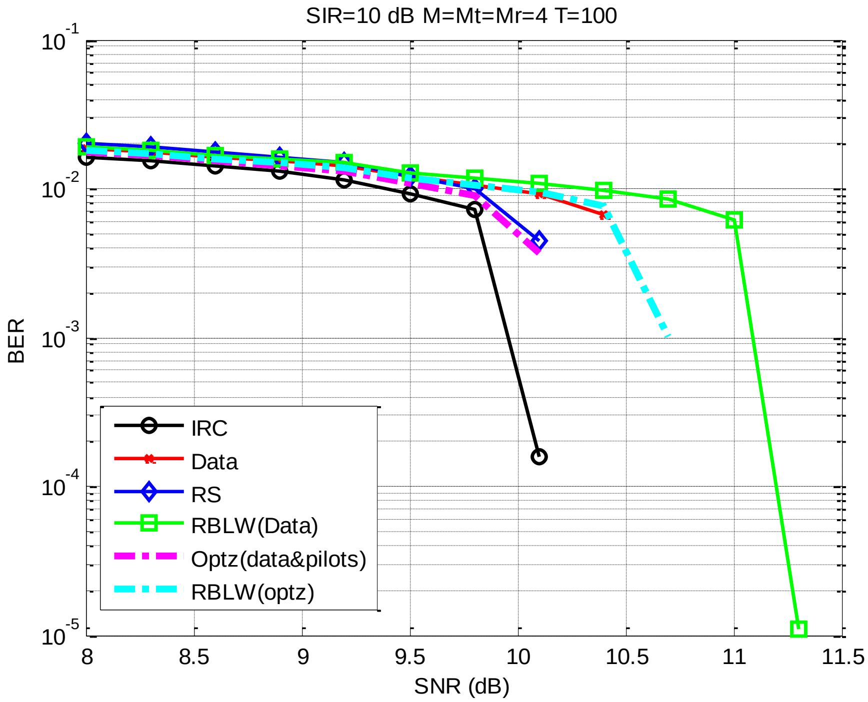

Figure 6. BER for True channel, T=100, SIR=10 dB, 4 QAM, with coding

4.2 Estimated Channel at the Receiver

Using the method, presented in Section 3.1.1, an estimate of the desired channel matrix is obtained according to Equation 3.3. In the simulations, performance for the case when T = 10, and T_p = 4 is evaluated. Normally covariance matrix estimation is done over the pilots alone, leaving the rest of the data block to waste; this is suitable for perfect channel estimates, as depicted in the previous section. However for not so accurate channel estimates, the data carrying symbols alone perform better than the pilots under certain channel conditions as will be seen later. The results presented in this section are obtained using the same covariance matrix estimation methods applied in the previous section, with the exception that instead of assuming the true channel, a practical channel estimate-matrix is used. Coding using LDPC is also applied.

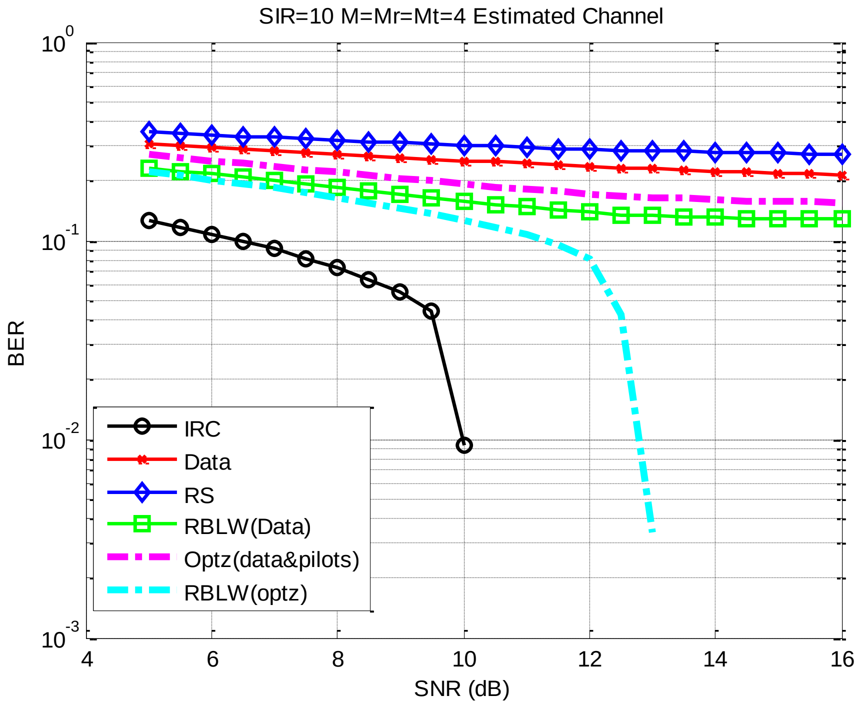

Figure 7. BER for the Channel Estimate, T=10, SIR=10 dB, 4-QAM with coding

In Figure 7 above, the SIR is fixed to 10 dB and a coherence time of 10 assumed. 4-QAM modulated symbols are used to simulate the BER of the system with the estimated channel.

