

Towards pilot overhead reduction in ultra-dense networks for 5G communication systems

Master of Science Thesis in Communication Engineering

ANASTASIOS KOLONIARIS

Department of Signals and Systems
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2014
Report No. EX043/2014

**Towards pilot overhead reduction in ultra-dense networks
for 5G communication systems**

ANASTASIOS KOLONIARIS



Department of Signals and Systems
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2014

Master Thesis for the Master Programme “Communication Engineering”
Towards pilot overhead reduction in ultra-dense networks for 5G communication
systems
ANASTASIOS KOLONIARIS

© ANASTASIOS KOLONIARIS, 2014

Technical report no. EX043/2014
Department of Signals and Systems
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone + 46 (0)31-772 1000

Cover: Abstract illustration of a 5G ultra-dense deployed network, with pilot and data symbols being transmitted from the user equipment to the base stations where the channel state information is derived.

Abstract

Considering the deployment of ultra-dense networks and several other technology enablers which are expected to boost the performance of 5G communication systems, the process of channel estimation should be re-considered. In such a network, the use of pilot symbol sequences is expected to result in an overhead which could potentially overwhelm the available time and frequency domain resources. On the contrary, blind and semi-blind channel estimation approaches have been considered as a good alternative, since they minimize the potential signalling congestion and use the air interface resources in a better way. The aim of this study is to evaluate the performance of the pilot, blind and semi-blind channel estimation methods, by using the Cramér-Rao lower bound, in terms of channel estimation accuracy and Binary Phase Shift Keying constrained capacity for a Rayleigh distributed channel model. The trade-offs between the resulting useful rate and the accuracy of the channel estimates are also analysed, along the air interface resources required for the realisation of this task. The acquired findings support the idea that upon assuming high SNR scenarios and an increased number of transmitted symbols, the channel estimates obtained by the blind and semi-blind channel estimation methods, are capable of achieving acceptable throughput and could also de-congest the network from the use of dedicated pilot symbols. As a result, a constructive analysis is realised and the results support the objective of pilots reduction for the future ultra-dense deployed 5G networks.

Keywords: 5G, blind, channel estimation, constrained capacity, Cramér-Rao lower bound, pilot, semi-blind, ultra dense networks, useful rate

Acknowledgements

Upon the realisation of this thesis, I would like to thank my family and friends for the plenty of courage offered, during the two years of my studies in Sweden and Chalmers University of Technology. Courage is an important quality in our life and as the Greek philosopher Aristotle once said, *“You will never do anything in this world without courage. It is the greatest quality of the mind next to honour”*. Furthermore, I would like to thank my examiner Tommy Svensson for offering me the opportunity to realise my Master thesis in his research group and my supervisor Tilak Rajesh Lakshmana for the countless hours of brainstorming and support on the work performed. Special credits should be also given to my classmate and friend Johan Östman for all his suggestions that helped me towards finalising this thesis. Last but not least, I would like to thank Reza Khanzadi and Rahul Devassy for their willingness to guide me through specific parts of this thesis.

Anastasios Koloniaris, Gothenburg 10/8/2014

Contents

Abstract	i
Acknowledgements	iii
List of Abbreviations	vii
List of Notations	ix
Preface	xi
1 Introduction	1
1.1 The future information society and 5G	1
1.2 Candidate technologies for 5G cellular networks	2
1.3 Ultra densification and time division duplex	4
1.4 Channel estimation considerations for 5G	6
1.5 The scope of this master thesis	8
1.6 Thesis outline	10
2 CRLB and constrained capacity for a deterministic channel	11
2.1 The Cramér-Rao lower bound	11
2.1.1 Pilot-only channel estimation	13
2.1.2 Blind channel estimation	17
2.1.3 Semi-blind channel estimation	21
2.2 The estimated channel	26
2.3 Constrained capacity upon the channel estimates	26
3 System level simulations and results analysis	31
3.1 Simulation assumptions	31
3.2 Channel estimation methods' CRLB comparison	32
3.3 Trade-offs for each channel estimation method	35
3.3.1 Constrained capacity over a quasi-static channel	35
3.3.2 Ergodic constrained capacity	37
3.3.3 Useful and signalling rate	38
4 Conclusions	43
4.1 Future work	45
Bibliography	48

List of Abbreviations

3GPP	- 3rd Generation Partnership Project
4G	- 4th Generation
5G	- 5th Generation
AWGN	- Additive White Gaussian Noise
BPSK	- Binary Phase Shift Keying
BS	- Base Station
CAPEX	- Capital Expenditure
CoMP	- Coordinated Multipoint
CQI	- Channel Quality Indicator
CRLB	- Cramér-Rao Lower Bound
CSI	- Channel State Information
CSIR	- Channel State Information on Receiver
DL	- Downlink
D2D	- Device to device
E2E	- End to End
FDD	- Frequency Division Duplex
ISI	- Intersymbol Interference
KPI	- Key Performance Indicator
LS	- Least Squares
LOS	- Line Of Sight
LTE-A	- Long Term Evolution Advanced
LTE	- Long Term Evolution
METIS	- Mobile and wireless communications Enablers for 2020 Information Society
MIMO	- Multiple Input Multiple Output
MMC	- Massive Machine Communication
mmW	- millimetre Wave
MMSE	- Minimum Mean Square Error

MVU	- Minimum Variance Unbiased
NSC	- Neighbourhood Small Cell
OFDM	- Orthogonal Frequency Division Multiplexing
OPEX	- Operational Expenditure
PDF	- Probability Density Function
PDPR	- Power to Data Power Ratio
SINR	- Signal to Interference and Noise Ratio
SNR	- Signal to Noise Ratio
SR	- Signalling Rate
TC2	- Test Case 2
TDD	- Time Division Duplex
UDN	- Ultra Dense Network
UE	- User Equipment
UL	- Uplink
UR	- Useful Rate
VNI	- Visual Network Index
WiFi	- Wireless Fidelity

List of Notations

Symbol	Description
x	Scalar parameter
\mathbf{x}	Vector parameter
\mathbf{X}	Matrix
$\mathbf{X}(i,j)$	Matrix row i and column j
\hat{h}	Estimated parameter
\mathbb{C}	Complex set of numbers
$\mathbb{C}^{N \times 1}$	Complex number set of a $N \times 1$ dimension
\mathbb{R}	Real set of numbers
\mathcal{CN}	Complex Normal
$f(x; y)$	Probability Density Function
$p(x)$	Probability Mass Function
σ^2	Variance of a random variable
$\ln(\cdot)$	Natural logarithm
$\log_2(\cdot)$	Natural logarithm with base 2
$\mathbb{E}[\cdot]$	Probabilistic expectation
\mathbf{C}	Covariance matrix

Preface

The ongoing project entitled “*Mobile and wireless communications Enablers for Twenty-twenty (2020) Information Society*” known by the acronym METIS, lays the foundation for a future mobile communication system for 2020 and beyond, which belongs to the fifth generation of communication systems. Its partnership is comprised of several vendors, operators and academic organizations in which Chalmers University of Technology is also participating. Independently, of the cooperation between Chalmers University of Technology and METIS, an academical study and investigation of a 5G scenario is realised in this report.

This thesis was conducted by Anastasios Koloniaris as a part of the Master programme in Communication Engineering at Chalmers University of Technology. The thesis was realised in the department of Signals and Systems, under the supervision of Tilak Rajesh Lakshmana and the examiner Tommy Svensson.

1

Introduction

SINCE the year 2009 and the launch of the first publicly available 4th generation (4G) network, operating according to the mobile broadband standard Long Term Evolution (LTE) which is specified with its eighth release, LTE has developed and matured in a manner of deployment and technology benefits. On the other hand, it should be taken into account that the majority of the users still relies on earlier generations of mobile networks for voice and data communication due to several reasons, e.g, financial reasons or limited usage of the network. Nonetheless, due to the growth of the mobile communication, the vendors and operators are aiming to the implementation of more advanced LTE networks with new integrated technologies, based on the expected future user's demands and several other requirements specified by the 3rd Generation Partnership Project (3GPP). As a result, new releases of LTE that lead into the so-called LTE-Advanced standard are under development, with the most recent one being release 12 and the forthcoming release 13. These releases are expected to further improve the network's performance, in order for it to be able to offer the best possible user experience, while always considering a sustainable development of the technology, which is of a high importance for the future industry.

1.1 The future information society and 5G

Conjointly with the further development of LTE and LTE-A, the METIS project is based on the vision of the future long-term networked information society and it takes under consideration the challenges that are expected to arise. The forthcoming 5th generation (5G) of mobile communications, must be considered as an integrative system, which will apparently combine the already deployed wireless communication systems [1], such as LTE and WiFi, under a brand new air interface. In addition, new network requirements are expected to emerge, since both a massive growth in connected devices and the traffic volume is expected in the near future. Since the aim is the constant access to information and sharing of data, METIS has identified the following key objectives in order to address the 5G requirements [2], with decreased cost and energy consumption than the already deployed 4G networks [3]:

- 1000 times higher mobile data volume per area,
- 10 to 100 times higher number of connected devices,
- 10 to 100 times higher typical user data rate,
- 10 times longer battery life for low power Massive Machine Communication (MMC) devices,

- 5 times reduced End-to-End (E2E) latency,

while furthermore, high reliability, low cost devices and minimized energy consumption should be always considered.

The METIS research activity is already on the course, while the standardization and commercialization of 5G will begin in a few years. Thus, several realistic test cases are under examination, based on the fundamental technical challenges derived from user-related concerns, which also address a wider range of relevant problems. As it is described in [2], end-user Key Performance Indicators (KPIs), are used in order to suggest candidate solutions and derive relevant solution-specific KPIs. The end-user KPIs are taken as a basis for assessing the radio link requirements, which are characterised with several specifications such as “traffic volume density”, “experienced end-user throughput”, “latency”, “reliability and availability” and also “retainability”. Furthermore, different KPIs are assigned on different test cases resulting in 12 test cases in total. This thesis can find application on different test cases with the most relevant being the test case TC2, entitled “*Dense urban information society*”, [2]. The TC2 test case takes under consideration the connectivity related problems in dense urban environments. Furthermore, different types of information traffic are taken into account for this test case. For this kind of scenarios, it is also of a high interest to consider the continuous increment of data traffic, as predicted by the visual network index (VNI) given in [4] and a recent mobility report [5]. Moreover, along with the useful data growth the pilot overhead is expected to increased due to the pilots that are traded between a base station (BS) and the user equipment (UE) in order to establish a connection between both sides, as presented in [6]. The issue of signalling traffic growth, is greater for smartphones since their applications require real time updates in order to maintain a connection with the network. Thus, the need of a study which compares different cases of data and signalling traffic for future densified networks and the degree in which these cases affect the network’s performance, is considered to be of a high importance.

1.2 Candidate technologies for 5G cellular networks

The objectives of the METIS project that were presented before, along with the demand for higher capacity and increased data rates, can be achieved upon emerging a wide range of new candidate technologies that might lead to rethinking many cellular principles. Potential solutions are being discussed and there have already been some ideas of what to expect in 5G. Generally speaking, a new system able to provide such an improvement will have to make clever use of evolved technologies similar to the already existing ones while it must also combine new techniques and technologies, in order to achieve the next big thing that 5G is looking for. The three most important candidate technologies which can lead to changes and boost the performance of the 5th generation cellular networks, as [7] suggests, are:

1. Extensively increased bandwidth - millimetre Wave (mmW)
2. Massively parallel communication - Massive MIMO
3. Ultra-dense networks (UDN)

with additional enhancements such as *Device to Device Communication* (D2D) and a new *Device-centric Architecture* suggested in [8], while potential alternatives on the signalling and multiple access formats are further suggested in [1]. A network which combines these technologies in a clever way, will have to reassure their compatibility in order to avoid inaccuracies that might occur. Regarding their performance, the above technologies are expected to offer a 1000× increased capacity, compared to the already existing networks.

The basic principle behind this number, as mentioned in [9], is that the $1000\times$ increased capacity can be achieved as a combination of a $10\times$ better performance, $10\times$ more available spectrum and a $10\times$ more air resources reuses, which corresponds to the use of the three key technologies of massive MIMO, mmW and UDN.

The usage of signals with extensively increased bandwidth highlights the need of more available spectrum. Since the wireless frequency spectrum is over occupied, the option of using millimetre waves for data transmission should be explored. The mmWave signals correspond to 30-300 GHz frequency bands, where larger bandwidth, e.g., 1 GHz can be allocated, resulting to higher data rates [10]. Furthermore, the interference level drops significantly since the majority of the beams do not interfere, while the communication in these frequencies is also noise limited, as suggested in [1]. However, due to the small length of the wave at these frequencies, the signals will be vulnerable to several wave propagation issues such as the near field path-loss and blocking, in different indoor and outdoor environments as mentioned in [10]. This leads to the conclusion that mmWave will require important changes in the system's design, such as a simplified radio interface that benefits from shorter range communications which will occur in ultra-dense networks [1].

On the other hand, advanced antenna solutions, such as massive MIMO, is based on the realisation of spatial dimension communication which was already embodied in earlier generations of mobile communication systems. However, the contemporary massive MIMO suggests the usage of an even more increased number of antennas. Approximately hundred antennas per BS will be used and along with accurate channel estimation, vast spatial diversity can be achieved. Moreover, spatial multiplexing will realise multiple data streams for several devices [3] and the ability of radiating the energy towards desired directions with narrow beams, i.e., beamforming, is going to minimize the intra and intercell interference and improve the performance [11]. Besides, with the combined solution of mmWave and massive MIMO, problems such as the near field path-loss and blocking can be overcome since these techniques increase the diversity due to the different versions of the received signal over spatial domain [12]. As a result the use of multiple antennas at the BS is considered as a key feature in the next generation cellular networks. Moreover it achieves higher data rates by realising parallel transmissions, hence it results in an overall improved performance.

Another option which suggests the dense deployment of smaller radius cells with low powered BSs, known as *cell densification*, can assist the network to overcome similar problems (for more information see section 1.3). However, on the other hand this technique could also come into contrast with the use of massive MIMO, since the mitigation of signal interference is considered challenging for massive MIMO and in conjunction with a multicell network it might lead to undesired issues [12]. It is easily understood that these technologies should be combined in a clever way, due to the fact that ultra densification and millimetre waves are complementary, since a dense network of cells can overcome the blocking problems upon offering Line-of-sight (LOS) communication. However massive MIMO is considered to work opposite to this fact and might oppose risks, as discussed in [7]. After briefly discussing the most important candidate technologies of 5G, the next section focuses on the advantages and disadvantages that can be derived from ultra densification in 5G networks, along with the operation under time division duplex mode (TDD), which is also a candidate operation mode for the future 5G communication systems.

1.3 Ultra densification and time division duplex

The idea of ultra densification is considered a key mechanism for the wireless evolution, hence some more information should be given about what it can offer to 5G. According to the expectations among researchers and engineers, a thousand-fold growth of data traffic will occur in the near future and this will drive the need for more spectrum and higher spectral efficiency. Since it is known that the traffic is not distributed evenly in the areas which are under cell coverage, the idea of network densification suggests the deployment and combined usage of an increased number of small coverage cells, e.g., pico & femto cells under the coverage of primitive cells, e.g., macro cells. The complementary low-power cell nodes can be deployed both from the operators and also from the users, e.g., Neighbourhood Small Cell (NSC) or WiFi and they are expected to provide increased capacity. Furthermore, cell densification will result in the reuse of spectrum between the cells and less resource competence among the UEs. A metric called *Base Station Densification Gain*, presented in [1], proves that densification achieves the above goals and is related to the effective increase of the data rate, along with the escalation of BSs deployment per square kilometres. On the other hand, due to the deployment of ultra-dense networks, aspects such as the handover caused due to the user mobility, the required transmission power level, the cell coordination management and the signal to interference and noise ratio (SINR) should be taken into consideration in order to overcome any unwanted impediments, as mentioned in [1].

As an example on the importance of the key points that affect the network's performance, consider the equation (1.1) which is given in [13]. It is seen that the throughput of a user in a cellular system is upper bounded by the capacity of an additive white Gaussian noise (AWGN) channel, such that

$$R < C = m \left(\frac{W}{n} \right) \log_2 \left(1 + \frac{S}{I + N} \right) \quad (1.1)$$

in which W denotes the signal bandwidth offered from the BS, the load factor n is used to denote the number of users that share the same BS, the spatial multiplexing factor m denotes the number of spatial streams between the BS and the UE, while the argument of the logarithm corresponds to the capacity of the AWGN channel. In addition, S stands for the signal power and $I + N$ is the sum of the interference and noise power at the receiver side, resulting in the SINR. From the above equation, it can be seen that cell densification can benefit the network since it increases the amount of air interface resources due to the fact that the factor n is decreased, given that with more cells the traffic will be distributed more evenly among the UEs of each cell, as suggested in [13]. Furthermore, the same source [13] suggests that cell splitting also reduces the path loss between the transmitter and the receiver, which results in the increased signal level S and interference I , while on the contrary the impact of the noise N is diminished. Finally, in order to further reduce the interference, different techniques can be applied both on the transmitters and receivers. As a result, it is seen that cell densification can have a great impact on the improvement of the cell user's throughput.

As mentioned in [7], the already deployed cellular networks are characterised by an uneven SINR allocation, with the UEs near the BS having a desired increment in their SINR while the UEs near the cell edge have a lower SINR which is an outcome of the poor signal coverage and the increased interference on the cell edges. Focusing in interference and taking under consideration that the number of users and base stations using the network is expected to grow rapidly in 5G, this will result in increased interference levels. Hence, the interference should be either reduced thus leading to higher SINR, or the system should accept its presence and manage to use it in a way that it will

contribute to the desired signal, as mentioned in [14]. Examining the way that cell densification affects the SINR and as it was shown before in (1.1), it can be said that in contradiction with what it would have been believed, the SINR can increase with densification. This happens because in noise limited cell edges the received signal power will be greater and since the cells are serving less UEs, the resulting interference will be mitigated, as explained in [1]. Other kind of advantages achieved upon using this technique, suggest that the deployment of smaller and lower power BSs or even turning the WiFi access points into small BSs in femtocells, is a cost efficient solution. This is because it lowers the installation cost and maintains the required transmission power at low levels, thus it consequently reduces the required capital expenditure (CAPEX) and operational expenditure (OPEX), [13].

Even though cell densification promises an improvement in the network's capacity without reducing the SINR, it must be highlighted that everything comes at a cost, with several other challenges arising by using this technique. For example, mobility and handover should be reconsidered due to the fact that with many small deployed cells, the user is expected to change between cells rapidly while transferring data sessions from one cell to another. The most important problem occurs due to the interference between neighbouring cells. The problem is greater near the boundaries of pico and femto cells with macro cells [13], since different transmission power levels are used between them and this potentially leads to severe interference.

Suggested solutions that can mitigate this problem and use it in a beneficial way are the coordinated multipoint transmission and reception (CoMP), the cell range expansion technique and the usage of advanced interference cancellation receivers as explained in [13]. On the contrary, it is believed that the usage of traditional interference management techniques between cells operating under frequency division duplex (FDD) mode, will not be applicable when the number of UEs and BSs is going to increase. This is due to the fact that in FDD the traffic of the uplink (UL) and downlink (DL) is considered asymmetric and results in inefficient resources usage [7]. As it was stated before, upon considering an ultra-dense deployed network with massive MIMO, different cases of signals exchange between the BSs and the UEs should be examined. For example, in order to achieve the desired outcomes, the 5G communication networks should have a very good knowledge of the wireless channel, hence the techniques used for the channel estimation are of a high interest. Massive MIMO can in principle assist the system in such cases by transmitting orthogonal reference signals from each antenna element in different frequencies and time slots, however as a drawback, this will increase the signalling overhead since its value grows linearly with the number of transmit antennas, as mentioned in [12].

