



CHALMERS
UNIVERSITY OF TECHNOLOGY

Compression analysis of massive MIMO uplink

Master's thesis in Communication Engineering

BOJAN DRVENICA
GUILHERME LUZ

REPORT NO. 2016:NN

Compression analysis of massive MIMO uplink

BOJAN DRVENICA
GUILHERME LUZ



CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Signals and Systems
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2016

Compression analysis of massive MIMO uplink
BOJAN DRVENICA
GUILHERME LUZ

© BOJAN DRVENICA, GUILHERME LUZ, 2016.

Technical report no 2016:NN
Department of Signals and Systems
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Typeset in L^AT_EX
Gothenburg, Sweden 2016

Abstract

Massive MIMO is a crucial technology that will drive the data capacity in wireless communication systems by implementing a large number of antenna elements. This new type of system creates new challenges that have to be solved. One of them is the considerable energy consumption that arises from the RF-chain that will consist of potentially more than one hundred parallel channels. If the Analog-to-Digital Converters (ADCs) will need to work at very high sample rates then there will be an enormous amount of data that will be quantised and transported at a very high speed from the many antennas to the CPU.

The main objective with this thesis is to reduce the number of data streams from the RF-chains to the CPU. This problem was approached by first experimenting with a lossy compression algorithm that apply Fast Fourier Transform (FFT) and Discrete Cosine Transform (DCT) transforms to the received signals and then it discards low power frequency coefficients. Then a more effective model was implemented that can fully recover the compressed signal with the pseudo inverse. One of the challenges of this system is that the antenna array has to be divided into many groups and the processing has to be applied separately to each of them.

The compression algorithm takes the received signal as an input after the ADC blocks and then it is multiplied by the FFT matrix. The next step is to compress the data by selecting a certain ratio out of the output coefficients after the FFT block that will be sent to the CPU side. In total four selection strategies are explored in the thesis.

The performance of the system when compression is applied is evaluated in terms of performance loss of the recovered symbols. The selection strategies are analysed by the energy efficiency and the condition number.

The system was tested for different amount of users, antennas, compression level and group sizes and the results show that the performance loss gets reduced for higher energy efficiency and lower condition number.

Acknowledgements

We would like to thank our examiner, Thomas Eriksson, that proposed a very interesting thesis to us that constitutes a young research area within wireless communications. Throughout the entire project he has helped us a lot with understanding the technical aspects of this project which we are very grateful for.

Also we would like to thank our supervisor, Ulf Gustavsson, that contributed a lot to the discussion surrounding many of the problems that we encountered, especially the hardware aspects.

Bojan Drvenica, Guilherme Luz, Gothenburg, May 2016

Acronym List

4G Fourth Generation.

5G Fifth Generation.

ADC Analog-to-Digital Converter.

AoA Angle-of-Arrival.

AWGN Additive White Gaussian Noise.

BS Base Station.

CDMA Code Division Multiple Access.

CPU Central Processing Unit.

DCT Discrete Cosine Transform.

DL Downlink.

FFT Fast Fourier Transform.

IDCT Inverse Discrete Cosine Transform.

IFFT Inverse Fast Fourier Transform.

IID Independent and Identically Distributed.

IoT Internet of Things.

LoS Line-of-Sight.

LTE Long Term Evolution.

MIMO Multiple-Input Multiple-Output.

MSE Mean Squared Error.

MTC Machine-Type-Communication.

PCB Printed Circuit Board.

QAM Quadrature Amplitude Modulation.

RF Radio Frequency.

SIMO Single-Input Multiple-Output.

SVD Singular Value Decomposition.

UE User Equipment.

UL Uplink.

Wi-Fi Wireless Fidelity.

ZF Zero-Forcing.

Symbol List

\mathbb{C} the field of complex numbers.

\mathbb{Q} the field of rational numbers.

a the path gain.

C the compression level.

c the speed of light.

d_{mk} the distance between the m -th receive antenna and the k -th User Equipment (UE).

η the energy efficiency.

f_c the carrier frequency.

G the group size.

\mathbf{H} the channel matrix.

$\tilde{\kappa}$ the condition number.

K the number of UE.

L the number of groups.

λ the wavelength.

M the number of antennas at the Base Station (BS).

P_{tot} the total power.

\mathbf{T} the Transform Matrix.

θ the angle-of-arrival.

T the number of time instances.

t a time instance.

t_s the time symbol.

V a vector space.

\mathbf{w} the noise vector.

W the transmission bandwidth.

\mathbf{x} the transmitted symbols.

\mathbf{y} the received signal vector by the BS.

Contents

List of Figures	xv
List of Tables	xvii
1 Introduction	1
1.1 Background and motivation	1
1.1.1 Massive MIMO	1
1.1.2 Digital backbone of massive MIMO	2
1.2 Aim and outline	2
1.3 Previous work	3
2 Massive MIMO	5
2.1 Deterministic MIMO channel	5
2.1.1 Spatial multiplexing	6
2.1.2 Channel modelling	7
2.1.3 Receiver	8
3 Transform Coding	9
3.1 Introduction	9
3.2 Transforms	10
3.2.1 Fast Fourier Transform	10
3.2.2 Discrete Cosine Transform	11
4 Compression Techniques	13
4.1 Correlated channels in massive MIMO	13
4.2 Lossy compression algorithm	14
4.3 Lossless compression algorithm	16
4.3.1 Implementation of lossless compression algorithm	16
4.3.2 Compression matrix selection	17
4.3.2.1 Power Selection algorithm	18
4.3.2.2 Random Selection algorithm	19
4.3.2.3 Correlation Selection algorithm	20
4.3.2.4 Hybrid Selection algorithm	22
5 Simulation Results	23
5.1 Massive MIMO model	23
5.2 Metrics	23

5.2.1	Energy efficiency	23
5.2.2	Mean squared error	24
5.2.3	Condition number	25
5.3	Lossy compression of massive MIMO uplink	26
5.3.1	One UE	26
5.3.2	Multiple UEs with varying group size	27
5.4	Lossless compression of massive MIMO uplink	31
5.4.1	Group size analysis	31
5.4.2	Number of UEs analysis	33
5.4.3	Number of antennas analysis	34
5.4.4	Compression level analysis	35
5.4.5	Hybrid-25 and Hybrid-75 algorithm	36
6	Discussion	39
7	Conclusions	43
	Bibliography	45
A	Appendix 1	I

List of Figures

2.1	M BS antennas and K UEs are placed in a two dimensional Cartesian coordinate system.	5
4.1	Single-Input Multiple-Output (SIMO) model with user placed at an AoA of θ	13
4.2	The lossy compression model with L number of groups. Each block takes four input elements of the received signal \mathbf{y} . The \mathbf{T}_l blocks implement a compression level of $C = \frac{2}{4}$ and thus reduce the number of streams to two that are sent to the Central Processing Unit (CPU) side for decompression. Note that the inverse transformation block \mathbf{T}_l^{-1} , adds zeros to the corresponding elements that were discarded (not transmitted).	15
4.3	The lossless compression model with L number of groups. Each block takes four input elements of the received signal \mathbf{y} and implements $C = \frac{2}{4}$, reducing the number of streams to two that are sent to the CPU side for decompression.	18
5.1	Energy efficiency of the compressed signal for different Angle-of-Arrival (AoA) with $C = 0.5$ and 128 antennas. Left shows results for FFT compression and the right shows for DCT compression.	26
5.2	Energy efficiency of the compressed signal for increasing number of UEs with compression level $C = 0.5$ and 128 antennas. Left shows results for FFT compression and the right shows for DCT compression	27
5.3	Average power for DCT and FFT coefficients for antenna group size 128 and 8 with $K = 10$ UEs. Notice that 128 group size has many coefficients close to zero whereas 8 group size has a more uniform shape.	28
5.4	Average power for DCT coefficients for antenna group size 8 for 1, 2, 5 and 10 UEs. More UEs yield a higher spread of the power in the DCT coefficients.	29
5.5	Average power for FFT coefficients for antenna group size 8 for 1,2,5 and 10 UEs. More UEs yield a higher spread of the power in the FFT coefficients.	30
5.6	Δ Mean Squared Error (MSE) versus group size for different selection methods for 128 antennas. The number of UEs is $K=10$ at $C = 0.5$. The red curve is the upper bound with 64 antennas. The group sizes are $G = 2, 4, 8, 16$ and 32	32

5.7	Energy efficiency η versus group size is plotted for different selection methods. Number of UEs is $K=10$ and is $C = 0.5$ and the number of antennas is 128. The power selection algorithm has clearly the best energy efficiency compared to the others. Hybrid selection yields a gain in energy efficiency compared to random and correlation selection algorithms.	33
5.8	Δ MSE vs number of UEs for 128 antennas at $C = 0.5$. Power selection tends to have higher Δ MSE for small group sizes and the opposite holds for the other selection methods. The Δ MSE with hybrid selection has least dependence on group size.	34
5.9	Δ MSE vs number of antennas $100 \leq N \leq 500$ at $C = 0.5$ and $K = 10$ UE. Power selection has the highest dependence on group size. In subfigures b-c only groups of 2 and 32 are plotted since the errors are converging with high number of antennas. With hybrid selection (d) all of the group sizes give errors below the upper bound with the largest group size having the smallest error.	35
5.10	Δ MSE vs C for 128 antennas with 10 UEs. Power selection tends to have higher Δ MSE for small group sizes, however when the compression level is close to one, the Δ MSE of power selection is better than the other types of selection.	36
5.11	η vs the compression level with 128 antennas and 10 UEs. Power selection has the capacity to keep a higher percentage of the energy in the original signal. On the other side, the random and correlation selection have a percentage of energy similar to the compression level.	36
5.12	Δ MSE with respect to the number of UE for different hybrid algorithms. The subfigure on the left shows the case when only 25% of the groups are selected with regard to power and the right subfigure shows the case when 75% of the groups are selected according to power.	37

List of Tables

4.1	List of steps of power selection algorithm	19
4.2	Sequence of the chosen rows with $G = 2$ and $C=0.5$	20
4.3	Sequence of the chosen rows with $G = 4$ and $C = 0.5$	20
4.4	Sequence of the chosen rows with $G = 8$	21
4.5	Sequence of the chosen rows with $G = 16$	21
4.6	Sequence of the chosen rows with $G = 32$	21
4.7	List of steps of power selection algorithm	22
5.1	Average condition number $\tilde{\kappa}$ for different selection algorithms and group sizes of $G = 2, 4, 8, 16$ and 32	32
A.1	Combinations for $G=4$ and $C=0.25$	I
A.2	Combinations for $G=4$ and $C=0.75$	I
A.3	Combinations for $G=16$ and $C=0.125$	I
A.4	Combinations for $G=16$ and $C=0.25$	II
A.5	Combinations for $G=16$ and $C=0.75$	II
A.6	Combinations for $G=16$ and $C=0.875$	II

1

Introduction

1.1 Background and motivation

It is estimated that the wireless data traffic will increase by a thousand fold over the next decade. This will be driven by higher user demand and the estimated 50 billion connected devices [1].

Today most of the data traffic is generated by smartphones, tablets and video streaming, but even the applications that are associated with these devices will be used more progressively which will add to the wireless data volume in the future network such as e-banking, e-health and on demand entertainment. These human centric devices will be largely complemented by a massive amount of communicating machines forming the Internet of Things (IoT) which will widen the range of wireless applications to include the automotive industry, security, health care, among others [1, 2].

The fulfilment of these new demands will require adequate technological advancements that will create a paradigm shift in comparison to the existing Fourth Generation (4G) system. Some of the key technologies will be to move to mm-wave frequencies, ultra-densification and massive Multiple-Input Multiple-Output (MIMO) that will be part of the new Fifth Generation (5G) standard that is expected to be rolled out around 2020 [2].

It is expected that 5G will have to serve a data volume that is 1000 times higher compared to 4G, partly due to the IoT consisting of an enormous amount of Machine-Type-Communication (MTC) low rate devices such as actuators and sensors. The mobile devices will have data rates down to 100 Mbps at the edge of the network which is 100 times higher than for 4G. The latency in 5G will be reduced in the order of one magnitude (to 1 ms or less) compared to 4G. 5G is also prospected to provide ultra-high service reliability including a guaranteed availability with a minimum success packet delivery of 99.9999%. Furthermore, the system should be as energy-efficient as possible, including devices, the base stations and the whole communications involved between stations and data centres which will be handled by the telecom operators, in order to achieve low operations costs and energy sustainability [2, 3].

1.1.1 Massive MIMO

Massive MIMO systems have an excessive number of antennas at the BS compared to the number of active UEs and can thus improve the spectral efficiency enormously compared with current Long Term Evolution (LTE) standard that supports up to

four antennas. This is accomplished by exploiting the spatial domain which enables all of the UEs to use the same time-frequency slot under the condition that accurate channel estimation can be done and that the channel is sufficiently well conditioned [2, 9].

Massive MIMO has the potential for a multitude of improvements such as better link reliability, coverage and energy efficiency. However, one of the challenging aspects is channel estimation in the Downlink (DL). Especially important is the potential to reduce the latency thanks to the possibility to use beamforming in massive MIMO in order to overcome fading dips [1, 4].

Another positive impact of Massive MIMO is that it can be built on less expensive components since the hardware, noise and fading deficits can be averaged out due to the combined effect of the signals from many antennas [2]. However, there are also new challenges that need to be considered that are mainly based on the energy consumption and hardware complexity of building a system of this scale [5].

1.1.2 Digital backbone of massive MIMO

5G will utilise more antennas and larger bandwidths over communication links which will significantly increase the amount of internal data traffic that has to be quantised, transported and sent over the digital backbone to the central node.

Thus, a modern communication link might consist of hundreds of parallel RF-chains with ADCs that operate in the Gsample/s region which produce several terabits of data that have to be transported to the central node. This will shift the balance of cost and power consumption in communication networks from power amplifiers, that consume most of the power today, towards the digital backbone [5].