It is observed that the pilots do not perform as well as they did with the true channel, in fact they experience the worst performance, especially at low SIR and SNR. The data based method also performs not so well, but can be improved by combining it with the RBLW method. Also the combination of the data based method and the reference symbols based method coupled with optimization offers slightly better results. Interestingly, it is observed that applying RBLW on optimized covariance matrix decreases the system error rate significantly, unlike the case with the true channel. The reason for this change in behavior is because, in case of the true channel (Figure 5), the optimized covariance matrix is largely made up of the reference symbols based covariance matrix, which explains the good performance. So combining the optimized covariance matrix with the RBLW method based matrix introduces more error as it is made up of the shrinkage target which is a scaled identity matrix, $\boldsymbol{F}$, Equation 3.21, and the sample covariance matrix. While in the case of the estimated channel, the poor performance of the optimized covariance matrix is inherited from both the data and the reference symbols covariance matrices. In this case, applying the RBLW method to the optimized covariance matrix, which has a high variance leads to better performance, as the RBLW method reduces the variance. The IRC receiver is used as the upper bound, and the gap between the IRC curve and the rest of the curves suggests that the channel estimate is not as good.

However at higher SIR, (20 dB in this case) it is possible to obtain a more accurate channel estimate which in turn leads to better covariance matrix estimates, and hence better performance as depicted in Figure 8 below.

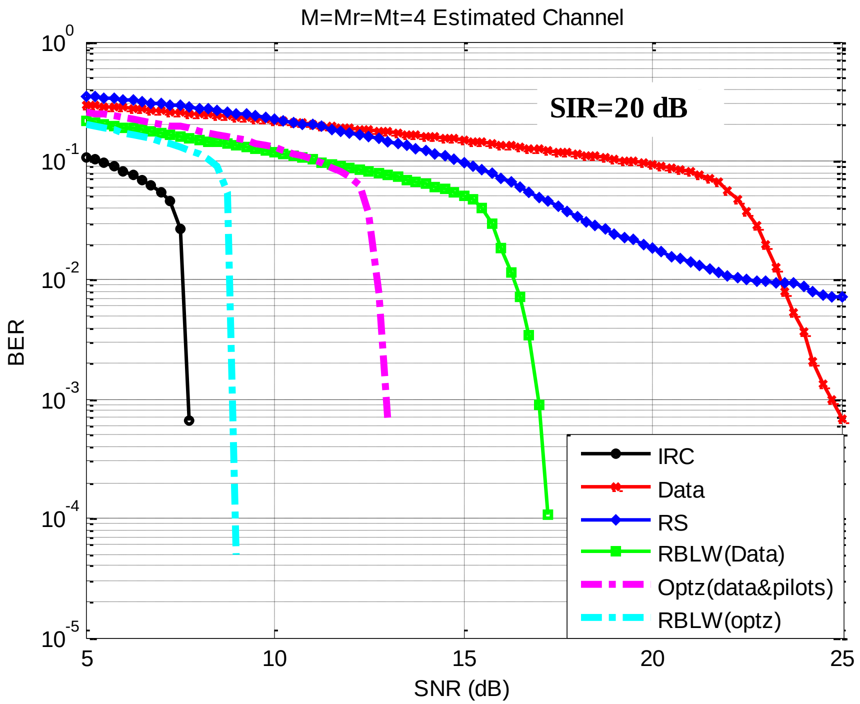

Figure 8. BER for Channel Estimate, T=10, SIR=20 dB, 4-QAM with coding

Figure 8 shows that at higher SIR (20 dB in this case), the pilots perform better than the data based method for SNR higher than 10 dB. But even better, is the performance of the RBLW and the optimization method. BERs of zero can be obtained by these two methods at SNRs of approximately 17 dB and 13 dB respectively, which is a huge improvement compared to the previous results at 10 dB of SIR.

By the optimization technique explained in Section 3.5, we obtain an optimal covariance matrix from the combination of the data based covariance matrix and the reference symbols based covariance matrix, which minimizes the MSE. This is because the data method and the reference symbols method vary in power due to the random nature of the channel, that is, at different SNR one of the two methods could have higher power than the other, as can be seen from the above figure. The optimized matrix is always allocated the highest power from the combination of the pilot and the data based covariance matrix which explains the good performance of the optimization method. Hence applying RBLW method to this well natured optimized covariance matrix boosts its performance as only a small number of samples has been considered.

Therefore high performance gains can be registered by these algorithms at high SIR, since the accuracy of the covariance matrix estimate increases with SIR. This further implies that the method of channel estimation by pilot boosting is not suited for interference limited systems.

4.3 Capacity Based Evaluation

In this section, results for the outage capacity of the various covariance matrix estimation methods, taking into account both the true channel and the estimated channel are presented. Furthermore the cumulative distribution function of the system capacity is also evaluated and presented in this section.

4.3.1 Outage Capacity

Outage capacity is an important measure especially in case where the transmitter lacks perfect channel state information, a case we assume, as we concentrate more on the receiver side (5). The *h%* outage capacity can be defined as the capacity which can be guaranteed for *h%* of all the channel realizations (9). The capacity evaluation is based on the 5% outage capacity, which is also considered as the cell-edge user throughput (14).