To overcome this kind of problems, according to the same source, 5G is expected to operate only under TDD. TDD operation assigns different time slots for the communication between the receiver and the transmitter, hence it uses the same frequency for DL and UL in contrast with LTE and LTE-A which support both FDD and TDD operation where the communication between the BS and UEs takes place over the frequency and time domain, as mentioned in [15] and [16]. One of the advantages of TDD usage, is that channel reciprocity between the BSs and UEs is better utilized allowing full knowledge of the channel information. This assists into attaining full precoding gains via UL channel estimation signals, something that does not happen when FDD is used because of the different carrier frequencies of UL and DL. In addition, with the usage of massive MIMO the time needed for the transmission of channel estimation signals (for more information see section 1.4) in TDD operating networks, is proportional to the number of the served UE antennas and independent of the BS antennas, thereupon it allows the number of the BS antennas to be increased without negative consequences on the channel estimation process, [12].

On the other hand, upon considering cell densification and under TDD operation, the channel estimation issue is still considered as a complicated task, due to the interference that might be caused between the reference signals from different UEs belonging to different cells. According to [12] this interference occurs due to the usage of non-orthogonal pilot signals between adjacent cells or even because of the reuse of orthogonal pilots, resulting in limited interference rejection performance.

1.4 Channel estimation considerations for 5G

As it was briefly explained in section 1.3, it is of a high importance to obtain accurate channel state information (CSI) which correspond to the properties of the multipath channel, i.e., fading, scattering and power decay between BSs and UEs. Due to the multipath propagation, the received signal is comprised of several copies from the originally transmitted signal arriving from different paths, where each of them faces different delay time and it is affected from the channel fading coefficient. Since the structure of the channel changes over time and frequency, i.e., time and frequency selective fading, the knowledge of these channel properties can assist the network into realizing adaptive transmissions based on the channel conditions, e.g., equalization process. The aim of the equalization process is to mitigate the channel impairments and minimize the detection error by subsequently adapting and evolving with the channel changes [17]. Thus with this way, the network can achieve the desired reliability of the communication link and eliminate the intersymbol interference (ISI) that might occur. The network can furthermore improve its performance upon changing the transmission data rate and coding, with the usage of specific channel quality indicators (CQI).

In order to succeed in the above challenges, there exist different methods that can be utilized to estimate the channel. These are the *pilot channel estimation*, the *blind channel estimation* method and the *semi-blind channel estimation* method. The existing mobile communication networks have knowledge of the CSI on the receiver side and feedback with the CSI to the transmitter is required. This process occurs on the DL where the UE derives the CSI and feeds back the information to the BS. Once the BS acquires knowledge about the CSI, it adapts its transmission characteristics and transmits the useful information data. In 5G however, assuming operation under TDD mode, the system can highly benefit from obtaining CSI on the UL receiver side, i.e., CSIR on the BS, without the need of CSI feedback. Furthermore, the shared knowledge of the reciprocal channel also on the BS side is expected to improve the network's performance during DL transmissions, by allowing beamforming, multi user precoding and more accurate transmissions on the spatial, time and frequency domain. Thus in 5G, a combination of massive MIMO with TDD leads to channel reciprocity due to the fact that the DL and UL use the same frequency. Hence, the effort of obtaining the channel properties is decreased and the gains are increased due to the use of more antennas, as suggested in [18]. Figure 1.1 illustrates the differences in the channel estimation process between FDD mode and TDD mode. In FDD mode, the channel is estimated after transmitting pilot signals on the DL frequency and the CSI is transmitted back on the different UL frequency. In contrast, under TDD operation the CSI feedback is not required, due to the channel reciprocity.

Further discussing the properties and performance of the methods used for channel estimation, the transmission of cell-specific reference signals, known as pilot sequences, is the most well known method which allows to gain knowledge of the channel properties. The reference signals are comprised of pilot symbols of predefined values which are located in specific orthogonal frequency division multiplexing (OFDM) symbols per resource block [15]. The OFDM transmission is used in order to combat the multipath

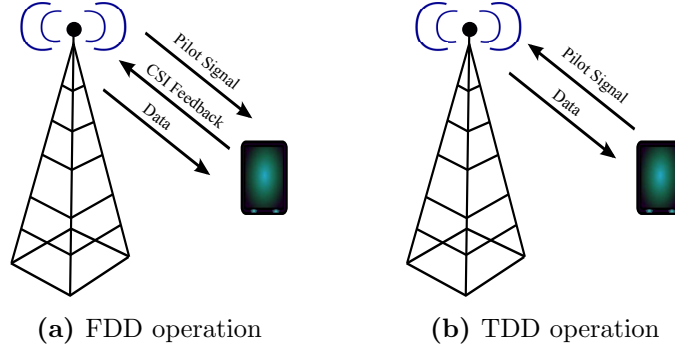


Figure 1.1: Comparison of the channel estimation process between BS and UE for FDD and TDD operation for downlink transmission. Operation on FDD mode requires the transmission of feedback signals, while TDD does not use feedback signals.

scattering [17]. This results in a pilot sequence which spreads over the time and frequency domain. As mentioned in [16], there exist 504 different types of pilot sequences in LTE and LTE-A, obtained from different frequency shifts and based on the available resource blocks. The pilot symbols can be either transmitted at the beginning of the data bursts or they can be superimposed on top of the information data [15]. Finally, since the UE has prior knowledge of the transmitted pilot sequence, it correlates the received set of symbols with the known sequence.

In addition and as it was previously mentioned, the usage of non-orthogonal pilot sequences in TDD might cause interference between neighbouring cells [19]. Especially upon assuming 5G scenarios, with the deployment of UDNs, resources are expected to be over occupied from pilot signals, hence a better alternative should be examined. Besides, as [19] mentions, under multi-cell scenarios the usage of orthogonal pilot sequences will require an extended length of pilot symbols equal to the number of cells times the number of single antenna users. This leads to the need of non-orthogonal pilot sequences utilization. This decision is taken considering the fact that long pilot sequences are not a good option due to the mobility of the UE. To visualise the above facts, an abstract comparison over a number of time and frequency domain resources and the occupation caused due to pilot traffic on the already deployed networks versus the potential pilot traffic for ultra-dense 5G networks is presented in Figure 1.2. Assuming uncoordinated BS deployment and viewing the air interface from the BS side, it can be seen that the usage of pilot signals from an extended number of UEs can decrease the spectral efficiency of the system and over occupy the time and frequency resources. Considering the above facts, alternative types of channel estimation methods should be considered and evaluated in order to examine if the system can become more spectral efficient in such a way that it will consume less time and frequency resources and it will also allow the channel estimation process to be performed properly.

The second channel estimation method considered in this report, is the blind channel estimation also known as self-recovering. This method does not require prior knowledge of a pilot sequence at the receiver side and it is based on extracting the channel information from the received data symbols only [20]. Two types of blind channel estimation methods exist, the *stochastic maximum likelihood estimation* [20] where the received random symbols are modelled with their known distribution and the *deterministic maximum likelihood* [20] where no statistical model is assumed for the received symbols. As it will be seen later in Chapter 2, only the stochastic maximum likelihood type of blind estimation is considered in this report. Blind channel estimation is proved to be useful for example during on-line transmissions, where the transmission of pilot signals for channel

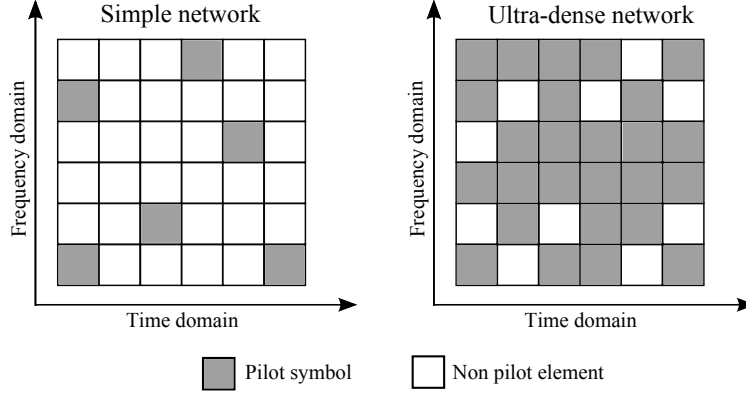


Figure 1.2: The effect of pilot signals viewed on the BS side, over time and frequency resources in an ultra-dense network, assuming uncoordinated BS deployment.

estimation is not possible while constantly transmitting information data. As a result, during the online data transmission, the receiver can monitor the channel from the data itself, upon using their statistics. This leads to a better utilization of the bandwidth since the usage of extra transmitted pilot symbols is avoided, thus allowing more time and frequency resources to be used for data transmission. However, the extraction of the CSI is considered to be a more difficult task than with the usage of pilot sequences and furthermore the channel estimates are of a lower quality when compared to pilot based estimation. More information about how blind channel estimation operates can be found in [20].

The final considered method of channel estimation is entitled *semi-blind channel estimation*. This method aims into estimating the channel characteristics with a combination of prior known symbols and the observation which corresponds to the unknown information data [20]. This method has gained popularity due to the fact that blind channel estimation methods have been proven to give deficient results while on the other hand pilot based channel estimation leads to poor bandwidth utilization. The main idea behind semi-blind channel estimation is based on the fact that in each transmission there exist symbols in each transmitted packet which are known to the receiver and they can be used in order to estimate the channel without the need of transmitting additional pilot symbols, as mentioned in [21]. Furthermore, it has been seen that the selection of the portions which correspond to the transmitted number of pilot and data symbols in every burst, is a key factor since it allows good channel exploitation and must also be bandwidth and energy efficient as suggested in [22] and [23]. The same source suggests the use of different power levels for the transmission of the pilot and data symbols, known as pilot to data power ratio (PDPR), which can assist the system into providing better accuracy of channel estimates.

1.5 The scope of this master thesis

After briefly discussing the three channel estimation methods based on the assumption of TDD operation for 5G ultra-dense networks, this section aims into providing information about the project's scope and the outcomes that have been identified. The three channel estimation methods are first compared in terms of the accuracy of the estimates they provide and the results are extended in order to derive the corresponding constrained capacity of each method. The achievable useful rate by each method is finally compared. The main target is to examine if the traditional channel estimation obtained by pilot sequences can be revised with semi-blind, or even better blind channel estimation meth-

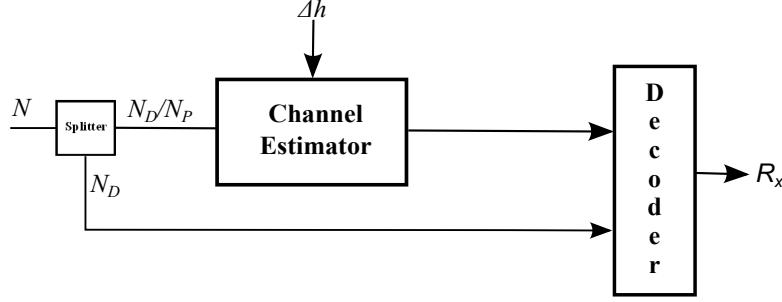


Figure 1.3: Block diagram of the channel estimation process and the derived rate from pilot P or data D symbols from each channel estimation method.

ods, in order to use the air interface in a more efficient way. The issue of inefficient use of time and frequency resources, due to the overhead created from signalling, is of a high importance for the 5G ultra-dense networks. Thus, channel access with minimized signalling resulting in zero-overhead communication is investigated. As follows, we discuss the trade off between insufficient channel training, which results in poor channel estimation and lower achievable constrained capacity, while on the contrary when more training is used than typically required, less time and frequency resources are left for transmission of the actual data before the channel fading coefficients changes.

In order to be able to evaluate the channel estimation methods, a system model is first defined. This model takes under consideration the transmitted and received symbols, the effect of the channel fading on them and the random noise that occurs. The channel fading coefficient of the model is assumed to be an accurate estimate of the channel, which is available on the receiver with the usage of a channel estimator and detector. The channel estimator and detector can either be a Least-Squares (LS) method, which is simple to implement but also sensitive to noise, or a Minimum Mean Square Error (MMSE) method which is more robust to noise, however it requires knowledge of the covariance. The consideration and implementation of a channel estimator is beyond the scope of this thesis, thus the channel estimate is considered deterministic. The channel's fading coefficients, can be considered as Rician, Nakagami- m or Rayleigh distributed, depending on the scenario, while new channel models are under development in order to consider all the effects that might occur with the usage of millimetre waves and massive MIMO in 5G [10]. It must be also noted that the Rician fading can consider LOS transmission [24] as a result of the increased cell deployment. On the other hand, the Rayleigh distributed channel assumes no LOS communication.

The block diagram of Figure 1.3 illustrates the general procedure followed in order to derive the channel estimate and the rate obtained by each channel estimation method. Assuming a received symbol vector of length N , the symbols are fed in a splitter which splits the pilot symbols N_P and the data symbols N_D . The pilot symbols are correlated with a known sequence and the channel estimate is obtained. Depending on the channel estimation method, the data symbols can also be used in order to obtain the channel estimates. The only factor that affects the channel estimate is Δh , a random variable with a variance obtained by the Cramér-Rao lower bound (CRLB), a method which calculates the lowest possible bound on the variance of the channel estimates obtained by unbiased estimators. Once the CRLB is derived, a new channel characterization is obtained, including the original deterministic channel along with the new channel variance given by the CRLB. Upon considering this new channel we are able to calculate the achievable constrained capacity from each channel estimation method for a quasi-static channel and examine how it changes for different signal to noise ratio (SNR) values. The whole process is repeated in order to also calculate the ergodic constrained capacity

of a channel with frequency selective and fast fading conditions.

However, during the channel estimation process, the pilot symbols do not provide a useful rate, i.e., transmission of information data, thus create a pilot overhead which is expected to increase in future 5G ultra-dense networks. The corresponding useful rates of the pilot based, semi-blind and blind channel estimation method are calculated upon considering the number of pilot and data symbols which are used by each method. As a result, we are able to compare these rates and conclude if the usage of other channel estimation methods, rather than the pilot only method, is worth using in order to reduce the overhead.

1.6 Thesis outline

The rest of this thesis is structured as follows. In Chapter 2, we present and discuss the mathematical methods and calculations which were used for deriving the CRLB of the estimate variance. The same chapter presents the mathematical expressions that were used in order to calculate the corresponding constrained capacity of each method, over a quasi-static channel. The corresponding ergodic constrained capacity of each method is also discussed. Chapter 3 analyses the system level simulations and presents the results and findings with relevant illustrations and tables. The three methods are compared, based on the accuracy of their channel estimates and the corresponding constrained capacity for one random channel realisation. The capacity is compared along different values, such as the SNR level and the number of received symbols. The same chapter discusses the useful and signalling rate obtained for each method, thus the three channel estimation methods are once more compared with respect to these rates. Finally, in Chapter 4 the conclusions of this thesis are presented and we provide our remarks towards which channel estimation method can be potentially used for future 5G ultra-dense deployed networks. Furthermore, relevant future work is also suggested.

2

CRLB and constrained capacity for a deterministic channel

THIS chapter presents to the reader the calculations that were realised in order to obtain any mathematical characterizations needed for the actual simulations. The first aim of this chapter is to explain how the Cramér-Rao Lower Bound assists in deriving a comparison on the accuracy of the channel estimates provided by each channel estimation methods and thus a metric for their performance is obtained. Furthermore, an analysis of the mathematical formulas used for calculating the constrained channel capacity is realised. The derived equations of this chapter are used in Chapter 3, in order to simulate at a system level and obtain the results required for the scope of the thesis.

2.1 The Cramér-Rao lower bound

The CRLB serves as a useful mathematical tool that can provide a benchmark from which the performance of different estimators can be compared. This performance is measured in terms of the variance that the estimated parameter can achieve. The lower bound of the variance that is obtained from the CRLB, corresponds to the Minimum Variance Unbiased estimator (MVU). Assuming that the estimators are unbiased and defined as $E[\hat{\theta}] = \theta$, where θ is the unknown parameter, $\hat{\theta}$ is the estimated parameter and $E[\cdot]$ represents the probabilistic expectation. The true value of the parameter can be obtained when the estimates are averaged, as [25] suggests. In addition, given that the estimator is unbiased, the next step is to minimize its variance and derive the MVU which minimizes the estimation error $\hat{\theta} - \theta$. The MVU attains the CRLB, for each value of θ and it is proven to be the most efficient estimator of the unknown parameter.

In order to calculate the CRLB and evaluate the performance of the three channel estimation methods, the first step is to define a single input single output (SISO) system model which considers the transmitted and received symbols, the channel fading coefficient, i.e., the channel gain, and the noise that occurs. The system model [17] that will be used in the following pages, expressed in scalars form, i.e., assuming the transmission of only one symbol, is given such that,

$$y = hx + n \tag{2.1}$$

where $y \in \mathbb{C}^{1 \times 1}$ corresponds to the received complex valued sample and $x \in \{\pm 1\}$ is the transmitted symbol for which Binary Phase Shift Keying (BPSK) modulation is as-

sumed. Depending on the channel estimation method, each symbol at the receiver side can be considered to be either known, corresponding to a pilot symbol, or unknown, corresponding to a data symbol, hence different probabilities are assigned for each method. Furthermore, $h \in \mathbb{C}^{1 \times 1}$ denotes the quasi-static channel fading coefficient between the transmitter and the receiver. On the following pages, the channel h is a random sample that follows a Rayleigh distribution and it is derived by taking the square root of two added independent and identically distributed (*iid*) zero-mean Gaussian random variables. Finally, $n \sim \mathcal{CN}(0, \sigma_n^2)$ corresponds to *iid* circularly symmetric complex Gaussian random noise, with zero mean and variance σ_n^2 . It must be also highlighted that the system model can be given in vector form, where a column vector of symbols is assumed to be transmitted. For each channel estimation method, we first provide the mathematical analysis of the model in scalars form and then we extend the results into vector form. More analytical models and calculations for each case are given in the following sections.

Since one part of this thesis focuses on the study of the channel estimates accuracy provided by each method, the parameter of interest and the one to be estimated is the channel fading coefficient. Given that the channel is considered as complex valued, it can be represented as the sum $h = h_r + jh_i$, where h_r denotes the real part of the channel and h_i the imaginary part respectively, while the conjugate is given by $h^* = h_r - jh_i$. Consequently, in order to be able calculate the CRLB for both parts of the complex channel, the vector $\boldsymbol{\theta}$ is considered. This vector is comprised of two values, the real and the imaginary part of the channel scalar h , such as $\boldsymbol{\theta} = [\text{Re}\{h\} \text{Im}\{h\}]^T = [h_r \ h_i]^T$, hence an extension of the CRLB for the estimation of a vector parameter will be used.