1.2 Aim and outline

This thesis is focused on the digital backbone which transports the bits from the RF-chains to the CPU. By principle, each antenna generates one data stream which in total becomes power consuming. Then, the idea is to reduce this number of data streams such that the loss of quality of the signal is minimised. The ideal case would occur when the number of streams can be reduced without any loss of quality caused by noise and/or interference.

Specifically, it will be analysed in the Uplink (UL) channel, from UE to BS. This will be done by dividing the antennas into groups and then apply transform coding with FFT and DCT. Two models will be studied, one where lossy compression is applied and one where lossless compression is applied.

The purpose of the lossy compression method is to study the problem of compression and try to assess the most effective way to handle the problem. The next part is to study the lossless compression method which is more efficient and the aim is to minimise the loss of the signal quality.

The system performance will be evaluated in terms of MSE where it is desired to reduce this number as much as possible.

1.3 Previous work

There are a few studies about ways to reduce the number of antennas minimising the loss of performance. The techniques are based on selection of antennas or channels according the received power at the antennas or the channel gain between the transmit and receive antennas. It can be done by adding a simple switch to the system, where the antennas can be turned off and turned on, but it can be done as well, by doing a combination with different weights of different antennas [6, 7]. Another technique is based on grouping a set of highly correlated antennas [8].

2

Massive MIMO

2.1 Deterministic MIMO channel

The MIMO channel has multiple receive and transmit antennas, the receive antennas are at the BS and the transmit antennas are referred as UEs. The difference between conventional MIMO and massive MIMO is that massive MIMO uses an excessive number of antennas at the BS compared to the number of active UEs.

If there is more than one active UE it will constitute a multiuser channel that can have both an uplink channel and a downlink channel. Figure 2.1 illustrates a MIMO channel.

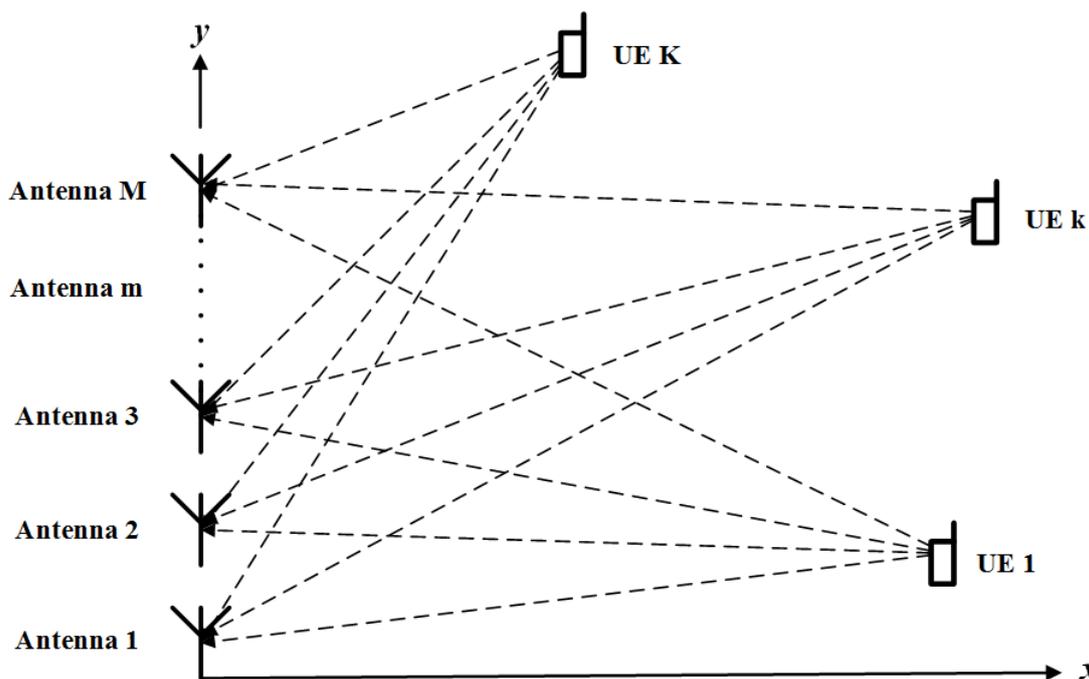


Figure 2.1: M BS antennas and K UEs are placed in a two dimensional Cartesian coordinate system.

The time-invariant channel at a symbol time can be expressed as:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w} = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1K} \\ h_{21} & h_{22} & \cdots & h_{2K} \\ \vdots & \vdots & & \vdots \\ h_{M1} & h_{M2} & \cdots & h_{MK} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_K \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_M \end{bmatrix} \quad (2.1)$$

where $\mathbf{x} \in \mathbb{C}^{K \times 1}$ is the transmitted signal, $\mathbf{y} \in \mathbb{C}^{M \times 1}$ is the received signal and $\mathbf{w} \sim \mathcal{CN}(0, N_0 \mathbf{I}_M)$ is Independent and Identically Distributed (IID) Additive White Gaussian Noise (AWGN) with power spectral density N_0 .

The channel $\mathbf{H} \in \mathbb{C}^{M \times K}$ is assumed to be deterministic and known at the receiver where h_{mk} is the channel gain from m -th receive antenna to the k -th UE. The transmitted signal \mathbf{x} corresponds to the symbols that are sent at time instances $t \in 0, \dots, (T-1)t_s$, where T is the number of time instances and t_s is the symbol time which can be expressed as the inverse of the transmission bandwidth W :

$$t_s = \frac{1}{W} \quad (2.2)$$

2.1.1 Spatial multiplexing

Multiple transmit and receive antennas form a MIMO channel that has the ability to increase the capacity due to the added *degrees-of-freedom* that arises in the spatial dimension. This is done by spatially *multiplexing* several data streams onto the MIMO channel [9].

The channel matrix \mathbf{H} can be divided into several independent sub-channels with the Singular Value Decomposition (SVD):

$$\mathbf{H} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^* \quad (2.3)$$

where $\mathbf{U} \in \mathbb{C}^{M \times M}$ and $\mathbf{V} \in \mathbb{C}^{K \times K}$ are unitary matrices whereas $\mathbf{\Lambda} \in \mathbb{C}^{M \times K}$ is a diagonal matrix [9]. The diagonal elements $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{n_{min}}$ are the ordered *singular values* of \mathbf{H} where $n_{min} = \min(M, K)$. The number of nonzero singular values correspond to the *rank* of \mathbf{H} . Since

$$\mathbf{H}\mathbf{H}^* = \mathbf{U}\mathbf{\Lambda}\mathbf{\Lambda}^t\mathbf{U}^* \quad (2.4)$$

the squared singular values λ_i^2 are the *eigenvalues* of the matrix $\mathbf{H}\mathbf{H}^*$. The SVD can be rewritten as:

$$\mathbf{H} = \sum_{i=1}^{n_{min}} \lambda_i \mathbf{u}_i \mathbf{v}_i^* \quad (2.5)$$

which is the sum of rank-one matrices. A parallel decomposition can be done with transmit precoding and receiver shaping [9]:

$$\tilde{\mathbf{x}} = \mathbf{V}^* \mathbf{x} \quad (2.6)$$

$$\tilde{\mathbf{y}} = \mathbf{U}^* \mathbf{y} \quad (2.7)$$

$$\tilde{\mathbf{w}} = \mathbf{U}^* \mathbf{w} \quad (2.8)$$

so that the channel in eq.(2.1) can be rewritten as:

$$\tilde{\mathbf{y}} = \Lambda \tilde{\mathbf{x}} + \tilde{\mathbf{w}} \quad (2.9)$$

where $\tilde{\mathbf{w}}$ has the same distribution as \mathbf{w} . The energy is thus preserved and the parallel decomposed channel can be expressed as:

$$\tilde{y}_i = \lambda_i \tilde{x}_i + \tilde{w}_i \quad (2.10)$$

Now, the capacity of the time-invariant MIMO channel can be expressed as:

$$C = \sum_{i=1}^{n_{min}} \log\left(1 + \frac{P_i^* \lambda_i^2}{N_0}\right) \text{ bits/s/Hz} \quad (2.11)$$

where P_i^* are the waterfilling power allocations:

$$P_i^* = \left(\mu - \frac{N_0}{\lambda_i^2}\right)^+ \quad (2.12)$$

where μ is chosen to satisfy the total power constraint. Every singular value λ_i corresponds to a *eigenmode* of the channel and can support one data stream, since the MIMO channel can have several eigenmodes it can enable spatial multiplexing of several data streams.

The number of non-zero singular values is the number of spatial degrees-of-freedom, with full rank the MIMO channel will have n_{min} spatial degrees-of-freedom [9].

2.1.2 Channel modelling

For a line-of-sight channel there are no obstructions between each antenna pair, thus forming a direct signal path. The continuous time response $h_{mk}(\tau)$ from k -th UE to the m -th receive antenna is given by [9]:

$$h_{mk}(\tau) = a \delta(\tau - d_{mk}/c), \quad m = 1, \dots, M \quad \text{and} \quad k = 1, \dots, K \quad (2.13)$$

where d_{mk} is the distance between the k -th transmit antenna and the m -th receive antenna, c is speed of light and a is the attenuation of the path. Considering $d_{mk}/c \ll 1/W$, where W is the transmission bandwidth, the baseband Line-of-Sight (LoS) channel gain is given by [9]:

$$h_{mk} = a \exp\left(\frac{-j2\pi f_c d_{mk}}{c}\right) = a \exp\left(\frac{-j2\pi d_{mk}}{\lambda_c}\right) \quad (2.14)$$

where f_c is the carrier frequency. The received signal at each antenna can be expressed relative to the first antenna which is set as a reference such that the remaining antennas receive a time-shifted version of the signal relative to the signal

received at the reference antenna, see Figure 2.1. The distance between k -th UE and the m -th receive antenna is given by:

$$d_{mk} = d_{1k} + (d_{mk} - d_{1k}) = d_{1k} + \Delta d_{mk} \quad (2.15)$$

Now, the channel $\mathbf{h}_k \in \mathbb{C}^{M \times 1}$ vector of the k -th UEs is represented in the following form:

$$\mathbf{h}_k = a \exp\left(\frac{-j2\pi d_{1k}}{\lambda_c}\right) \begin{bmatrix} 1 \\ \exp(-j2\pi \Delta d_{2k}/c) \\ \vdots \\ \exp(-j2\pi \Delta d_{Mk}/c) \end{bmatrix} \quad (2.16)$$

The channel matrix \mathbf{H} is then:

$$\mathbf{H} = [\mathbf{h}_1 \quad \mathbf{h}_2 \quad \cdots \quad \mathbf{h}_K] \quad (2.17)$$

2.1.3 Receiver

When the signal is received by the BS, the signal vector \mathbf{y} is formed by a sum of different symbols that were sent through the channel \mathbf{H} . So, the received signal \mathbf{y} needs to be separated. So, multiplying the vector \mathbf{y} by a matrix \mathbf{J} , such that:

$$\hat{\mathbf{x}} = \mathbf{J}\mathbf{y} = \mathbf{J}(\mathbf{H}\mathbf{x} + \mathbf{w}) = \mathbf{JH}\mathbf{x} + \mathbf{Jw} \quad (2.18)$$

and assuming the channel is known, the Zero-Forcing (ZF) receiver can be used, which means that the matrix \mathbf{J} is equal to the pseudo inverse (Moore-Penrose inverse) of \mathbf{H} :

$$\mathbf{H}^+ = (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H \quad (2.19)$$

Then, (2.18) can be written as:

$$\hat{\mathbf{x}} = \mathbf{H}^+ \mathbf{H}\mathbf{x} + \mathbf{H}^+ \mathbf{w} \quad (2.20)$$

And since $\mathbf{H}^+ \mathbf{H}$ is the identity matrix \mathbf{I} , the reconstructed signal is given by:

$$\hat{\mathbf{x}} = \mathbf{x} + \mathbf{H}^+ \mathbf{w} \quad (2.21)$$

3

Transform Coding

3.1 Introduction

The basic idea of transform coding is to represent a sequence in a different form, such that the output sequence contains the information in a reduced number of coefficients and then, these coefficients can be transmitted [10]. Later in the receiver, the inverse transform of the reduced number of coefficients is taken and a reconstructed sequence is obtained [10]. This should originate a sequence similar to the original sequence. If a sequence given by $\{x_n\}$ is assumed and $\{\theta_n\}$ is its transformed with $n = 1, 2, \dots, N$, the amount of errors is equal in both sequences:

$$\sum_{i=1}^N (x_i - \hat{x}_i)^2 = \sum_{i=1}^N (\theta_i - \hat{\theta}_i)^2 \quad (3.1)$$

Since x_n should be different than \hat{x}_n , this is a form of lossy compression. Linear transforms are one of the most popular type of transforms. These generate the sequence $\{\theta_i\}$ from $\{x_i\}$ through the following expression:

$$\theta_n = \sum_{i=1}^N x_i a_{n,i} \quad (3.2)$$

And to obtain the sequence $\{x_i\}$ again, the inverse transform is applied as:

$$x_n = \sum_{i=1}^N \theta_i b_{n,i} \quad (3.3)$$

Then, this expression can be represented in matrix form:

$$\boldsymbol{\theta} = \mathbf{A}\mathbf{x} \quad (3.4)$$

$$\mathbf{x} = \mathbf{B}\boldsymbol{\theta} \quad (3.5)$$

And the elements in position (i,j) are given by

$$[\mathbf{A}]_{i,j} = a_{i,j} \quad (3.6)$$

$$[\mathbf{B}]_{i,j} = b_{i,j} \quad (3.7)$$

where \mathbf{A} and \mathbf{B} are $N \times N$ matrices and \mathbf{A} and \mathbf{B} are the inverse of each other, such that, $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$, where \mathbf{I} is the identity matrix.

