OPEN ACCESS

University of BRISTOL

This electronic thesis or dissertation has been downloaded from Explore Bristol Research, http://research-information.bristol.ac.uk

Author:
Chitambira, Benny

Title:
Towards Localisation in Next-generation Wireless Systems

General rights
Access to the thesis is subject to the Creative Commons Attribution - NonCommercial-No Derivatives 4.0 International Public License. A copy of this may be found at https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode This license sets out your rights and the restrictions that apply to your access to the thesis so it is important you read this before proceeding.

Take down policy
Some pages of this thesis may have been removed for copyright restrictions prior to having it been deposited in Explore Bristol Research. However, if you have discovered material within the thesis that you consider to be unlawful e.g. breaches of copyright (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please contact collections-metadata@bristol.ac.uk and include the following information in your message:

• Your contact details
• Bibliographic details for the item, including a URL
• An outline nature of the complaint

Your claim will be investigated and, where appropriate, the item in question will be removed from public view as soon as possible.

Towards Localisation in Next-generation Wireless Systems

By

BENNY CHITAMBIRA

Department of Electrical & Electronic Engineering
UNIVERSITY OF BRISTOL

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of DOCTOR OF PHILOSOPHY in the Faculty of Engineering.

MARCH 2021

Word count: 42611

Abstract

Localisation in wireless communication systems is an important topic that has many applications. Accurate geolocation in urban environments is a well understood challenge and Global Navigation Satellite System (GNSS) based technologies like GPS yield poor accuracy in non-line-of-sight (NLOS) scenarios. Use of multiple GNSS systems improves location accuracy but is still affected by urban canyons. Mobile network-based schemes offer an alternative, particularly when antenna arrays are deployed. Massive antenna arrays in Massive Multiple-Input-Multiple-Output (Massive MIMO) present an opportunity because of the possibility to use inexpensive, low-power and low-precision components with greatly reduced complexity/cost, in addition to Massive MIMO being a core component of 5G.

This thesis reviews the problem of localisation in urban environments and investigates the techniques that are relevant towards achieving localisation in next generation wireless systems. It evaluates the NLOS problem and proposes schemes that use machine leaning techniques in the form of Least-Squares Support Vector Machines (LSSVMs), to address the challenges. It also investigates Direction of Arrival (DOA) estimation, which is a step towards localisation, using the Bristol Massive MIMO testbed.

The proposed location specific approach presented in this thesis is a new framework which is shown to achieve a best case NLOS identification accuracy of greater than 98 percent. This result exceeds reported accuracies for existing NLOS identification techniques. Another proposed direct localisation approach achieves an 80th percentile probability location accuracy of 10 metres without utilising NLOS identification and mitigation. DOA estimation experiments demonstrate the possibility of performing estimation using subsets of antennas on a single base station antenna array. An outdoor experiment demonstrated a best case DOA RMSE of 2 degrees achieved for 3 different User Equipment (UE) positions.

Overall, the results in this thesis demonstrate techniques that can be employed to solve the intermediate challenges towards device localisation and can contribute to the design and operation of next-generation systems. Significant benefits for mobile wireless systems can be derived from localisation. New handover strategies, new dynamic resource, power and pilot allocation schemes that take advantage of the location information, can also be developed.

Acknowledgements

I would like to thank, first, my supervisors, Dr. Simon Armour and Professor Mark Beach, and my industrial supervisor, Stephen Wales, for their immense support and invaluable insight throughout the course of my PhD research. I would also like to thank my colleagues Paul Harris, Wael Boukley and Henry Brice for their support during my experimentation with the massive MIMO testbed.

I would also want to extend my gratitude to academic colleagues at Lund University for their help with the testbed experiments, and also BT (formerly British Telecom) for collaborating with our research team and making possible, the experiments that were conducted at their site in Ipswich. My thanks also go to the Bristol City Council, who, through their Bristol-is-Open Initiative, allowed me access to the Massive MIMO testbed system.

I would also like to acknowledge the funding from the EPSRC via the CDT in Communications at the University of Bristol, together with additional financial support from Roke Manor Research, which enabled me to undertake my research.

Lastly, but not least, I would like to thank my family for their moral and material support, in addition to their continuous encouragement.

Author's declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: ..... DATE: .....

Table of Contents

List of Tables

List of Figures

Figure **Page**

Acronyms

3G third generation mobile cellular system

3GPP 3rd Generation Partnership Project

4G fourth generation mobile cellular system

5G fifth generation mobile cellular system

ADMA angle division multiple access

AIC akaike's information criterion

ALPA adaptive likelihood particle filter

APML alternating projection maximum likelihood

AOA angle of arrival

AOD angle of departure

AWGN additive white gaussian noise

BIO Bristol is open

BS base station

CDMA code-division multiple access

CDF cumulative distribution function

CFO carrier frequency offset

CIR channel impulse response

CSI channel state information

CSN communication systems and network

DAC digital to analogue converter

DF direction finding

DOA direction of arrival

DSP digital signal processing

eCID enhanced cell-ID

eNB evolved NodeB

E-SMLC evolved serving mobile location centre

EKF extended Kalman filter

ESPRIT estimation of signal parameters via rotational invariance techniques

FCC federal communications commission

FDD frequency division duplex

FLEX-RIO flexible reconfigurable I/O

FPGA field-programmable gate array

GDOP geometric dilution of precision

GLONASS globalnaya navigazionnaya sputnikovaya sistema, or global navigation satellite system

GNSS global navigation satellite systems

GPS global positioning system

GR ground reflection

HDF hybrid data fusion

IDE integrated development environment

ITU international telecommunications union

LBS location-based services

LIDAR laser illuminated detection and ranging

LOS line-of-sight

LPP lte positioning protocol

LTE long-term evolution

mmWave millimeter wave

MDL minimum description length

MIMO multiple input and multiple output

ML maximum likelihood

MRC mobile radio cellular

MRN mobile radio network

MS mobile station

MU-MIMO multi-user MIMO

MUSIC multiple signal classification

MVB merchant ventures building

NLOS non-line of sight

NOMA non-orthogonal multiple access

O-TDOA observed time difference of arrival

OCXO oven controlled crystal oscillator

OFDM orthogonal frequency division multiple access

OMA open mobile alliance

OTA over the air

PCIe peripheral component interconnect express

PDP power delay profile

PF particle filter

PKF position Kalman filter

PLL phase-locked loop

PSS primary synchronisation signal

PXI PCI extensions for instrumentation

QP quadratic programming

QPSK quadrature phase shift keying

RB resource block

RBF radial basis function

RF radio frequency

RFID radio frequency identification

ROC receiver operating curve

RRH remote radio head

RSRP reference signals received power

RSTD reference signal time difference

RSS received signal strength

RSSI received signal strength indicator

RTT round trip time

SAGE space alternating generalised expectation maximisation

SDR software defined radio

SISO single input and single output

SLP SUPL Location platform

SPDT single pole double throw

SMA SubMiniature version A

SNR signal to noise ratio

SRM structural risk minimisation

SRS sounding reference signals

SUPL secure user plane location

SVM support vector machine

TADV timing advance

TDD time division duplex

TDE time delay estimation

TDOA time difference of arrival

TOA time of arrival

ToF time of flight

TTFF time to first fix

UE user equipment

U-TDOA uplink time difference of arrival

UKF unscented Kalman filter

USRP universal software radio peripherals

UWB ultra wide band

WLSE weighted least squares estimator

WLAN wireless local area network

NOTATION

F_c Carrier frequency (Hz) $R1$ Range (distance) of MS or UE from base station 1

$BS1$ First base station Base station 1

R_i Range (distance) of MS or UE from base station i (m)

$R_{i,1}$ Distance difference between BS i and BS1

τ_i Time delay as seen on BS i

$\tau_{i,1}$ Time difference of arrival between first BS and BS i

x_i Known x-coordinate for BS i (or MS i where appropriate)

y_i Known y-coordinate for BS i (or MS i where appropriate)

x Unknown x-coordinate for the target or MS/UE

y Unknown y-coordinate for the target or MS/UE

B Total number of base stations

c Speed of propagation (m/s)