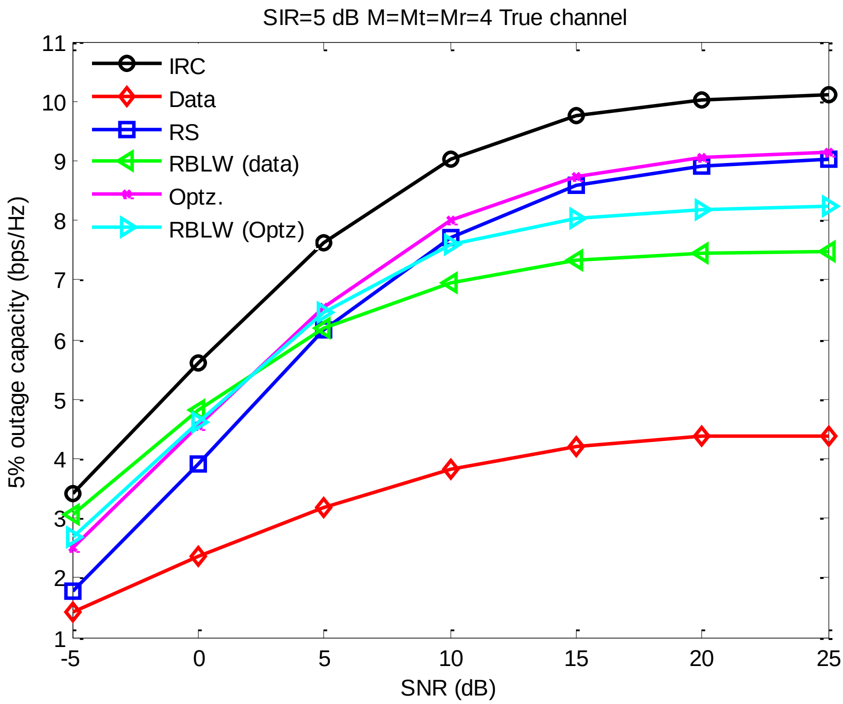

Figure 9. 5% outage capacity for 4-QAM, true channel

In Figure 9, above, the SIR is fixed to 5 dB, for data modulation scheme of 4-QAM, and the 5% outage capacity evaluated at various SNR for the true channel with a coherence interval of 10. As expected the outage capacity increases with SNR, till a point when it gets saturated, i.e., beyond this point, the capacity remains almost constant even with increasing SNR.

The outage capacity of the reference symbols based estimation method achieves more gain than the data based method. But by applying the RBLW algorithm, the outage capacity of the data based method can be improved, this is especially prevalent at low SNR (0-5 dB) where the RBLW method achieves a higher outage capacity than the reference symbols method as shown in Figure 9.

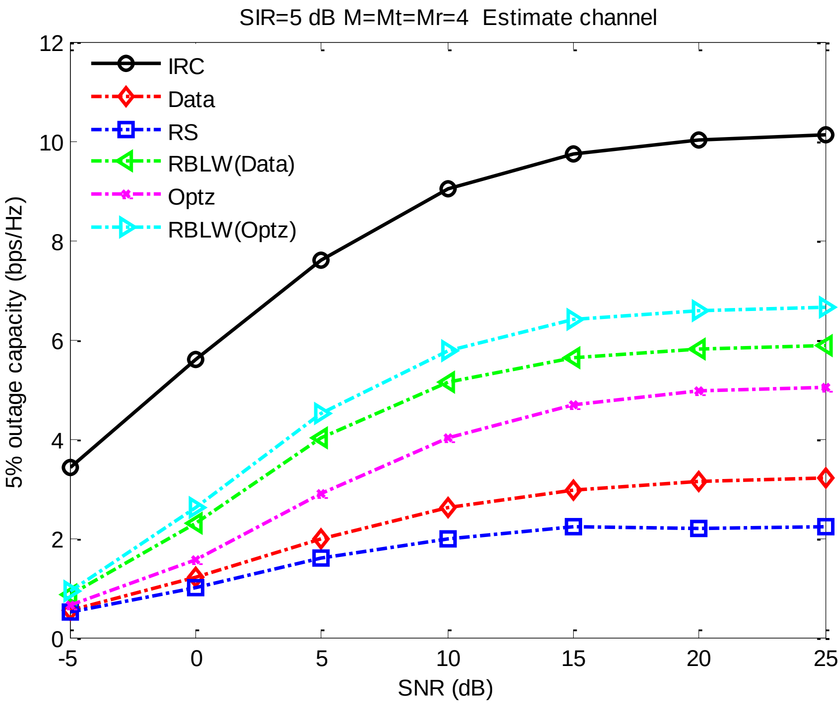

Figure 10. 5% Outage Capacity for 4-QAM, channel estimate

In Figure 10, we present the 5% outage capacity using the estimated channel matrix, for a coherence interval of 10, while SIR is 5 dB.

The reference symbols based covariance matrix estimation method performs the worst, because of a matrix mismatch between the true channel and the channel estimate as a result of the channel estimation error.
The data based method performs better because it offers a more accurate covariance matrix estimate since more samples (the entire block of 10 symbols) were used in the estimation. By using optimization between the data based method and the reference symbols based method, user throughput can be improved significantly.

However applying the RBLW algorithm to the data based method shows a significant gain in outage capacity, proving further that with the RBLW method, we are able to obtain a more accurate estimate of the covariance matrix, than relying just on the data based estimation method.

The results above shows that the system performance based on the estimated channel matrix is degraded for all covariance matrix estimation methods in comparison to the true channel based results. This is attributed to the poor channel estimation using pilot boosting in an interference limited system.