If the transform is orthonormal, the inverse transform corresponds to the inverse of the transform matrix, which also corresponds to the transpose of the transform matrix, as in

$$\mathbf{B} = \mathbf{A}^{-1} = \mathbf{A}^T \quad (3.8)$$

Moreover, orthonormal transforms preserve the energy. In other words, the sum of the squares of the transformed sequence is equal to the sum of the squares of the original sequence

$$\begin{aligned} \sum_{i=1}^N \theta_i^2 &= \boldsymbol{\theta}^T \boldsymbol{\theta} \\ &= (\mathbf{A}\mathbf{x})^T (\mathbf{A}\mathbf{x}) \\ &= \mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} \end{aligned} \quad (3.9)$$

If \mathbf{A} is orthonormal, $\mathbf{A}^T = \mathbf{A}^{-1} = \mathbf{I}$, so

$$\begin{aligned} \mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} &= \mathbf{x}^T \mathbf{x} \\ &= \sum_{n=1}^N x_n^2 \end{aligned} \quad (3.10)$$

then

$$\sum_{n=1}^N x_n^2 = \sum_{n=1}^N \theta_n^2 \quad (3.11)$$

In summary, the orthonormal transform preserves the energy of the original signal as the Parseval's Theorem.

3.2 Transforms

3.2.1 Fast Fourier Transform

The FFT is a linear and orthonormal transform that converts a sequence from its original domain to a representation in the frequency domain. Consider \mathbf{y} with N elements:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} \quad (3.12)$$

The matrix \mathbf{T}_{FFT} is the FFT matrix with size $N \times N$ and the vector \mathbf{z} is the transform of vector \mathbf{y} and it is applied such that

$$\mathbf{z} = \mathbf{T}_{\text{FFT}} \mathbf{y} = \begin{bmatrix} \omega(1,1) & \omega(2,1) & \cdots & \omega(N,1) \\ \omega(1,2) & \omega(2,2) & \cdots & \omega(N,2) \\ \vdots & \vdots & & \vdots \\ \omega(1,N) & \omega(2,N) & \cdots & \omega(N,N) \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_N \end{bmatrix} \quad (3.13)$$

where

$$\omega(n,k) = e^{-2\pi j(k-1)(n-1)/N}, \quad n = 1, \dots, N \quad \text{and} \quad k = 1, \dots, N \quad (3.14)$$

In order to recover the original sequence, it has to be inverse transformed by the Inverse Fast Fourier Transform (IFFT) which is given by:

$$\begin{aligned} \mathbf{y} &= \mathbf{T}_{\text{IFFT}} \mathbf{z} \\ &= \frac{1}{N} \begin{bmatrix} \omega^{-1}(1,1) & \omega^{-1}(1,2) & \cdots & \omega^{-1}(1,N) \\ \omega^{-1}(2,1) & \omega^{-1}(2,2) & \cdots & \omega^{-1}(2,N) \\ \vdots & \vdots & & \vdots \\ \omega^{-1}(N,1) & \omega^{-1}(N,2) & \cdots & \omega^{-1}(N,N) \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_N \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} \end{aligned} \quad (3.15)$$

3.2.2 Discrete Cosine Transform

As the FFT, the DCT is a linear and orthonormal transform that converts a sequence to a representation in the frequency domain. If \mathbf{T}_{DCT} is the DCT matrix and the vector \mathbf{z} is the transform of \mathbf{y} :

$$\mathbf{z} = \mathbf{T}_{\text{DCT}} \mathbf{y} = \begin{bmatrix} \omega(1,1) & \omega(2,1) & \cdots & \omega(N,1) \\ \omega(1,2) & \omega(2,2) & \cdots & \omega(N,2) \\ \vdots & \vdots & & \vdots \\ \omega(1,N) & \omega(2,N) & \cdots & \omega(N,N) \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} \quad (3.16)$$

where

$$\begin{aligned} \omega(n,k) &= r(k) \cos\left(\frac{\pi}{2N}(2n-1)(k-1)\right), \\ n &= 1, \dots, N \quad \text{and} \quad k = 1, \dots, N. \end{aligned} \quad (3.17)$$

and

$$r(k) = \begin{cases} \frac{1}{\sqrt{N}}, & k=1 \\ \sqrt{\frac{2}{N}}, & 2 \leq k \leq N \end{cases}$$

The vector \mathbf{y} is obtained through the Inverse Discrete Cosine Transform (IDCT):

$$\mathbf{y} = \mathbf{T}_{\text{IDCT}} \mathbf{z} = \begin{bmatrix} \omega(1,1) & \omega(1,2) & \cdots & \omega(1,N) \\ \omega(2,1) & \omega(2,2) & \cdots & \omega(2,N) \\ \vdots & \vdots & & \vdots \\ \omega(N,1) & \omega(N,2) & \cdots & \omega(N,N) \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_N \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} \quad (3.18)$$

4

Compression Techniques

4.1 Correlated channels in massive MIMO

Equation (2.1) refers that the received signal \mathbf{y} is a linear combination of \mathbf{H} and \mathbf{x} with AWGN noise added to it, thus the matrix \mathbf{H} is the element in this equation which is going to determine the amount of correlation in the channel.

In the case of the LoS channel, \mathbf{H} adds a phase shift to the incoming signal at each receive antenna and scales it by a factor a that is the path loss, but this factor is neglected since it is assumed that the path loss is same for all receive antennas. So the channel matrix \mathbf{H} is just adding a phase shift to the receive antennas that correspond to the distance to the transmit antenna.

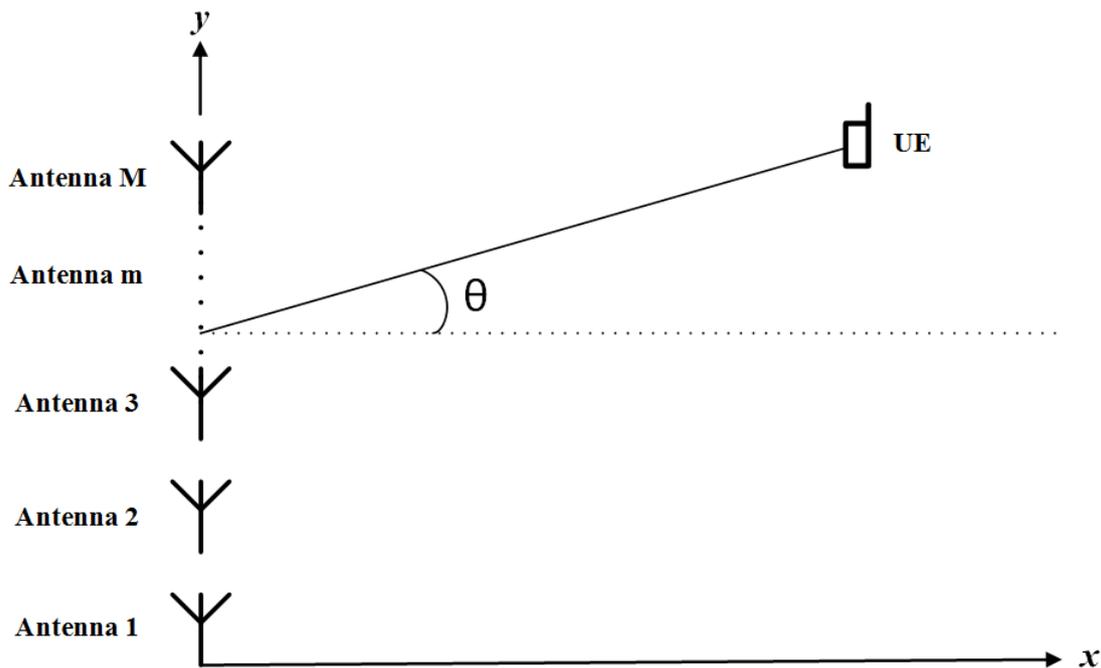


Figure 4.1: SIMO model with user placed at an AoA of θ

As can be seen from Figure 4.1 the user can be placed at different angles relative to the antenna array. For certain channel realisations there will be strong correlation between the channel vectors. This implies in turn that the system is sparse and that lossless compression may be applied. A frequently used tool for this purpose is transform coding, where transforms like FFT and DCT can transform the receive

signal vector \mathbf{y} at the instant t and then discard low power coefficients in order to compress the signal, because they will not cause large errors in the decompressed signal and most of the power will be concentrated to a few coefficients if the receive signal vector \mathbf{y} is correlated. One of the reasons for using the FFT and DCT is that they have relatively low computational complexity.

4.2 Lossy compression algorithm

The signal vector \mathbf{y} will be transformed into the frequency domain and then low power frequency coefficients will be discarded which means that it applies lossy compression and throws away part of the information in the original signal. The receive signal vector \mathbf{y} will not be compressed directly since this would require a complex hardware implementation.

The vector \mathbf{y} will instead be divided into several smaller groups that are connected to integrated Radio Frequency (RF)-chains in the system and compression will be applied to each group independently. Specifically, if the receive signal vector \mathbf{y} is divided into L number of groups, it can be written as:

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_L \end{bmatrix} \quad (4.1)$$

where each group is given by $\mathbf{y}_l \in \mathbb{C}^{G \times 1}$ in which $G = \frac{M}{L}$ is referred as the *group size*.

The next step is to apply the transform to each of these groups. The transform matrix is denoted by $\mathbf{T} \in \mathbb{C}^{G \times G}$ and the transformation of the received signal is then:

$$\mathbf{z} = \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \\ \vdots \\ \mathbf{z}_L \end{bmatrix} = \begin{bmatrix} \mathbf{T}_1 \mathbf{y}_1 \\ \mathbf{T}_2 \mathbf{y}_2 \\ \vdots \\ \mathbf{T}_L \mathbf{y}_L \end{bmatrix} \quad (4.2)$$

where $\mathbf{z}_l \in \mathbb{C}^{G \times 1}$. \mathbf{z}_l will thus have $\frac{M}{L} = G$ coefficients. Compression is done by only sending a fraction of the elements to the CPU, specifically the compression level C is defined as:

$$C = \frac{G'}{G}, \quad G' \leq G \quad (4.3)$$

Since each element in the \mathbf{z}_l is sent over one data stream to the CPU, it will have G' elements $\mathbf{z}_l \in \mathbb{C}^{G' \times 1}$. A scheme over the system is illustrated in Figure 4.2. At the CPU side the \mathbf{z}_l is taken as an input to the inverse transformation matrix (\mathbf{T}_l^{-1}) and zeros are added for the elements that were not sent, then, the output signal is decompressed into the $\hat{\mathbf{y}}_l$ vector with $\frac{M}{L}$ elements as:

$$\hat{\mathbf{y}}_l = \mathbf{T}_l^{-1} \mathbf{z}_l \quad (4.4)$$

The last step is to take the decompressed $\hat{\mathbf{y}}$ as an input to the pseudo inverse of the channel \mathbf{H}^+ as in (2.19) to decode the transmitted symbols as:

$$\hat{\mathbf{x}} = \mathbf{H}^+ \hat{\mathbf{y}} \quad (4.5)$$

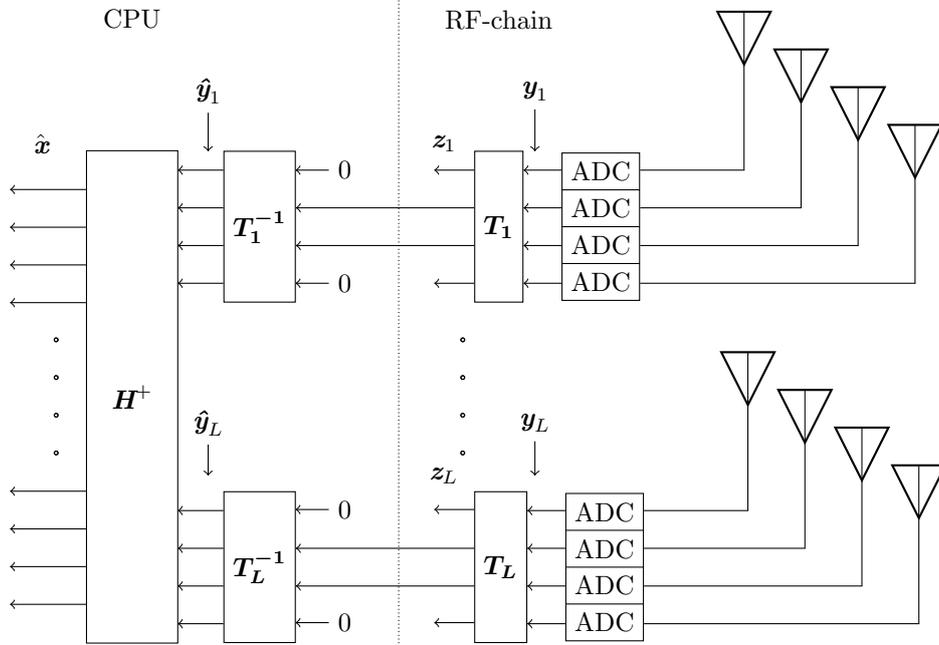


Figure 4.2: The lossy compression model with L number of groups. Each block takes four input elements of the received signal \mathbf{y} . The \mathbf{T}_l blocks implement a compression level of $C = \frac{2}{4}$ and thus reduce the number of streams to two that are sent to the CPU side for decompression. Note that the inverse transformation block \mathbf{T}_l^{-1} , adds zeros to the corresponding elements that were discarded (not transmitted).

There is also a trade-off between the size of the groups and the concentration of the power among the frequency coefficients, thus smaller groups will yield more uniformly distributed power over the coefficients and it will be increasingly difficult to compress the data without losing a large amount of signal power. A group size of 4 or 8 antennas is considered realistic because this can fit on one single Printed Circuit Board (PCB).