T_{ij} Cotangent of the bearing angle between BS i and BS j

$y(t)$ Time domain continuous received signal

$x(t)$ Time domain continuous transmitted signal

M Total number of co-channel signals in DOA estimation

N Number of antennas / antenna elements on an array

L Number of sampled signal snapshots

λ_i Eigenvalue corresponding to signal source i

$P(\theta)$ MUSIC spectrum

Ξ Array error matrix

XX DOA beamscan

e_i Location error (m)

X_k k^{th} input for the LSSVM

Y_k LSSVM Output corresponding to the X_k input

w Vector normal to the SVM separating hyper-plane

b SVM classifier parameter

h_{BS} Height of the BS

h_{MS} Height of the MS

α_i Received signal power for MS position i

CHAPTER 1

INTRODUCTION

Location of an object, a machine, an animal or a person, is an essential task that can be used to aid asset management, emergency services and recently, location based services and gaming [1]. Common methods that are used to determine the location or position in different systems or applications make use of radio waves (Radio Frequency (RF) signals) or sound waves (Sonar). Examples of methods that are RF signal based are RAdio Detection And Ranging (Radar), Radio Frequency Identification (RFID) and RF baseband signals (e.g. in Wireless sensor networks), Wi-Fi based, Ultra-Wide Band (UWB) Radio techniques and Global Navigation Satellite Systems (GNSS) methods such as the Global Positioning System (GPS). Geolocation or localisation applications can be found in navigation (naval or air), animal tracking and in communications where Location-Based-Services (LBS) have become a real business opportunity and can aid or enhance the operation of a wireless network through better management of resources [2].

In communication systems, Mobile Radio Cellular (MRC) systems such as Long Term Evolution (LTE) systems, and GNSS systems are used to determine the location of mobile devices. The location-based services that use these systems range from emergency services, to commercial applications like local advertising and simple finding of a friend or business. While mobile users have the option to enable GNSS on their devices for LBSs, often enjoying improved accuracy through the use of assisted GNSS (A-GNSS), these may not be enabled at times of emergencies.

For emergency services, there are often stipulated statutory location accuracy that is required of a network operator, utilising only their Mobile Radio Network (MRN). The Federal Communications Commission (FCC) in the USA has got the Enhanced 911 mandate which requires network operators to avail to emergency call dispatchers, the location of a cell phone in a 911 call [3], at the same time meeting the defined regulatory and mandatory accuracy requirements. In 2003,

the European Union (EU) issued the E112 directive for cellular network operators which requires them to avail to emergency services, the location from where an emergency call is received using any information available to them.

The accuracy requirements prescribed by the regulatory authorities must be met and also be guaranteed regardless of the underlying positioning technique or the mobile device's operating environment or capability. This poses some environmental and operational challenges to the engineers designing these systems, particularly in urban areas where the signals can be highly shadowed, or when users are within buildings. GNSS based solutions in urban areas are affected by the urban canyon effect [4], making it difficult for the mobile device to have a clear line-of-sight (LOS) to the required minimum of three satellites, most of the time. They are also not able to provide indoor localisation. Older mobile devices may not have GNSS or A-GNSS technology. Multi-GNSS systems improve location accuracy, but still have been reported to have achieved a best-case 2-D accuracy of 147m, in a worst-case urban canyon environment [5]. Their results suggest that multi-GNSS using GPS+GLONASS+Galileo improves horizontal accuracy by about 18% on average, for the 2 urban environments considered.

On the other hand, MRN localisation had been primarily based on Enhanced Cell ID (eCID), sometimes in conjunction with Angle-of-Arrival (AOA), time delay or received signal strength (RSS) measurements at the cell of origin (COO). From 3GPP LTE Release 9, time difference of arrival (TDOA) methods started to become the preferred technique [6]. AOA and TDOA techniques require specially calibrated hardware for super-resolution accuracy. With massive multiple-input-multiple-output (Massive MIMO) technology being a core component in next generation wireless systems and already a key technology in the fifth generation (5G) systems [7], it presents an opportunity for improved accurate positioning because of the possibility to use inexpensive, low-power and low-precision components, with greatly reduced complexity/cost. The need to reduce dependency on calibrated antennas and taking advantage of simple and inexpensive components (instead of specialised localisation or direction-finding equipment) means that technologies like Massive MIMO may be able to offer that hardware requirement relaxation. Massive MIMO also offers increased degrees of freedom when direction of arrival (DOA) estimation is performed using different subsets of antennas with the best result selected according to set confidence criterion.

The drive for mobile radio network positioning is further motivated by the significant potential system benefits that can be derived from mobile user location information, towards the next generation systems like millimetre wave (mmWave) technology and Massive MIMO. If these systems can perform location estimation during user detection or channel estimation, that information can be used to aid user scheduling by addressing some of the challenges of these next-generation systems, like reducing pilot contamination, power allocation, designing new handover techniques or new dynamic resource allocation techniques. These potential benefits are discussed in detail in section 1.1 below. Recent activity towards availability of practical testbeds [8], [9] allows these ideas to be explored and evaluated.

1.1 5G and Massive MIMO

5G refers to the 5th Generation cellular technology which offers lower latency and better speeds than the third (3G) and fourth generation (4G) technologies. 5G also offer better and increased wireless connectivity, owing to greater spectral efficiency. Figure. 1.1 shows the download speeds using the three cellular technologies.

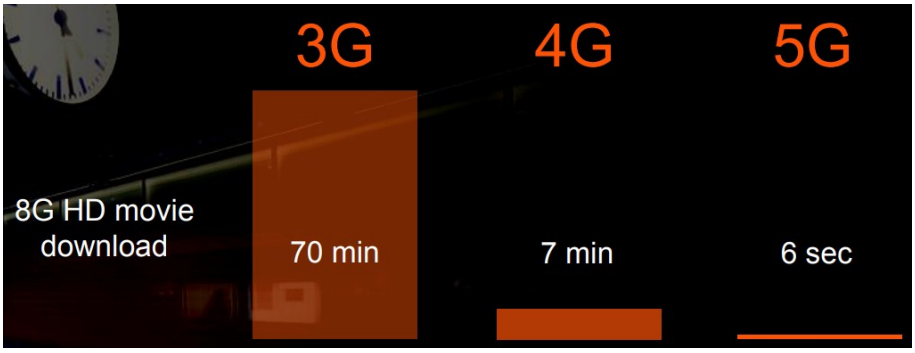


Figure 1.1: 3G to 5G movie download time comparisons [10].

5G networks are envisaged to achieve latencies that are less that a millisecond, whereas in comparison, current 4G networks have approximately 50ms [10]. Samsung contend that theoretically 5G is 100 times faster than 4G [11]. About 12 billion mobile devices are expected to be connected to the internet by 2022, according to a report by Cisco [12]. This ever-increasing number of connected devices places an increase on the bandwidth requirements. 5G is envisaged to enable the evolution of new technologies in machine-to-machine (M2M), vehicle-to-everything (V2X), medical robotics, Internet of Things (IoT) and industrial applications. Economically, the impact of 5G, according to World Economic Forum, should reach $12 trillion dollars by 2035 [13].

Massive MIMO is a candidate technology for evolved 5G mobile networks, and next generation wireless systems. There is a lot of research around massive MIMO owing to the significant benefits it has to mobile systems in terms of capacity and data rates. It can therefore be inferred that the next generation base-stations will employ many antennas in some way. Figure 1.2 below shows a representation of a multi-cell, Multi-user massive MIMO network. Massive MIMO employs Spatial Multiplexing (SMX) and takes the concept of Multi-user MIMO (MU-MIMO) to the highest level, by deploying hundreds of antennas at the BS, each with its own individual Radio Frequency (RF) chain. This allows the BS to serve tens of users or User-Equipment (UEs) at the same time, using the same frequency resource, with greater reliability. This makes the higher spectral efficiencies and improved energy efficiencies possible. The theoretical benefits and challenges of massive MIMO are discussed in [7], [15] and [16].