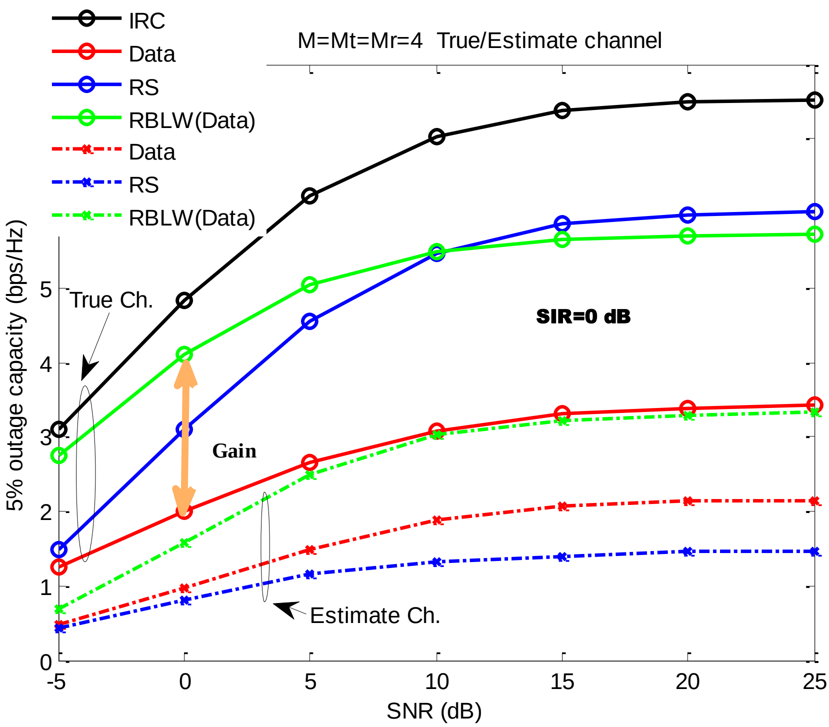

Figure 11. Comparison of the 5% Outage Capacity for the true channel and the channel estimate using 4-QAM

We also evaluate the 5% outage capacity performance of these covariance matrix estimation methods at 0 dB SIR, for 4-QAM, while maintaining a coherence interval of 10 symbols, the results of which are depicted in Figure 11 above with both the true channel and the channel estimate for comparison. As expected generally the 5% outage capacity is less than it was at 10 dB of SIR which was illustrated in Figure 10, increased interference and noise causes more error even with the true channel, hence the decrease in capacity.

However the RBLW method using the estimate channel performs very close to the data symbols covariance matrix estimation method with the true channel. Therefore in extreme interference scenarios, such as for users at the cell edge, a combination of the data symbols covariance matrix estimation and the RBLW method is a more viable option.

On the other hand, for the true channel matrix, the pilot based method outperforms the data based method, however at low SNR $\leq 10\ dB$, the combination of the RBLW method and the data based covariance matrix estimation method significantly outperforms the pilots based method. However at high SNR, pilots based method is able to accurately estimate the covariance matrix better than the RBLW method leading to a better outage capacity performance.

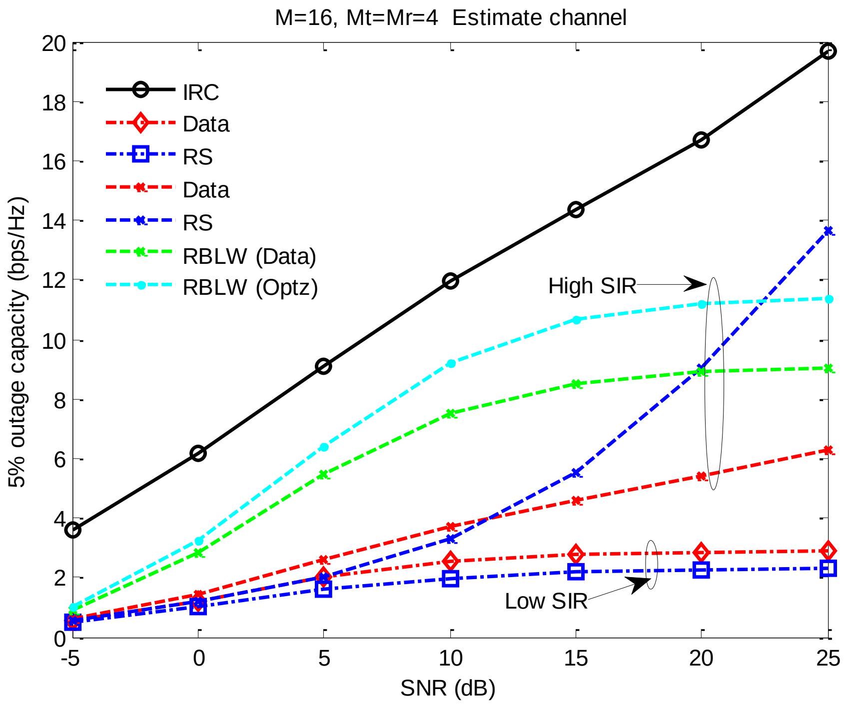

Figure 12. Comparison of 5% outage capacity at high/low SIR for the channel estimate using 16-QAM

Figure 12 shows the comparison in performance of the covariance matrix estimation methods at high and low SIR, for 16-QAM, with the estimated channel matrix.

At high SIR, it is observed that applying the RBLW method provides better performance in comparison to the reference symbols method until a specific SNR, but as the SNR increases beyond 20 dB the reference symbols perform best. In a low SNR region, even though large SIR is available, the performance curves for the data based and the RS based methods are almost the same. This is because of the low SINR, which leads to poor a channel estimate and the mismatch between the covariance matrix and the multiplied channel estimate is the same for both schemes. But as the SNR increases, so does the SINR, translating into a more accurate channel estimate. Hence the reference symbols achieve higher capacity than the data symbols based method owing to the reduced mismatch with a more accurate channel estimate, a case similar to the true channel simulation results.

Furthermore, since the covariance matrix includes the true channel matrix of the serving cell regardless of whether the true or the estimate channel matrix is used, the mismatch of the channel matrix between the covariance matrix and the multiplied channel matrix becomes more predominant when practical channel estimation is performed. This mismatch explains the deterioration in performance of the data symbols based covariance matrix estimation method, whereas there is no mismatch when using the reference symbols. This is because the mismatch is eliminated as shown in Equation 3.10 (14).

A slight increase in the outage capacity can observe using the 16-QAM instead 4-QAM and this can be seen more clearly in Tables 1 and 2.