To start with, the initial information that can be used, is a mathematical model of the observed data described by their probability density function (PDF). In order to obtain accurate estimates of the channel fading coefficient from (2.1), the PDF should be dependent on the parameter to be estimated, $\boldsymbol{\theta}$. Generally speaking, the accuracy of the estimates is affected by the degree in which the unknown parameter influences the PDF. Thus, given that the noise n is modelled as circularly symmetric complex Gaussian and considering that the input symbol x is affected by the channel fading coefficient h , the CRLB as introduced in [25], requires the derivation of the likelihood function $f(y; \boldsymbol{\theta})$, which is simply the PDF of the observed sample y parametrized by $\boldsymbol{\theta}$, as mentioned in [25]. Considering that we are interested in estimating the vector parameter $\boldsymbol{\theta}$, the CRLB is defined as

$$\text{var}(\hat{\theta}_i) \geq [\mathbf{I}^{-1}(\boldsymbol{\theta})]_{ii} \quad (2.2)$$

for $i = [1, 2]$ and $\mathbf{I}(\boldsymbol{\theta})$ is the Fisher information matrix of a 2×2 dimension, due to the fact that the vector $\boldsymbol{\theta} \in \mathbb{R}^{2 \times 1}$ and the CRLB is derived from the diagonal elements $[i, i]$ of the inverse Fisher information matrix. The Fisher information matrix measures the information for the unknown parameter $\boldsymbol{\theta}$ which is included in a random variable and affects its probability. It is defined such that [25],

$$[\mathbf{I}(\boldsymbol{\theta})]_{ij} = -\mathbb{E} \left[\frac{\partial^2 \ln f(y; \boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} \right] \quad (2.3)$$

and since $\boldsymbol{\theta} = [h_r \ h_i]^T$, it follows that the Fisher information matrix is given,

$$\mathbf{I}(\boldsymbol{\theta}) = \begin{bmatrix} -\mathbb{E} \left[\frac{\partial^2 \ln f(y; \boldsymbol{\theta})}{\partial h_r^2} \right] & -\mathbb{E} \left[\frac{\partial^2 \ln f(y; \boldsymbol{\theta})}{\partial h_r \partial h_i} \right] \\ -\mathbb{E} \left[\frac{\partial^2 \ln f(y; \boldsymbol{\theta})}{\partial h_i \partial h_r} \right] & -\mathbb{E} \left[\frac{\partial^2 \ln f(y; \boldsymbol{\theta})}{\partial h_i^2} \right] \end{bmatrix} \quad (2.4)$$

where the negative expectation is taken with respect to $f(y; \boldsymbol{\theta})$. In addition the calculation of the first and the second derivative of the likelihood function's logarithm should

be first derived. Furthermore, if the following condition is satisfied [25],

$$\frac{\partial \ln f(y; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \mathbf{I}(\boldsymbol{\theta}) (\mathbf{g}(y) - \boldsymbol{\theta}) \quad (2.5)$$

then the unbiased estimator attains the CRLB and the MVU estimator is $\hat{\boldsymbol{\theta}} = \mathbf{g}(y)$, while the minimum variance equals to $\mathbf{I}(\boldsymbol{\theta})^{-1} = \mathbf{C}_{\hat{\boldsymbol{\theta}}}$, which is the covariance matrix.

The following sections present the calculations that were carried out, in order to derive the CRLB for each channel estimation method. At the beginning of each section the system model is analysed, along with its components and the assumptions that were used. The calculations are first based on the system model expressed by (2.1) which is given in scalar form and then an extension of the calculations and the results corresponding to the same system model in vector form is also provided.

2.1.1 Pilot-only channel estimation

Scalar parameters model

The method of pilot channel estimation assumes the reception of known symbols at the receiver side, as discussed in the introduction of this thesis. Upon the reception of only one symbol, the system model is given in scalars form by (2.1). The transmitted symbol is modelled as $x \in \{+1\}^{1 \times 1}$ and constant, due to the fact that x corresponds to a known symbol. The channel entry h is modelled as zero mean complex Gaussian with unit variance. Moreover and as previously explained, the contribution of the noise n is assumed to be *iid* circularly symmetric complex Gaussian additive noise, $n \sim \mathcal{CN}(0, \sigma_n^2)$ where the noise variance σ_n^2 depends on the SNR, such that

$$\text{SNR} = \frac{E_b}{N_0} = \frac{1}{\sigma_n^2}, \quad (2.6)$$

where E_b corresponds to the energy per bit. By using (2.6), we are able to derive the CRLB for different values of SNR, corresponding to different noise variances. For the sake of simplicity, in the following calculations we do not assume a specific SNR thus the variance notation σ_n^2 is not changed to a specific value.

The likelihood function derived from the PDF of the observed value y , which depends on the parameter to be estimated h , is given for circularly symmetric Gaussian scalars and according to [26] it will be

$$f(y; h) = \frac{1}{\pi \sigma_n^2} \exp \left(-\frac{|y - hx|^2}{\sigma_n^2} \right) \quad (2.7)$$

while the log likelihood function of (2.7), is given by

$$\begin{aligned} \ln f(y; h) &= \ln \left(\frac{1}{\pi \sigma_n^2} \right) - \left(\frac{|y - hx|^2}{\sigma_n^2} \right) \\ &= \ln \left(\frac{1}{\pi \sigma_n^2} \right) - \left(\frac{1}{\sigma_n^2} (|y|^2 - y^* h x - h^* x^* y + |h|^2 x^2) \right) \end{aligned} \quad (2.8)$$

where $\ln(\cdot)$ corresponds to the natural logarithm while it is given that $|y|^2 = y^* y$ and $|h|^2 = h^* h$. Upon writing the channel h as $h = h_r + j h_i$ with its conjugate being

$h^* = h_r - jh_i$, (2.8) can be written as

$$\ln f(y; h) = \ln \left(\frac{1}{\pi \sigma_n^2} \right) + \left(-\frac{1}{\sigma_n^2} (|y|^2 + h_r^2 x^2 + h_i^2 x^2 - y^* h_r x - j y^* h_i x - y h_r x^* + j y h_i x^*) \right). \quad (2.9)$$

Since the channel fading coefficient h , is complex valued, for the rest of the report, we consider the estimation parameter vector $\boldsymbol{\theta} = [\text{Re}\{h\} \text{Im}\{h\}]^T = [h_r \ h_i]^T$. Upon using an extension of the CRLB estimation for vector parameters and the Fisher information matrix given in (2.4), as explained in [25], the differentiation should be taken twice, with respect to the real part of the channel, such that $\frac{\partial^2}{\partial^2 h_r}$ and the imaginary part, such that $\frac{\partial^2}{\partial^2 h_i}$ and also with respect to both the real and imaginary part, such that $\frac{\partial^2}{\partial h_r \partial h_i}$ and $\frac{\partial^2}{\partial h_i \partial h_r}$. Differentiating once (2.9) with respect to the real part of the channel h_r , gives

$$\frac{\partial \ln f(y; \boldsymbol{\theta})}{\partial h_r} = -\frac{1}{\sigma_n^2} (2h_r x^2 - y^* x - y x^*) \quad (2.10)$$

while the second differentiation, again with respect to the real part of the channel, derives

$$\frac{\partial^2 \ln f(y; \boldsymbol{\theta})}{\partial h_r^2} = \frac{-2x^2}{\sigma_n^2}. \quad (2.11)$$

Repeating the above calculations with respect to the imaginary part of the channel h_i , will also result in

$$\frac{\partial \ln f(y; \boldsymbol{\theta})}{\partial h_i} = -\frac{1}{\sigma_n^2} (2h_i x^2 - y^* x - y x^*) \quad (2.12)$$

and

$$\frac{\partial^2 \ln f(y; \boldsymbol{\theta})}{\partial h_i^2} = \frac{-2x^2}{\sigma_n^2}. \quad (2.13)$$

Since the differentiation of (2.9) with respect to both the real and imaginary part of the channel results in zero, it is obtained

$$\frac{\partial^2 \ln f(y; \boldsymbol{\theta})}{\partial h_r \partial h_i} = 0 \quad \text{and} \quad \frac{\partial^2 \ln f(y; \boldsymbol{\theta})}{\partial h_i \partial h_r} = 0 \quad (2.14)$$

while upon taking the negative expectation of equations (2.11) and (2.13) with respect to y , it results in

$$-\mathbb{E} \left[\frac{\partial^2 \ln f(y; \boldsymbol{\theta})}{\partial h_r^2} \right] = \frac{2x^2}{\sigma_n^2} \quad \text{and} \quad -\mathbb{E} \left[\frac{\partial^2 \ln f(y; \boldsymbol{\theta})}{\partial h_i^2} \right] = \frac{2x^2}{\sigma_n^2} \quad (2.15)$$

hence the Fisher information matrix of (2.4) will result in a diagonal matrix, such that

$$\mathbf{I}(\boldsymbol{\theta}) = \begin{bmatrix} \frac{2x^2}{\sigma_n^2} & 0 \\ 0 & \frac{2x^2}{\sigma_n^2} \end{bmatrix}. \quad (2.16)$$

Finally, in order to derive the CRLB, the Fisher information matrix is inverted such that

$$\text{var}(\hat{h}) = \mathbf{I}^{-1}(\boldsymbol{\theta}) = \begin{bmatrix} \frac{\sigma_n^2}{2x^2} & 0 \\ 0 & \frac{\sigma_n^2}{2x^2} \end{bmatrix} \quad (2.17)$$

and the lower bound of the estimated parameter's variance is found by taking the diagonal elements, as shown in (2.2), where the fraction in the first column and the first row corresponds to the variance of the real part of the estimated channel while the second fraction on the second row and second column corresponds to the variance of the imaginary part. Finally, as it can be seen in (2.17), for the pilot sequence channel estimation method, the CRLB is the same for the real and imaginary part of the channel and it depends only on the variance of the noise which is the only random variable, since the channel is assumed to be deterministic.

Extension to vector parameters

For vector parameters it is generally considered that for a number N of *iid* observations, the CRLB is given by $1/N$ times the CRLB of a single observation, obtained by the inverse Fisher information matrix. Thus, when assuming vector parameters, where the length of the vector is given by the number N , the resulting CRLB is $1/N$ of the CRLB obtained for a scalar parameter. It can generally be understood that when more symbols are received, the lower bound of the variance decreases and the resulting channel estimates are more accurate.

Once again the system model is the same as in (2.1) but as it was explained before, this time it is comprised of vector elements instead of scalars. In vector form it is given such that $\mathbf{y} = \mathbf{h}\mathbf{x} + \mathbf{n}$, where the vector $\mathbf{y} = [y_1, y_2, \dots, y_N]^T$ corresponds to the received samples, $\mathbf{h} = [h, h, \dots, h]^T$ is the deterministic channel fading vector with each element of the vector being the same, since the channel is considered quasi-static. Furthermore, the vector \mathbf{x} contains the known pilot sequence comprised of BPSK symbols such that $\mathbf{x} \in \{+1\}^{N \times 1}$ and $\mathbf{n} = [n_1, n_2, \dots, n_N]^T$ is the noise vector whose entries are modelled as *iid* circularly symmetric complex Gaussian and additive, with zero mean and σ_n^2 variance. Thus $n \sim \mathcal{CN}(0, \sigma_n^2)$, where the variance can be obtained by (2.6). Moreover \mathbf{y} and $\mathbf{h} \in \mathbb{C}^{N \times 1}$.

Following the same procedure as in [25], the likelihood function for this method is derived again from the PDF of the observable and it is first given in vectors form (2.18), as suggested in [26] and afterwards in an explicit scalar form (2.19) by

$$f(\mathbf{y}; \boldsymbol{\theta}) = \frac{1}{\pi^N \det(\mathbf{C})} \exp(-(\mathbf{y} - \mathbf{h}\mathbf{x})^H \mathbf{C}^{-1} (\mathbf{y} - \mathbf{h}\mathbf{x})) \quad (2.18)$$

$$= \prod_{n=0}^{N-1} \frac{1}{\pi \sigma_n^2} \exp\left(-\frac{|y[n] - h[n]x[n]|^2}{\sigma_n^2}\right) \quad (2.19)$$

where \mathbf{C} in (2.18) corresponds to the covariance matrix which equals to $\sigma_n^2 \mathbf{I}_{N \times N}$ and the superscript H denotes Hermitian transpose. In addition, it can be seen that the vector elements can be written into scalars form, by using the product of all the N elements. The log-likelihood function becomes,

$$\begin{aligned} \ln f(\mathbf{y}; \boldsymbol{\theta}) &= \sum_{n=0}^{N-1} \left[\ln\left(\frac{1}{\pi \sigma_n^2}\right) - \frac{(y[n] - h[n]x[n])^* (y[n] - h[n]x[n])}{\sigma_n^2} \right] \\ &= \sum_{n=0}^{N-1} \left[\ln\left(\frac{1}{\pi \sigma_n^2}\right) - \frac{(|y[n]|^2 - y[n]^* h[n]x[n] - y[n]h[n]^* x[n]^* + |h[n]|^2 |x[n]|^2)}{\sigma_n^2} \right] \end{aligned} \quad (2.20)$$

where the product changes to a summation due to logarithmic operation on *iid* samples, [25]. By changing $h = h_r + jh_i$ and its conjugate in (2.20) and after taking the first derivative with respect to the real part of the channel, it is found that

$$\frac{\partial \ln f(\mathbf{y}; \boldsymbol{\theta})}{\partial h_r} = \sum_{n=0}^{N-1} \frac{-2h_r[n]x[n]^2 + y[n]^*x[n] + y[n]x[n]^*}{\sigma_n^2} \quad (2.21)$$

and the second differentiation again with respect to the channel's real part derives

$$\frac{\partial^2 \ln f(\mathbf{y}; \boldsymbol{\theta})}{\partial h_r^2} = \sum_{n=0}^{N-1} -\frac{2x[n]^2}{\sigma_n^2}. \quad (2.22)$$

Upon taking the negative expectation with respect to \mathbf{y} and since both the expectation operator and the summation are linear operations, it is seen that

$$-\mathbb{E} \left[\frac{\partial^2 \ln f(\mathbf{y}; \boldsymbol{\theta})}{\partial h_r^2} \right] = -\sum_{n=0}^{N-1} \mathbb{E} \left[-\frac{2x[n]^2}{\sigma_n^2} \right] \quad (2.23)$$

and the Fisher information is scaled by the number of observations N , since the observations are assumed to be identically distributed, thus it becomes

$$I(\theta) = N \frac{2\mathbf{x}^2}{\sigma_n^2}. \quad (2.24)$$

Now if the steps in equations (2.21), (2.22) and (2.23) are repeated with respect to the imaginary part of the channel, it will be derived

$$-\mathbb{E} \left[\frac{\partial^2 \ln f(\mathbf{y}; \boldsymbol{\theta})}{\partial h_i^2} \right] = N \frac{2\mathbf{x}^2}{\sigma_n^2} \quad (2.25)$$

and as mentioned above the differentiation with respect to $\frac{\partial^2}{\partial h_r \partial h_i}$ and $\frac{\partial^2}{\partial h_i \partial h_r}$ results in zero, hence the Fisher information matrix becomes

$$\mathbf{I}(\boldsymbol{\theta}) = \begin{bmatrix} \frac{2N\mathbf{x}^2}{\sigma_n^2} & 0 \\ 0 & \frac{2N\mathbf{x}^2}{\sigma_n^2} \end{bmatrix}. \quad (2.26)$$

As a result, the CRLB for the N length pilot symbols sequence, is found similar to (2.2) with the only difference being the scaling factor $1/N$, which is given by

$$\text{var}(\hat{\boldsymbol{\theta}}) = \mathbf{I}^{-1}(\boldsymbol{\theta}) = \begin{bmatrix} \frac{\sigma_n^2}{2N\mathbf{x}^2} & 0 \\ 0 & \frac{\sigma_n^2}{2N\mathbf{x}^2} \end{bmatrix}. \quad (2.27)$$

The result of (2.27) shows that the CRLB is not affected by the value of the BPSK symbols while furthermore it is only depending on the noise variance, as it was previously explained.

2.1.2 Blind channel estimation

Scalar parameters model

The blind channel estimation method considers the stochastic maximum likelihood estimation, thus the received symbol is assumed to be following a specific distribution. Based on this fact and upon considering the same system model of (2.1), the transmitted data symbol is given such that $x \in \{+1, -1\}^{1 \times 1}$ where x can take two values, due to the fact that it corresponds to a BPSK modulated symbol which is uniformly distributed with probability of occurrence $\Pr(x = +1) = \Pr(x = -1) = \frac{1}{2}$. Since the received symbol is considered random, it can be characterised by its distribution, which is the key point of the blind channel estimation method. Furthermore, the noise n is assumed as *iid* circularly symmetric complex Gaussian additive noise $n \sim \mathcal{CN}(0, \sigma_n^2)$, with the variance obtained from (2.6) as before.

The likelihood function for the blind channel estimation method, parametrized by the channel h according to [26] and [20] will be

$$\begin{aligned} f(y; \boldsymbol{\theta}) &= \sum_x f(y; \boldsymbol{\theta}|x) p(x) \\ &= \frac{1}{2} f(y; \boldsymbol{\theta}|x = +1) + \frac{1}{2} f(y; \boldsymbol{\theta}|x = -1) \end{aligned} \quad (2.28)$$

where $p(x)$ corresponds to the probability mass function of x and for the sake of simplicity the symbol scalar x is not changed to its corresponding value, in order to examine how the symbol affects the CRLB estimation. The PDF is parametrized by the vector $\boldsymbol{\theta}$ and the likelihood function is given such that

$$\begin{aligned} f(y; \boldsymbol{\theta}) &= \frac{1}{\pi \sigma_n^2} \left[\frac{1}{2} \exp \left(\frac{-|y - hx|^2}{\sigma_n^2} \right) + \frac{1}{2} \exp \left(\frac{-|y + hx|^2}{\sigma_n^2} \right) \right] \\ &= \frac{1}{\pi \sigma_n^2} \left[\frac{1}{2} \exp \left(-\frac{|y|^2 - h^* x^* y - y^* h x + x^* x |h|^2}{\sigma_n^2} \right) \right] \\ &\quad + \frac{1}{\pi \sigma_n^2} \left[\frac{1}{2} \exp \left(-\frac{|y|^2 + h^* x^* y + y^* h x + x^* x |h|^2}{\sigma_n^2} \right) \right]. \end{aligned} \quad (2.29)$$

Since the logarithm of a sum requires complicated calculations, it is preferred to bring the inner product of the brackets into a form that will allow easier calculations. Thus, by expressing the inner arguments of the two exponentials of equation (2.29) in a different way, it can be derived

$$\begin{aligned} f(y; \boldsymbol{\theta}) &= \frac{1}{\pi \sigma_n^2} \frac{1}{2} \left[\exp \left(-\frac{\overbrace{|y|^2}^a}{\sigma_n^2} + \frac{\overbrace{h^* x y}^b}{\sigma_n^2} + \frac{\overbrace{y^* h x}^c}{\sigma_n^2} - \frac{\overbrace{x^2 |h|^2}^d}{\sigma_n^2} \right) \right] \\ &\quad + \frac{1}{\pi \sigma_n^2} \frac{1}{2} \left[\exp \left(-\frac{\overbrace{|y|^2}^a}{\sigma_n^2} - \frac{\overbrace{h^* x y}^b}{\sigma_n^2} - \frac{\overbrace{y^* h x}^c}{\sigma_n^2} - \frac{\overbrace{x^2 |h|^2}^d}{\sigma_n^2} \right) \right] \end{aligned} \quad (2.30)$$

where the conjugate of the symbol scalar x^* , is treated as the symbol itself x , since there exists no imaginary part for BPSK symbols. Consequently, by using the fractions marked by the letters a , b , c and d of (2.30) in the exponentials, the above equation can

be written as

$$\begin{aligned}
 f(y; \boldsymbol{\theta}) &= \frac{1}{\pi\sigma_n^2} \frac{1}{2} \left[e^{-a} \cdot e^{-[-b-c+d]} + e^{-a} \cdot e^{-[b+c+d]} \right] \\
 &= \frac{1}{\pi\sigma_n^2} \frac{1}{2} \left[e^{-a-d} \cdot e^{b+c} + e^{-a-d} \cdot e^{-b-c} \right] \\
 &= \frac{1}{\pi\sigma_n^2} \frac{1}{2} \left[e^{-a-d} \left(e^{b+c} + e^{-(b+c)} \right) \right] \\
 &= \frac{1}{\pi\sigma_n^2} \left[e^{-a-d} \cosh(b+c) \right]
 \end{aligned} \tag{2.31}$$