Traditional communication channels consisted of a single transmitter and a single receiver. This is what is now commonly referred to as a single-input-single-output (SISO) channel. Early research shows that multiple-input-multiple-output (MIMO) systems, where there are multiple

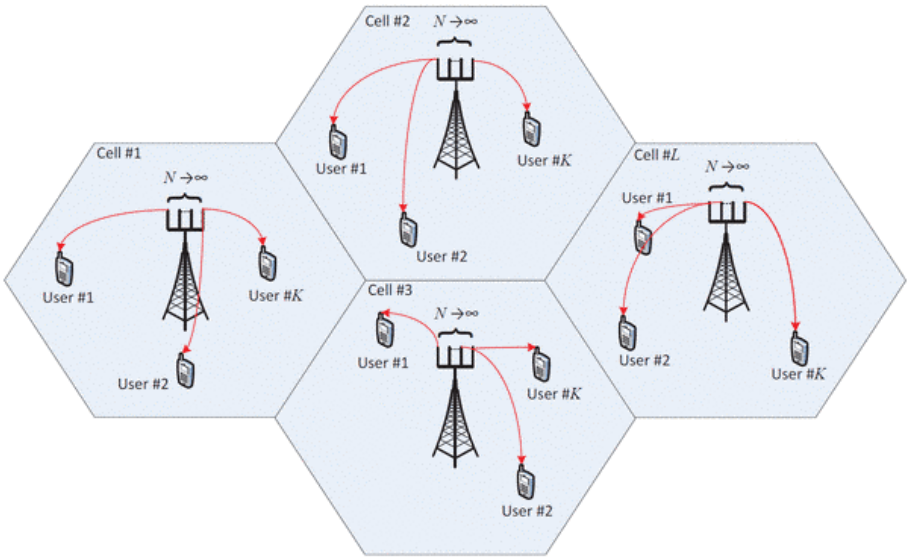


Figure 1.2: A representation of a multi-cell, Multi-user Massive MIMO network [14].

transmitters and multiple receivers within the same system, started to get considered in attempts to address problems of mutual interference [17]. Consequently, there was realisation that MIMO systems can actually take advantage of multipath propagation and OFDM processing [18], with Bell Labs demonstrating practical results [19] which proved the viability of the theory. During the time of their inception, MIMO systems, as suggested by Gerard Foschini et al. [20], around late 90s, appeared to violate the Shannon limit, but it was noted that MIMO actually transforms a SISO link into multiple parallel channels and thus deriving a multiplicative effect on the channel capacity. The channel capacity of a SISO link is a function of the SNR as shown in the equation in Figure. 1.3.

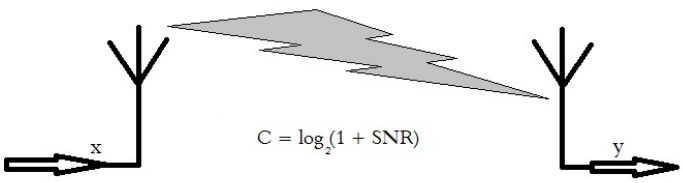


Figure 1.3: Typical SISO link and its channel capacity.

It can be shown that doubling the power does not double the channel capacity. A brief mathematical review of MIMO below helps to conceptualise the key benefits of Massive MIMO. For a SISO channel with received signal y and transmitted signal x, the system can be written as in (1.1) where h is the impulse response of the channel and n is the noise.

$$y = hx + n \tag{1.1}$$

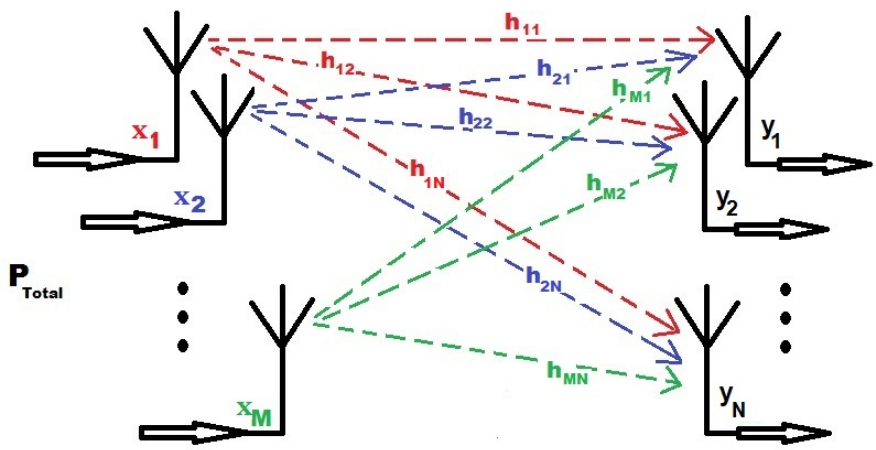


Figure 1.4: $M \times N$ MIMO channel.

From the Figure 1.4 above, it can be noted that for for MIMO, the individual channel components between each transmit antenna and receive antenna can form a channel matrix $\boldsymbol{H}$ and the MIMO relationship can be described in matrix form as

$$\boldsymbol{Y} = \boldsymbol{HX} + \boldsymbol{N} \tag{1.2}$$

where, for an $M \times N$ MIMO channel, with M denoting the number of transmitting antennas and N denoting the number of receiving antennas (note the difference between this and $\boldsymbol{N}$ which denotes the noise matrix), the channel matrix can be expressed as

$$\boldsymbol{H} = \begin{bmatrix} h_{11} & h_{12} & \dots & h_{1N} \\ h_{21} & h_{22} & \dots & h_{2N} \\ . & & & \\ . & & & \\ . & & & \\ h_{M1} & h_{M2} & \dots & h_{MN} \end{bmatrix} \tag{1.3}$$

The advantage of a MIMO system can be realised in two ways, diversity gain from transmitting the same information over multiple channels, and spatial multiplexing where each channel carries different data. Also one can imagine a situation where there are enough antennas to

gain from both diversity and spatial multiplexing. The channel capacity of a MIMO link can be expressed as [21] [20]

$$C = \max_{tr(\boldsymbol{R_{XX}})=P_T} log_2 det\left(\boldsymbol{I_M} + \frac{1}{\sigma_m^2}\boldsymbol{H}\boldsymbol{R_{XX}}\boldsymbol{H}^H\right) \quad bits/s/Hz \tag{1.4}$$

where $\boldsymbol{R_{XX}}$ is the covariance matrix of the transmitted signal and σ^2 is the signal noise variance.

If N_t is the number of transmit antennas (to differentiate it with the notation for noise), and M is the number of receive antennas, the signal from the transmitter travels over $N_t \times M$ multiple channels and is then recombined in the receiver to realise the channel gains. It is therefore easy to see why an increase in the number of antennas at both ends result in an increase in channel gains hence capacity.

With Multi-user MIMO (MU-MIMO) in modern communication systems, where multiple users access the system via a single base-station, as shown in Figure 1.2, it can easily be noticed that the system benefits from spatial multiplexing, where different data is sent to different users simultaneously. As discussed above, it can be realised that increasing the number of antennas on the BS increases the spectral efficiency of the system. Furthermore, the sensitivity of the system to the propagation environment is much reduced since MU-MIMO benefits from multi-user diversity.

The received signal for a point to point link can be expressed as

$$\boldsymbol{Y} = \sqrt{\rho}\boldsymbol{H}\boldsymbol{X} + \boldsymbol{N} \tag{1.5}$$

where $\boldsymbol{X} \in \mathbb{C}^{N_t \times 1}$, $\boldsymbol{Y} \in \mathbb{C}^{M \times 1}$ and ρ is the total power when the transmit signal is normalised. Assuming the transmit signal is complex Gaussian i.i.d, with perfect channel-state-information (CSI) available, no correlation between the individual MIMO links, and zero-mean Gaussian noise, whose covariance matrix can be expressed as an identity matrix $\boldsymbol{I}$, the channel capacity for channels with equal power can be expressed as [22]

$$C = log_2 det\left(\boldsymbol{I} + \frac{\rho}{N_t}\boldsymbol{H}\boldsymbol{H}^H\right) \quad bits/s/Hz \tag{1.6}$$

Of particular interest, from MU-MIMO systems where it is expected that the number of antennas on the base station (M) is much larger than the number of users simultaneously served by that base station ($N_t = K$, where K is the common notation for number of single antenna user equipment in massive MIMO). It is shown in [23] that for cases where single antenna users, i.e. $K \leq M$, the BS is capable of spatially multiplexing the users, there by allowing simultaneous communication between the users and the BS. MU-MIMO systems used in LTE for example, have managed to employ $M \leq 8$ [23]. Massive MIMO expands the concept to employ hundreds of antennas at the BS. The large number of antennas on the BS also enhances channel hardening which is discussed in [24].