4.3.2 Over all System Capacity

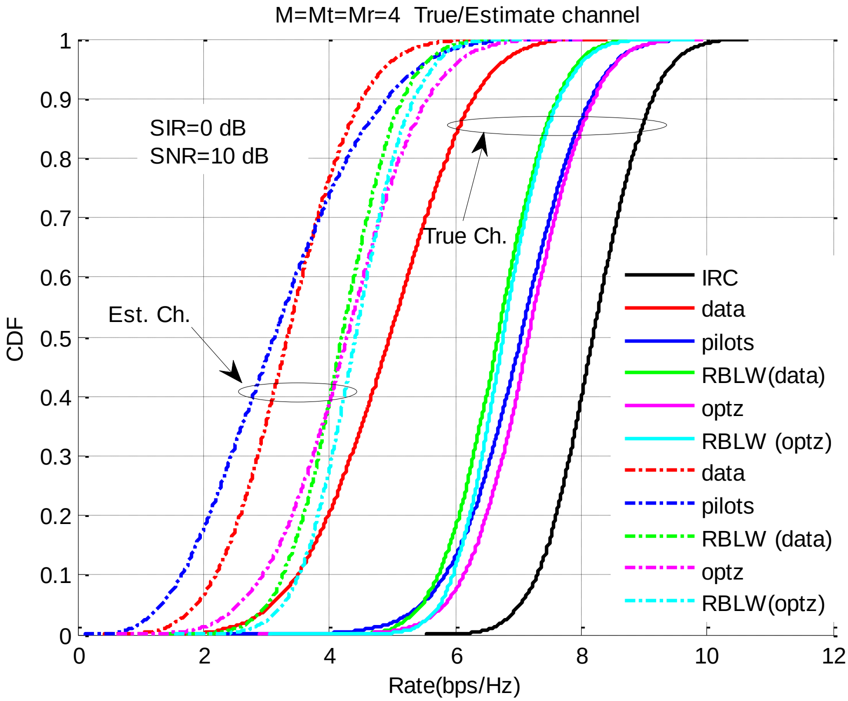

Figure 13. Comparison of the CDF of the overall user throughput for both the true channel and the channel estimate, using 4-QAM

In Figure 13 above, the SIR is fixed to 0 dB, for a channel coherence interval of 10, and the CDF of the overall user throughput is evaluated, for the various covariance matrix estimation methods, considering both the true channel and the channel estimate at a signal to noise ratio of 10 dB. The constellation used is 4-QAM.

As depicted, the pilots based method exhibits the lowest capacity with the channel estimate, up to nearly CDF = 70%. This poor performance is attributed to the high channel estimation error between the estimated channel matrix and the true channel matrix. On the other hand applying the RBLW method to the optimized covariance matrix achieves the highest throughput up to nearly CDF = 70%, beyond which the optimized covariance matrix method takes over.

Using the true channel, its observed that the data based covariance matrix estimation method achieves the lowest capacity, followed by the RBLW method, then the pilots based method, while the optimization between the data based covariance matrix and the reference symbols covariance matrix estimate the achieves the highest data rate. However in the lower part of the curve (which includes the 5% outage capacity), the gain from the RBLW method in comparison with the reference symbols method can be observed. The IRC is shown for comparison and is considered as the upper bound. The results still show an improvement in data rate when the RBLW method is applied to the data based covariance estimation method, proving further that the data based covariance method can in fact be improved using the RBLW method. The optimization between the data symbols and the reference symbols covariance matrix estimation performs even better than the reference symbols based estimation method.

4.3.2 Outage Capacity and Ergodic Capacity

Tables 1 and 2 below depict the summary of the 5% outage capacity and the ergodic (average) capacity for each of the methods investigated, using both signal constellations with respect to both the true channel and the estimated channel matrices respectively. The ergodic capacity of the channel is the expected value of the capacity taken over all the channel realizations (9).

Table 1. User Throughput (5% outage and average capacity) For the True Channel

SIR=5 dB SNR=10 dB	Signal Constellation	True Channel					
		IRC	Optz. (Data + Pilots)	RS	RBLW (Optz.)	RBLW (Data)	Data
5%outage capacity (bps/Hz)	4- QAM	7.02	5.84	5.45	5.76	5.47	3.08
	16 QAM	8.93	8.04	7.78	7.53	6.81	3.66
Average capacity (bps/Hz)	4- QAM	7.80	7.16	7.01	6.79	6.67	4.92
	16 QAM	10.52	9.74	9.61	8.64	8.44	5.98

As shown in the Tables, the 5% outage capacity, which is defined as the user throughput at the cell edge, can be improved using the RBLW method and the optimization method. The same is true for the average capacity.
Generally, more gain is achieved by using 16-QAM as compared to 4-QAM.

From the point of view of the estimated channel matrix, the RBLW and the optimization method still dominate the system performance whereas the reference symbols based method is degraded in performance, for reasons explained earlier.

Table 2. User Throughput (5% outage and average capacity) for the Estimated Channel

SIR=5 dB SNR=10 dB	Signal Constellation	Estimated Channel					
		IRC	RBLW (Optz)	RBLW (Data)	Optz. (Data + Pilots)	Data	RS
5%outage capacity (bps/Hz)	4- QAM	7.02	3.27	3.02	2.57	1.88	1.33
	16 QAM	8.93	5.74	5.10	4.05	2.54	1.97
Average capacity (bps/Hz)	4- QAM	7.80	4.42	4.21	4.27	3.35	3.21
	16 QAM	10.52	7.17	6.54	6.32	4.43	4.88

4.4 Performance of the RBLW method for different coherence intervals

Figure 14 below illustrates the performance of the RBLW with different number of data symbols taken into consideration. It is observed that when *T* is small, the RBLW improves the performance of the data based method significantly, but as the *T* increases, in other words, as the number of data samples considered increases, the RBLW does not make any improvement and it converges to the data symbols covariance estimation method in performance.