4.3 Lossless compression algorithm

It is possible to implement a compression matrix \mathbf{T} and reconstruct a signal with the same power. This can be achieved if the transformed signals \mathbf{z}_l are decoded with respect to the product of $(\mathbf{TH})^+$ instead of \mathbf{H}^+ . Consider the received signal:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w} \quad (4.6)$$

where \mathbf{y} is the receive signal vector as before. The next step is to multiply this matrix by the compression matrix \mathbf{T} in order to compress the data.

$$\begin{aligned} \mathbf{z} &= \mathbf{T}\mathbf{y} \\ &= \mathbf{T}(\mathbf{H}\mathbf{x} + \mathbf{w}) \\ &= \mathbf{TH}\mathbf{x} + \mathbf{T}\mathbf{w} \end{aligned} \quad (4.7)$$

In order to recover the transmitted symbols \mathbf{x} , \mathbf{z} is multiplied by the factor $(\mathbf{TH})^+$:

$$\begin{aligned} \hat{\mathbf{x}} &= (\mathbf{TH})^+ \mathbf{z} \\ &= (\mathbf{TH})^+ (\mathbf{TH}\mathbf{x} + (\mathbf{TH})^+ \mathbf{T}\mathbf{w}) \\ &= \mathbf{I}\mathbf{x} + (\mathbf{TH})^+ \mathbf{T}\mathbf{w} \\ &= \mathbf{x} + (\mathbf{TH})^+ \mathbf{T}\mathbf{w} \end{aligned} \quad (4.8)$$

The recovered symbols $\hat{\mathbf{x}}$ of this compression method will according to (4.8) be degraded by noise that is scaled by the factor $(\mathbf{TH})^+ \mathbf{T}$. This is also true even without implementing the \mathbf{T} matrix, in that case the noise \mathbf{w} would just be scaled by \mathbf{H}^+ . By adding \mathbf{T} in the system and thus scaling the noise by $(\mathbf{TH})^+ \mathbf{T}$, it could potentially degrade the system performance even further.

4.3.1 Implementation of lossless compression algorithm

The received signal will be divided in groups in order to enable a simple hardware implementation (see section 4.3). This receive signal vector will then be multiplied with \mathbf{T}_l in order to compress the data. If the antenna group size is $G = 4$ and the compression level is $C = \frac{2}{4}$ then the \mathbf{T}_l matrix will have only two rows! Consider the example below with $C = \frac{2}{4}$:

$$\mathbf{z}_l = \mathbf{T}_l \mathbf{y}_l = \begin{bmatrix} t_1 & t_2 & t_3 & t_4 \\ t_5 & t_6 & t_7 & t_8 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \quad (4.9)$$

where \mathbf{z}_l is the compressed signal with two elements: z_1 and z_2 , that will be transmitted over two data streams. Without the \mathbf{T}_l the entire \mathbf{y}_l would have to be transmitted to the CPU that has four elements: y_1, y_2, y_3 and y_4 . This implies that

the compression matrix \mathbf{T}_l enables 50% compression, since the number of elements is reduced by half. The C that the \mathbf{T}_l matrix introduces, is solely determined by the number of rows it has, on the other hand it has to have the same number of columns as the antenna group size in order to match with the \mathbf{y}_l matrix. Thus by varying the number of rows of the \mathbf{T}_l matrix, an arbitrary C can be chosen.

This process is applied to each antenna group individually, where each \mathbf{T}_l matrix can be either the same for each group, or adaptively chosen for every group independently. In order to decode the decompressed data correctly, all of the \mathbf{T}_l matrices have to be appended into a total \mathbf{T} matrix that contains information of the whole system. Each \mathbf{T}_l will be appended diagonally into a big matrix and all other elements will be set to zero.

$$\mathbf{z} = \mathbf{T}\mathbf{y} = \begin{bmatrix} t_1 & t_2 & t_3 & t_4 & 0 & 0 & 0 & 0 & \cdots & \cdots \\ t_5 & t_6 & t_7 & t_8 & 0 & 0 & 0 & 0 & \cdots & \cdots \\ 0 & 0 & 0 & 0 & t_9 & t_{10} & t_{11} & t_{12} & & \\ 0 & 0 & 0 & 0 & t_{13} & t_{14} & t_{15} & t_{16} & & \\ & \vdots & \vdots & & & & & & \ddots & \\ & \vdots & \vdots & & & & & & & \ddots \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ \vdots \\ \vdots \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ \vdots \end{bmatrix} \quad (4.10)$$

Each group of 4 antennas will be multiplied by a group \mathbf{T}_l matrix corresponding to the l -th antenna group. The first group \mathbf{y}_1 (antennas y_1 - y_4) will be multiplied by \mathbf{T}_1 and the second group \mathbf{y}_2 (antennas y_5 - y_8) will be multiplied by \mathbf{T}_2 and so on.

$$\mathbf{z} = \mathbf{T}\mathbf{y} = \begin{bmatrix} \mathbf{T}_1 & \mathbf{0} & \cdots & \\ \mathbf{0} & \mathbf{T}_2 & & \\ \vdots & & \ddots & \\ & & & \mathbf{T}_L \end{bmatrix} \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_L \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_L \end{bmatrix} \quad (4.11)$$

The system model can be seen in the figure below:

4.3.2 Compression matrix selection

In order to reduce the number of data streams only a fraction of the elements in \mathbf{z}_l will be sent to the CPU. By selection a fraction of the elements in \mathbf{z}_l is the same thing as removing rows in the \mathbf{T}_l matrix so the choice can be seen as which rows to chose from the $G \times G$ \mathbf{T}_l matrix that will produce a new transformation matrix with the size $G' \times G$ where $G' \leq G$.

But before the rows are removed, it has to be established first, which rows should be kept and which ones should be removed. Consider an example system with a group size $G = 4$ where it is desired to reduce 2 streams for each group, such that $C = \frac{2}{4}$. In this case 2 out of 4 rows will be chosen in order to obtain a compression matrix as in (4.9).

The rows have to be chosen such that the resulting MSE of the recovered symbols is minimised. There are basically two factors that have to be taken into consideration, one is the amount of power that is kept after compression (power efficiency) and the other is the resulting correlation in the \mathbf{TH} matrix.

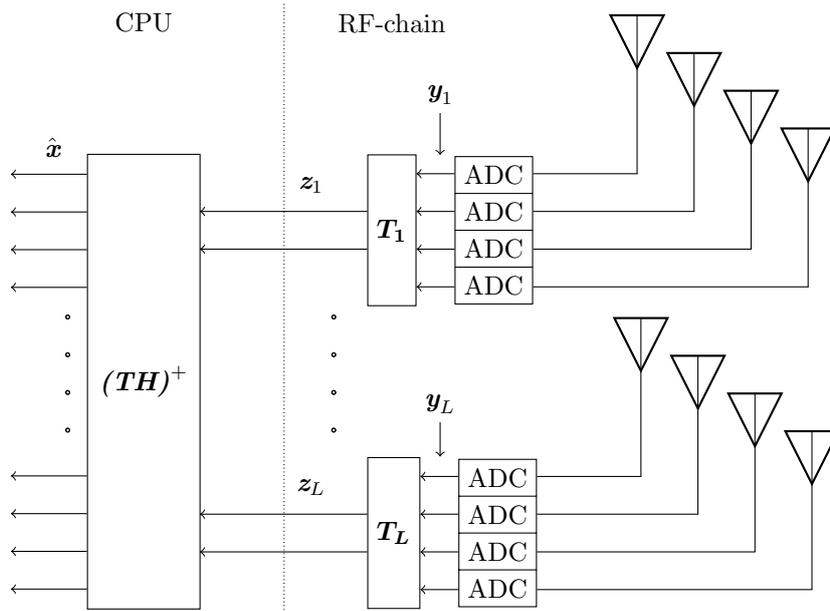


Figure 4.3: The lossless compression model with L number of groups. Each block takes four input elements of the received signal \mathbf{y} and implements $C = \frac{2}{4}$, reducing the number of streams to two that are sent to the CPU side for decompression.

So, with selection based on power, each \mathbf{T}_l matrix has to keep the rows that keep most of the power. With selection based on correlation, each \mathbf{T}_l matrix has to select the rows so that the total \mathbf{TH} has as low correlation as possible. Four different selection strategies are described in further detail below.

4.3.2.1 Power Selection algorithm

The basic idea of power selection is to choose the rows in the \mathbf{T}_l matrix that produces the z_l coefficients with highest power and then send these to the CPU. This is done calculating the power of the elements in z_l according to (5.1) and then, a sorting by power is done. The C determines how many of these coefficients will be sent to the CPU. If compression level $C = \frac{2}{4}$ is desired, then this means that the 2 sorted coefficients with highest power are selected out of 4 rows (remember that the denominator in C specifies the group size).

This method will keep most of the energy rich coefficients and thus it will have the highest possible power efficiency η . Keep in mind that this method does not do anything about the amount of correlation in the system. If the received signals are highly correlated, then the power distribution might be similar among the antenna groups after the transformation matrix. The power selection algorithm will then choose the same rows for each group which might introduce correlation to the system.

Table 4.1: List of steps of power selection algorithm

Step 1	Send Q known symbols
Step 2	Every group applies the transform
Step 3	Every group measures the average power of each element of \mathbf{z}_l
Step 4	Select the G' rows of \mathbf{T}_l that correspond to the \mathbf{z}_l elements with the highest average power.
Step 5	Reduce the \mathbf{T}_l to a $G' \times G$ according to the selection in step 4

4.3.2.2 Random Selection algorithm

A random selection method is set up to choose the rows in each \mathbf{T}_l matrix randomly. The advantage of this is that, it will not induce correlation in the \mathbf{TH} matrix since the same each row will appear as often as any other on average.

4.3.2.3 Correlation Selection algorithm

This selection algorithm is very similar to the random selection, however it is imposed that each row in \mathbf{T}_l has to be selected the same number of times in the whole system and each two groups cannot select the same combination of rows. Evidently, if the number of combinations is lower than the number of groups, which is the case where there is a high number of antennas, the combination of rows is going to repeat itself.

Therefore, the same rows and/or combinations should be repeated a few times as possible to reduce correlation. Another important aspect can be the order of the combinations to avoid periodic sequences of chosen rows.

The number of possible combinations of different rows is given by:

$$\binom{G}{G'} = \frac{G!}{G'!(G - G')!} \quad (4.12)$$

When $G=2$, considering a system where $M=128$, there are only 2 combinations $\{1; 2\}$. That implies that there are not many choices to obtain \mathbf{T}_l , so an alternative is to find a sequence of these group that is random. The assumed sequence is reported in Table 4.2.

Table 4.2: Sequence of the chosen rows with $G = 2$ and $C=0.5$

Group	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	...
Rows	1	2	1	2	2	1	1	2	2	1	2	2	1	2	1	1	...

where after the sixteenth group, the sequence repeats until the last group, L .

When $G=4$, there are 6 possible combinations $\{1\ 2; 1\ 3; 1\ 4; 2\ 3; 2\ 4; 3\ 4\}$, so the first six groups have the sequence shown in Table 4.3.

Table 4.3: Sequence of the chosen rows with $G = 4$ and $C = 0.5$.

Group	1	2	3	4	5	6	...
Rows	1 2	3 4	1 3	2 4	1 4	2 3	...

where the sequence repeats after the sixth group, until the last group, L . When $G=8$, there are 70 possible combinations and only 16 groups, so in this case, the number of possible combinations is bigger than the number of groups and it will not be necessary to have repeated sequences. However, the system is limited to select each line the same number of times. The assumed selection is given by Table 4.4:

With $G = 16$, the number of possible combinations is 12870, so much higher than $L=8$. The same condition should be imposed as explained for $G = 8$, as shown in Table 4.5 :

Table 4.4: Sequence of the chosen rows with $G = 8$

Group	1	2	3	4	5	6	7	8
Rows	1 2 3 4	5 6 7 8	1 2 3 5	4 6 7 8	1 2 3 6	4 5 7 8	1 2 3 7	4 5 6 8
Group	9	10	11	12	13	14	15	16
Rows	1 2 3 8	4 5 6 7	1 2 4 5	3 6 7 8	1 2 4 6	3 5 7 8	1 2 4 7	3 5 6 8

Table 4.5: Sequence of the chosen rows with $G = 16$

Group	Rows
1	1 2 3 4 5 6 7 8
2	9 10 11 12 13 14 15 16
3	2 3 4 5 6 7 9
4	8 10 11 12 13 14 15 16
5	1 2 3 4 5 6 7 10
6	8 9 11 12 13 14 15 16
7	1 2 3 4 5 6 7 11
8	8 9 10 12 13 14 15 16

Table 4.6: Sequence of the chosen rows with $G = 32$

Group	Rows
1	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
2	17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32
3	2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32
4	1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31

Analogous to the the last two situations, when $G=32$, the combination of rows for this case is on Table 4.6.

This exact combinations occur for the system with $M= 128$ and $C = 0.5$, but it is achievable to find the sequences with different M and C in a similar manner according to the same conditions.

4.3.2.4 Hybrid Selection algorithm

Since both energy and correlation are factors that can affect the system performance, a new selection algorithm has been designed that chooses rows according to both criteria, which will be referred as Hybrid selection. Rather than achieving the best results in terms of power efficiency or correlation, this selection strategy aims to combine the benefits of both criteria.

The Hybrid selection algorithm calculates first the average power efficiency of all groups and then applies power selection to a certain ratio of the groups whereas the remaining groups will be selected according to the correlation selection algorithm. In this way the power efficiency and amount of correlation will be balanced with regard to each other. The Hybrid algorithm is described in detail in Table 4.7.