The concept of having more antennas on the BS is a positive proposition for location systems. This is discussed in great detain in Chapter 3 section 3.3 and further theoretic and experimental results that look at the effect of number of antennas for location techniques, are also presented in Chapters 3 and 5.

But the promise of Massive MIMO comes with potential problems of scheduling users when multiple adjacent cells are considered, especially for cellular communications. Channel estimation involves sending a predetermined pilot signal, from the user equipment, to the base station. Pilot contamination [22] occurs when a pilot signal meant for a user terminal's controlling BS, is also received on the adjacent cell base station, together with its own user's pilots. Power consumption

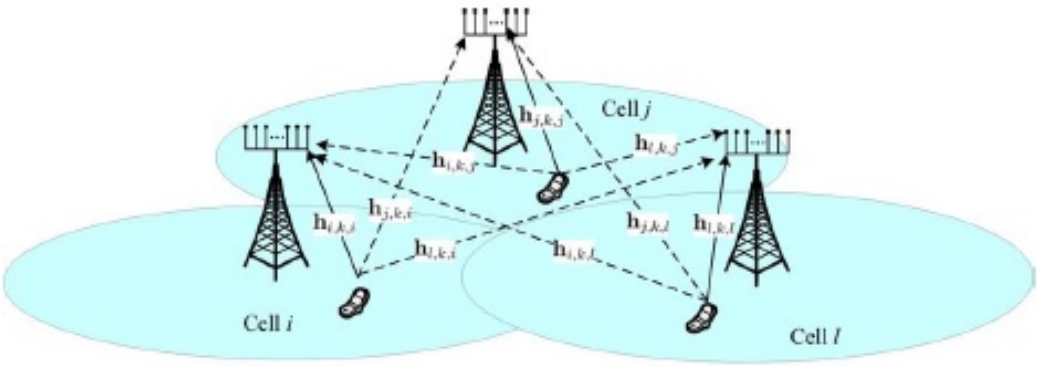


Figure 1.5: Pilot contamination during uplink transmission [22].

in Massive MIMO is determined by the radio chains that drive the large number of antennas on the base station. In such a system, during periods when the number of users is very low, no gain from spectral efficiency is realised. This leads to poor energy efficiency. This issue can be addressed by adapting the array configuration to the load in terms of number of users connected [25].

The opportunity that localisation presents, towards tackling pilot contamination is discussed in section 1.2. This author believes that while localisation can benefit from an increased number of antennas on the base station, the communication system can also benefit from localisation when user location results are fed back to inform and possibly adapt the system to take into consideration, the user position information, thereby alleviating some of the system challenges. Other system challenges, in addition to pilot contamination, are user grouping in massive MIMO. A brief discussion of these challenges, and the opportunities provided by localisation, is presented in the following section.

1.2 Localisation and next-generation wireless systems: Opportunities and benefits

Massive MIMO provides an opportunity for superior positioning schemes based on the highly accurate Angle of Arrival or Angle of Departure (AOA/AOD) information that can be obtained via the use of massive antenna arrays. In addition to that, it presents an opportunity for much simpler, single BS localisation by utilizing the LOS AOA and mobile station range information obtained either by Time of Arrival (TOA) or Time Difference of Arrival (TDOA). These techniques are discussed in detail in Chapter 2. Next generation wireless systems like Massive MIMO and mmWave can benefit greatly from employing localisation (in form of angular information, range information, or geolocation), or tracking, to address some of their challenges. User/mobile detection is usually done as a first step in modern systems. If localisation can be done at the same time, there are potential benefits such an approach can bring, to the next-generation wireless systems.

- **Pilot contamination**
  Pilot contamination is one of the major issues with mobile cellular systems. Dynamic pilot allocation may be a solution that reduces pilot contamination. Pilots can be allocated in such a way that all the UEs with similar AOAs are prevented from sharing the same pilot [26].

- **Channel estimation**
  Recent research [27] proposed a practical channel estimation for "massive MIMO mmWave" that exploits location. It proved that the true DOAs for each uplink multipath can be extracted using their efficient technique for array signal processing, and also information pertaining the gain for the channel can be obtained linearly utilising only small amounts of training resources, and both these two parameters can then be used to build the channel estimation. One of the challenges of Massive MIMO is channel estimation, so if localisation, like in this reported case, or as proposed in [26], can aid or improve channel estimation, then the true benefits of massive MIMO can be realised.

- **NOMA, ADMA and User grouping**
  Existing cellular networks allocate user resources based on orthogonal multiple access (OMA). Orthogonality amongst users requires high bandwidth and also low latency if a massive number of devices has to be supported. Non-Orthogonal Multiple Access (NOMA) is a multiple access technique that addresses the spectral efficiency challenge of OMA techniques. NOMA techniques can generally be categorised under either code domain or power domain approaches. Under code domain approaches, multiple users can be multiplexed over the same time and frequency resource by allocating them different codes [28] [29] [30]. Under power domain approaches, different users are multiplexed by using different power

coefficients which are determined by their channel condition. This allows superimposition of multiple users' data at the base station and the receivers employ successive interference cancellation (SIC) to decode the signals [31].

One of the challenges of the large number of antennas as in Massive MIMO, manifest in form of the large size of the channel matrices. The computational complexity posed by such large matrices, can be reduced by reducing the dimensions of the channel matrices if low-ranking channel characteristics like direction of arrival (DOA) are considered [32]. The spatial based expansion model (SBEM) [33], which is essentially Angle Division Multiple Access (ADMA), demonstrates that the angular information directly corresponds to the directions of the users, and such direction information can be exploited together with uplink/downlink reciprocity to reduce the downlink channel estimation overheard and complexity [34].

Localisation can be used to extend the application of NOMA systems. AOA knowledge for users within a cell, may be used to group them, and those users with similar AOAs but significantly different power levels, can then employ NOMA to further improve capacity. Clustering of users in co-operative NOMA can be informed by location information. ADMA can similarly be developed, by considering the UEs AOAs.

User grouping and scheduling in Massive MIMO systems is a hot topic, that can benefit from user location information. A colleague of this author has already demonstrated the need for user grouping in Massive MIMO [35] , and has presented an adaptive user grouping algorithm which maximises spectral efficiency in their publications [36]. Any user grouping algorithm that takes account of location information will have reduced complexity and will benefit a system that has to handle a large number of users at the same time.

- **Resource allocation**
  In multi-user systems, performance of dynamic radio resource allocation techniques, including precoding techniques, depends on the channel characteristics on the users involved. Location information together with knowledge of channel characteristics like the location specific NLOS identification schemes, such as those discussed in [37] can be used to design new resource allocation strategies. Received signal strength heat-map style tools can be developed, to be used in implementing dynamic and adaptive modulation and coding schemes.

- **New handover strategies**
  If the base station could build a picture of the environment, with respect to the way mobile devices are moving, this could be used to design new handover strategies. A desirable and reliable hand-off point can be predicted by taking use of the UE's geolocation and inertial measurements. This could be achieved by employing map-based radio prediction techniques such as the one proposed in [38].

- **Power control**
  Power Control in communications, is key to good performance. The effect of power control errors has been studied since third generation (3G) systems [39]. When a UE moves from a LOS position to a highly shadowed, NLOS position in an urban area, that change is most likely to be abrupt because of building edges. Closed-loop power control algorithms struggle to cope with these changes and this leads to power control errors. If power control algorithms could take into consideration, the location information, together with prior knowledge of the environment, the power control errors can be reduced.