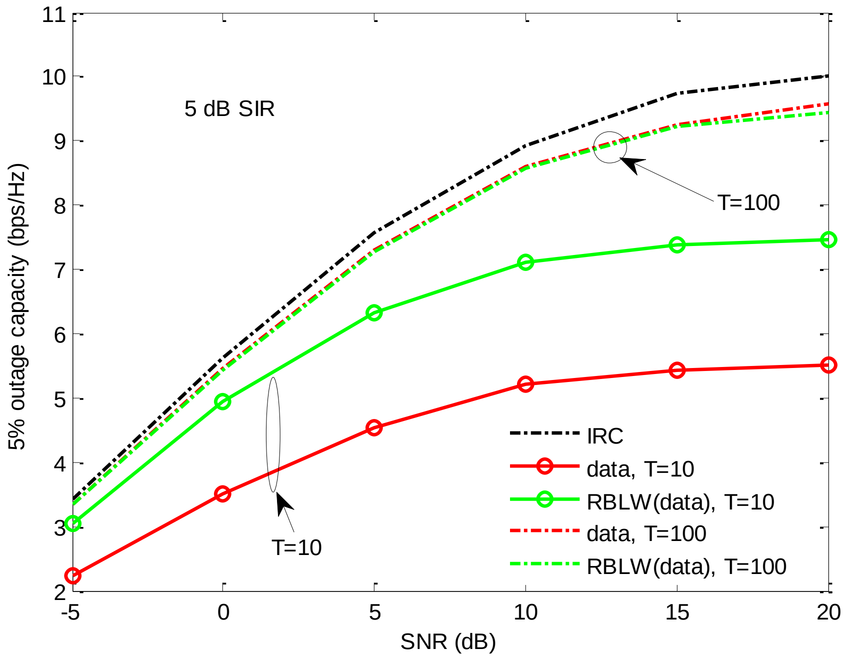

Figure 14. Performance of the RBLW method for different block lengths

This is because of the high variance of the covariance matrix of the data based estimation method due to few samples (i.e. T=10, in our case). The RBLW method aims to reduce this high variance by shrinking the covariance matrix towards the shrinkage target $\boldsymbol{F}$ (Equation 3.21), by using the shrinkage coefficient ρ that minimises the MSE. The RBLW method is hence well conditioned for small sample sizes.
As expected, the variance of the covariance matrix reduces with increasing number of samples, rendering the RBLW method rather unsuitable with large samples. This explains why for T =100, i.e. large number of samples under consideration, the RBLW method behaves the same as the unbiased data based covariance method. The reader is refered to (13) for more details.

Considering the estimated channel, a gain in performance can be achieved by applying the RBLW method over the optimized covariance matrix whereas for the true channel, this technique does not show much gain in performance.

4.5 Optimal Training Interval

One piece of information that could come in handy for system designers is knowledge about the optimal number of pilots to use, for various channel conditions and signal modulation schemes, see Section 3.3.1. In this section the results for the optimal number of pilots that can be used to achieve maximum capacity with 16-QAM and 4-QAM modulation schemes and a channel of coherence interval of 10 symbols are presented.

The maximum capacity is determined according to Equation 3.13, which depicts that the capacity increases with increasing number of pilots up until a certain point, where the capacity starts to decrease with increasing number of pilots. This relationship is further illustrated in Figure 15 below. In this particular case, we consider 4-QAM symbols while the SIR and the SNR are both fixed to 10 dB. Maximum capacity is achieved with 3 reference symbols. Therefore, for our system model under these mentioned channel parameters, 3 would be the optimal number of reference symbols.

Using the same procedure we estimate the optimal number of pilots that would be needed under various channel conditions (SIR and SNR), the results of which are presented in Tables 3 and 4.

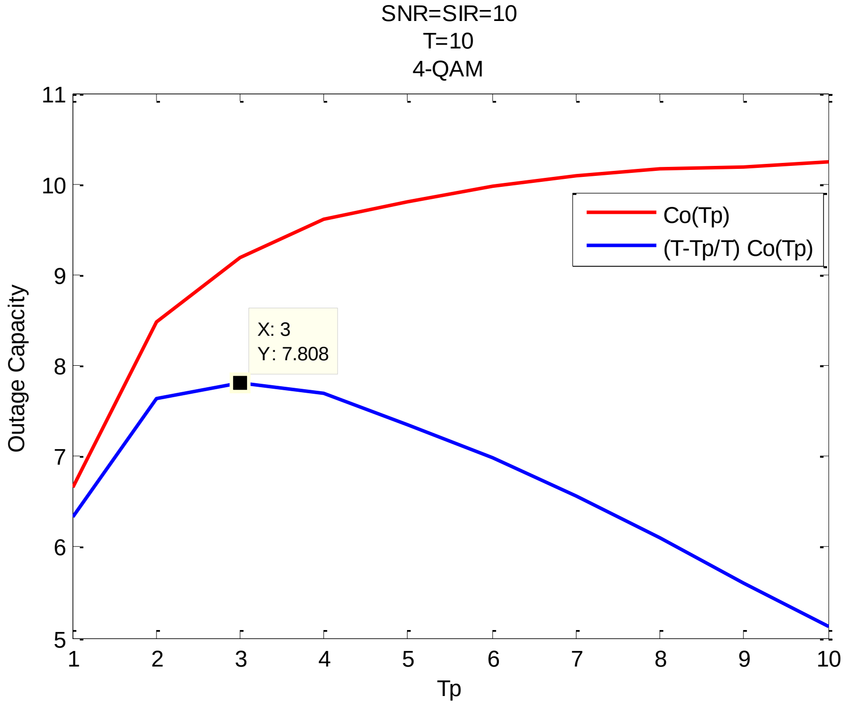

Figure 15. How to find the optimal number of pilots

Table 3. Optimal training interval for coherence time T =10 (in symbols) using 16-QAM

SNR / **SIR**	0 dB	5 dB	10 dB	15 dB	20 dB
0 dB	4	3	3	3	3
5 dB	3	3	2	2	2
10 dB	3	3	2	2	2
15 dB	3	3	2	2	1
20 dB	3	3	2	1	1

In Table 3 above, the coherence time of the channel is assumed to be 10, and the modulation type used is 16-QAM, for which the optimal number of pilots needed in the range of 0-20 dB for both SIR and SNR is evaluated.