Upon substituting again a , b , c and d with the corresponding fractions of (2.30) in to (2.31), it yields

$$f(y; \boldsymbol{\theta}) = \frac{1}{\pi\sigma_n^2} \left[\exp \left(\frac{-|y|^2 - x^2|h|^2}{\sigma_n^2} \right) \cosh \left(\frac{h^*xy + y^*hx}{\sigma_n^2} \right) \right]. \tag{2.32}$$

Thus, if we consider the natural logarithm, the log-likelihood function is then derived as

$$\ln f(y; \boldsymbol{\theta}) = \ln \left(\frac{1}{\pi\sigma_n^2} \right) + \left(\frac{-|y|^2 - x^2|h|^2}{\sigma_n^2} \right) + \ln \left[\cosh \left(\frac{h^*xy + y^*hx}{\sigma_n^2} \right) \right]. \tag{2.33}$$

Accordingly, by writing h as $h_r + jh_i$ and its conjugate h^* as $h_r - jh_i$, the above equation will result in

$$\begin{aligned}
 \ln f(y; \boldsymbol{\theta}) &= \ln \left(\frac{1}{\pi\sigma_n^2} \right) + \left(\frac{-|y|^2 - x^2(h_r^2 + h_i^2)}{\sigma_n^2} \right) \\
 &\quad + \ln \left[\cosh \left(\frac{(h_r - jh_i)xy + y^*(h_r + jh_i)x}{\sigma_n^2} \right) \right].
 \end{aligned} \tag{2.34}$$

Subsequently, by differentiating once with respect to the real part of the parameter to be estimated h_r , produces

$$\frac{\partial \ln f(y; \boldsymbol{\theta})}{\partial h_r} = \frac{-2x^2 h_r}{\sigma_n^2} + \tanh \left(\frac{(h_r - jh_i)xy + y^*(h_r + jh_i)x}{\sigma_n^2} \right) \left(\frac{xy + y^*x}{\sigma_n^2} \right) \tag{2.35}$$

while the second derivative, also with respect to h_r , will derive

$$\frac{\partial^2 \ln f(y; \boldsymbol{\theta})}{\partial h_r^2} = \frac{-2x^2}{\sigma_n^2} + \text{sech}^2 \left(\frac{(h_r - jh_i)xy + y^*(h_r + jh_i)x}{\sigma_n^2} \right) \left(\frac{xy + y^*x}{\sigma_n^2} \right)^2. \tag{2.36}$$

Upon taking the negative expectation of (2.36) with respect to y again, the result is found to be

$$- \mathbb{E} \left[\frac{\partial^2 \ln f(y; \boldsymbol{\theta})}{\partial h_r^2} \right] = \frac{2x^2}{\sigma_n^2} - \mu_r \tag{2.37}$$

where μ_r is derived by moving the expectation inside (2.36) and it corresponds to a positive number, which is found such that

$$\mu_r = \mathbb{E} \left[\text{sech}^2 \left(\frac{(h_r - jh_i)xy + y^*(h_r + jh_i)x}{\sigma_n^2} \right) \left(\frac{xy + y^*x}{\sigma_n^2} \right)^2 \right]. \tag{2.38}$$

Due to the complexity of the calculations required for obtaining the result of (2.38), μ_r is calculated by realising a Monte Carlo simulation of L iterations. The resulting value of μ_r is derived by considering the real and imaginary part of a quasi-static channel coefficient along with the noise variance. Hence, different values of μ_r are derived for different

channel realisations while the noise variance might be considered constant. Once the differentiation procedure is repeated with respect to the imaginary part of the channel, it produces

$$\frac{\partial^2 \ln f(y; \boldsymbol{\theta})}{\partial h_i^2} = \frac{-2x^2}{\sigma_n^2} + \text{sech}^2\left(\frac{(h_r - jh_i)xy + y^*(h_r + jh_i)x}{\sigma_n^2}\right) \left(\frac{-jxy + jy^*x}{\sigma_n^2}\right)^2 \quad (2.39)$$

and the negative expectation also derives

$$-E\left[\frac{\partial^2 \ln f(y; \boldsymbol{\theta})}{\partial h_i^2}\right] = \frac{2x^2}{\sigma_n^2} - \mu_i. \quad (2.40)$$

The same procedure is followed, with the realisation of a Monte Carlo simulation where the mean value of μ_i is derived from

$$\mu_i = E\left[\text{sech}^2\left(\frac{(h_r - jh_i)xy + y^*(h_r + jh_i)x}{\sigma_n^2}\right) \left(\frac{-jxy + jy^*x}{\sigma_n^2}\right)^2\right]. \quad (2.41)$$

In addition, after following the instructions for the generation of the Fisher information matrix (2.4) of the blind channel estimation method, the derivative of the log likelihood function (2.34) with respect to both h_r and h_i , $\frac{\partial^2}{\partial h_r \partial h_i}$ results in

$$\begin{aligned} \frac{\partial^2 \ln f(y; \boldsymbol{\theta})}{\partial h_r \partial h_i} &= \frac{\partial}{\partial h_r} \left(\frac{\partial \ln f(y; \boldsymbol{\theta})}{\partial h_i} \right) \\ &= \frac{\partial}{\partial h_r} \left(\frac{-2x^2 h_i}{\sigma_n^2} + \tanh\left(\frac{(h_r - jh_i)xy + y^*(h_r + jh_i)x}{\sigma_n^2}\right) \left(\frac{-jxy + jy^*x}{\sigma_n^2}\right) \right) \\ &= \text{sech}^2\left(\frac{(h_r - jh_i)xy + y^*(h_r + jh_i)x}{\sigma_n^2}\right) \left(\frac{-jxy + jy^*x}{\sigma_n^2}\right) \left(\frac{yx + y^*x}{\sigma_n^2}\right) \end{aligned} \quad (2.42)$$

while the differentiation with respect to the reverse order, $\frac{\partial^2}{\partial h_i \partial h_r}$, will also result in

$$\begin{aligned} \frac{\partial^2 \ln f(y; \boldsymbol{\theta})}{\partial h_i \partial h_r} &= \frac{\partial}{\partial h_i} \left(\frac{\partial \ln f(y; \boldsymbol{\theta})}{\partial h_r} \right) \\ &= \text{sech}^2\left(\frac{(h_r - jh_i)xy + y^*(h_r + jh_i)x}{\sigma_n^2}\right) \left(\frac{-jxy + jy^*x}{\sigma_n^2}\right) \left(\frac{yx + y^*x}{\sigma_n^2}\right) \end{aligned} \quad (2.43)$$

which is the same result as in (2.42). Since the result is the same for both cases, the negative expectation of (2.42) and (2.43) with respect to y is called μ and given such that

$$-E\left[\frac{\partial^2 \ln f(y; \boldsymbol{\theta})}{\partial h_r \partial h_i}\right] = -E\left[\frac{\partial^2 \ln f(y; \boldsymbol{\theta})}{\partial h_i \partial h_r}\right] = -\mu \quad (2.44)$$

where it is defined that

$$\mu = E\left[\text{sech}^2\left(\frac{(h_r - jh_i)xy + y^*(h_r + jh_i)x}{\sigma_n^2}\right) \left(\frac{-jxy + jy^*x}{\sigma_n^2}\right) \left(\frac{yx + y^*x}{\sigma_n^2}\right)\right] \quad (2.45)$$

and the value of μ is derived once more by using a Monte Carlo simulation of L iterations for randomly generated channel realisations. Thus, the 2×2 Fisher information matrix can be derived by using (2.37), (2.40) and (2.44) similarly to (2.4), resulting in

$$\mathbf{I}(\boldsymbol{\theta}) \leq \begin{bmatrix} \frac{2x^2}{\sigma_n^2} - \mu_r & -\mu \\ -\mu & \frac{2x^2}{\sigma_n^2} - \mu_i \end{bmatrix}. \quad (2.46)$$

The inverse of the above Fisher information matrix is derived similar to

$$\mathbf{I}^{-1}(\boldsymbol{\theta}) = \frac{1}{\left(\frac{2x^2}{\sigma_n^2} - \mu_r\right) \left(\frac{2x^2}{\sigma_n^2} - \mu_i\right) - \mu^2} \begin{bmatrix} \frac{2x^2}{\sigma_n^2} - \mu_i & \mu \\ \mu & \frac{2x^2}{\sigma_n^2} - \mu_r \end{bmatrix}. \quad (2.47)$$

Finally, the CRLB for the blind channel estimation method of scalar parameters, is found by taking the diagonal elements $[\mathbf{I}^{-1}(\boldsymbol{\theta})]_{11}$ and $[\mathbf{I}^{-1}(\boldsymbol{\theta})]_{22}$ of (2.47), as previously shown in (2.2), such that

$$\text{var}(\hat{h}) \geq \left[\frac{2x^2 - \mu_i \sigma_n^2}{\frac{4x^4}{\sigma_n^2} - 2x^2(\mu_i + \mu_r) + \sigma_n^2(\mu_r \mu_i - \mu^2)} \quad \frac{2x^2 - \mu_r \sigma_n^2}{\frac{4x^4}{\sigma_n^2} - 2x^2(\mu_i + \mu_r) + \sigma_n^2(\mu_r \mu_i - \mu^2)} \right] \quad (2.48)$$

where the first element of the vector corresponds to the variance obtained by the CRLB with respect to the real part of the channel while the second element corresponds to the variance of the imaginary part of the channel. Consequently, μ_r , μ_i and μ can be replaced by the values that have been obtained from the Monte Carlo simulations for one channel realisation and randomly generated noise. Thus, the result is depending on the variance of the noise, which differs for different SNR scenarios (2.6) and the channel fading coefficient h .

Extension to vector parameters

In order to extend the results of the CRLB to vector parameters for the blind channel estimation method, approximately the same procedure as previously is followed, thus only a summary of the main calculations that were realised is provided. The system model given once more by $\mathbf{y} = \mathbf{h}\mathbf{x} + \mathbf{n}$ where \mathbf{y} , \mathbf{h} and $\mathbf{n} \in \mathbb{C}^{N \times 1}$. This time, the N length vector of the transmitted symbols, $\mathbf{x} \in \{+1, -1\}^{N \times 1}$, is comprised by BPSK symbols and it is not treated as a known pilot sequence on the receiver side. Hence we only consider the fact that the symbols are uniformly distributed, with each symbol having a probability of occurrence $\Pr(s = +1) = \Pr(s = -1) = \frac{1}{2}$.

The likelihood function of the blind channel estimation method in vector form, according to [26], is given by

$$\begin{aligned} f(\mathbf{y}; \boldsymbol{\theta}) &= \sum_x f(\mathbf{y}; \boldsymbol{\theta} | x) p(x) = \frac{1}{2} \left[f(\mathbf{y}; \boldsymbol{\theta} | x = +1) + f(\mathbf{y}; \boldsymbol{\theta} | x = -1) \right] \\ &= \frac{1}{2} \frac{1}{\pi^N \det(\mathbf{C})} \exp \left(-(\mathbf{y} - \mathbf{h}\mathbf{x})^H \mathbf{C}^{-1} (\mathbf{y} - \mathbf{h}\mathbf{x}) \right) \\ &\quad + \frac{1}{2} \frac{1}{\pi^N \det(\mathbf{C})} \exp \left(-(\mathbf{y} + \mathbf{h}\mathbf{x})^H \mathbf{C}^{-1} (\mathbf{y} + \mathbf{h}\mathbf{x}) \right) \end{aligned} \quad (2.49)$$

which can be also written in scalars form by using the product operator, given that the observations are *iid*, such that

$$f(\mathbf{y}; \boldsymbol{\theta}) = \prod_{n=0}^{N-1} \frac{1}{2} \frac{1}{\pi \sigma_n^2} \left[\exp \left(-\frac{|y[n] - h[n]x[n]|^2}{\sigma_n^2} \right) + \exp \left(-\frac{|y[n] + h[n]x[n]|^2}{\sigma_n^2} \right) \right]. \quad (2.50)$$

As previously done in (2.30), the above equation is written in a form similar to (2.31), which does not require the calculation of the sum of logarithms in order to derive the

log likelihood function. Thus, upon taking the log likelihood function for *iid* scalar parameters, it will result in

$$\ln f(\mathbf{y}; \boldsymbol{\theta}) = \sum_{n=0}^{N-1} \left[\ln \left(\frac{1}{\pi \sigma_n^2} \right) + \left(\frac{-|y[n]|^2 - x[n]^2 (h_r[n]^2 + h_i[n]^2)}{\sigma_n^2} \right) \right. \\ \left. + \ln \left[\cosh \left(\frac{(h_r[n] - j h_i[n]) x[n] y[n] + y[n]^* (h_r[n] + j h_i[n]) x[n]}{\sigma_n^2} \right) \right] \right]. \quad (2.51)$$

After obtaining the log-likelihood function, the first and second derivatives with respect to the real, the imaginary part and both the real and imaginary parts of the channel fading coefficients can be calculated by using μ_r , μ_i and μ from (2.38), (2.41) and (2.45) respectively. Consequently, the same process as before is followed. Finally, the CRLB of the blind channel estimation method for vector parameters is found by the diagonal elements of the inverse Fisher information matrix, such that

$$\text{var}(\hat{h}) \geq \left[\frac{2x^2 - \mu_i \sigma_n^2}{N \left(\frac{4x^4}{\sigma_n^2} - 2x^2 (\mu_i + \mu_r) + \sigma_n^2 (\mu_r \mu_i - \mu^2) \right)} \right. \\ \left. \frac{2x^2 - \mu_r \sigma_n^2}{N \left(\frac{4x^4}{\sigma_n^2} - 2x^2 (\mu_i + \mu_r) + \sigma_n^2 (\mu_r \mu_i - \mu^2) \right)} \right]. \quad (2.52)$$

which is similar to the result obtained in (2.48), parametrised by the inverse number of the received data symbols $1/N$ since the observations are considered *iid*.

2.1.3 Semi-blind channel estimation

Scalar parameters model

The semi-blind channel estimation method can be considered as a hybrid method, which combines the transmission of data symbols while it also profits from the usage of a known pilot symbols sequence during the exchange of packets between the BS and the UE. A good set up of the ratio of data symbols or pilot symbols to the number of the total transmitted symbols, can prove to be a good technique for efficiently obtaining accurate channel estimations, without the need of extensive pilot overhead generation.

The system model for scalar parameters is given once more by (2.1). Furthermore, the noise scalar is assumed to be zero mean *iid* circularly symmetric complex Gaussian additive noise such that $n \sim \mathcal{CN}(0, \sigma_n^2)$ where the variance is derived from (2.6). For the semi-blind method, instead of a single transmitted symbol, a column vector comprised of one pilot and one data symbol is defined, such that $\mathbf{x} = [x_p \ x_d]^T$. Thus, we first consider the transmission of $N = 2$ BPSK symbols, where the first symbol x_p is a known pilot symbol such that $x_p = \{+1\}^{1 \times 1}$. The second symbol x_d , is a uniformly distributed random data symbol which can take two values, thus $x_d \in \{+1, -1\}$ with probability of occurrence $\Pr(x_d = +1) = \Pr(x_d = -1) = \frac{1}{2}$. The channel is also assumed to be slow fading, hence the channel fading coefficient remains constant over the transmission of the two symbols, x_p and x_d . Moreover, assuming that the symbols are *iid*, both are considered to be independent of the channel. The received symbols are also given by a column vector comprised of a pilot and a data symbol, such that $\mathbf{y} = [y_p \ y_d]^T$, following that $y_d = h x_d + n$ and $y_p = h x_p + n$. Finally, as it was previously stated, the unknown vector parameter is given such that $\boldsymbol{\theta} = [\text{Re}\{h\} \ \text{Im}\{h\}]^T = [h_r \ h_i]^T$.

The likelihood function based on the PDF of the observable [26], this time can be expressed by

$$\begin{aligned}
 f(\mathbf{y}; \boldsymbol{\theta}) &= f(y_p; \boldsymbol{\theta})f(y_d; \boldsymbol{\theta}) \\
 &= f(y_p; \boldsymbol{\theta}) \left(\sum_x f(y_d; \boldsymbol{\theta} | x_d = \pm 1) p(x_d) \right) \\
 &= f(y_p; \boldsymbol{\theta}, x_p) \left(f(y_d; \boldsymbol{\theta} | x_d = +1) \frac{1}{2} + f(y_d; \boldsymbol{\theta} | x_d = -1) \frac{1}{2} \right) \quad (2.53)
 \end{aligned}$$

which as it can be seen is a combination of the pilot channel estimation method, where the received symbol x_p is considered known and the blind channel estimation method, with the received data symbol x_d considered random along with its probability of occurrence. Upon combining the PDFs of the pilot (2.7) and blind channel estimation methods (2.29), the likelihood function will then result in

$$\begin{aligned}
 f(y; \boldsymbol{\theta}) &= \frac{1}{\pi \sigma_n^2} \exp \left(\frac{-|y_p - hx_p|^2}{\sigma_n^2} \right) \\
 &\quad \left[\frac{1}{\pi \sigma_n^2} \frac{1}{2} \left(\exp \left(\frac{-|y_d - hx_d|^2}{\sigma_n^2} \right) + \exp \left(\frac{-|y_d + hx_d|^2}{\sigma_n^2} \right) \right) \right] \\
 &= \frac{1}{\pi \sigma_n^2} \exp \left(\frac{-(y_p^* y_p + h^* h x_p^* x_p - y_p^* h x_p - y_p h^* x_p^*)}{\sigma_n^2} \right) \\
 &\quad \left[\frac{1}{\pi \sigma_n^2} \left(\exp \left(\frac{-y_d^* y_d - x_d^2 h^* h}{\sigma_n^2} \right) \cosh \left(\frac{h^* x_d y_d + y_d^* h x_d}{\sigma_n^2} \right) \right) \right] \quad (2.54)
 \end{aligned}$$

where it is considered that $x_d^* = x_d$ and $x_p^* = x_p$, as previously explained, while the same procedure as in (2.30) and (2.31) is followed. After taking the logarithm, the log likelihood function yields

$$\begin{aligned}
 \ln f(y; \boldsymbol{\theta}) &= \ln \left(\frac{1}{\pi \sigma_n^2} \right) - \frac{|y_p|^2 + |h|^2 x_p^2 - y_p^* h x_p - y_p h^* x_p}{\sigma_n^2} \\
 &\quad + \ln \left(\frac{1}{\pi \sigma_n^2} \right) + \frac{-|y_d|^2 - x_d^2 |h|^2}{\sigma_n^2} + \ln \left[\cosh \left(\frac{h^* x_d y_d + y_d^* h x_d}{\sigma_n^2} \right) \right]. \quad (2.55)
 \end{aligned}$$