Table 4.7: List of steps of power selection algorithm

Step 1	Send 100 known symbols
Step 2	Every group applies the transform
Step 3	Every group measures the average power of each element of \mathbf{z}_l
Step 4	Select the G' rows of \mathbf{T}_l that correspond to the \mathbf{z}_l elements with the highest average power.
Step 5	Reduce the \mathbf{T}_l to a $G' \times G$ according to the selection in step 4
Step 6	Every group calculates the sum of the average power of the elements of \mathbf{z}_l
Step 7	The half of the groups with the highest average power keeps the Power selection algorithm
Step 8	The rest of the groups uses the Correlation selection algorithm

It is also possible to choose a different trade-off for the Hybrid selection. Two additional Hybrid algorithms will be tested where one chooses 25% of the rows according to power and the remaining 75% according to correlation, which will be called Hybrid-25. The other algorithm will choose 75% of the rows according to power and the remaining 25% according to correlation, which will be called Hybrid-75.

5

Simulation Results

5.1 Massive MIMO model

Starting with the transmitter, the symbols \mathbf{x} are generated according to a 16-Quadrature Amplitude Modulation (QAM) constellation with in-phase and quadrature components: $(-3, -1, 1, 3)$. The channel is assumed to be in LoS conditions without multipath, which means there is a single LoS path between each UE k and each receive antenna at the BS.

Moreover, the receive antennas form an array with a linear configuration and a uniform separation of $\lambda/2$ (half-wavelength). An operating frequency $f_c = 30$ GHz is assumed, yielding a wavelength in the order of one centimetre, so the UEs, with a single antenna, are located at a distance from the BS in the order of meters which is much larger than the antenna separation at the BS. The noise $\mathbf{w} \in \mathbb{C}^{M \times 1}$ denotes AWGN, whose elements are IID zero-mean Gaussian random variables with unit variance and channel independent.

The positions of the UEs were generated randomly and the signal attenuation a was assumed to be equal for all antenna paths. This system model was simulated in MATLAB for at least 100 channel realisations in order to obtain a good mean for the measurements.

5.2 Metrics

5.2.1 Energy efficiency

Since some coefficients that were set to zero may have been nonzero from the beginning, it will lead to a reduced amount of power in the decompressed signal and it will cause distortion compared to the original receive signal \mathbf{y} . This will degrade the performance of the system at the cost of simplifying the hardware by reducing the number of data streams to the CPU.

The amount of lost energy will consequently be the key factor that determines the efficiency of compression and this value is desired to be as small as possible. It is therefore important to calculate the power ratio of the received signal and the compressed signal so that it can be determined how much power is lost after compression. The total power of the received signal \mathbf{y} can be computed with regard to the frequency coefficients $z(k)$ [11]:

$$\sum_{m=1}^M |y(m)|^2 = \frac{1}{M} \sum_{k=1}^M |z(k)|^2 \quad (5.1)$$

which is Parseval's Theorem that says that the sum of the squared signal is the sum of the square of its transform. Applied to the received signal it becomes:

$$P_{total} = \sum_{t=0}^{T-1} \sum_{l=1}^L \left(\frac{1}{M} \sum_{k=1}^M |z(k, l, t)|^2 \right) \quad (5.2)$$

The sum of P_{total} is taken over M coefficients for each group repeated over l groups and t time epochs. The power ratio relative to the compressed and uncompressed signals will be defined as the *energy efficiency*:

$$\eta = \frac{P_{total \text{ after compression}}}{P_{total}} \quad (5.3)$$

where $P_{total \text{ after compression}} \leq P_{total}$. If $\eta = 1$, it means that no energy has been lost.

5.2.2 Mean squared error

Another important metric that can serve as a tool to evaluate the system performance is the Mean Squared Error (MSE) of the decoded signal which is intimately connected with the power loss due to compression because lost power will lead to increased error in the received signal. The MSE is expressed below:

$$\text{MSE} = \frac{\sum_{t=1}^T (\mathbf{x} - \hat{\mathbf{x}})^T (\mathbf{x} - \hat{\mathbf{x}})}{TK} \quad (5.4)$$

ΔMSE is a metric that is relative to the upper and lower bound. MSE of the lower bound will always correspond to a system with M BS antennas *without* compression and the upper bound will correspond to the same system, but with a *reduced* number of antennas. The antennas for the upper bound will be decreased according to the compression level. If $C = 0.5$, the number of antennas will be reduced by half.

This is an important comparison with the case when compression is done because the reduced number of data streams sent to the CPU will correspond to the reduced number of antennas for the system *without* compression. The bounds are expressed below:

$$\text{MSE}_{\text{lower bound}} \rightarrow N \text{ BS antennas} \quad (5.5)$$

and

$$\text{MSE}_{\text{upper bound}} \rightarrow (N \cdot C) \leq N \text{ BS antennas} \quad (5.6)$$

Thus the ΔMSE can be expressed as:

$$\Delta\text{MSE}_{\text{compression}} = \text{MSE}_{\text{compression}} - \text{MSE}_{\text{lower bound}} \quad (5.7)$$

where $\text{MSE}_{\text{compression}}$ is the mean square error of the recovered signals after compression, the upper bound has also to be expressed in an analogous way:

$$\Delta\text{MSE}_{\text{upper bound}} = \text{MSE}_{\text{upper bound}} - \text{MSE}_{\text{lower bound}} \quad (5.8)$$

If compression is to be effective then $\Delta\text{MSE}_{\text{compression}}$ should be lower than $\Delta\text{MSE}_{\text{lower bound}}$. $\Delta\text{MSE} = 0$ means that the system does not introduce higher errors with compression applied to it ($\Delta\text{MSE}_{\text{compression}} = \Delta\text{MSE}_{\text{lower bound}}$).

The upper bound serves as a reference to evaluate the performance loss of the system that applies compression. This means that if the ΔMSE is above the upper bound then it is considered that the system introduces too much degradation. The optimal goal is to keep the ΔMSE close to the lower bound.

5.2.3 Condition number

The condition number evaluates the condition of the matrix and it is given by:

$$\tilde{\kappa}(\mathbf{TH}) = \|(\mathbf{TH})^+\| \cdot \|\mathbf{TH}\| \quad (5.9)$$

and it satisfies that

$$\tilde{\kappa}(\mathbf{TH}) = \|(\mathbf{TH})^+\| \cdot \|\mathbf{TH}\| \geq \|(\mathbf{TH})^+ \cdot \mathbf{TH}\| = 1 \quad (5.10)$$

When the condition number tends to 1, the changes due to condition of the matrix tends to reduce.

In a Massive MIMO system, one of the things that affects the performance is the correlation of \mathbf{H} , take into account that in LoS conditions, \mathbf{H} is given by the distance to the antennas. Then, it is easy to understand that two UEs, k_1 and k_2 , in the same position generate two columns in \mathbf{H} , such that $\mathbf{h}_{\mathbf{k}_1} = \mathbf{h}_{\mathbf{k}_2}$, which makes \mathbf{H} non-invertible. However, in the scenario where two UEs are not in the same place, but close to each other, $\mathbf{h}_{\mathbf{k}_1} \approx \mathbf{h}_{\mathbf{k}_2}$, then \mathbf{H} becomes invertible, but ill-conditioned. In that case, a small change (like the noise added in the channel) in \mathbf{y} can cause a large change in the recovered symbols. The problem appears, for example, when the symbols are recovered where the ZF Receiver, according to (2.20), requires the pseudo inverse of \mathbf{H} . The same is going to happen when \mathbf{T} is added to the system. If the \mathbf{TH} is ill-conditioned, the changes in the recovered symbols will increase in presence of noise, in other words, the errors and MSE will be bigger.

5.3 Lossy compression of massive MIMO uplink

In this section, the results for the LoS massive MIMO model are presented with lossy compression by the implementation of FFT and DCT.

5.3.1 One UE

The model was tested first with the case of only one UE that transmits its signal to the BS antenna array consisting of 128 antennas. The energy efficiency of the decompressed signal is plotted for AoA in the interval $[0^\circ, 90^\circ]$ with respect to different group size at the base station. The group sizes that were tested for were 2, 4, 8, 16 and 32. The angle is between the UE and a parallel line to the horizontal x-axis placed in the center of the antenna array (as in Figure 4.1). Therefore, at 0° the UE is placed in front of the antenna array adjacent to its centre point and at 90° the UE is positioned along the vertical line of the antenna array that is aligned with the y-axis as well.

It can be seen in Figure 5.1 how the energy efficiency depends with the AoA for different group sizes for a compression level of $C = 0.5$. Notice the difference in shape between Figure 5.1a and Figure 5.1b. The FFT yields a more regular dependence on the AoA unlike the DCT. For a group size of two, 50% of the energy is lost at an angle of 30° which means that there is no compression gain.

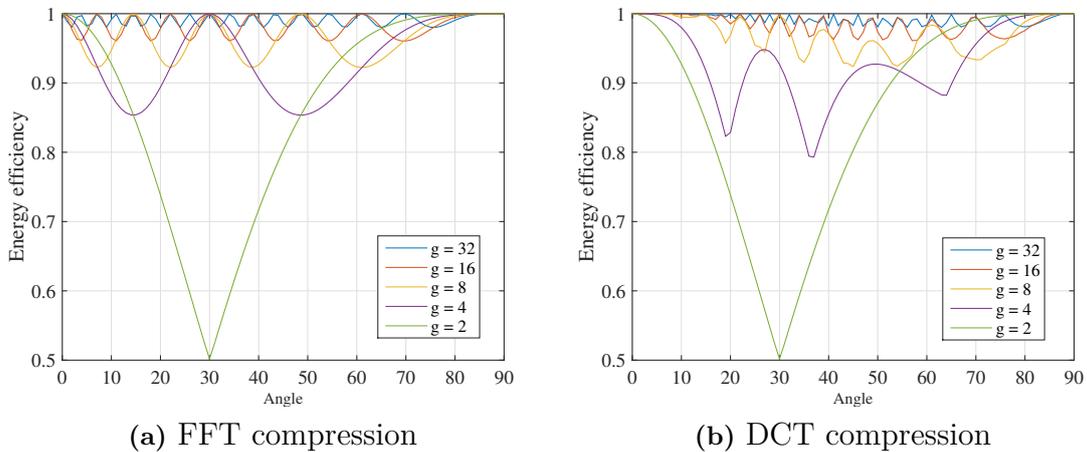


Figure 5.1: Energy efficiency of the compressed signal for different AoA with $C = 0.5$ and 128 antennas. Left shows results for FFT compression and the right shows for DCT compression.

5.3.2 Multiple UEs with varying group size

The following system setup consists of a variable number of UEs that are positioned randomly around the antenna array of 128 antennas. The number of UEs are set to vary from 1 to 10. The results are presented as energy efficiency with regard to group size and the number of UEs. The frequency coefficients are also presented with respect to group size and the number of UEs.

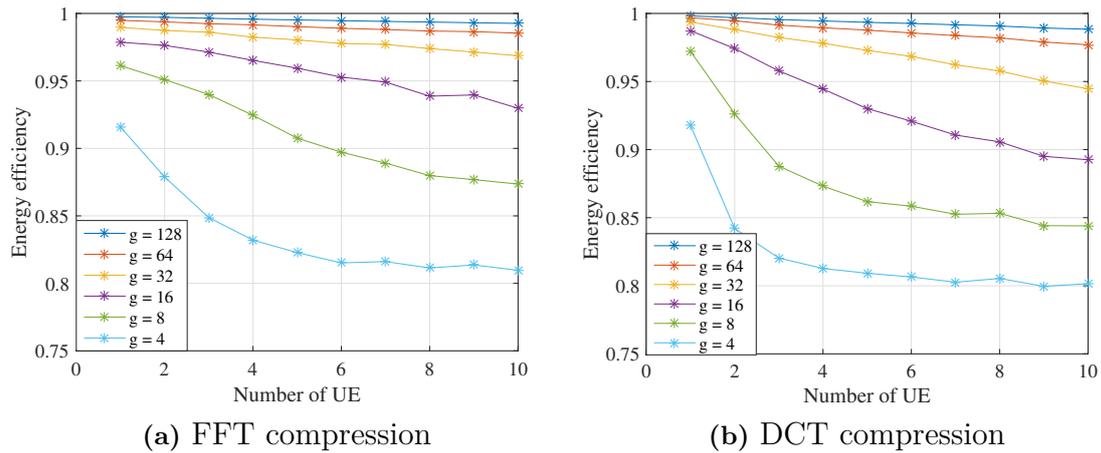
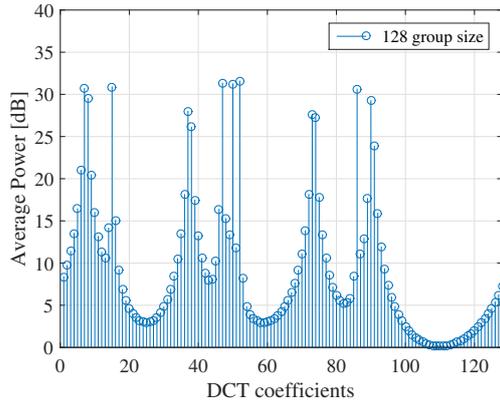


Figure 5.2: Energy efficiency of the compressed signal for increasing number of UEs with compression level $C = 0.5$ and 128 antennas. Left shows results for FFT compression and the right shows for DCT compression

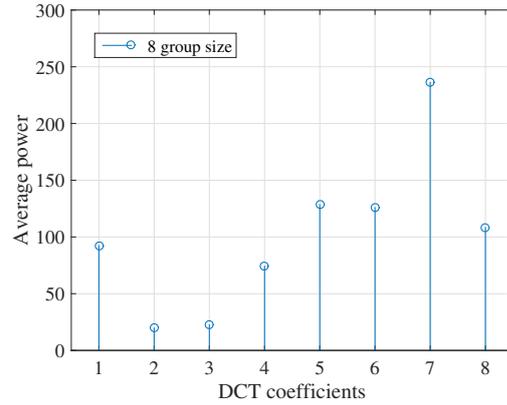
The results of the energy efficiency with multiple UEs can be seen in Figure 5.2. Systems with larger antenna grouping yields higher energy efficiency and thus better performance. FFT compression produces slightly better results than for DCT compression, a few percent.