As shown in Table 3, the optimal number of pilots decreases with increase in both SIR and SNR, and at high SIR and SNR, a single pilot is all that is needed. In fact one could argue that at high SIR and SNR, pilots are not necessary, since the channel conditions are already excellent hence sending more information carrying symbols is more appropriate and also leads to better bandwidth efficiency.

Table 4 below also depicts the optimal number of pilots needed for the same conditions as before, except in this case 4-QAM modulation is used. Clearly there is a difference between these two tables, as the 16-QAM based evaluation shows less number of pilots needed for similar channel conditions as compared to the 4-QAM based evaluation. Hence the optimal number of pilots not only depends on the channel state conditions, but also on the signal modulation type used.

It is also important to keep in mind that for this evaluation, pilot boosting was not used, but rather pilots which have the same energy as the data carrying symbols. The information presented in Tables 3 and 4 can be used to build devices that can adapt to the channel state to vary the number of pilots thereby ensuring effective spectral efficiency.

Table 4. Optimal training interval for coherence time T =10 using 4-QAM

SNR / SIR	0 dB	5 dB	10 dB	15 dB	20 dB
0 dB	6	5	5	5	5
5 dB	5	4	4	4	4
10 dB	4	4	3	2	2
15 dB	4	4	3	2	2
20 dB	4	4	3	2	1

CHAPTER 5

5 Conclusions

In this piece of work we investigate some of the methods that can be employed to deal with interference rejection suppression in MIMO systems. These methods all have one thing on common, their performance is very much dependent on the interference covariance matrix estimation, of course the more accurate the matrix estimate, the better the performance. Some of the methods studied include, the reference symbols covariance estimation method, the data symbols covariance matrix estimation, RBLW method plus the data covariance matrix estimation method, optimization of data covariance estimation method and the reference symbols covariance estimation method, and applying RBLW method to the optimized matrix. The performance of the IRC receiver has also been evaluated and shown for comparison, and is considered as the upper bound in this thesis work.

From the BER point of view, assuming ideal channel estimation (true channel) the data covariance matrix estimation method incurs the highest error rate, while the reference symbols covariance matrix estimation method incurs the least error rate. The RBLW method leads to improved BER performance of the data symbols covariance estimation method as a result of a more accurate estimate of the covariance matrix, though it still performs worse than the reference symbols based estimation method.
However using the practical channel estimate obtained using pilot boosting, the pilots encounter a higher BER than the data symbols estimation method, while the RBLW still leads to improved performance of the data based method, especially at low SNR. At higher SNR, optimization between the data covariance matrix and the reference symbols based covariance matrix method outperforms the RBLW method. It is also observed that the data based covariance method encounters less BER with increasing block size.

With LDPC codes, the BER of all the covariance matrix estimation methods was improved.

From the perspective of the outage capacity, assuming the true channel, the reference symbols based estimation method offers the highest outage capacity, followed by the RBLW method, while the data based method has the lowest outage capacity. However with the channel estimate, the reference symbols based method achieves the lowest outage capacity, followed by the optimization method (between the reference symbols and the data based covariance estimate methods), the data based method and the RBLW achieves the highest outage capacity.

However in terms of general capacity, with increasing SIR, the pilot based method outperforms the data based method, though the RBLW method still performs better. The RBLW method also exhibits excellent performance with small coherence intervals but as the number of data symbols considered increases, the RBLW method shows little or no improvement at all to the data based estimation method.

Therefore at low SIR, it is better to use the received data symbols for channel estimation other than using the reference symbols. The data based estimation method can further be improved by using RBLW method.
We also investigate how much training is needed and how it is affected by SIR, SNR, and signal constellation. The results show that higher order modulation schemes, such as 16-QAM, generally require less training in comparison to lower order modulation schemes, such as 4-QAM. This is something we are yet to understand, since it higher order modulation schemes have higher bit rates and hence require accurate However with better channel conditions, less training is required to achieve maximize capacity.

CHAPTER 6

6 Future Work

From the results it can be observed that applying the RBLW method to perfect channel estimation does more harm than good, contrary to its effect on the not so perfect channel estimate. Unfortunately due to time constraints we were unable to deeply investigate the reason for this ambiguous behavior. Therefore we propose a further investigation into the nature of conditions under which the RBLW thrives and why it is unsuitable for the perfect channel estimation.

Also as can be observed from the results, in the low SINR region, for instance, the RBLW method outperforms the reference symbols method. It would therefore be interesting to build a system that switches between the different covariance matrix estimation methods, depending on the performance and channel state conditions.