Since the channel fading coefficient is comprised by a real and imaginary part, such that $h = h_r + jh_i$ and $h^* = h_r - jh_i$, upon differentiating once with respect to the channel's real part h_r , it is derived

$$\begin{aligned}
 \frac{\partial \ln f(y; \boldsymbol{\theta})}{\partial h_r} &= - \frac{2h_r x_p^2 - y_p^* x_p - y_p x_p}{\sigma_n^2} - \frac{2h_r x_d^2}{\sigma_n^2} \\
 &\quad + \tanh \left(\frac{(h_r - jh_i)x_d y_d + y_d^*(h_r + jh_i)x_d}{\sigma_n^2} \right) \left(\frac{x_d y_d + y_d^* x_d}{\sigma_n^2} \right) \quad (2.56)
 \end{aligned}$$

and after taking the second derivative, again with respect to h_r , it is obtained

$$\frac{\partial^2 \ln f(y; \boldsymbol{\theta})}{\partial h_r^2} = - \frac{2x_p^2}{\sigma_n^2} - \frac{2x_d^2}{\sigma_n^2} + \text{sech}^2 \left(\frac{(h_r - jh_i)x_d y_d + y_d^*(h_r + jh_i)x_d}{\sigma_n^2} \right) \left(\frac{x_d y_d + y_d^* x_d}{\sigma_n^2} \right)^2. \quad (2.57)$$

At this point, it must be highlighted that since the symbol vector \mathbf{x} is equally comprised of one pilot symbol and one data symbol, we consider the same symbol energy for both types of symbols, such that $\|\mathbf{x}\|^2 = \|x_d\|^2 = \|x_p\|^2$. For the interested reader, the channel estimation accuracy achieved for different energy levels of pilot and data symbols

is discussed in [23]. Upon taking the negative expectation of the above equation, it is derived that (2.57) converges to

$$-E \left[\frac{\partial^2 \ln f(y; \boldsymbol{\theta})}{\partial h_r^2} \right] = \frac{4x^2}{\sigma_n} - \mu_r \quad (2.58)$$

where μ_r is given by equation (2.38). Finally, when the above process is repeated by differentiating twice with respect to the imaginary part of the channel h_i , it will result in

$$-E \left[\frac{\partial^2 \ln f(y; \boldsymbol{\theta})}{\partial h_i^2} \right] = \frac{4x^2}{\sigma_n} - \mu_i \quad (2.59)$$

in which μ_i is given by equation (2.41). It must be mentioned that both μ_r and μ_i will have the same value as previously computed for the blind channel estimation method, since both methods are compared on the estimation accuracy they provide for the same channel fading coefficient. The derivatives of (2.55) with respect to both the real and imaginary channel are calculated such that

$$\begin{aligned} \frac{\partial^2 \ln f(y; \boldsymbol{\theta})}{\partial h_r \partial h_i} &= \frac{\partial}{\partial h_r} \left(\frac{\partial \ln f(y; \boldsymbol{\theta})}{\partial h_i} \right) \\ &= \text{sech}^2 \left(\frac{(h_r - jh_i)x_d y_d + y_d^*(h_r + jh_i)x_d}{\sigma_n^2} \right) \left(\frac{-jx_d y_d + jy_d^* x_d}{\sigma_n^2} \right) \\ &\quad \left(\frac{x_d y_d + y_d^* x_d}{\sigma_n^2} \right) \end{aligned} \quad (2.60)$$

which upon taking the negative expectation is similar to (2.45) and thus results in

$$-E \left[\frac{\partial^2 \ln f(y; \boldsymbol{\theta})}{\partial h_r \partial h_i} \right] = -E \left[\frac{\partial^2 \ln f(y; \boldsymbol{\theta})}{\partial h_i \partial h_r} \right] = -\mu \quad (2.61)$$

where the result is the same for the differentiation with $\frac{\partial^2}{\partial h_r \partial h_i}$ and $\frac{\partial^2}{\partial h_i \partial h_r}$. The resulting Fisher information matrix for the semi-blind method, will be

$$\mathbf{I}(\boldsymbol{\theta}) \leq \begin{bmatrix} \frac{4x^2}{\sigma_n^2} - \mu_r & -\mu \\ -\mu & \frac{4x^2}{\sigma_n^2} - \mu_i \end{bmatrix} \quad (2.62)$$

while the inverse Fisher information is given similar to

$$\mathbf{I}^{-1}(\boldsymbol{\theta}) = \frac{1}{\left(\frac{4x^2}{\sigma_n^2} - \mu_r \right) \left(\frac{4x^2}{\sigma_n^2} - \mu_i \right) - \mu^2} \begin{bmatrix} \frac{4x^2}{\sigma_n^2} - \mu_i & \mu \\ \mu & \frac{4x^2}{\sigma_n^2} - \mu_r \end{bmatrix}. \quad (2.63)$$

Since the CRLB is given by the diagonal elements $[\mathbf{I}^{-1}(\boldsymbol{\theta})]_{11}$ and $[\mathbf{I}^{-1}(\boldsymbol{\theta})]_{22}$ of the inverse Fisher information matrix, the lower bound of the variance will equal to

$$\begin{aligned} \text{var}(\hat{h}) &\geq \left[\frac{4x^2 - \mu_i \sigma_n^2}{\frac{16x^4}{\sigma_n^2} - 4x^2 (\mu_i - \mu_r) + \sigma_n^2 (\mu_r \mu_i - \mu^2)} \right. \\ &\quad \left. \frac{4x^2 - \mu_r \sigma_n^2}{\frac{16x^4}{\sigma_n^2} - 4x^2 (\mu_i - \mu_r) + \sigma_n^2 (\mu_r \mu_i - \mu^2)} \right] \end{aligned} \quad (2.64)$$

where the first element of the vector corresponds to the lower bound of the variance calculated for the real part of the channel, while the second element corresponds to the imaginary part of the channel. As previously stated, μ_r , μ_i and μ can be replaced by the values that have been obtained from the Monte Carlo simulation for a deterministic channel and randomly generated noise.

Extension to vector parameters

In contradiction to the previous scalar model and in order to derive an expression for the semi-blind channel estimation method for vector parameters, the vector of the transmitted symbols is given $\mathbf{x} = [\mathbf{x}_p^T \mathbf{x}_d^T]$. The symbol vector is comprised of a pilot symbols vector $\mathbf{x}_p \in \mathbb{R}^{N_P \times 1}$, of length N_P and a vector of random data symbols $\mathbf{x}_d \in \mathbb{R}^{N_D \times 1}$, of length N_D , where $N = N_P + N_D$ is the number of the total received symbols. As a result, the ratio of pilot to total received symbols, is given such that

$$\alpha = N_P/N \quad (2.65)$$

where α is defined for $0 \leq \alpha \leq 1$. Following this definition, when $\alpha = 0$ no pilot symbols are used and the channel estimation is achieved similar to the blind method. On the other hand, for $\alpha = 1$ the channel estimate is achieved similar to the pilot method. The definition of this ratio will allow us to achieve different accuracy levels of channel estimates, as it is discussed in [22]. In addition, since data are also transmitted during the channel estimation process, the useful rate is expected to increase. A further step would have been to define the proportionate pilot to data power ratio (PDPR), presented also in [22] and [23], which also results to different channel estimates accuracy. In this thesis, only three different values of α ratio (2.65) will be examined, $\alpha = 0.2$, $\alpha = 0.5$ and $\alpha = 0.8$ which provide different numbers of pilot and data symbols depending on the length N of the transmitted symbol vector. The rest of the vectors for this method are defined as previously.

The likelihood function of the semi-blind method, presented in vectors form and according to [26] is given by

$$\begin{aligned} f(\mathbf{y}; \boldsymbol{\theta}) &= f(\mathbf{y}_p; \boldsymbol{\theta}) f(\mathbf{y}_d; \boldsymbol{\theta}) \\ &= \frac{1}{\pi^N \det(\mathbf{C})} \exp \left(-(\mathbf{y}_p - \mathbf{h}\mathbf{x}_p)^H \mathbf{C}^{-1} (\mathbf{y}_p - \mathbf{h}\mathbf{x}_p) \right) \\ &\quad \left[\frac{1}{2} \frac{1}{\pi^N \det(\mathbf{C})} \exp \left(-(\mathbf{y}_d - \mathbf{h}\mathbf{x}_d)^H \mathbf{C}^{-1} (\mathbf{y}_d - \mathbf{h}\mathbf{x}_d) \right) \right. \\ &\quad \left. + \frac{1}{2} \frac{1}{\pi^N \det(\mathbf{C})} \exp \left(-(\mathbf{y}_d + \mathbf{h}\mathbf{x}_d)^H \mathbf{C}^{-1} (\mathbf{y}_d + \mathbf{h}\mathbf{x}_d) \right) \right] \end{aligned} \quad (2.66)$$

where \mathbf{y}_p and \mathbf{y}_d correspond to the received vectors of pilot and data symbols of length N_P and N_D respectively. Furthermore, \mathbf{C} is the $\sigma_n^2 I_{N \times N}$ covariance matrix and the superscript H denotes Hermitian transpose. The above equation can also be written in scalars form by using the product operator and upon considering the observations as *iid*, such that

$$\begin{aligned} f(\mathbf{y}; \boldsymbol{\theta}) &= \prod_{n=0}^{N-1} \frac{1}{\pi \sigma_n^2} \left[\exp \left(-\frac{|y_p[n] - h[n]x_p[n]|^2}{\sigma_n^2} \right) \right. \\ &\quad \left. \left(\frac{1}{2} \frac{1}{\pi \sigma_n^2} \left(\exp \left(-\frac{|y_d[n] - h[n]x_d[n]|^2}{\sigma_n^2} \right) + \exp \left(-\frac{|y_d[n] + h[n]x_d[n]|^2}{\sigma_n^2} \right) \right) \right) \right] \end{aligned} \quad (2.67)$$

where the vectors are changed to scalar parameters. In order to obtain the log-likelihood function a similar procedure as in (2.55) is followed which results in the sum of scalars. Subsequently, we differentiate the log-likelihood function twice with respect to the real and imaginary part of the channel. For the real part of the channel h_r , the result of the second derivative for vector parameters is given by

$$\begin{aligned} \frac{\partial^2 \ln f(\mathbf{y}; \boldsymbol{\theta})}{\partial h_r^2} = \sum_{n=0}^{N-1} \left[-\frac{2x_p[n]^2}{\sigma_n^2} - \frac{2x_d[n]^2}{\sigma_n^2} + \left(\frac{x_d[n]y_d[n] + y_d[n]^*x_d[n]}{\sigma_n^2} \right)^2 \right. \\ \left. \operatorname{sech}^2 \left(\frac{(h_r[n] - jh_i[n])x_d[n]y_d[n] + y_d[n]^*(h_r[n] + jh_i[n])x_d[n]}{\sigma_n^2} \right) \right] \end{aligned} \quad (2.68)$$

while the second derivative with respect to the imaginary part of the channel h_i , will be

$$\begin{aligned} \frac{\partial^2 \ln f(\mathbf{y}; \boldsymbol{\theta})}{\partial h_i^2} = \sum_{n=0}^{N-1} \left[-\frac{2x_p[n]^2}{\sigma_n^2} - \frac{2x_d[n]^2}{\sigma_n^2} + \left(\frac{-jx_d[n]y_d[n] + jy_d[n]^*x_d[n]}{\sigma_n^2} \right)^2 \right. \\ \left. \operatorname{sech}^2 \left(\frac{(h_r[n] - jh_i[n])x_d[n]y_d[n] + y_d[n]^*(h_r[n] + jh_i[n])x_d[n]}{\sigma_n^2} \right) \right]. \end{aligned} \quad (2.69)$$

The differentiation with respect to h_r and h_i , such that $\frac{\partial^2}{\partial h_r \partial h_i}$ and $\frac{\partial^2}{\partial h_i \partial h_r}$, will result in

$$\begin{aligned} \frac{\partial^2 \ln f(y; \boldsymbol{\theta})}{\partial h_r \partial h_i} &= \frac{\partial^2 \ln f(y; \boldsymbol{\theta})}{\partial h_i \partial h_r} \\ &= \sum_{n=0}^{N-1} \left[\operatorname{sech}^2 \left(\frac{(h_r[n] - jh_i[n])x_d[n]y_d[n] + y_d[n]^*(h_r[n] + jh_i[n])x_d[n]}{\sigma_n^2} \right) \right. \\ &\quad \left. \left(\frac{-jx_d[n]y_d[n] + jy_d[n]^*x_d[n]}{\sigma_n^2} \right) \left(\frac{x_d[n]y_d[n] + y_d[n]^*x_d[n]}{\sigma_n^2} \right) \right] \end{aligned} \quad (2.70)$$

where the result depends only on data symbols. Upon taking the negative expectation the above equations will converge to

$$-E \left[\frac{\partial^2 \ln f(\mathbf{y}; \boldsymbol{\theta})}{\partial h_r^2} \right] = -\sum_{n=0}^{N-1} E \left[-\frac{2x_p[n]^2}{\sigma_n^2} - \frac{2x_d[n]^2}{\sigma_n^2} + \mu_r \right] \quad (2.71)$$

and

$$-E \left[\frac{\partial^2 \ln f(\mathbf{y}; \boldsymbol{\theta})}{\partial h_i^2} \right] = -\sum_{n=0}^{N-1} E \left[-\frac{2x_p[n]^2}{\sigma_n^2} - \frac{2x_d[n]^2}{\sigma_n^2} + \mu_i \right] \quad (2.72)$$

while

$$-E \left[\frac{\partial^2 \ln f(y; \boldsymbol{\theta})}{\partial h_r \partial h_i} \right] = -E \left[\frac{\partial^2 \ln f(y; \boldsymbol{\theta})}{\partial h_i \partial h_r} \right] = \sum_{n=0}^{N-1} -\mu \quad (2.73)$$

where μ_r is given by (2.38), μ_i by (2.41) and μ by (2.45). Furthermore, upon assuming the same symbol energy for pilot and data symbols, such that $\|x\|^2 = \|x_p\|^2 = \|x_d\|^2$, while also considering the sum of *iid* symbols, the Fisher information matrix for the semi-blind method is derived such that

$$\mathbf{I}(\boldsymbol{\theta}) = \begin{bmatrix} \frac{2x^2}{\sigma_n^2}N - \mu_r N_D & -N_D \mu \\ -N_D \mu & \frac{2x^2}{\sigma_n^2}N - \mu_i N_D \end{bmatrix} \quad (2.74)$$

in which the numbers of pilot symbols N_P and data symbols N_D have been assigned and can be adjusted by using (2.65). The CRLB is finally derived by the diagonal elements of the inverse Fisher information matrix, hence the result for the semi-blind method extended to vector parameters will be

$$\text{var}(\hat{h}) \geq \left[\begin{array}{c} \frac{2x^2N - \mu_i\sigma_n^2N_D}{\frac{4x^4}{\sigma_n^2}N^2 - N_D(2x^2N(\mu_i + \mu_r) - N_D\sigma_n^2(\mu_r\mu_i - \mu^2))} \\ \frac{2x^2N - \mu_r\sigma_n^2N_D}{\frac{4x^4}{\sigma_n^2}N^2 - N_D(2x^2N(\mu_i + \mu_r) - N_D\sigma_n^2(\mu_r\mu_i - \mu^2))} \end{array} \right]. \quad (2.75)$$

From the above result of the CRLB, it can be seen that regarding on the selection of the ratio α (2.65), the semi blind method is able of obtaining different results. For example, when $\alpha = 0$, it is derived that $N_D = N$ thus no pilot symbols are used and the resulting CRLB is same to (2.52). On the other hand, when $\alpha = 1$ then $N_D = 0$ and the resulting CRLB is given by (2.27). The final result thus is shaped by the noise variance, the values of μ_r , μ_i and μ , which depend on the channel fading coefficient and the selection of the pilot to total received symbols ratio given in (2.65).

2.2 The estimated channel

Given that the variance of the channel estimates has been determined by using the CRLB, a new channel can now be considered as the true channel which is used in order to derive the corresponding constrained capacity. This channel is comprised of a deterministic channel fading coefficient, along any estimation uncertainties which occur from each channel estimation method. The uncertainty is derived by the CRLB and the corresponding lower bound of variance of the estimates, where a decreased level of variance leads to more accurate estimates. Hence, the true channel fading coefficient $h \in \mathbb{C}^{1 \times 1}$ according to [27] is given such that

$$h = \hat{h} + \Delta h \quad (2.76)$$

where \hat{h} corresponds to the estimated channel, comprised of a deterministic Rayleigh distributed channel fading coefficient and Δh is the uncertainty of the channel estimate, modelled as a complex Gaussian random variable $\Delta h \sim \mathcal{CN}(0, \sigma_{\text{CRLB}}^2)$, where $\sigma_{\text{CRLB}}^2 = \text{var}(\hat{h})$ is the corresponding variance obtained by the CRLB derived for each channel estimation method according to (2.27), (2.52) and (2.75). Consequently, each channel estimation method introduces a different degree of CSI uncertainty due to estimation inaccuracies, hence the channel h used by the system deviates from the original accurate estimate \hat{h} . In addition, several more uncertainties and errors can be added on the channel estimate, however for the sake of simplicity, these errors are not considered in this report and we only focus on the CSI errors caused due to the corresponding variance of the CRLB of each channel estimation method.

2.3 Constrained capacity upon the channel estimates

Since a different considered channel is derived based on the results of each channel estimation method (2.76), we can now compute the corresponding achievable channel capacities. Channel capacity is defined as the highest achievable rate of reliably transmitted data over a wireless channel [28]. Two types of wireless channels can be considered.

A quasi-static channel, where its fading coefficient remains constant in time over the transmission and reception of a packet without defining the coherence interval which is frequency, speed and environment dependent [24]. This channel can be used in order to derive the corresponding channel capacity based on the channel estimates which are affected by the CRLB (2.76) with a number of N pilot or data symbols. In addition, the useful and signalling rate can be easily compared for a single channel realisation. On the other hand, the wireless channel might be assumed as frequency selective and fast fading, thus an ergodic capacity should be derived and averaged over independent channel realisations. Hence, the derivation process of the quasi-static channel constrained capacity, constraint due to BPSK modulated symbols, and the ergodic channel constrained capacity based on the channel estimates and the CRLB, is discussed and analysed.