- **Beamforming**
  mmWave systems can benefit from localisation by using the location information of the UE to aid downlink beamforming. AOA estimation can be performed using sub-6GHz transmission and the beamforming that exploits the AOAs done at mmWave frequencies.

- **Other considerations**
  MIMO systems in general offer better signal detection than SISO systems as discussed above in section 1.1 and this benefit is even greater for Massive MIMO systems. This means devices can transmit at very low levels of power. Low power devices like those used for sensor networks and IoT can be located easily using a Massive MIMO base station. Also, the robustness of Massive MIMO detection can be used to keep devices transmitting at their minimum levels and thus improving energy efficiency.

This research seeks to explore, investigate and validate some of the identified and proposed techniques relevant towards localisation in next generation wireless system. Traditional techniques like TDOA and AOA are explored using ray tracing data, an emerging machine learning technique is investigated and real-world measurements are used to explore a proposed direction of arrival estimation technique.

1.3 Resources for this research

This research utilises 2 main tools, the first one being the Bristol Ray-tracing tool (Prophecy), which is introduced and discussed in detail in Chapter 3, but used throughout the research to investigate traditional localisation techniques, to develop NLOS identification and mitigation schemes, and to validate outdoor DOA results produces using the Massive MIMO Testbed. The second tool is the Bristol Massive MIMO testbed itself, which is described in detail in Appendix C. All algorithms used for data processing were written in MATLAB, except for code that was used to write reciprocity calibration data to disk, from the Massive MIMO testbed, which was written in LabVIEW, as is the whole Massive MIMO Framework.

1.4 Contributions

This section summarises the contributions of this thesis, to the state of the art in wireless localisation. A list of publications and achievements gained during the course of the project are also presented, together with a description of the organisation of this thesis towards the end.

1.4.1 Summary of contributions

- **An evaluation of traditional techniques for urban localisation using ray-tracing data.** Traditional techniques like TDOA and AOA are evaluated in an urban environment using ray-traced data. The techniques considered, together with the data processing used, demonstrate that classical TDOA algorithm generally produces better localisation accuracy than AOA when considering the 80th percentile level. This is presented in Chapter 3. This work demonstrates a new framework for evaluating traditional localisation techniques, using the same data under the same framework, by exploiting ray-traced databases. The key outcomes are that TDOA is more suitable in NLOS environments while AOA is more suitable in LOS environments. Existing comparison of these techniques, as detailed in section 3.5 have focused of system differences without considering the environments in which the techniques are employed.

- **Methodologies for NLOS identification and mitigation, together with their evaluation.** Methodologies for NLOS identification and mitigation are designed, and evaluated for a general case, and for a location specific case. Usage scenarios for these approaches are also provided. These are discussed in Chapter 4. While the general case described in section 4.3 is comparable to existing techniques, the novel location specific approach demonstrates that AOA measurements in addition to received power and time delay, are key to achieving an NLOS identification accuracy that beats existing reported results as outlined in section 4.3.3. This scheme is based on ray-traced data and specially designed pre-preprocessing of individual rays. Such a framework had not been used in this manner before. To the best knowledge of the author, these proposals and data processing frameworks do not exist in literature.

- **Direct localisation using Ray-tracing and Least-squares support vector machines.** A direct approach for localisation using least-squares support vector machines and ray-tracing, is described, and evaluated for an urban environment. usage scenarios together with recommendations on practical application are also presented in Chapter 4. This novel framework exposes a completely new way of exploiting ray-traced data for localisation while at the same time demonstrating improved localisation performance in multipath environments. It exposes the relationship between the extent of multipath, as quantified by

the azimuth spread and the localisation performance. These outcomes are outlined from section 4.4.2 to section 4.4.4.

- **Direction of arrival estimation with a massive MIMO testbed.** A maximum likelihood algorithm is used for direction of arrival estimation with a massive MIMO testbed. The benefit of having multiple antennas is demonstrated in real world, and the algorithm is evaluated in both indoor and outdoor environments. The massive MIMO testbed that is used in this study was not developed for direction finding but still managed to achieved good accuracies comparable with some receivers that are specifically designed for direction finding although they use a much lower number of antennas. This work is presented in Chapter 5.

1.4.2 Published work

The following is a list of conference publications and a published journal article, all which stemmed off the work that was done towards the contributions described above.

- P. Harris, W. B. Hasan, H. Brice, B. Chitambira, M. Beach, E. Mellios, A. Nix, S. Armour, and A. Doufexi, "An overview of massive MIMO research at the University of Bristol," in *Radio Propagation and Technologies for 5G*, pp. 1–5, Durham, Oct 2016.

- B. Chitambira, S. Armour, S. Wales, and M. Beach, 'NLOS Identification and Mitigation for Geolocation Using Least-squares Support Vector Machines', in *2017 IEEE Wireless Communications and Networking Conference (WCNC) (IEEE WCNC 2017)*, San Francisco, USA, March 2017.

- B. Chitambira, S. Armour, S. Wales, and M. Beach, 'Direct Localisation using Ray-tracing and Least-Squares Support Vector Machines', in *2018 8th International Conference on Localisation and GNSS (ICL-GNSS 2018)*, Guimaraes, Portugal, June 2018.

- B. Chitambira, S. Armour, S. Wales, and M. Beach, 'Employing Ray-Tracing and Least-Squares Support Vector Machines for Localisation', *Sensors*, vol. 18, no. 11, Nov 2018.

1.4.3 Contributions to European Cooperation in Science & Technology (COST)

The author participated in the technical meetings of the latest European Cooperation in Science & Technology (COST) Action CA15104 (IRACON), on localisation related activities. The following paper was presented and discussed on the 6th technical meeting.

- B. Chitambira, S. Armour, S. Wales, and M. Beach, 'Localisation and Massive MIMO: Opportunities and Benefits', *COST Action CA15104, TD(18)06033*, Nicosia, Cyprus, Jan 2018.

1.4.4 Technical reports and white papers

This author was a contributing author to the following published white paper. Their contributions can be found in Sections V (D) and VI (C) of that white paper. The white paper was a deliverable of the Localisation group of the mentioned COST Action CA15104 (IRACON).

- Editors: Klaus Witrisal and Carles Antón-Haro, 'Whitepaper on New Localization Methods for 5G Wireless Systems and the Internet-of-Things', *COST Action CA15104 Technical report*, April 2018.

1.4.5 Other contributions

In addition to the above contributions, the author presented an invited IEEE Communications Society (ComSoc) webinar titled 'Addressing Direction Finding in Over-the-Air Testbeds and Prototypes[1]'. The author was a part of the Bristol team that set new world records in spectral efficiency[2]. The author contributed in setting up of the hardware and spatial recording of the experimental environment during the massive MIMO indoor measurement campaigns, and also during the demonstration campaign with British Telecom (BT), National Instruments (NI) and Lund University (Lund) at BT's research and development (R&D) headquarters in Adastral Park, Ipswich. For that work, the Bristol, Lund and BT teams were recipients of the 2017 Collaborate to Innovate (C2I) Award[3] in the Information, Data and Connectivity category.

1.5 Thesis structure

Chapters 2 to 5 of this thesis are organised in a manner that seeks to show the progression of the research from the original research motivations. Chapter 2 proffers a detailed literature review of the state of the art in radio localisation, outlining the traditional techniques and the science behind them. Advances in mobile radio network localisation are discussed and emerging research focusing on next-generation technologies like massive MIMO and mmWave, is presented. An account of how multiple antennas improve localisation is given and selected algorithms are discussed in detail. Section 2.5 provides literature on AOA localisation approaches which tie together the foreground DOA estimation work in Chapter 5 and how this is connected to ultimate device localisation.