In Figure 5.3, the frequency coefficients are depicted in average power for antenna group size of 8 and 128. With 128 group size there are many coefficients with power close to zero unlike the coefficients for 8 in group size.

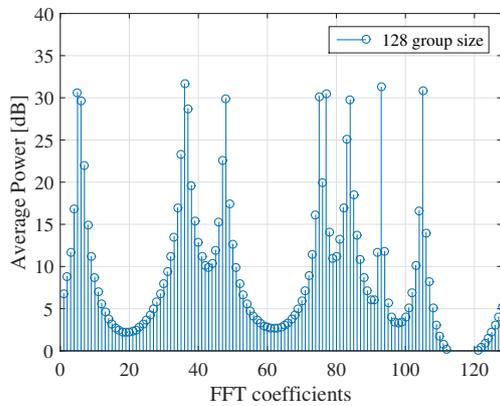
5. Simulation Results



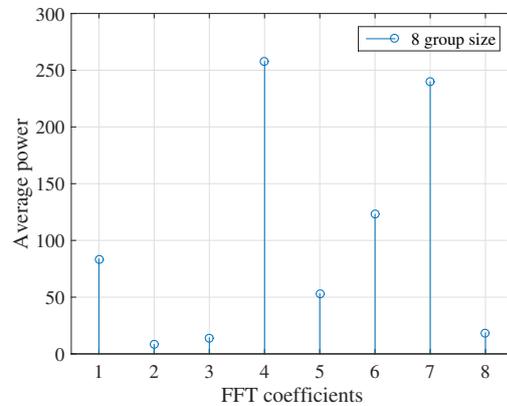
(a) 128 group size and $K = 10$



(b) 8 group size and $K = 10$



(c) 128 group size and $K = 10$



(d) 8 group size and $K = 10$

Figure 5.3: Average power for DCT and FFT coefficients for antenna group size 128 and 8 with $K = 10$ UEs. Notice that 128 group size has many coefficients close to zero whereas 8 group size has a more uniform shape.

In Figures 5.4 and 5.5 the frequency coefficients are shown for a group size of 8 antennas with 1,2,5 and 10 UEs. The difference is most clearly observed in Figure 5.5 for the FFT coefficients where the case of one UE yields only 1 nonzero coefficient and for more UEs the energy is spread more evenly among the coefficients. If these plots are compared with Figure 5.2, it can be seen that a higher number of UEs leads to reduced energy efficiency for both transforms.

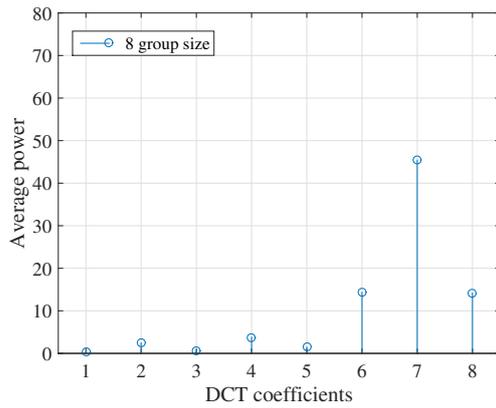
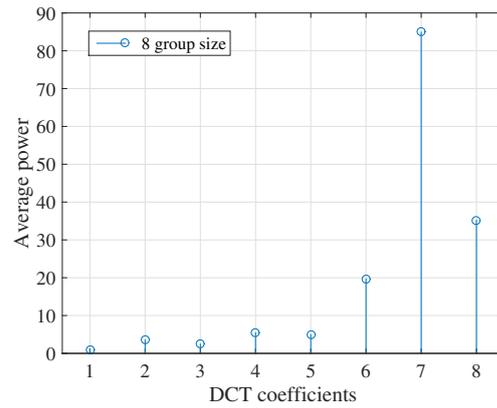
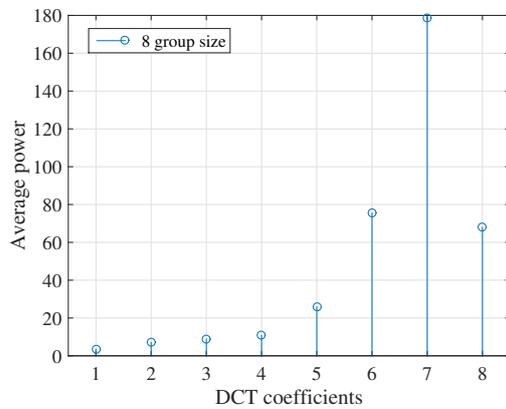
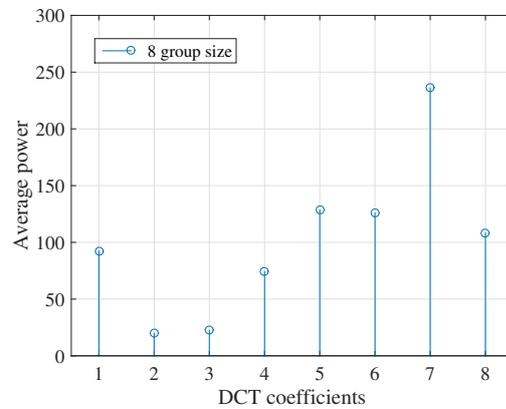
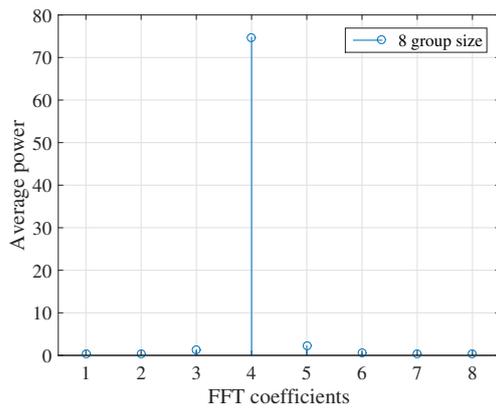
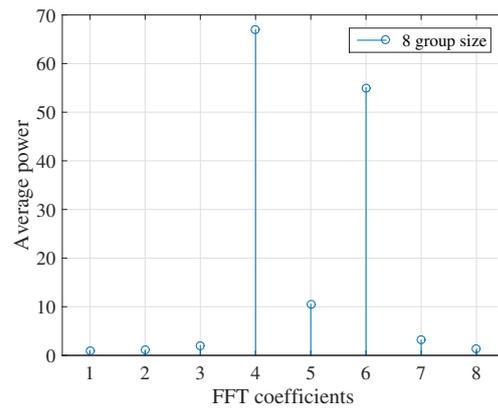
(a) 8 group size and $K = 1$ (b) 8 group size and $K = 2$ (c) 8 group size and $K = 5$ (d) 8 group size and $K = 10$

Figure 5.4: Average power for DCT coefficients for antenna group size 8 for 1, 2, 5 and 10 UEs. More UEs yield a higher spread of the power in the DCT coefficients.

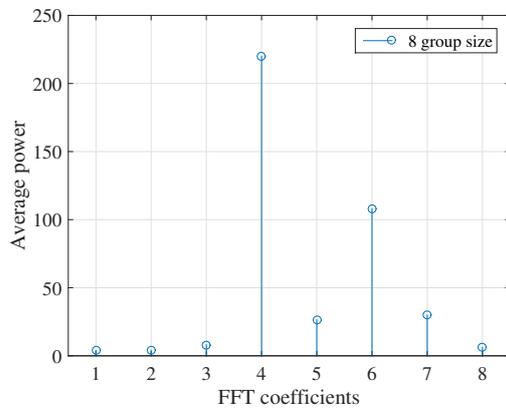
5. Simulation Results



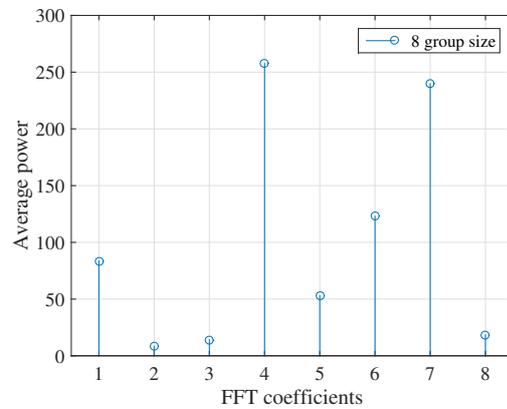
(a) 8 group size and $K = 1$



(b) 8 group size and $K = 2$



(c) 8 group size and $K = 5$



(d) 8 group size and $K = 10$

Figure 5.5: Average power for FFT coefficients for antenna group size 8 for 1,2,5 and 10 UEs. More UEs yield a higher spread of the power in the FFT coefficients.

5.4 Lossless compression of massive MIMO up-link

The metrics that were evaluated here are: ΔMSE , energy efficiency (η) and average condition number ($\tilde{\kappa}$). While the number of UEs (K) and number of antennas (M) at the BS, the group size (G) and the compression level (C) are the studied parameters.

The condition number, κ , is taken with regard to the \mathbf{TH} matrix and an average is formed based on the simulation of at least 100 channel realisations.

5.4.1 Group size analysis

The ΔMSE is plotted versus the group size in Figure 5.6. The subfigures represent the result for the the different selection strategies. The power selection algorithm has large ΔMSE for group sizes 2,4 and 8 but for sizes 16 and 32 it is very small compared to the rest of the selection algorithms. The random selection algorithm has the opposite behaviour compared to power selection.

Correlation selection manages to have ΔMSE under the upper bound for all groups whereas the hybrid selection algorithm has even smaller ΔMSE . Table 5.1 illustrates the average condition number of the \mathbf{TH} matrix in terms of selection for each group size and selection algorithm. Furthermore, both the ΔMSE and average condition number decreases or increases together with the group size. This holds for all selection algorithms except for hybrid selection since the condition number increases for group size 16 and 32 while the ΔMSE decreases.

5. Simulation Results

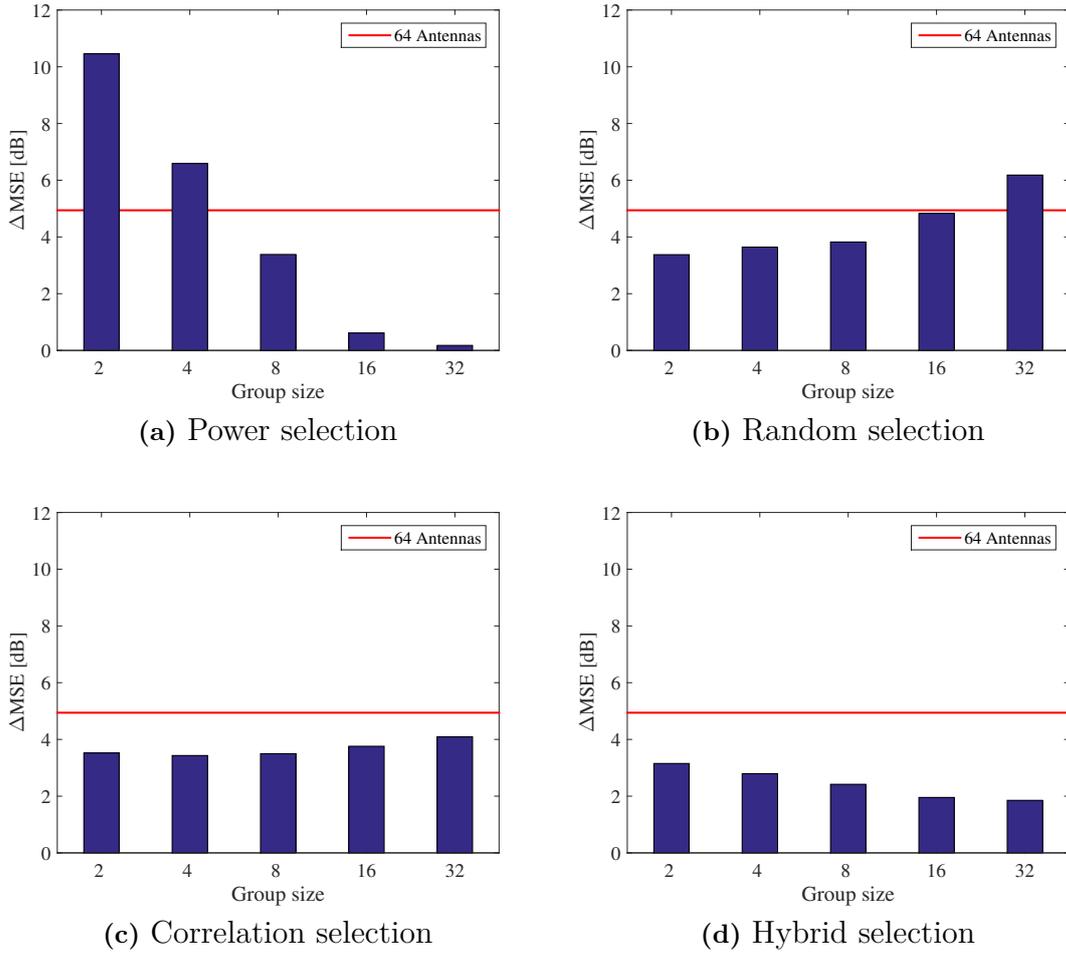


Figure 5.6: ΔMSE versus group size for different selection methods for 128 antennas. The number of UEs is $K=10$ at $C = 0.5$. The red curve is the upper bound with 64 antennas. The group sizes are $G = 2, 4, 8, 16$ and 32

Table 5.1: Average condition number $\tilde{\kappa}$ for different selection algorithms and group sizes of $G = 2, 4, 8, 16$ and 32 .