The system model is initially given by equation (2.1) with the effect of the channel fading coefficient h on the transmitted symbols. According to the previous discussion, the channel estimate is assumed to be the most accurate estimate given by (2.76), hence the model in scalar parameters (2.1) will be modified such that

$$\begin{aligned} y &= hx + n \\ &= (\hat{h} + \Delta h)x + n \\ &= \hat{h}x + \Delta hx + n \end{aligned} \quad (2.77)$$

which includes the channel estimate \hat{h} and the estimate uncertainties Δh that vary for each channel estimation method. The aim is to remove the effect of the channel fading coefficients from the received symbols. As a result a new expression of the system model is obtained, which is mainly affected by the random noise term. Since we are considering complex quantities and the CRLB for the blind and semi-blind channel estimation method differs for the real and the imaginary part of the channel due to μ_r and μ_i , as shown in (2.52) and (2.75), an extension to (2.77) is given by

$$\begin{bmatrix} y_r \\ y_i \end{bmatrix} = \begin{bmatrix} \hat{h}_r & -\hat{h}_i \\ \hat{h}_i & \hat{h}_r \end{bmatrix} \begin{bmatrix} x_r \\ 0 \end{bmatrix} + \begin{bmatrix} \Delta h_r & -\Delta h_i \\ \Delta h_i & \Delta h_r \end{bmatrix} \begin{bmatrix} x_r \\ 0 \end{bmatrix} + \begin{bmatrix} n_r \\ n_i \end{bmatrix} \quad (2.78)$$

where the subscript r corresponds to the real part and i is used to denote the imaginary part. Furthermore, since there is no imaginary part considered for the BPSK symbols, only x_r is used and x_i is set to zero. Upon compensating for the channel fading by using the inverse of the estimated channel \hat{h} and the estimation uncertainty Δh , the above equation yields

$$y' = \frac{y}{\hat{h}} = x + \frac{\Delta hx + n}{\hat{h}} \quad (2.79)$$

or in matrix form

$$\mathbf{Y}' = \hat{\mathbf{H}}^{-1} \mathbf{Y} = \mathbf{X} + \underbrace{\left(\hat{\mathbf{H}}^{-1} \Delta \mathbf{H} \mathbf{X} + \hat{\mathbf{H}}^{-1} \mathbf{N} \right)}_{N'}. \quad (2.80)$$

The inner of the parenthesis in (2.80), denoted by N' , is comprised of the uniformly distributed symbols, $x \in \{\pm 1\}$, the complex Gaussian distributed random variable considered as the channel $h \sim \mathcal{CN}(0,1)$, the complex Gaussian random noise term $n \sim \mathcal{CN}(0, \sigma_n^2)$ and the complex Gaussian random variable $\Delta h \sim \mathcal{CN}(0, \sigma_{\text{CRLB}}^2)$. As a result, we can consider N' as a new random variable which will equal to

$$N' = \frac{1}{|\hat{h}|^2} \begin{bmatrix} \hat{h}_r & \hat{h}_i \\ -\hat{h}_i & \hat{h}_r \end{bmatrix} \begin{bmatrix} \Delta h_r & -\Delta h_i \\ \Delta h_i & \Delta h_r \end{bmatrix} \begin{bmatrix} x_r \\ 0 \end{bmatrix} + \frac{1}{|\hat{h}|^2} \begin{bmatrix} \hat{h}_r & \hat{h}_i \\ -\hat{h}_i & \hat{h}_r \end{bmatrix} \begin{bmatrix} n_r \\ n_i \end{bmatrix} \quad (2.81)$$

while focusing on the real part of N' due to BPSK modulation, denoted as N'_r , it is derived that

$$N'_r = \frac{1}{|\hat{h}|^2} \left(\hat{h}_r \Delta h_r x_r + \hat{h}_i \Delta h_i x_r \right) + \frac{1}{|\hat{h}|^2} \left(\hat{h}_r n_r + \hat{h}_i n_i \right) \quad (2.82)$$

which according to the property of the variance under linear transformation, $\text{var}(aX) = a^2 \text{var}(X)$, gives a corresponding variance equal to

$$\begin{aligned} \sigma^2 &= \text{var}(N'_r) = \frac{1}{|\hat{h}|^4} \left(\hat{h}_r^2 \sigma_{\text{CRLB},r}^2 + \hat{h}_i^2 \sigma_{\text{CRLB},i}^2 \right) + \frac{1}{|\hat{h}|^4} \left(\hat{h}_r^2 \frac{\sigma_n^2}{2} + \hat{h}_i^2 \frac{\sigma_n^2}{2} \right) \\ &= \frac{\left(\hat{h}_r^2 \sigma_{\text{CRLB},r}^2 + \hat{h}_i^2 \sigma_{\text{CRLB},i}^2 \right)}{|\hat{h}|^4} + \frac{\sigma_n^2}{2|\hat{h}|^2} \end{aligned} \quad (2.83)$$

where $\sigma_{\text{CRLB},r}^2$ and $\sigma_{\text{CRLB},i}^2$ correspond to the variances obtained by the CRLB with respect to the real and imaginary part of the channel, for each channel estimation method. Furthermore, it is derived that $\text{E}[N'_r] = 0$ due to the addition of zero mean random variables, which results in a new zero mean random variable with variance σ^2 , as mentioned in [29]. In addition, as shown in (2.6), by using different values of SNR we are able to shape the complex noise variance in (2.83) without considering a coding rate $R = 1$ [28].

According to the new system model given in (2.79) and since the achieved capacity is related to the channel fading coefficient, we consider the input to the channel X , i.e., the transmitted symbols and the channel output Y , i.e., the received symbols, [23]. The equivalent mutual information $I(X; Y)$ according to [30] is given

$$\begin{aligned} I(X; Y) &= h(Y) - h(Y|X) \\ &= h(Y) - h(X + Z|X) \\ &= h(Y) - h(Z|X) \\ &= h(Y) - h(Z) \end{aligned} \quad (2.84)$$

where $h(Y)$ corresponds to the differential entropy of the continuous channel output, $h(Y|X)$ denotes the differential conditional entropy, while $h(Z) = \log_2(\pi e \sigma^2)$ corresponds to the entropy of the complex noise term [30] with the new variance σ^2 given by (2.83), which is independent of the channel input X . In order to consider the effect of the pilot symbols in the mutual information, (2.84) can be written in terms of the data and pilot symbols, as presented in [23], where the subscripts p and d are used to notate pilot and data symbols respectively. Thus, the mutual information will be now given such that

$$I(X_p, Y_p, X_d; Y_d) = I(X_d; Y_d | X_p, Y_p) \quad (2.85)$$

where the observations of the symbols X_p , Y_p , X_d are considered known, while Y_d corresponds to the unknown received symbols. In addition, the symbols Y_d are considered to be independent of X_p and Y_p , [23]. In the case of blind channel estimation, the symbols X_p and Y_p are discarded, hence the mutual information is given as previously in (2.84).

The constrained channel capacity can be obtained by the maximum mutual information where the maximization occurs over the channel input probability distribution

$\Pr(x) = \frac{1}{2}(x = \pm 1)$, as suggested in [28], defined as

$$\begin{aligned}
C_c &= \max_{\Pr(x)} I(X_d; Y_d | X_p, Y_p) \\
&= h(Y_d | X_p, Y_p) - h(Z | X_p, Y_p) \\
&= \int_{-\infty}^{+\infty} p(y') \log_2(p(y')) dy - \log_2(\pi e \sigma^2) \\
&= E[-\log_2(p(y'))] - \log_2(\pi e \sigma^2)
\end{aligned} \tag{2.86}$$

where the integral can be calculated by using the expectation which subsequently is estimated with a Monte Carlo simulation as suggested in [28], such that

$$E[-\log_2(p(y'))] \simeq \frac{1}{L} \sum_{l=1}^L \log_2(p(y'_l)) \tag{2.87}$$

in which $l = [1, \dots, L]$ corresponds to the number of Monte Carlo iterations. In order to calculate the two different types of capacities, we should consider two different techniques. For the ergodic constrained capacity, in every Monte Carlo iteration a new channel realisation is generated along with randomly generated complex Gaussian noise which both affect the received symbols, as suggested in [28]. As a result the expectation in (2.87) is taken with respect to both the random channel and the noise. On the other hand, in order to calculate the constrained capacity of a quasi-static channel, the channel realisation is considered constant, over the transmitted symbols and it is only the complex Gaussian noise which varies for every Monte Carlo iteration, thus the expectation in (2.87) is only taken with respect to the random noise.

In the above equation, $p(y')$ corresponds to the PDF of the received symbol and the noise, thus it is given such that

$$p(y' | x = \pm 1) = \frac{1}{\pi \sigma^2} \exp\left(\frac{-|y \pm 1|^2}{\sigma^2}\right) \tag{2.88}$$

where the channel effect has been removed according to (2.79) and considering the symbol's probability of occurrence it is derived

$$p(y') = \frac{1}{2} \left[\frac{1}{\pi \sigma^2} \exp\left(\frac{-|y - 1|^2}{\sigma^2}\right) + \frac{1}{\pi \sigma^2} \exp\left(\frac{-|y + 1|^2}{\sigma^2}\right) \right] \tag{2.89}$$

since the symbols are BPSK modulated and uniformly distributed. In addition, the variance σ^2 is derived from (2.83) and differs for every channel estimation method, since it considers the CRLB of the pilot, blind and semi-blind channel estimation method.

The above procedure can be repeated for different SNR levels, defined by (2.6) and it can consider different Rayleigh distributed channel realisations. However, in order to calculate the corresponding constrained capacity (2.86) and derive the resulting useful or signalling rate for each method, a quasi-static channel is generally considered, with a fixed channel fading coefficient. The system level simulations that use the formulas presented in this chapter, such as the CRLB of each channel estimation method and the corresponding capacities are presented and compared on the following chapter.

3

System level simulations and results analysis

THE aim of this chapter is to use numerical evaluations and system level simulations in order to compare the performance and advantages of using each channel estimation method. This comparison is realised with respect to the quality of the channel estimates, the corresponding constrained capacity and the useful rate obtained by each method. By providing our remarks in an illustrated form, a more efficient comparison of the performance is achieved. In conclusion, useful remarks can be derived which can find application in future 5G network scenarios, such as an ultra-dense deployed network.

3.1 Simulation assumptions

The following simulations are based on the mathematical equations which were derived and presented at Chapter 2. In order to apply these equations and obtain the desired outcomes, while also for the sake of simplicity, several assumptions and set ups have been considered. In addition, in order to assist on the generation of the desired equations, some assumptions have been presented and discussed on the previous section, such as the noise's distribution and the symbols' constellation. The assumptions used for the actual simulations are listed below:

- The Monte Carlo simulations use $L = 10^6$ iterations, in order to ensure accurate results, based on the *law of large numbers* [29] where the mean value from a large number of samples converges to the expected value.
- The estimated channel fading coefficient is given such that $\hat{h} = 0.4937 + 0.1939i$ and it is a random sample from a Rayleigh distributed sample space.
- The SNR is assumed to be equal to 0 dB, unless noted otherwise.
- The noise, the symbols and the channel fading coefficients are considered as independent and identically distributed parameters.
- An arbitrary function is used for mapping one symbol in each time and frequency element of the network resources.

3.2 Channel estimation methods' CRLB comparison

The first comparison of the three channel estimation methods, is realised with respect to the lowest achievable variance of the channel estimate (2.2), derived by the CRLB method. The CRLB for the channel estimation methods is given by equations (2.27), (2.52) and (2.75) for the pilot based, blind and semi-blind method in vector form. The estimated channel fading coefficient is given such that $\hat{h} = 0.4937 + 0.1939i$ which is used in order to calculate the mean values of equations (2.38), (2.41) and (2.45), upon using a Monte Carlo simulation, such that $\mu_r = 0.9258$, $\mu_i = 1.1972$ and $\mu = -0.1261$. The histograms which present the resulting values of μ_r , μ_i and μ over the Monte Carlo simulation, can be seen in Figure 3.1.

Figure 3.2 illustrates the lower bound on the variance of the estimated real part of the channel, obtained by the CRLB of each channel estimation method with respect to the number of the received symbols, in a logarithmic scale. The number of the received symbols for each method, extends from $N = 4$ symbols up to $N = 100$ symbols. The lowest number of comparison is selected to be $N = 4$, since the comparison for only one received symbol, such that $N = 1$, is not considered a fair comparison due to the fact that it is not equally defined for all the channel estimation methods if this symbol is a pilot or data symbol. Furthermore, three cases of semi-blind channel estimation methods with different α ratios (2.65) are taken into consideration, along with the pilot and blind channel estimation methods. The noise variance is consider constant and equal to $\sigma_n^2 = 1$, thus resulting in an SNR of 0 dB, according to (2.6).

As it can be seen in Figure 3.2, the variance of the channel estimate decays expo-

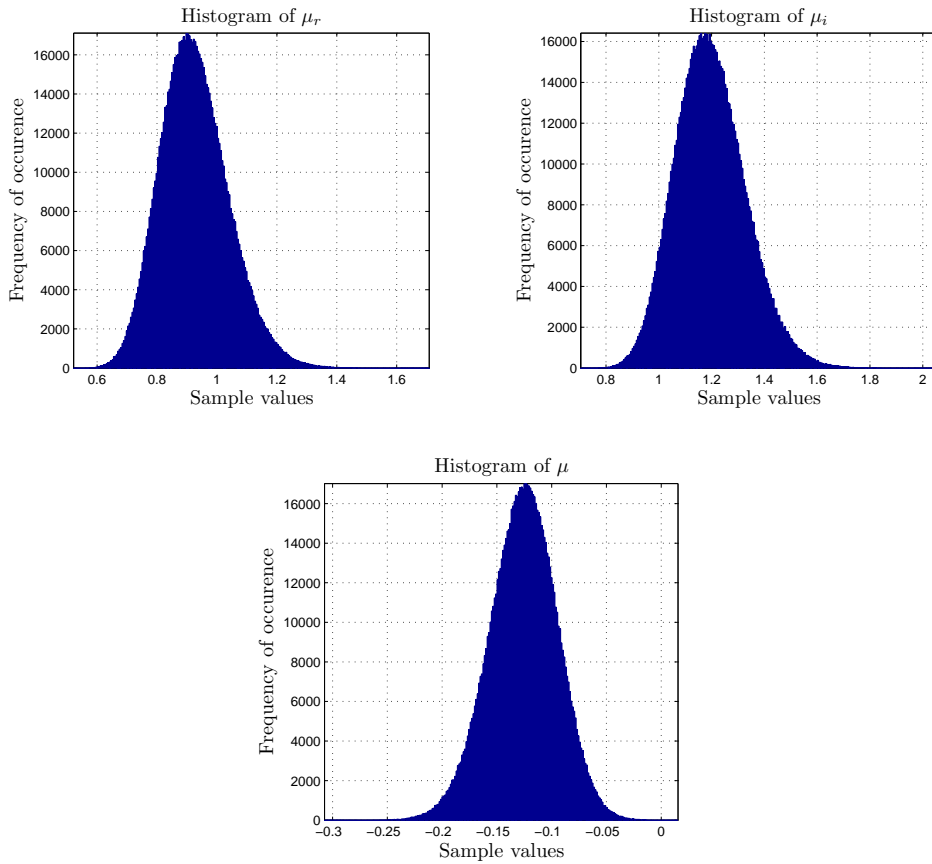


Figure 3.1: Histograms of μ_r , μ_i and μ obtained by a Monte Carlo simulation of 10^6 iterations, for the channel fading coefficient \hat{h} .

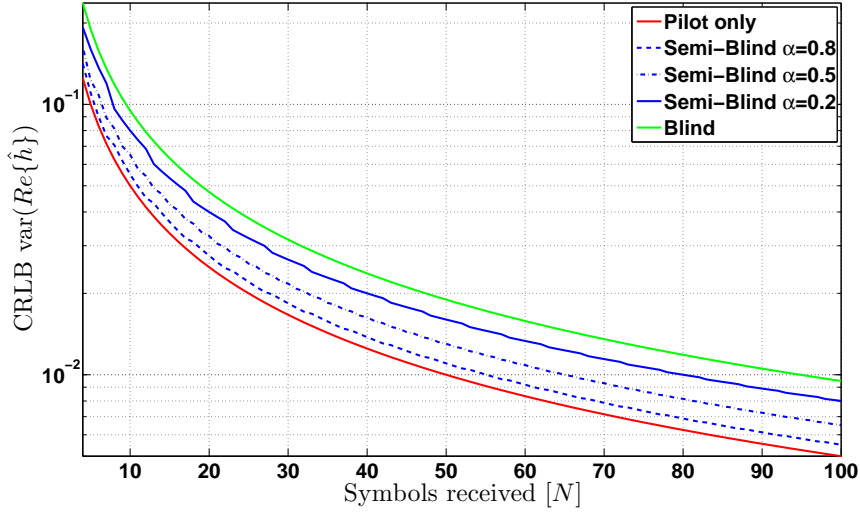


Figure 3.2: Comparison of the estimate variance \hat{h} , obtained by the CRLB of each channel estimation method plotted along the number of received symbols N , assuming an SNR=0 dB.

nentially when the number of received symbols N is increased. For a small number of received symbols, the performance of the estimation is low, since the variance of the channel estimate is high, thus deviating from the true channel fading coefficient, as shown in (2.76). However, when the number of the received symbols is slightly greater, e.g., $N = 10$, the variance reduces significantly. This proves that there can exist a minimum threshold on the number of received symbols in order to obtain accurate channel estimates by each method. On the other hand, a maximum threshold on the number of pilot symbols used for channel estimation can also be considered in order to avoid overwhelming the network resources, especially when no useful information is transmitted. In total, it can be seen that for an SNR of 0 dB, the pilot method gives the best estimates. The semi-blind method for $\alpha = 0.8$ also performs well, while the other methods give results of a slightly lower performance.

This proves that the signalling overhead which is expected to overwhelm the networks due to the increased number of transmitted pilot symbols, can be reduced depending on the SNR level and the selected channel estimation method, while still being able of achieving acceptable channel estimates. Table 3.1 compares the variances of the three channel estimation methods in a numerical way, for three different numbers of received symbols.

Since the performance of the three channel estimation methods changes for different

Table 3.1: Approximate variance of the real channel estimate, obtained by the CRLB for a fixed number of symbols of each channel estimation method and SNR=0dB .

	N=4	N=50	N=100
Blind	0.2371	0.0190	0.0095
Semi-Blind $\alpha = 0.2$	0.1927	0.0160	0.0080
Semi-Blind $\alpha = 0.5$	0.1629	0.0130	0.0065
Semi-Blind $\alpha = 0.8$	0.1414	0.0110	0.0055
Pilot only	0.1250	0.01	0.005

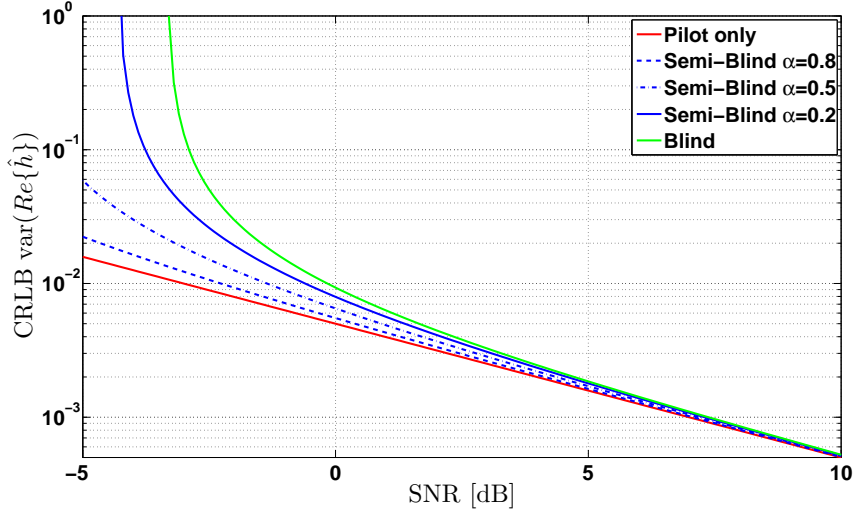


Figure 3.3: Comparison of the estimate variance obtained by the CRLB of each channel estimation method plotted along different SNR values, considering $N = 100$ received symbols.

SNR levels, a comparison of the channel estimate's variance obtained from the CRLB for different SNR levels and for a fixed number of received symbols, such that $N = 100$, is given in Figure 3.3. For low SNR levels the best performance is achieved by the pilot based channel estimation method since the known pilot sequence at the receiver side can be retrieved from signals affected by increased noise, resulting in low SNR. On the contrary, the blind method provides the least accurate estimates since it is only able of obtaining channel estimates, above approximately -3.3dB of SNR, while furthermore, the semi-blind $\alpha = 0.2$ has also limited performance for lower SNR levels. For higher SNR levels however, all the methods provide approximately similar estimates with a decreased variance.