Bibliography

1. **Vega, Luis Felipe Del Carpio.** *System Level Modeling and Evaluation of Advanced Interference aware Receivers.* s.l. : Aalto University, School of Electrical Engineering, 2012.
2. **Reed, J. D. Laster& J. H.** Interfernece Rejection in Digital Wireless Communications. *IEEE Siganal Processing Magazine.* 1997.
3. **Lindell, Göran.** *Introduction to Digital Communications.* Lund : s.n., 2006.
4. **Goldsmith, Andrea.** *Wireless Communications.* Cambridge : Cambridge University Press, 2005.
5. **Arogyaswami Paulraj, Rohit Nabar, Dhananjay Gore.** *Introduction to space-time Wireless Communications.* Cambridge : Cambridge University Press, 2003.
6. **Mohider, Jankiraman.** *Space-Time Codes and MIMO Systems.* Boston : Artech House, 2004.
7. **Kjetil Fagervik, Arne Sjøthun Larssen.** *Performance and Complexity Ccomparison of Low Density Check Codes and Turbo Codes.* s.l. : Schlumberger/Stavanger University College.
8. **Ralf Koetter, Andrew C. Singer,and Micheal Tuchler.** Turbo Equalization. *IEEE Signal Processing Magazine.* January 2004, pp. 67-80.
9. **Molisch, Andreas F.** *Wireless communications.* 2nd . West Sussex : Wiley, 2011.
10. *Achievable rates of ID Gaussian symbols on the non-coherent block-fading channel without channel distribution knowledge at the receiver.* **Rusek, Fredrik.** 2012.
11. **Chen, Yilun.** *Regularized Estimation of High-dimensional Covariance Matrices.* s.l. : (Electrical Engineering Systems) in The University of Michigan, 2011.
12. **L. Hanzo, M. Munster, B.J. Choi and T. Keller.** *OFDM and MC-CDMA.* s.l. : Wiley.

13. *Shrinkage Algorithms for MMSE Covariance Estimation.* **Yilun Chen, Ami Wiesel,Yonina C.Eldar, Alfred O. Hero.** 2010, Vol. 58, pp. 5016-5029.
14. **Yusuke Ohwatari, Nobuhiko Miki,Takahiro Asai,Tetsushi Abe,and Hidekazu Taoka.** Performance of Advanced Receiver Employing Interference Rejection. Munich : s.n., 2011.
15. *On The Capacity Of Multi-Antenna Gaussian Channels.* **Nicolae Chiurtu, Bixio Rimoldi and Emre Telatar.** Washington, DC : s.n., 2001.
16. **Babak Hassibi, Bertrand M. Hochwald.** *How Much Training is Needed in Multiple-Antenna Wireless Links?* s.l. : Bell Laboratories, Lucent technologies, 2010.
17. **Trees, H.L Van.** *Detection, Estimation and Modelling Theory.* Part 1. New York : Wiley, 1971.
18. *A well-conditioned estimatorfor large-dimensional covariance matrices* . **O. Ledoit, M. Wolf.** 2004, Journal of Multivariate Analysis 88, pp. 365-411.
19. http://www.radio-electronics.com/info/rf-technology-design/pm-phase-modulation/what-is-qam-quadrature-amplitude-modulation-tutorial.php. [Online]
20. **Xiaoyi Tang, Mohamed-Slim Alouini, Member, IEEE, and Andrea J. Goldsmith.** Effect of Channel Estimation Error on M-QAM. *IEEE TRANSACTIONS ON COMMUNICATIONS.* December 12, 1999, pp. 1856-1864.

List of Figures

List of Tables

Appendix

A. Derivation of scaling factors

From

$$y = kHx + lSz + n \tag{A.1}$$

$$SIR \triangleq \frac{E\|kHx\|^2}{E\|lSz\|^2} = \frac{k^2 \sigma_x^2 \, E\left[Tr\,(HH^H)\right]}{l^2 \, \sigma_z^2 \, E[Tr(SS^H)]} \tag{A.2}$$

But

$$E[TrHH^H] = E\left\{\sum_{i=1}^{N} \sum_{j=1}^{N} \left|h_{ij}\right|^2\right\} \tag{A.3}$$

$$N^2 E\left\{\left|h_{ij}\right|^2\right\} = N^2 \tag{A.4}$$

Since the channel is normalized to have unit energy (the same applies for the interference channel)

$$E\{Tr(SS^H)\} = N^2 \tag{A.5}$$

Substituting for A.4 and A.5 in A.2, we obtain

$$SIR = \frac{k^2 \, \sigma_x^2 \, N^2}{l^2 \, \sigma_z^2 \, N^2} = \frac{k^2}{l^2} \tag{A.6}$$

The SNR is also computed as

$$SNR = \frac{k^2 \, \sigma_x^2 \, N^2}{E\{nn^*\}} = \frac{k^2 \, \sigma_x^2 \, N^2}{N_o N} \tag{A.7}$$

However the noise is also normalized to have unit energy

Therefore

$$SNR = k^2 \, \sigma_x^2 \, N \tag{A.8}$$

For $\sigma_x^2 = \sigma_z^2 = 1$

Equation (A.8) reduces further to

$$SNR = k^2 N \tag{A.9}$$

A.9 can be expressed in dB as

$$SNR_{|dB} = 10 \log (k^2 N)$$

$$10^{\frac{SNR_{|dB}}{10}} = k^2 N$$

Hence

$$k = \sqrt{\frac{1}{N} 10^{\frac{SNR_{|dB}}{10}}} \tag{A.10}$$

A.6, can also be expressed in dB as

$$SIR_{|dB} = 10 \log\left(\frac{k^2}{l^2}\right) = 20 \log(k) - 20 \log(l)$$

$$\log(l) = \frac{1}{20} \left(20 \log(k) - SIR_{|dB}\right) = \log(k) - \frac{SIR_{|dB}}{20} \tag{A.11}$$

$$l = 10^{\left(\log(k) - \frac{SIR_{|dB}}{20}\right)}$$

Therefore

$$l = k\, 10^{\left(-\frac{SIR_{|dB}}{20}\right)} \tag{A.12}$$