Group	$\tilde{\kappa}(\mathbf{TH})$ vs selection			
	Power	Random	Correlation	Hybrid
2	27.9	2.3	2.61	2.92
4	9.1	2.4	2.69	2.82
8	4.6	2.6	2.80	2.70
16	2.1	3.2	3.07	2.74
32	1.8	11.4	3.37	2.77

The energy efficiency, η , of the selection algorithms can be seen in Figure 5.7. The power selection algorithm has the highest power efficiency whereas random and correlation selection algorithms have an efficiency of $\eta \approx 0.5$. Note also that the hybrid selection algorithm has a better power efficiency compared to random and

correlation selection.

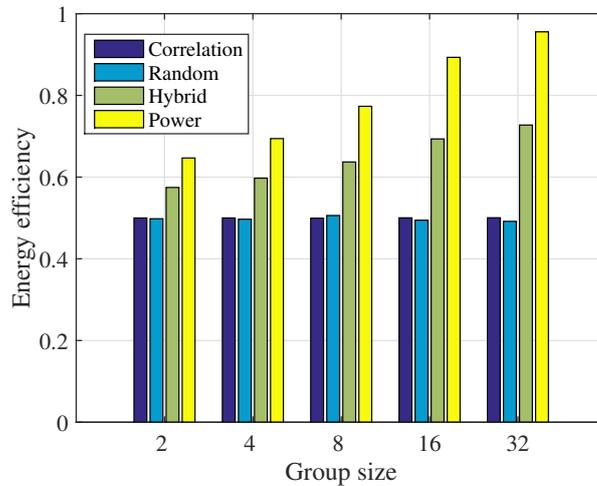


Figure 5.7: Energy efficiency η versus group size is plotted for different selection methods. Number of UEs is $K=10$ and is $C = 0.5$ and the number of antennas is 128. The power selection algorithm has clearly the best energy efficiency compared to the others. Hybrid selection yields a gain in energy efficiency compared to random and correlation selection algorithms.

5.4.2 Number of UEs analysis

ΔMSE is depicted in Figure 5.8 versus number of UEs ranging from 1-20. The number of antennas is 128 and $C = 0.5$. Note that the ΔMSE of the upper bound is increasing non-linearly with number of users.

In Figure 5.8a it can be seen that the error is large for small group size and then it decreases rapidly with growing group size. The selection algorithms of random and correlation selection in Figure 5.8b-c have an opposite effect. The error increases instead with larger group size (groups of 8, 16 and 32). The group sizes of 2 and 4 are not included since they behave in similar way as for the case with group size of 8.

The error of the correlation selection algorithm is clearly better than the random selection algorithm since the ΔMSE is within the bounds. The hybrid selection algorithm is even better than correlation selection since the error is even smaller, especially for a smaller amount of UEs. But one other advantage with this algorithm is that the error does not increase with the group size.

Also, the ΔMSE of the upper bound is increasing non-linearly with number of users.

5. Simulation Results

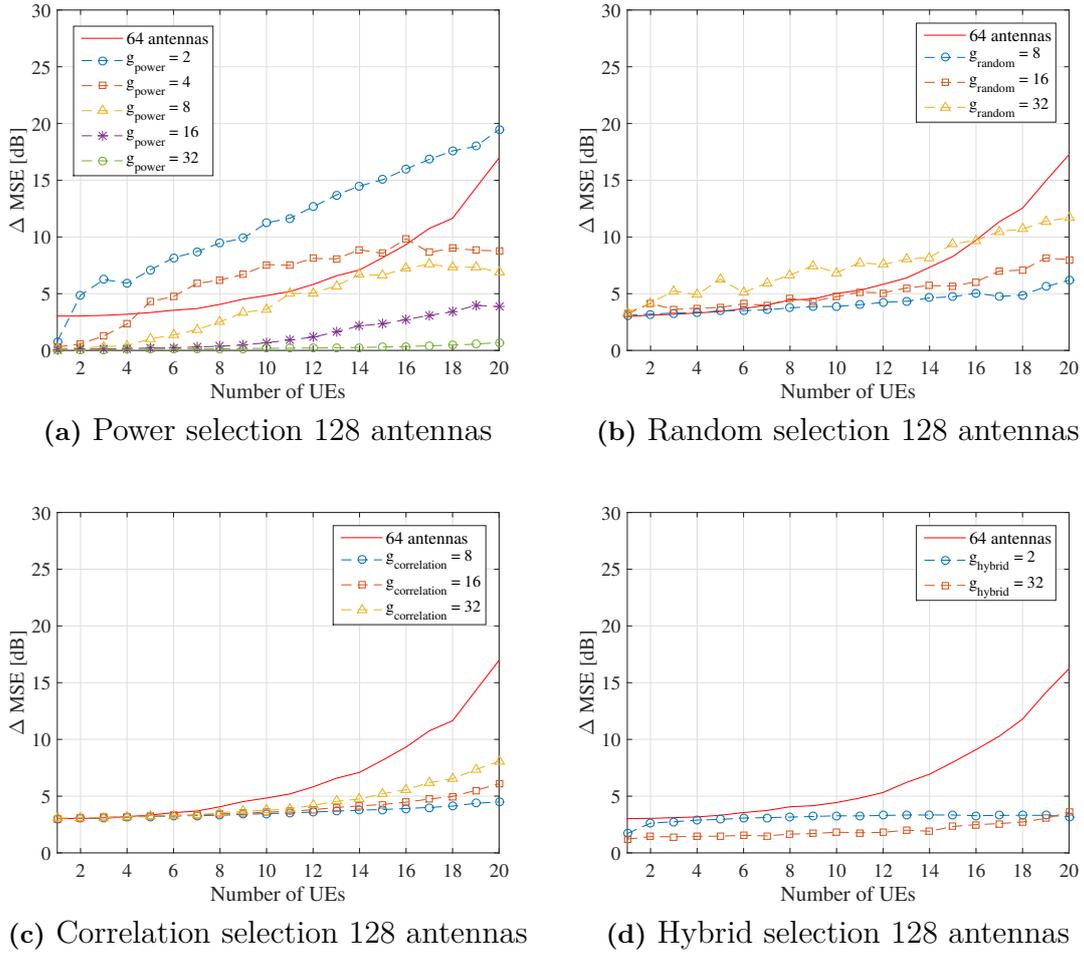


Figure 5.8: Δ MSE vs number of UEs for 128 antennas at $C = 0.5$. Power selection tends to have higher Δ MSE for small group sizes and the opposite holds for the other selection methods. The Δ MSE with hybrid selection has least dependence on group size.

5.4.3 Number of antennas analysis

The Δ MSE is depicted in Figure 5.9 versus the number of antennas for different selection algorithms. In Figure 5.9a with power selection, the Δ MSE is the highest for the smallest group size of 2 and then it is decreasing with higher group size. This method has very small errors for group sizes of 16 and 32 compared to the other algorithms.

For random and correlation selection in Figure 5.9b-c the errors are converging with high number of antennas for group sizes of 2-32. On the other hand, the errors are diverging with small number of antennas. The divergence is much higher for random selection since the error with group of 32 is $\Delta\text{MSE}_{g=32}(100) \approx 10 \text{ dB}$ whereas for correlation selection $\Delta\text{MSE}_{g=32}(100) \approx 5.5 \text{ dB}$.

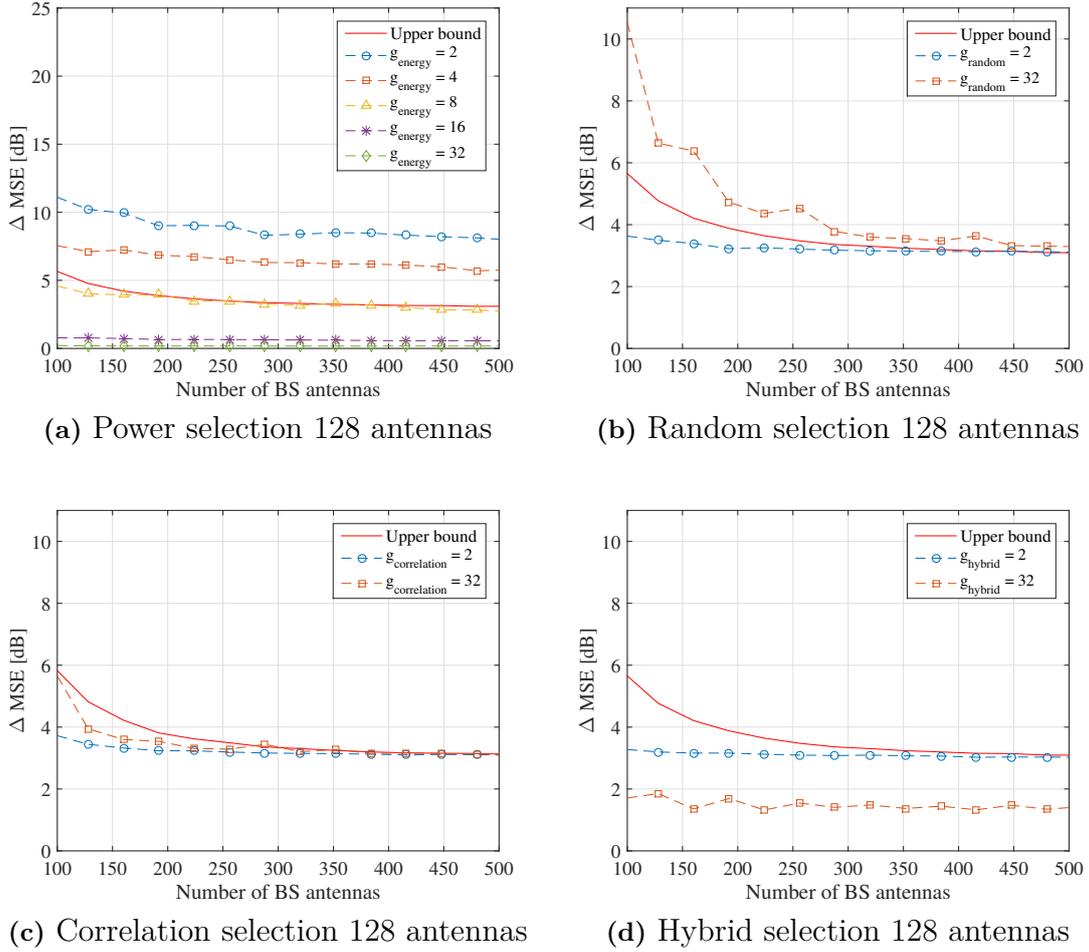


Figure 5.9: Δ MSE vs number of antennas $100 \leq N \leq 500$ at $C = 0.5$ and $K = 10$ UE. Power selection has the highest dependence on group size. In subfigures b-c only groups of 2 and 32 are plotted since the errors are converging with high number of antennas. With hybrid selection (d) all of the group sizes give errors below the upper bound with the largest group size having the smallest error.

The hybrid selection algorithm in Figure 5.9d the errors are $1.5 \text{ dB} \leq \Delta \text{MSE} \leq 3.5 \text{ dB}$ for group size from 2-32.

5.4.4 Compression level analysis

Figure 5.10 shows the Δ MSE according to the compression level C for $G=4$ and 16. In most of the cases, the Hybrid Selection is the method which has smallest Δ MSE. Also, as shown in the previous sections, power selection behaves better when the G is higher.

In Figure 5.11, the energy efficiency is shown with group size of 4 and 8 for the different selection algorithms. Power selection leads to higher energy efficiency, while correlation selection has energy efficiency equivalent to the compression level.

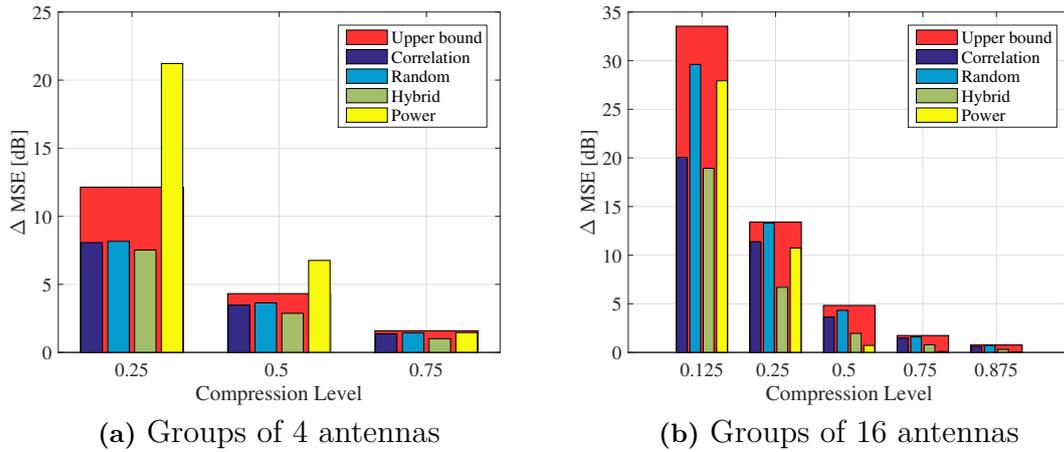


Figure 5.10: ΔMSE vs C for 128 antennas with 10 UEs. Power selection tends to have higher ΔMSE for small group sizes, however when the compression level is close to one, the ΔMSE of power selection is better than the other types of selection.