The third comparison is a combination of the previous, where the variance obtained by the CRLB of the real part of the channel is presented along three SNR levels and different number of received symbols, as illustrated in Figure 3.4. Once more, for low

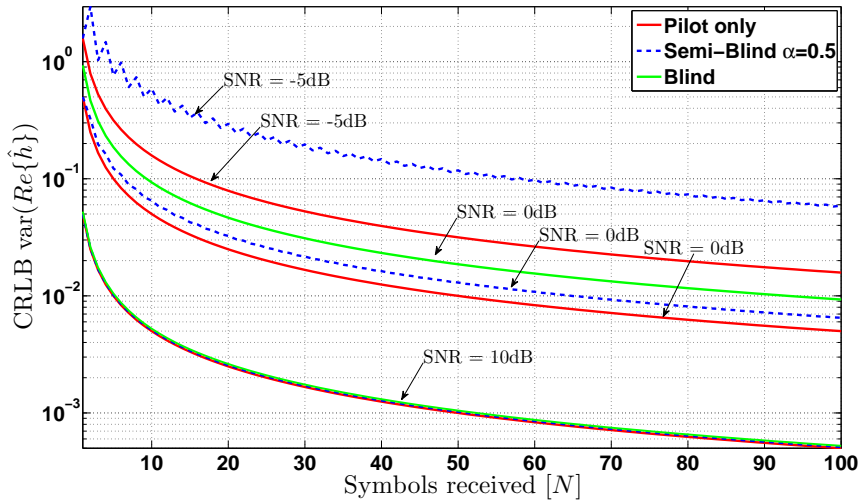


Figure 3.4: CRLB comparison of three channel estimation methods plotted for SNR=-5dB, SNR=0dB and SNR=10dB, along the number of the received symbols N .

SNR scenarios, e.g., -5 dB, the blind channel estimation method fails into providing channel estimates. For 0dB SNR, the results are the same as in Figure 3.2, while for 10dB SNR all the channel estimation methods provide similar performance and obtain accurate channel estimates since the variance is significantly decreased.

3.3 Trade-offs for each channel estimation method

As it has been briefly explained in section 1.5 of the introduction, in order to derive the trade-offs of each channel estimation method, two performance metrics are considered. First, the corresponding constrained capacity should be evaluated and compared (2.86), based on the channel estimates derived upon considering the CRLB according to (2.76). On the other hand, we consider the number N of the symbols used to achieve the desired accuracy of the channel estimate and if useful information is transmitted through this process. The transmission of information data and a resulting useful rate during the channel estimation process is a desired outcome and it is expected to reduce the signalling overhead which occurs by using only dedicated symbols for channel estimation. The consideration that generally occurs, is that for a long number of pilot symbols, an accurate channel estimate can be obtained, thus allowing the network to achieve higher rates due to the resulting constrained capacity which is obtained after the channel estimation process. However, less time and frequency resources are left for actual data transmission, before the channel changes again. On the contrary, poor channel estimates will lead to lower corresponding achievable rates, thus the ability to transmit more information bits per second per Hertz will diminish. Therefore, the trade-off between the need of an accurate channel estimate and the transmission of useful data during the channel estimation process, should be examined upon considering ultra-dense networks deployment in 5G communication systems.

3.3.1 Constrained capacity over a quasi-static channel

A real time scenario, considers the channel estimate \hat{h} obtained by a channel estimator and detector, as it was suggested in the introduction of this thesis. For the simulation

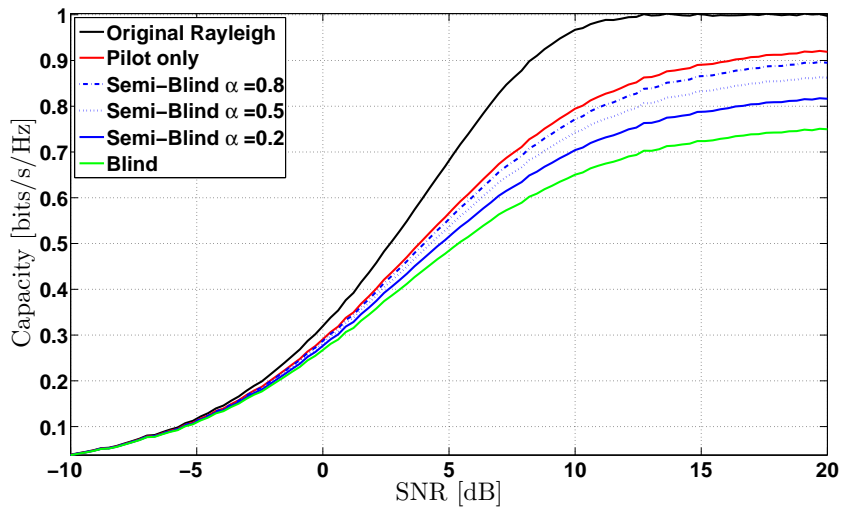


Figure 3.5: Constrained capacity over a quasi-static channel, \hat{h} , along different SNR values. The maximum constrained capacity is compared with the constrained capacities which correspond to each channel estimation method for $N = 4$ received symbols

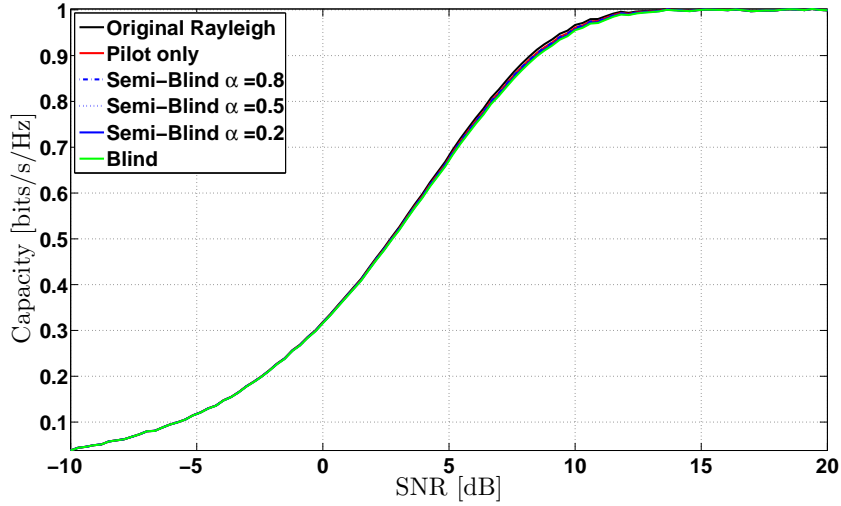


Figure 3.6: Constrained capacity over a quasi-static channel, \hat{h} , along different SNR values. The maximum constrained capacity is compared with the constrained capacities which correspond to each channel estimation method for $N = 100$ received symbols

process, \hat{h} is taken to be a single random channel sample which follows a Rayleigh distribution. Assuming that each transmitted symbol of the N -length vector faces the same channel fading coefficient over a coherence time period, the channel is considered quasi-static while the noise varies for every transmitted symbol. Based on the CRLB equations of each method (2.27), (2.52) and (2.75) where the results are used in order to derive Δh in (2.76) and upon considering the channel \hat{h} in the same equation, the constrained capacity of a quasi-static channel can be calculated according to (2.86). The Monte Carlo simulation considers the deterministic and constant channel estimate \hat{h} , thus the expectation in (2.87) is taken only with respect to the random noise. The corresponding constrained capacities are plotted in Figure 3.5 and Figure 3.6. for $N = 4$ and $N = 100$ received pilot or data symbols.

Figure 3.5 shows that for a small number of received symbols, the channel estimates do not achieve high accuracy, since the estimate's variance is high as shown in Figure 3.2. Thus, the maximum achievable constrained channel capacity, denoted by the black line, cannot be achieved even for high SNR scenarios. In order to be able to achieve the maximum constrained channel capacity, the number of the symbols used to estimate the channel, should be increased. This is done upon using $N = 100$ symbols for the channel estimation process, as it can be seen in Figure 3.6. All the channel estimation methods give sufficient estimates with decreased variance, as previously shown in Figure 3.2, thus they are able of achieving increased capacity. As a result, the maximum constrained capacity is achieved and the resulting constrained capacities of each channel estimation method overlap.

On the other hand, it must be also considered, that the usage of $N = 100$ symbols dedicated only for channel estimation, will potentially overwhelm the time and frequency domain resources, especially in a scenario which assumes ultra-dense deployed networks. The pilot based channel estimation, which requires the transmission and reception of $N = 100$ dedicated pilot symbols, will lead into inefficient usage of the air resources. On the contrary, the results of Figure 3.6 prove that the semi-blind or the blind channel estimation method can de-congest the system and assist into transmitting more useful data.

Finally, Figure 3.7 illustrates the corresponding constrained capacity of each method

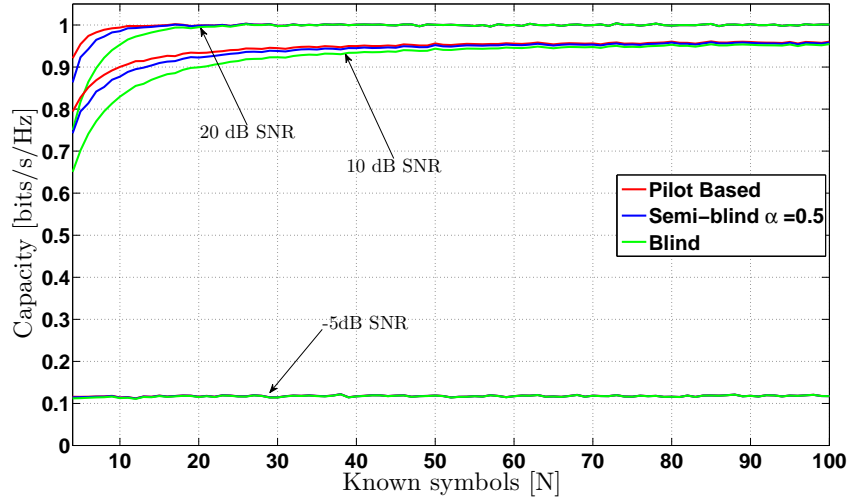


Figure 3.7: Constrained channel capacities over a quasi-static channel along the number of received symbols N , for different SNR levels, equal to -5dB, 10dB and 20dB.

over a quasi-static channel for different numbers of received symbols N and for three SNR levels. It can be seen that when $\text{SNR} = -5\text{dB}$, the corresponding constrained capacity is reduced, while for higher SNR levels the capacity reaches approximately the maximum value of 1 bit per second per Hertz. It should also be highlighted that especially for high SNR scenarios, the maximum capacity is approximately achieved by slightly increasing the number of symbols used for channel estimation, e.g., from $N = 4$ to $N = 20$. Thus, the need of a threshold on the number of the symbols required for accurate channel estimates, if the pilot channel estimation is used, should be also considered.

3.3.2 Ergodic constrained capacity

For the ergodic case, it is assumed that N symbols, comprised of N_D data or N_P pilot symbols, depending on the channel estimation method, are transmitted over many *iid* channel fading realisations which belong in a Rayleigh distribution. As a result, the transmitted symbols in a form of packets, extend over many channel coherence periods and the channel can be considered fast fading. In order to get the ergodic capacity, new values of μ_r (2.38), μ_i (2.41) and μ (2.45) are derived for every channel iteration, upon using a Monte Carlo simulation. This is due to the fact that the resulting CRLB for the blind and semi-blind method is based on these values which also depend on each channel realisation. Afterwards, the constrained capacity is simulated based on the CRLB and the channel realisation, as shown in equations (2.76) and (2.86). This process is repeated for many Rayleigh distributed channel fading coefficients and complex Normal distributed noise realisations, in a second Monte Carlo simulation, while the final result of the ergodic constrained capacity is averaged over all these realisations.

The only difficulty on obtaining results of high accuracy for the ergodic capacity, lies in the fact that two nested Monte Carlo simulations are used. Since this kind of simulation is considered time consuming, the iterations which were finally used, were slightly lower than the ones mentioned in the simulation assumptions, i.e., $L = 10^2$ for the simulation of μ_r , μ_i and μ , while the results were averaged over $L = 10^5$ channel and noise realisations. As a result, the ergodic constrained capacity obtained for every channel estimation method is compared in Figure 3.8 and Figure 3.9 for $N = 4$ and $N = 100$ received symbols respectively. For the case where $N = 100$ symbols are received, the

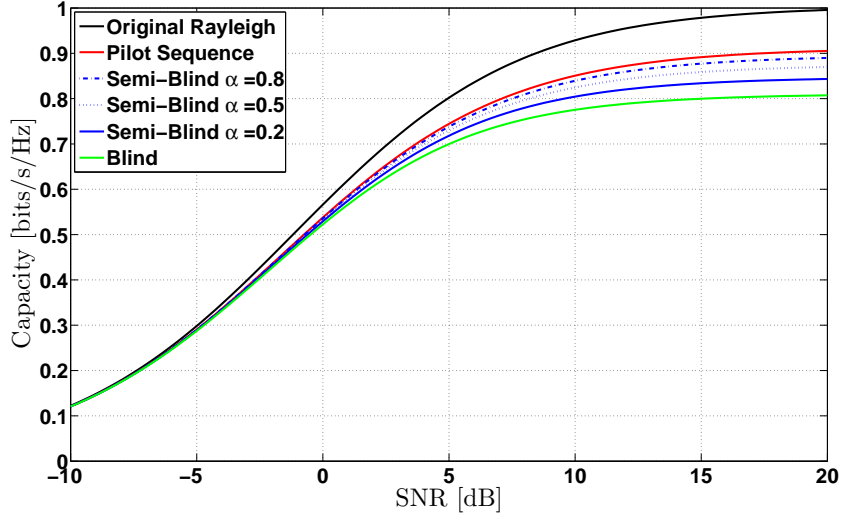


Figure 3.8: Ergodic constrained capacities over Rayleigh fading channel realisations, with $N = 4$ received symbols for each channel estimation method, plotted along different SNR values.

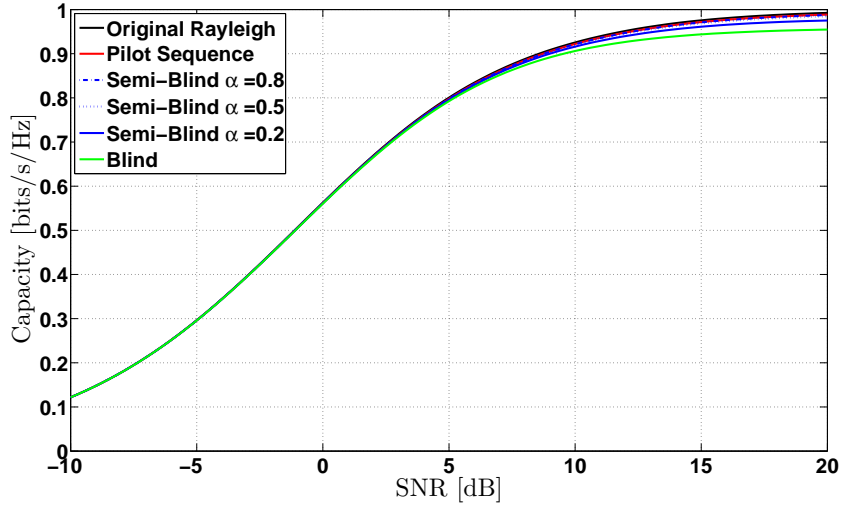


Figure 3.9: Ergodic constrained capacities over Rayleigh fading channel realisations, with $N = 100$ received symbols for each channel estimation method, plotted along different SNR values.

blind channel estimation method is not able of achieving the maximum channel capacity and this is due to the significantly lower number of two Monte Carlo simulation for μ_r , μ_i and μ , which affects the CRLB result and introduces higher uncertainty in Δ_h of (2.76).

3.3.3 Useful and signalling rate

Upon obtaining the corresponding constrained capacity for a quasi-static channel realisation, the useful rate (UR) and the signalling rate (SR) are defined, such that

$$\text{UR} = \frac{N - N_P}{N} R_x \quad \text{and} \quad \text{SR} = \frac{N_P}{N} R_x \quad (3.1)$$

where N corresponds to the number of total received symbols, N_P is the number of pilot symbols and $N - N_P = N_D$ is the number of data symbols which are used by each

channel estimation method. Furthermore $R_x \in \{R_P, R_B, R_{SB}\}$ denotes the achievable rate that corresponds to each channel estimation method as it is derived in terms of the constrained capacity in section 3.3.1 and as presented in Figures 3.5 and 3.6.

One of the goals for the future ultra-dense networks, is to decrease the signalling rate, while still being able to obtain accurate channel estimates and achieve high capacity. When the useful rate is also increased, the system will be able to achieve higher data rates since data symbols will be transmitted during the channel estimation process. Furthermore, by assuming an arbitrary symbol mapping function, we are able to derive the useful and signalling rate which can find application in a real network scenario. In addition, since the channel is assumed quasi static, the estimated channel fading coefficients can be used also for the adjacent frequency and time resources of each element, as mentioned in [15].

The same numbers of received symbols are considered, $N = 4$ and $N = 100$, which as it was previously shown provide different quality of channel estimates. These symbols are mapped in two time and frequency grids, similar to Figure 1.2. For the first case, four different frequency elements extending over one time slot are considered, such that a 1×4 grid is obtained, where pilot or data symbols of each method can be mapped. Subsequently, when $N = 100$, a 10×10 grid is assumed, occupying respectively 10 time and 10 frequency slots. Table 3.2 presents the number of pilot and data symbols which are mapped into the time and frequency grid with respect to each channel estimation method.

Table 3.2: Resulting numbers of pilot and data symbols for each channel estimation method on a 2×2 and 10×10 time and frequency grid.

	$N = 4$	$N = 100$
Blind	$N_P = 0 \ \& \ N_D = 4$	$N_P = 0 \ \& \ N_D = 100$
Semi-Blind $\alpha = 0.2$	$N_P = 1 \ \& \ N_D = 3$	$N_P = 20 \ \& \ N_D = 80$
Semi-Blind $\alpha = 0.5$	$N_P = 2 \ \& \ N_D = 2$	$N_P = 50 \ \& \ N_D = 50$
Semi-Blind $\alpha = 0.8$	$N_P = 3 \ \& \ N_D = 1$	$N_P = 80 \ \& \ N_D = 20$
Pilot only	$N_P = 4 \ \& \ N_D = 0$	$N_P = 100 \ \& \ N_D = 0$

We first examine the pilot based method which provides the most accurate channel estimates, thus achieves higher corresponding constrained capacity, as it was previously illustrated in Figures 3.2, 3.5 and 3.6. On the contrary however, during the channel estimation process no useful information is transmitted which as a result potentially creates a signalling overhead. When $N = 4$ symbols are used, the useful rate obtained by the pilot based method is given, such that

$$\text{UR}_P = \frac{4-4}{4}R_P = 0 \quad (3.2)$$

due to the fact that all the symbols are pilot symbols and no data are transmitted during the channel estimation method. On the other hand, the signalling rate is found to be

$$\text{SR}_P = \frac{4}{4}R_P = R_P \quad (3.3)$$

which reaches the maximum value, thus can possibly overwhelm the network resources when ultra-dense networks are considered. Consequently, extending the results for $N = 100$ received symbols, it is derived

$$\text{UR}_P = \frac{100-100}{100}R_P = 0 \quad \text{and} \quad \text{SR}_P = \frac{100}{100}R_P = R_P \quad (3.4)$$

which is the similar result as before.