Chapter 3 presents simulations and preliminary experiments. First a detailed account of the Bristol ray-tracing tool is provided after-which an evaluation of traditional techniques in the form of TDOA and AOA in an urban environment, which utilises ray-tracing data, is presented. Direction of arrival estimation simulation follows and the proposed technique is modeled using

[1]Webinar available at https://www.comsoc.org/webinars/addressing-direction-finding-over-air-testbeds-and-prototypes

[2]News article available at https://techxplore.com/news/2016-05-world-5g-wireless-spectrum-efficiency.html

[3]News article available at http://www.bris.ac.uk/news/2017/september/massive-mimo.html

an antenna array and the simulation results are presented. The ray-tracing simulations are a precursor to the work presented in Chapter 4 and the DOA simulations are a precursor to the work that is presented in Chapter 5.

Chapter 4 discusses 2 approaches to localisation in NLOS environments, first the NLOS identification plus mitigation approach and second, the direct localisation approach which avoids need for NLOS identification and mitigation. A brief outline of methods for NLOS identification and mitigation is supplied before the proposed methods are introduced. A detailed model of Least-squares Support Vector Machines (LSSVMs) is presented and an account of how they are employed for both NLOS identification and mitigation is supplied together with the results. A direct localisation approach that uses LSSVMs and Ray-tracing data is also presented, and the results are analysed. A comparison of the traditional localisation techniques after NLOS mitigation, and under the Direct approach, is performed using the same data set, and the results are presented.

Chapter 5 presents the results from different measurement experiments. Detailed experimental setup for each campaign is provided, and the results are discussed. Chapter 5 should be read with Appendix C which presents the University of Bristol's Massive MIMO testbed. It provides details of the massive MIMO testbed at the University of Bristol, that was used for DOA estimation. It also details the design of the base station with respect to the hardware components, synchronisation and the antenna array design and characterisation. It also provides details on the design of the user equipment. The thesis ends with conclusions which set out some suggestions and recommendations for further study on issues raised in the thesis. A discussion of key deductions from the work is also provided.

1.6 Summary and conclusions

This chapter has discussed the motivations for this research. Work that is already published is presented. The motivations and opportunities discussed in section 1.2 are not tested in this research, but are identified so that next generation systems can consider them. A discussion of the future work related to this research is outlined in Chapter 6. The next chapter discusses in detail, the techniques explored and proposed in this research.

CHAPTER 2

LOCALISATION IN WIRELESS COMMUNICATION SYSTEMS

The general motivations for localisation have been discussed in Chapter 1. Of particular importance and interest are the potential benefits of localisation to communication systems. These potential benefits are discussed in Chapter 1, section 1.2. These demonstrate the interest in developing techniques that are specifically applicable to wireless communication systems. This chapter reviews the state of the art in wireless localisation and links to background review of positioning in 3G and LTE systems (See Appendix A). It discusses in detail, the traditional localisation techniques of TOA, TDOA and AOA, providing a thorough comparison of the TDOA and AOA approaches. A relevant model and algorithm for each of them is supplied. Because these techniques are used as localisation algorithms in proposals that are presented throughout this thesis, the theory of time delay estimation and AOA estimation is provided, together with a forward-looking discussion of these topics. The proposed Maximum Likelihood (ML) algorithm for AOA estimation is then presented, together with its simulation results. The University of Bristol ray-tracing tool which was used extensively in this research, is introduced before some preliminary evaluation of the traditional techniques using the City of Bristol ray-traced data. The chapter is then concluded with a flow diagram of the relationships between various sections of this chapter, to some sections within the succeeding chapters.

2.1 Localisation using RF signals

Localisation using RF signals is traditionally achieved by two main mechanisms, range estimation and direction (angular domain based) methods. Direction finding methods mainly utilise angular or geometrical information of the RF signals as they reach a particular sensor or receiver. The common technique used in this case is Angle of Arrival (AOA). Distance/range estimation methods utilise two most common techniques that are propagation delay (time delay-based

methods) like Time of Arrival (TOA) and Time Difference of Arrival (TDOA), and signal strength based techniques like Received Signal Strength (RSS) method.

The Received Signal Strength Indicator (RSSI) technique uses units based on processed signal strength measurements, unlike in RSS method where the actual signal strength measurement is used. RSSI has been used widely in wireless sensor networks. The technique relies on the fact that radio signals decay as they propagate. No time synchronisation is required but a radio propagation model [40] is required to be able to determine the path loss and hence the distance or range of the receiving station from the transmitting source. Various path loss models are discussed in [41]. RSS techniques suffer heavily from the multipath problem because the transmitted signal can decay quite rapidly over short distances in highly shadowed environments like dense urban centres. Many path-loss models use empirically obtained path loss coefficients or path loss indexes which are location specific. However, depending on the environmental and weather conditions, the path-loss coefficient varies greatly, which makes location estimation very difficult or inaccurate. This technique therefore has not found much application in mobile communication systems. This thesis will therefore focus more on TOA/TDOA and AOA techniques.

TOA, which is sometimes referred to as time of flight (ToF) uses the absolute time the signal arrives at a receiving station, from a source. The distance of the target or receiving station, from the source, is inferred from the known propagation velocity of the signals. This means that this technique is applicable in LOS scenarios. TOAs measured at 2 BSs narrows the location to an intersection of 2 circles and a third BS measurement will be required to ascertain the position of the target. Many GNSS technologies like GPS utilise algorithms that make use of TOA or TDOA [42] [43] [44].

The TDOA technique uses at least 2 receivers. In the simplest 2-D geolocation setup, the difference in arrival time between two stations, is the measurement that is used, as shown in Figure 2.1. It is assumed the sensors/receivers or BSs are time synchronised, and with known positions/locations. The time difference in the arrival of the signal, from the transmitter to the two receivers, is then used to determine the location of the transmitter.

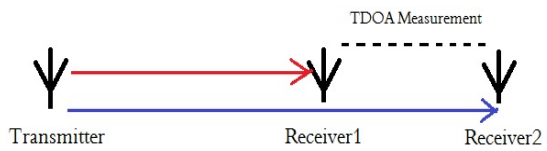


Figure 2.1: Range Detection Principle of TDOA

Common direction-finding systems generate bearings, which are actually angles of arrival of the signal from the target if a LOS propagation is assumed. Strictly speaking, AOA refers to the angle at which the signal from a target, arrives at the receiver, which could be an angle caused by the last reflection or diffraction in its propagation. Because the term AOA is often

used to describe the algorithm(s) that are used to calculate the actual location of the target or UE, the AOA Direction Finding (AOA-DF) that has just been described, is distinguished as the Direction of Arrival (DOA). AOA-DF or DOA may be used interchangeably. DOA determination is achieved by determining the power of the received signal relative to angular directions at the receiver. DOAs measured at different receiver stations or BSs, can then be used in an AOA algorithm (usually triangulation) to obtain the location coordinates of the UE. AOA does not require synchronisation of the receivers and works with all signal types. For arrays, accuracy of DOA estimation increases with the number of antennas as demonstrated in Chapter 3 sections 3.3 and 3.4. Appendix A provides an overview of the state of localisation in LTE. This provides the key references to literature on specifications as well as the background literature that builds into the choices made by this author with regard to TDOA and AOA.

2.2 A detailed look into TDOA and AOA

Common methods supported in LTE utilise AOA and TDOA algorithms, for example, enhanced Cell ID (eCID) and Uplink Time Difference of Arrival (UTDOA) methods (see Table A.2 in Appendix A) use these algorithms, and are capable of BS-based (eNB-assisted) positioning, which is key for UE or handset agnostic approaches where the mechanism has to be transparent to the user, and also operable regardless of the hardware capability of the UE. Other emerging techniques such as those discussed in Chapter 4, also use these algorithms. For these reasons, the two algorithms are therefore discussed (including a detailed comparison), modelled and evaluated with ray-tracing data, in the remainder of this chapter. The mathematical models presented here, are the same models for TDOA and AOA, that are used throughout this thesis.