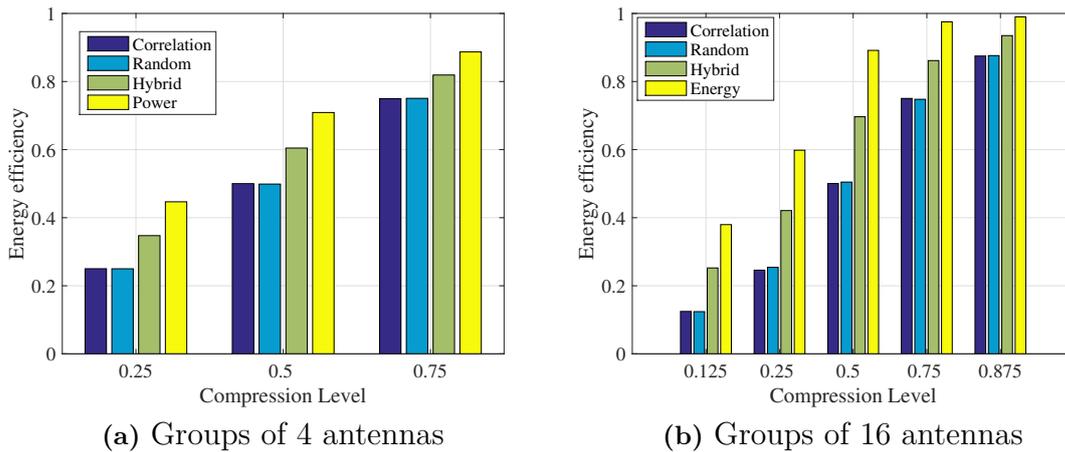


Figure 5.11: η vs the compression level with 128 antennas and 10 UEs. Power selection has the capacity to keep a higher percentage of the energy in the original signal. On the other side, the random and correlation selection have a percentage of energy similar to the compression level.

5.4.5 Hybrid-25 and Hybrid-75 algorithm

The hybrid algorithm has already been analysed for the case where it chooses half of the groups according to highest power and the rest according to the row selection combinations that minimises that reuse of each row. Here, the same algorithm will be studied but with different ratio of the selection with regard to power. In Figure 5.12 the ratio of 25% and 75% power selection cases are presented.

The main difference is that when more of the groups are chosen according to power as in Figure 5.12b, then the ΔMSE will be much lower for larger groups, whereas for the case when fewer groups are chosen to power then the ΔMSE will be higher.

However, the ΔMSE is largely reduced for small group sizes in both cases, specifically group sizes of 4 and 8 compared to the case of the Power selection algorithm in Figure 5.8a.

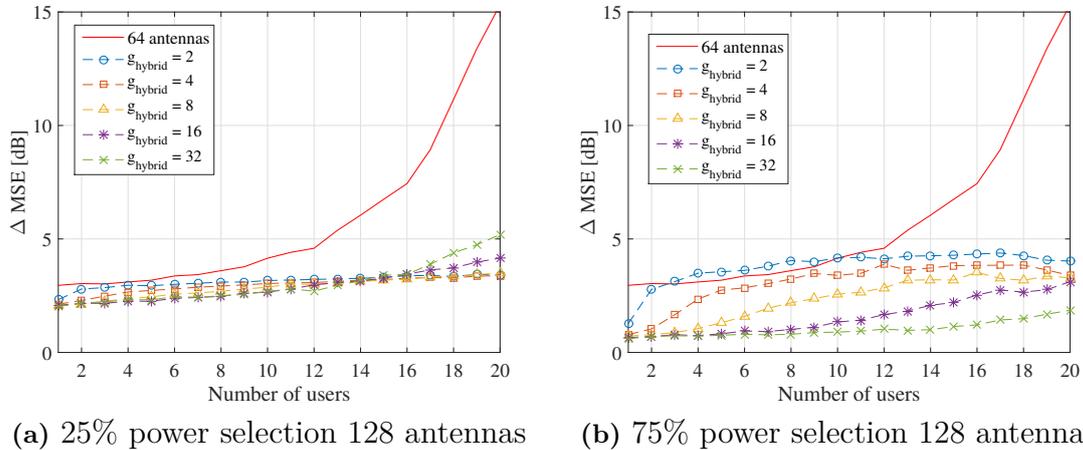


Figure 5.12: ΔMSE with respect to the number of UE for different hybrid algorithms. The subfigure on the left shows the case when only 25% of the groups are selected with regard to power and the right subfigure shows the case when 75% of the groups are selected according to power.

6

Discussion

Antenna group size

The optimum selection algorithm is the one that minimises the ΔMSE . It should be observed that the power selection algorithm has the smallest ΔMSE for antenna group sizes of 16 and 32 (see Figure 5.6a). This is because it has a low condition number ($\tilde{\kappa}_{g=16} = 2.1$), but also because it has higher energy efficiency for these group sizes as it is seen in Figure 5.7. For groups of 16 antennas, energy efficiency is close to 0.9 while for groups of 8 antennas is energy efficiency around 0.8, which does not seem to be a huge difference. It leads to the idea that power has a large impact on the ΔMSE , specially when η is closer to 1.

However, for small groups the power selection algorithm produced very large condition number ($\tilde{\kappa}_{g=4} = 9.1$) compared to the random selection algorithm ($\tilde{\kappa}_{g=4} = 2.4$), which in turn managed to obtain smaller ΔMSE despite of having even lower energy efficiency which proves that the higher correlation can degrade the ΔMSE .

For the hybrid selection algorithm (see Figure 5.6d) it is observable that the ΔMSE decreases with higher antenna group size compared to the correlation selection algorithm. One of the reasons for this is that it has lower condition number but also because it has higher energy efficiency. Since the energy efficiency is not as high as for the power selection algorithm it did not manage to get as low ΔMSE which again indicates the importance of the energy efficiency.

Number of UEs

From Figure 5.8 the ΔMSE can be analyzed with regard to the number of UEs. The ΔMSE tends to increase with the number of UEs. For smaller antenna group sizes, power selection has higher ΔMSE , because, similarly to the previous discussion for different group sizes, the power selection algorithm introduces a lot of correlation since it tends to select the same rows and thus increase the correlation of the system. Keep in mind that the receive antennas are uniformly spaced and it generates correlated received signals. Then the power distribution among the rows for each group will be similar and so the power selection algorithm will tend to select the same rows which in turn increases the correlation.

Moreover, when the number of UE is increased, correlation of the systems increases, but also each received signal is a sum of different signals sent by the UEs which makes the power distribution less compacted in the frequency components. For example, with power selection of groups of 4 antennas, with up to 4 UEs, the ΔMSE is below 3 dB and the upper bound, but with more than 4 UEs the upper

bound of 64 antennas is crossed.

The remaining selection techniques (random, correlation and hybrid selection) reduce the correlation of the system and power assumes reduced importance (in correlation and random methods no importance at all), thus are less sensitive to the increase of the number of UEs. This is observed in Figure 5.8b-d where it is clear that the ΔMSE increases much slower with number of UEs.

However, these algorithms are much closer to the upper bound than for power selection (small group sizes) which again indicates that the correlation of the system is not the only factor that affects the ΔMSE . This can be verified by comparing the hybrid selection algorithm with the correlation selection algorithm (see 5.8c-d), both of them have very similar condition number when the number of UEs is 10 but the hybrid technique offers a smaller ΔMSE by approximately 2 dB for the case of group size of 32. This can be explained by the higher energy efficiency of the hybrid technique (see Figure 5.7).

When the groups contain higher number of antennas, energy selection tends to achieve very good performance, as seen in Figure 5.8 with groups of 32 antennas, where the ΔMSE is smaller than 1 dB even with 20 UEs.

Degree of Compression

According to the Figure 5.10, the Hybrid selection is the method which in most of the cases minimizes more the ΔMSE . It occurs with group size of 4 antennas and 16 antennas (when the compression level is below $C=0.25$). Again, a first look would lead to conclude that energy selection should have better performance, however the correlation is increased and it conduces to poorer performance.

Also, energy selection tends to have higher ΔMSE than any other selection algorithms when compression level goes to 0, specially for small groups, and when compression level is close to 1, energy selection tends to be the best option. Since the channel is correlated, every group tends to select the same coefficients, specially those with the highest energy, while the lower power coefficients tend to be different group to group. The hybrid selection reduces the amount of correlation when compared to energy selection and then, it has generally better performance than the energy selection.

Hybrid selection algorithm

The hybrid selection algorithm uses power selection for half of the groups and the remaining groups use correlation selection. This is a very a simple strategy and has some parameters that could be more studied, so it is likely that there is a margin to improve the algorithm. One potential point of improvement is to adjust the weight of the groups that selects the rows according to power and correlation, in this way the ΔMSE may reduce since the hybrid selection has proven to produce smaller errors compared to selection purely based on correlation.

Moreover, this study was based on average, so it is possible that the ideal \mathbf{T} and its selection is dependent on the channel matrix \mathbf{H} , where the ideal weight of energy and correlation can change as well. In summary, it is assumed that there is a trade-

off between energy and correlation that should be a case study to understand better the role of each one on system's performance.

7

Conclusions

Compression in the LoS massive MIMO UL channel was conducted in order to reduce the number of data streams that are sent from the RF-chains to the CPU. A maximum compression level of $C = 0.125$ was conducted. The performance of the system was evaluated in terms of ΔMSE of the recovered symbols, energy efficiency η and the condition number $\tilde{\kappa}$.

The compression was done through four different algorithms which are referred as: power, random, correlation and hybrid selection. It was shown that lower condition number and higher energy efficiency lead to lower values of ΔMSE .

The results showed that there are different solution depending on the hardware limitations. Power was the first option to be studied and it is clear that for larger groups, power is the best criterion to minimise the loss of signal quality. On the other hand, larger groups are more difficult to realise due to hardware complexity, which in turn forced the system design to implement smaller groups that lead to a poor performance.

Then, through the condition number of \mathbf{TH} it was noticed that power selection suffered from a high amount of correlation. Most of the groups selected the exact same lines, which led to an increase of the condition number and consequently, worse performance. A solution to this problem was to test and design an algorithm based on correlation which proved to have improved performance compared with power selection for smaller groups, but this algorithm had a new drawback which was that it had worse results for bigger groups when compared with power selection.

At this point it was clear that both power and correlation could affect the system. So the hybrid selection was designed in order to have half of the groups select the rows according to power and the remaining ones avoid correlation between groups with correlation selection. This method achieved smaller losses than the other algorithms for small group sizes which made it evident that both power and correlation affect the performance of the system.

Bibliography

- [1] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, M. Schellmann, H. Schotten, H. Taoka, H. Tullberg, M. Uusitalo, B. Timus, M. Fallgren. "Scenarios for 5G m and wireless communications: the vision of the METIS project", *IEEE Communications Magazine*, vol.52, no. 5, pp. 26-35, May 2014.
- [2] J. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, C. Soong, J. Zhang. (2014) "What will 5G be?", *IEEE Journal On selected Areas In Communications*, vol. 32, no. 6, pp. 1065-1082, June 2014.
- [3] Ericsson, *Ericsson Mobility Report - On the pulse of the networked society*, Nov 2015. [Online] Available: <http://www.ericsson.com/res/docs/2015/mobility-report/ericsson-mobility-report-nov-2015.pdf>
- [4] A. Gupta, R. Jha, "A survey of 5G network: Architecture and emerging technologies", *IEEE Journals & Magazines*, vol. 3, pp. 1206-1232, August 2015.
- [5] T. Eriksson, "Analog-to-digital conversion and data transportation in smart antenna systems", 2014.
- [6] S. Sanayei and A. Nosratinia, "Antenna selection in MIMO systems", *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 68-73, 2004.
- [7] X. Gao, O. Edfors, F. Tufvesson and E. Larsson, "Massive MIMO in Real Propagation Environments: Do All Antennas Contribute Equally?", *IEEE Trans. Communications*, vol. 63, pp. 3917-3928, 2015.
- [8] B. Lee and B. Shim, "An Efficient Feedback Compression for Large-Scale MIMO Systems", *2014 IEEE 79th Vehicular Technology Conference (VTC Spring)*, 2014.
- [9] D. Tse, P. Viswanath, "Fundamentals of Wireless Communications"., New York, Cambridge University Press, 2005.
- [10] K. Sayood, "Introduction to Data Compression"., Waltham, Elsevier, 2012.
- [11] A. Oppenheim, A. Willsky, S. Nawab, "Signals and Systems"., Pearson Education Limited, Harlow, 2014.

A

Appendix 1

Table A.1: Combinations for $G=4$ and $C=0.25$

Group	1	2	3	4	...
Rows	4	1	3	2	...

Table A.2: Combinations for $G=4$ and $C=0.75$

Group	1	2	3	4	...
Rows	2 3 4	1 2 3	1 3 4	1 2 4	...

Table A.3: Combinations for $G=16$ and $C=0.125$

Group	1	2	3	4	5	6	7	8	...
Rows	1 2	3 4	5 6	7 8	9 10	11 12	13 14	15 16	...

Table A.4: Combinations for $G=16$ and $C=0.25$

Group	Rows
1	1 2 3 4
2	13 14 15 16
3	5 6 7 8
4	9 10 11 12
5	1 2 5 6
6	3 4 7 8
7	9 10 13 14
8	11 12 15 16

Table A.5: Combinations for $G=16$ and $C=0.75$

Group	Rows
1	1 2 3 4 5 6 7 8 9 10 11 12
2	5 6 7 8 9 10 11 12 13 14 15 16
3	1 2 3 4 9 10 11 12 13 14 15 16
4	1 2 3 4 5 6 7 8 13 14 15 16
5	3 4 7 8 9 10 11 12 13 14 15 16
6	1 2 5 6 9 10 11 12 13 14 15 16
7	1 2 3 4 5 6 7 8 11 12 15 16
8	1 2 3 4 5 6 7 8 9 10 13 14

Table A.6: Combinations for $G=16$ and $C=0.875$

Group	Rows
1	1 2 3 4 5 6 7 8 9 10 11 12 13 14
2	3 4 5 6 7 8 9 10 11 12 13 14 15 16
3	1 2 5 6 7 8 9 10 11 12 13 14 15 16
4	1 2 3 4 7 8 9 10 11 12 13 14 15 16
5	1 2 3 4 5 6 9 10 11 12 13 14 15 16
6	1 2 3 4 5 6 7 8 11 12 13 14 15 16
7	1 2 3 4 5 6 7 8 9 10 13 14 15 16
8	1 2 3 4 5 6 7 8 9 10 11 12 15 16