When the blind channel estimation method is evaluated, it has been shown that the obtained channel estimates are of slightly lower accuracy, as shown in Figure 3.2. However, this method is still capable of providing an acceptable level of corresponding constrained capacity, as presented in Figures 3.5 and 3.6. For $N = 4$ received symbols, the useful and signalling rate are found to be

$$\text{UR}_B = \frac{4-0}{4}R_B = R_B \quad \text{and} \quad \text{SR}_B = \frac{0}{4}R_B = 0 \quad (3.5)$$

where the useful rate is maximized while the signalling rate is minimized. The same results are obtained upon considering $N = 100$ received symbols, such that

$$\text{UR}_B = \frac{100-0}{100}R_B = R_B \quad \text{and} \quad \text{SR}_B = \frac{0}{100}R_B = 0 \quad (3.6)$$

The semi-blind channel estimation method, is capable of transmitting both pilot and data symbols, thus reducing the signalling overhead, compared to the pilot based method. Moreover it obtains more accurate channel estimates in comparison to the blind method. For $\alpha = 0.2$ and $N = 4$ received symbols, it is found that

$$\text{UR}_{\text{SB1}} = \frac{4-1}{4}R_{\text{SB1}} = \frac{3R_{\text{SB1}}}{4} \quad \text{and} \quad \text{SR}_{\text{SB1}} = \frac{R_{\text{SB1}}}{4} \quad (3.7)$$

while for $\alpha = 0.5$

$$\text{UR}_{\text{SB2}} = \frac{4-2}{4}R_{\text{SB2}} = \frac{R_{\text{SB2}}}{2} \quad \text{and} \quad \text{SR}_{\text{SB2}} = \frac{R_{\text{SB2}}}{2} \quad (3.8)$$

and for $\alpha = 0.8$

$$\text{UR}_{\text{SB3}} = \frac{4-3}{4}R_{\text{SB3}} = \frac{R_{\text{SB3}}}{4} \quad \text{and} \quad \text{SR}_{\text{SB3}} = \frac{3R_{\text{SB3}}}{4} \quad (3.9)$$

Finally, when $N = 100$ symbols are considered, the results for each semi-blind case are derived such that

$$\text{UR}_{\text{SB1}} = \frac{100-20}{100}R_{\text{SB1}} = \frac{4R_{\text{SB1}}}{5} \quad \text{and} \quad \text{SR}_{\text{SB1}} = \frac{20}{100}R_{\text{SB1}} = \frac{R_{\text{SB1}}}{5} \quad (3.10)$$

while

$$\text{UR}_{\text{SB2}} = \frac{100-50}{100}R_{\text{SB2}} = \frac{R_{\text{SB2}}}{2} \quad \text{and} \quad \text{SR}_{\text{SB2}} = \frac{50}{100}R_{\text{SB2}} = \frac{R_{\text{SB2}}}{2} \quad (3.11)$$

and finally

$$\text{UR}_{\text{SB3}} = \frac{100-80}{100}R_{\text{SB3}} = \frac{R_{\text{SB3}}}{5} \quad \text{and} \quad \text{SR}_{\text{SB3}} = \frac{80}{100}R_{\text{SB3}} = \frac{4R_{\text{SB3}}}{5} \quad (3.12)$$

The rate R_x is finally replaced by the corresponding constrained capacity of each method, $\{R_P, R_B, R_{\text{SB}}\}$, over a quasi-static channel as presented in Figures 3.5 and 3.6, for a fixed 5dB SNR level. The upper plot of Figure 3.10, presents the corresponding useful rate of each channel estimation method for $N = 4$ and $N = 100$ received symbols. It can be seen that the blind method provides the highest useful rate during the channel estimation process, with an achievable rate $R_B = 0.5320$ [bits/sec/Hz] for $N = 4$ received data symbols and $R_B = 0.6015$ [bits/sec/Hz] for $N = 100$ received data symbols. The pilot channel estimation method does not provide a useful rate, thus $R_P = 0$ for every number N of received symbols, since the received symbols are all pilot symbols dedicated for channel estimation, thus no useful data is transmitted. The semi-blind channel

Table 3.3:

	UR Slope (m)	SR Slope (m)
Blind	7.2356×10^{-4}	0
Semi-Blind $\alpha = 0.2$	3.1747×10^{-4}	1.0582×10^{-4}
Semi-Blind $\alpha = 0.5$	2.2205×10^{-5}	1.0508×10^{-4}
Semi-Blind $\alpha = 0.8$	1.1102×10^{-5}	3.3307×10^{-5}
Pilot only	0	0

estimation method provides different results depending on the α pilot to total symbols ratio value (2.65).

The signalling rate is presented in Figure 3.10, lower plot. This plot illustrates the

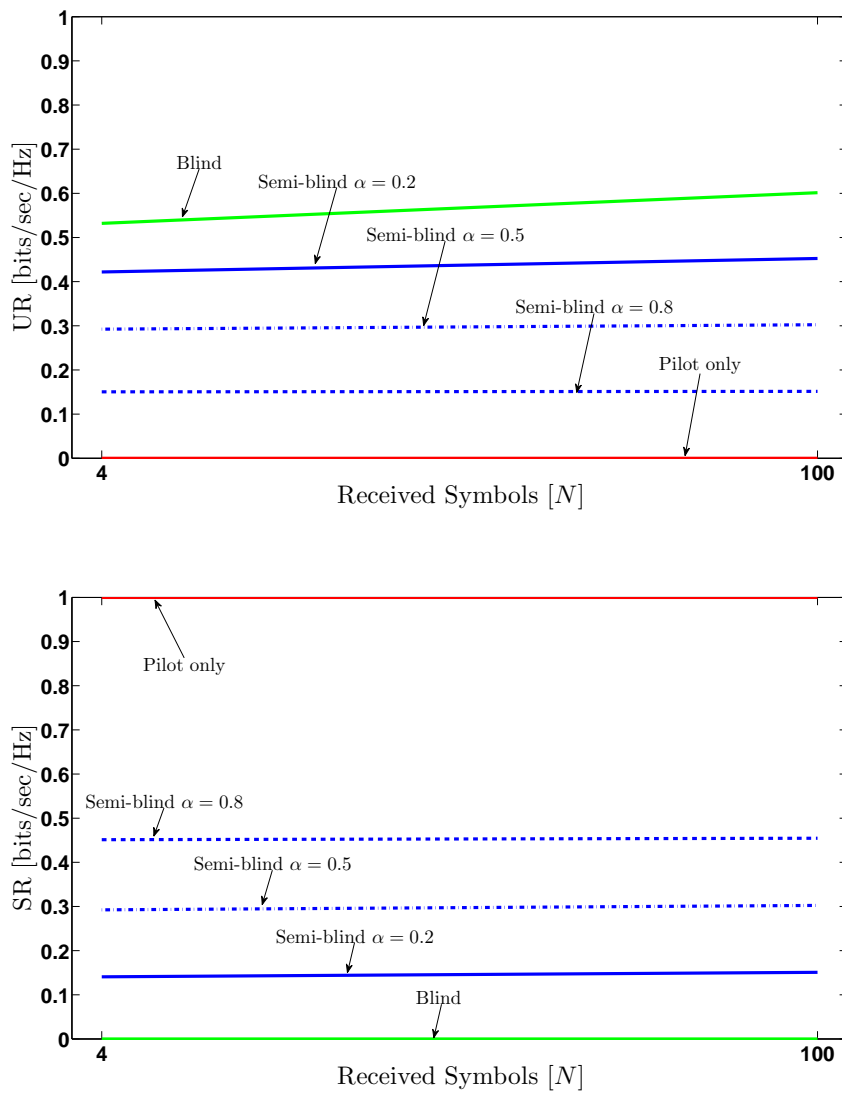


Figure 3.10: Comparison of the achievable useful and signalling rate for every channel estimation method, with respect to $N = 4$ and $N = 100$ received symbols along a fixed 5dB SNR level.

fact that the pilot based method generates the most signalling rate, as expected, with the semi-blind methods providing average signalling rates, while the blind channel estimation method does not generate any signalling, which is a desired outcome.

It must be also highlighted that the useful rate results are inversely proportional to the signalling rate results, since when the useful rate is increased the signalling rate is diminished, as seen in the above figure for the blind and pilot channel estimation method. Table 3.3 presents the slopes of the useful and signalling rate lines of each channel estimation method, as presented in Figure 3.10. As it can be seen, the blind channel estimation method has the highest slope of the useful rate, thus if more data symbols are used for channel estimation $N_D \geq 100$, the useful rate is expected to increase when the blind channel estimation method is used. The above outcome can find application in online transmissions or mobile broadcasting services which is expected to be of high importance for future networks and it can also decrease the signalling overhead.

4

Conclusions

Despite its exploratory nature, this thesis offers some insight into the performance of the channel estimation methods which can find application in future of 5G communication systems. Due to this nature of the thesis, the three channel estimation methods are evaluated by different aspects of their performance and the way the results are obtained. To support this idea, there are published several scientific papers which examine similar aspects, considering different scenarios and set-ups which provide a variety of results and findings. However, this thesis makes mostly noteworthy contribution to the following three aspects:

- The channel estimates accuracy for a number of received symbols and different SNR levels,
- The corresponding constrained capacity over a quasi-static channel assuming different SNR levels,
- The resulting useful and signalling rate of each method, for a number of transmitted symbols during the channel estimation process.

Upon using the CRLB, we were able to analyse and evaluate the accuracy of obtained the channel estimates, as shown in equations (2.27), (2.52) and (2.75). Likewise, the channel used for the rest of the evaluation process, considered these channel estimates, as shown in (2.76). Furthermore, the corresponding constrained capacity over a quasi-static channel and the ergodic constrained capacity over many channel realisations, were simulated. The final results were obtained upon using system level Monte Carlo simulations.

Channel estimates accuracy

This thesis has first shown with the usage of the CRLB, that upon considering the number of received pilot or data symbols and for an SNR of 0 dB, the most accurate channel estimates are obtained by the pilot based method, as shown in Figure 3.2. Besides, all methods provide more accurate estimates when the number of received symbols N is increased. When high SNR scenarios are considered, the performance of all methods is similar, thus the blind or semi-blind methods should be preferred due to the fact that they provide a useful rate thus increased throughput. As shown in Figure 3.4, for low SNR levels the best performance is obtained first by the pilot method and then by the semi-blind method with $\alpha = 0.2$ (2.65), which performs quite similar. In contrast, the semi-blind $\alpha = 0.8$ and the blind method fail into providing channel estimates for low

SNR values, thus their usage should be reconsidered for relevant scenarios where low SNR is expected to occur. To conclude, when low SNR scenarios are considered, the pilot method should be used since the channel estimation process provides results robust to the noise. On the other hand, for higher SNR scenarios the blind or semi-blind $\alpha = 0.2$ method should be preferred. This is due to the fact that the accuracy of the resulting channel estimates is improved and upon using the blind and semi-blind method, the pilot overhead is minimized.

Corresponding constrained capacity

The second remark of this thesis, is related to the accuracy of the channel estimates, which was previously discussed. It is found that the channel estimates obtained by the pilot based method, provide the highest corresponding constrained capacity when a small number of pilot symbols is used, i.e., $N = 4$, upon assuming high SNR scenarios and as shown in Figure 3.5. However, the semi-blind $\alpha = 0.8$ method is also able of providing channel estimates which correspond to a constrained capacity very close to the one obtained upon using the pilot based method. This fact makes it a very good alternative, when the reduction of signalling rate is required. On the contrary, the simulations have shown that when the number of the received symbols is increased, i.e., $N = 100$ and a high SNR level is assumed, there is no need of overwhelming the system with pilot symbols. This conclusion is derived since the estimates of the blind method are capable of providing the maximum achievable capacity on a Rayleigh distributed channel, as shown in Figure 3.6. Thus, the blind method should be preferred when an increased number of received symbols is used during the channel estimation process, or during online transmissions. On the other hand, when low SNR scenarios are assumed, independently of the number of received pilot or data symbols, the value of the corresponding constrained capacity is very low, no matter which channel estimation method is used.

Useful and signalling rate

Finally, the last remark of this thesis contributes on the resulting useful and signalling rate, obtained by each channel estimation method. In summary, it has been shown that the blind channel estimation provides the maximum useful rate which results in a total increased throughput, as presented in Figure 3.10. It should be mentioned that the resulting useful rate depends on the number of received data symbols and the SNR level. This result is expected since all the symbols are data symbols, which include information bits, thus more information is being transmitted and the signalling rate is diminished. On the other hand, the method which uses only pilot symbols, does not succeed into providing a useful rate, since the transmitted symbols do not carry any information. Furthermore it generates the most signalling rate which overwhelms the network's resources, as shown in Figure 3.10. In addition, the semi-blind methods give average results depending on the ratio α that is used.

Overall, it has been shown that the final decision on if the blind and semi-blind channel estimation methods should be used, in order to achieve the objective of pilot overhead reduction, depends both on the SNR level and on the number of the total received symbols N . For low SNR levels and a few received symbols, the classic pilot based method is proved to be the best option for obtaining accurate channel estimates, however resulting in pilot overhead. On the other hand for high SNR levels, the blind channel estimation method is able of providing channel estimates of high accuracy, given that the transmission of an increased number of symbols is expected, i.e., more than 100

symbols. In addition, the semi-blind method also achieves good performance by making use of several known symbols in each transmission.

Consequently, based on the key motivation of ultra-dense networks deployment in 5G communication systems, where the SNR is expected to be increased, [7], [1] and upon considering the transmission of a vast number of data packets, e.g., HD video content, the use of blind and semi-blind methods could be preferred opposite to the pilot based method. This decision can lead on the de-congestion of the network resources from the extensive use of pilot symbol sequences, while it further allows the transmission of more actual data symbols, which can potentially result in higher achievable data rates and throughput.

4.1 Future work

This thesis and its final results can be extended if further research is conducted in the future. As it has been shown in this report, several assumptions have been used in order to derive the obtained results. These assumptions can be changed and thus provide a different insight on how the channel estimation methods perform upon assuming 5G scenarios. For example, different channel models can be assumed, such as the Rician or the Nakagami- m channel model and new channel models that consider mmWave for 5G communication systems, which are under development as suggested in [10]. Furthermore, since this report assumes BPSK modulated symbols, higher order symbol constellations can be investigated in the future, which will result in different likelihood functions, thus different CRLB results. In addition the system model can be of higher spatial orders, in order to result in a MIMO system. Finally, since the set up of the semi-blind channel estimation method is adjustable, further than the already deployed different pilot and data symbol ratio α , the ratio of the power levels used for the transmission of the pilot symbols compared to the data symbols, i.e. PDPR, can be further investigated. To conclude, all these further assumptions and scenarios, can result in a better understanding of how the usage of different channel estimation methods can assist in the reduction of the pilot overhead, which is expected to occur in the future 5G networks.

Bibliography

- [1] J. G. Andrews, S. Buzzi, W. Choi, S. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, “What Will 5G Be?” *IEEE J. Sel. Areas Commun.*, vol. PP, no. 99, May 2014.
- [2] METIS, “Scenarios requirements and KPIs for 5G mobile and wireless system,” Deliverable D1.1 ICT-317669-METIS/D1.1, April 2013.
- [3] E. Dahlman, G. Mildh, S. Parkvall, J. Peisa, J. Sachs, Y. Selen, “Ericsson Review: 5G radio access,” June 2014.
- [4] Cisco, “Visual Networking Index: Global Mobile Data Traffic Forecast,” White paper, February 2014.
- [5] “Ericsson Mobility Report,” White Paper, June 2014.
- [6] Nokia Siemens Networks Smart Labs, “Understanding Smartphone Behaviour in the Network,” White paper, 2011.
- [7] J. G. Andrews, “Beyond 4G Cellular Networks: Is density all we need?” WNCG Dept. of Electrical and Computer Engineering The University of Texas at Austin, Tech. Rep., 2013.
- [8] F. Boccardi, R. W. Heath Jr., A. E. Lozano, T. L. Marzetta, and P. Popovski, “Five Disruptive Technology Directions for 5G,” *IEEE Commun. Mag.*, vol. abs/1312.0229, no. 2, pp. 74–80, February 2014.
- [9] Nokia Siemens Networks: Liquid Net, “LTE Release 12 and Beyond,” White paper, 2012.
- [10] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, J. Wong, G.N. Schulz, M. Samimi, and F. Gutierrez, “Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!” *IEEE Access*, vol. 1, pp. 335–349, 2013.
- [11] U. Jang, K. Y. Lee, K. Cho, and W. Ryu, “Downlink Transmit Beamforming for Inter-Cell Interference Mitigation with BS Cooperation,” in *IEEE Global Telecommunications Conference (GLOBECOM 2010)*, December 2010, pp. 1–5.
- [12] METIS, “Position of multi-node/multi-antenna technologies,” Deliverable D3.1 ICT-317669-METIS/D3.1, July 2013.
- [13] N. Bhushan, J. Li, D. Malladi, R. Gilmore, D. Brenner, A. Damnjanovic, R. Sukhavasi, C. Patel, and S. Geirhofer, “Network densification: The Dominant Theme for Wireless Evolution into 5G,” *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 82–89, February 2014.

- [14] M. Farber, "Densification in mobile networks and the potential evolution paths of the base station," in *8th International Symposium on Wireless Communication Systems (ISWCS)*, November 2011, pp. 181–185.
- [15] E. Dahlman, S. Parkvall, and J. Skold, *4G LTE/LTE-Advanced for Mobile Broadband*, 1st ed. United Kingdom: Elsevier Ltd, 2011.
- [16] S. Sessia, M. Baker, and I. Toufik, *LTE–The UMTS long term evolution : From theory to practice*, 1st ed. United Kingdom: John Wiley & Sons Ltd, 2009.
- [17] J. Proakis and M. Salehi, *Digital Communications*, 5th ed. New York: McGraw-Hill, 2008.
- [18] T. L. Marzetta and B. M. Hochwald, "Fast transfer of Channel State Information in wireless systems," on *IEEE Trans. Signal Process.*, vol. 54, no. 4, pp. 1268–1278, April 2006.
- [19] J. Jose, A. Ashikhmin, T. Marzetta, and S. Vishwanath, "Pilot contamination Problem in Multi-Cell TDD Systems," in *Information Theory IEEE International Symposium*, June 2009, pp. 2184–2188.
- [20] J. Tugnait, L. Tong, and Z. Ding, "Single-user channel estimation and equalization," *IEEE Signal Processing Mag.*, vol. 17, no. 3, pp. 16–28, May 2000.
- [21] J. Ayadi, E. De Carvalho, and D. T. M. Slock, "Blind and semi-blind maximum likelihood methods for fir multichannel identification," in *Acoustics, Speech and Signal Processing proceedings of the 1998 IEEE International Conference*, vol. 6, May 1998, pp. 3185–3188.
- [22] S. M. Omar, D. T. M. Slock, and O. Bazzi, "Bayesian semi-blind FIR channel estimation algorithms in SIMO systems," in *IEEE 12th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, June 2011, pp. 421–425.
- [23] B. Hassibi and B. Hochwald, "How much training is needed in multiple-antenna wireless links?" on *IEEE Trans. Inform. Theory*, vol. 49, no. 4, pp. 951–963, April 2003.
- [24] A. Goldsmith, *Wireless Communications*, 1st ed. UK: Cambridge University Press, 2005.
- [25] S. M. Kay, *Fundamentals of Statistical Signal Processing, Vol.1: Estimation Theory*, 1st ed. New Jersey: Prentice Hall Signal Processing Series, 1993.
- [26] R. G. Gallager, "Circularly-Symmetric Gaussian random vectors," <http://www.rle.mit.edu/rgallager/documents/CircSymGauss.pdf>, Accessed: 2014-05-12.
- [27] ARTIST4G D1.4, "Interference Avoidance Techniques and System Design," Deliverable, July 2012.
- [28] W. E. Ryan and S. Lin, *Channel Codes Classical and Modern*, 1st ed. The Edinburgh Building, Cambridge UK: Cambridge University Press, 2009.
- [29] S. L. Miller and D. Childers, *Probability and Random Processes*, 2nd ed. Oxford UK: Elsevier, 2012.
- [30] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New Jersey, USA: John Wiley and Sons, 2006.