2.2.1 TDOA vs AOA Comparison

A comprehensive comparison of these two traditional algorithms is published and maintained by the ITU in the form of ITU-Radiocommunication (ITU-R) reports under Spectrum Management (SM) and the latest of such, is from 2018 [45]. A summarised discussion of similar comparisons is presented below, with special highlights and emphasis to those aspects that are applicable to the localisation scenarios considered in this research. This means the particular focus is on how the two algorithms perform in NLOS or urban environments when applied to next-generation systems. The comparison is presented in the form of TDOA strengths and weaknesses over AOA. This allows the distinction of cases where each of these techniques is applicable. The weaknesses and strengths of TDOA over AOA are conversely the strengths and weaknesses of AOA over TDOA respectively.

2.2.1.1 TDOA Strengths over AoA

1. *Lower cost and complexity antennas:* TDOA allows the use of low cost antennas that are of lower complexity. The antennas are also generally smaller in size. TDOA receivers can utilise simple antennas like a monopole or a dipole. Equipment precision is somewhat relaxed for TDOA than AOA. This is also the case with the antenna calibration. AOA requires elaborate calibration schemes and can be very sensitive to calibration errors and also equipment imperfections [46]. Another key advantage of TDOA is the ability to utilise very small and subtle antennas, which can be useful when deploying the system at sites with space restrictions. Densification in 5G and distributed deployment of a massive number of antennas in urban areas may benefit from conspicuous antennas.

2. *Relaxed calibration and simple siting requirements:* Before considering the regulatory issues like planning approval, TDOA siting requirements are less restrictive than AOA which allows greater site choice flexibility, thus making TDOA installations/deployments faster. This is particularly beneficial in dense urban environments where additional TDOA receivers can be installed in order to overcome shadowing. In contrast, sites for installation of AOA receivers have to be chosen so that wave-front distortion from local scatters and ground reflections together with ground conductivity, is minimised. TDOA antennas on a single BS may require little or no calibration whereas antenna arrays for AOA systems require calibration, often after site installation in order to reduce errors that are dependent on frequency as well as direction. So calibration is thus a critical performance limiting factor for AOA systems [46]. This is also demonstrated in Chapter 5 where calibration is key to all the results presented.

3. *Complexity:* TDOA receivers and antennas are generally less complex than AOA antenna arrays. Whilst advanced TDOA processing may be appropriate for some challenging environments, a TDOA receiver generally requires at least a single RF chain. Because of TDOA's popularity with GPS systems, appropriate synchronisation techniques and data link interfaces are readily available. For the cases considered in this research, it means that it is easier to implement TDOA on 5G testbeds with enough bandwidth and in a system that achieves a high enough SNR like mmWave systems because of their use of directional antennas with beam steering [47].

4. *Signals:* As discussed in Chapter 3 and also highlighted above, TDOA critically performance depends on bandwidth for a given SNR. This means that TDOA will perform very well with with next-generation systems which may use complex modulation schemes, with wider bandwidths and short duration signals. On the other hand AOA systems are known to perform best with narrow-band signals [48], although the AOA algorithms themselves can be implemented for systems that use any type of signal, be it wide-band or narrow-band, long or short duration. For low SNR signals, correlation provides a processing gain which

makes it possible for TDOA schemes to detect signals that have negative SNR. In contrast, common AOA systems struggle to locate low SNR signals and are not able to detect negative SNR signals. It has to be noted, however, that whilst AOA schemes do not benefit from correlation processing gain, they do benefit from system gain which is derived from multiple antenna usage, as demonstrated later in this chapter.

5. *Interference and uncorrelated noise:* The benefits of TDOA correlation processing gain makes TDOA able to locate signals with low SINR because the TDOA correlation processing is able to suppress interference signals that are inter-site de-correlated [49]. All receivers make synchronised measurements and signals which are not common to at least two receivers can be suppressed and it is possible to run localisation using only correlations with best observed signal if advanced TDOA processing is employed [50]. While some advanced AOA processing techniques like MUltiple Signal Classification (MUSIC) may be robust to interference and uncorrelated noise, their other requirements like calibration computational complexity make them unsuitable for some systems [51]. Also if multiple sources are active, MUSIC can not separate those sources.

6. *Multipath conditions:* While both AOA and TDOA are affected by multipath, they are affected in a completely different way and hence the effects may be less severe in one method under any given environment. For example, as already pointed out, TDOA schemes have lesser sensitivity to wave-front distortion than AOA systems, assuming a sufficient bandwidth and AOA will require the manifold to be known and accurate in such scenarios. For an environment with distant scatters, TDOA may need additional outlier processing to improve the accuracy [52]. This means that in an environment with local scatters like in dense urban areas, TDOA is expected to perform better while in environments that cover large distances like parks, with no local scatters at the receivers, AOA, is expected to perform better. The effect of multipath on each method can therefore be said to be dependent on the position of the scatters and the environment in which the system is operating. Also advanced TDOA processing may be able to suppress time resolved multipath between sites which can result in a very good performance in dense urban environments [53].

7. *Indoor and stadium geolocation:* TDOA is capable of geolocating high bandwidth signals in outdoor, indoor and high multipath environments at short range of less than 100m if advanced processing techniques are used [54], whereas AOA systems are known not to perform poorly in these conditions. Also in indoor environments, the TDOA challenge of timing synchronisation is simplified by the shorter distances, meaning cable synchronisation schemes or synchronisation with Ethernet switches that are IEEE-1588 compatible becomes possible.

8. *Geometry considerations:* When the target is centred within the perimeter of the receivers, both AOA and TDOA location accuracy is improved, but for TDOA, the uncertainty of

the location is not related or affected by the distance between the TDOA receivers [55], a property which can be advantageous in certain geographical conditions. In AOA the location accuracy is dependent on the accuracy of the lines of bearings to the target and also the distance between the target and the receivers. However when the target is much further outside the perimeter of measurement, the location uncertainty increases in a similar way, with distance for both methods [56].

9. *Possibility of full offline analysis:* TDOA systems are able to store to a central server, the synchronised signal measurements from all receivers as described for LTE TDOA systems in Appendix A.1. This allows full offline analysis which include cross-correlation, spectral analysis and the ultimate geolocation of the target. Spectral analysis and cross- correlations are not typically stored for AOA systems because there is usually no back-hauling of data.

10. *Benefit from increased number of receivers:* Adding more receivers leads to better location accuracy for both TDOA and AOA as a consequence of improved statistics and reduction in path loss. However, TDOA is more suited to multiple receivers (BSs) because of its reduced complexity, ability to use reduced antenna sizes, reduced power requirements and relaxed siting requirements as already discussed above. Further to that, TDOA correlation processing gain allows additional TDOA receivers that may be receiving very low SNR, to participate in localisation [49].

2.2.1.2 TDOA Weaknesses

1. *Homing and standoff:* Homing as well as standoff are possible for AOA systems utilising just one receiver [57], whereas TDOA requires at least two receivers and at least one has to be mobile for homing purposes. A network is also required to have the measurements at the TDOA stations onto a single central server. This means that only AOA can allow for localisation in environments where the receivers cannot be networked due to technical, logistical or cost/benefit challenges.

2. *Narrowband signals:* Slowly varying signals, like narrowband signals and unmodulated carriers, may be difficult to geolocate with TDOA techniques. TDOA performance degrades as signal bandwidth decreases for a given SNR. This means that TDOA may be suitable for high bandwidth systems like mmWave systems or UWB systems, but may not be applicable on some sub 6GHz systems. This is also discussed in Chapter 5 on the choice of AOA over TDOA to use with the massive MIMO testbed. TDOA with narrowband signals can improve under high SNR and longer observation times. In contrast, AOA systems can operate well for both wideband and narrowband signals, and also with unmodulated signals [58].

3. *Higher data rate back-haul links:* Because TDOA systems generally require transmission of signal samples to the central processing server [59], high data rate links are needed even

though this is only primarily for uplink transmission of data to the server, and little data is downloaded from the server. This presents as an inefficient utilisation of the high data rate links. Some advanced prepossessing of the data like compression, may help to alleviate the huge volumes of data that has to be send over. On the other hand an AOA receiver typically has to send only a few parameters like bearing angle, time and frequency, to the central server or processor.

4. *Sensitivity to signal de-correlation and synchronisation:* TDOA receivers or base stations have to be carefully synchronised to reduce signal de-correlation between the base stations. They also have to mitigate the Doppler shift from both a mobile target or from mobile scatters. A discussion of these aspects with respect to NLOS mitigation is presented in Chapter 4. The performance of the TDOA system is therefore affected by the stability of the synchronisation clock and the channel dynamics in addition to bandwidth and SNR. Channel dynamics may be mitigated by use of tracking loops and exploiting any available inertial measurements. For standard AOA systems, signal de-correlation between base stations is usually not an issue. However, some advanced AOA systems that rely on correlation with a reference signal will also be affected by de-correlation between base stations. Time synchronisation requirements in AOA systems are less demanding than TDOA (can be as loose as a few seconds versus TDOA which requires nanosecond level synchronisation) [60], although some advanced AOA schemes that make use of hopping signals may require tighter synchronization.

5. *Periodic signals:* Signals that contain some periodic elements, as is the case for synchronization pulses or any repeating sequences may result in cross-correlation errors in some cases, so the choice of the signal used has to consider this effect [61]. On the other hand, standard AOA systems are not affected by signal periodicity because they do not perform cross-correlation.

6. *Positioning rate:* Positioning rate is the time period at which a position fix, i.e. the target location, is produced. TDOA systems may have only up to 1 position fix per second whereas AOA systems may have say 100 fixes per second. A key limitation in TDOA is usually the data links that relay the measurements to a central localisation server as the case in cellular networks [59]. Faster links plus the use of compression techniques can improve TDOA positioning rate.

7. *Single base station localisation not possible:* With TDOA, at least 3 receivers or base stations are needed for 2-D geolocation, while 4 are needed for 3-D geolocation. As demonstrated and discussed in Chapter 5, AOA can be utilised for single base station localisation. This means that all the multiple site challenges, for example, synchronisation, do not apply to the single BS scenario, thus simplifying the system.

8. *Multiple targets/sources:* Wideband direction finding (Wideband DF) allows AOA systems to concurrently geolocate multiple frequency separated signals, whereas this may not be practical in TDOA systems due to data backhauling requirements.

9. *Offline analysis with single BS measurements:* Offline analysis of the lines of bearing for AOA is possible with single base station measurements. On the other hand, analysis of single base station lines of position for TDOA is not useful.

10. *GDOP effect* For any position of target outside the perimeter of measurement, TDOA Geometric Dilution Of Precision (GDOP) (discussed in section 2.2.3) effect increases more rapidly than that of AOA as the distance of the target from the perimeter increases. The TDOA's line-of-position (LOP) approximates the AOA's line-of-bearing (LOB) as the target distance from the perimeter becomes very large [62].

2.2.2 TDOA localisation model

TDOA methods are becoming more attractive in cellular positioning for the following reasons;

- Availability of compact and inexpensive computing power on devices
- Availability of more advanced receiver technologies
- Availability of faster backhaul links with adequate capacity.
- Availability of accurate and distributed synchronisation signal

The major challenge with TDOA has been the nanosecond level synchronisation that is required. Geo-location of short duration signals needs networked receivers that are tightly synchronised to approximately a small portion of inverse of the signal bandwidth. In addition to the time synchronisation, the accuracy for TDOA systems is also governed by the GDOP (see section 2.2.3) and the accuracy/performance of the TDOA algorithm itself.

UTDOA processing is based on trilateration, with at least 3 base stations. Time differences between two base stations are measured and converted to a constant difference distance between them, as a foci, which defines a hyperbolic curve. Figure 2.2 below illustrates the setup. $BS1$ and $BS2$ have their hyperbolic curve $R2-R1$, and $BS1$ and $BS3$ have their hyperbolic curve $R3-R1$. $R1$,$R2$ and $R3$ are the distances between the target (handset/UE) and each respective base station. The intersection of the two hyperbola gives the position of the target.

Considering a 2D UTDOA estimation system using B base stations. All time differences are referenced to the first base station, which is the first to receive the transmitted signal. If the known 2D coordinates of the base stations are (x_i, y_i), for $i = 1,\ldots,B$, then the TDOA system can be modelled, in order to calculate the unknown UE/target coordinates (x, y), as follows

$$R_i = \sqrt{(x_i - x)^2 + (y_i - y)^2} = \sqrt{x_i^2 + y_i^2 - 2x_i x - 2y_i y + x^2 + y^2} \tag{2.1}$$

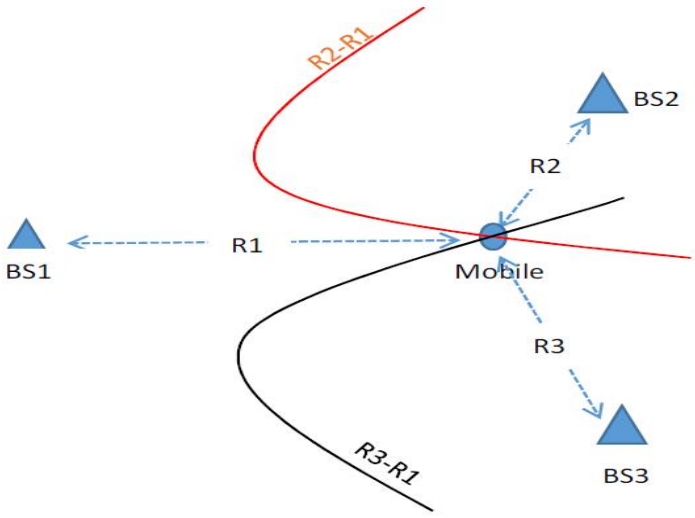


Figure 2.2: TDOA setup.

The distance difference between any base station and the first base station where the signal arrives first, is given as

$$R_{i,1} = c\tau_{i,1} = R_i - R_1 = \sqrt{(x_i - x)^2 + (y_i - y)^2} - \sqrt{(x_1 - x)^2 + (y_1 - y)^2} \tag{2.2}$$

where $R_{i,1}$ is the range/distance difference between the reference BS and the ith base station.
c is the propagation speed of the signal.
R_1 is the distance of the mobile station from the first base station.
$\tau_{i,1}$ is the estimated TDOA between the ith base station and the first base station.
From measured time delays, $\tau_{i,1} = \tau_i - \tau_1$, which are the difference between their times of arrivals. Equation (2.2) defines a set of nonlinear hyperbolic equations whose solution are the 2D coordinates of the mobile station [63].

It is worth noting that there is no explicit solution to the above nonlinear equations (as formed from (2.2)). There are ways of solving these equations that have been proposed in literature, which have varying complexity and performance. When the BSs are arranged linearly, position fixing for the target can be simplified and there are a number of good optimum processing techniques that have been proposed for such scenarios [64] [65] [66]. However, as typical in cellular systems, when the TDOA receivers or BSs are distributed arbitrarily, the situation becomes complex and the solution to the hyperbolic equations is not easy because the resulting equations are nonlinear. There are three main techniques; Fang's algorithm [67], Taylor-series methods [68] [55] and Chan's algorithm [69], which are generally considered for solving this scenario. Taylor-series method is an iterative process where a good initial guess is required. The choice of this initial guess is not straightforward and convergence is not guaranteed. Fang's solution has a limitation in that it cannot make use of extra measurements (beyond the basic 3) from other base stations, in order to improve the position accuracy. Chan proposed a solution that is valid for both distant and

close targets, which was demonstrated to come closer to the Cramer-Rao Lower Bound (CRLB). Chan's algorithm can utilise more than 3 base stations, unlike Fang's solution which is limited to 3. Chan's algorithm is therefore the preferred TDOA algorithm which was used throughout this thesis, and its model is presented below. Chan's algorithm transforms a set of nonlinear equations into a set of linear equations; first, by re-arranging the first part of (2.2) and squaring both sides, into

$$\begin{aligned} R_i^2 &= (R_{i,1} + R_1)^2 \\ &= R_{i,1}^2 + 2R_{i,1}R_1 + R_1^2 \end{aligned} \tag{2.3}$$

and then rewriting (2.1) as

$$R_{i,1}^2 + 2R_{i,1}R_1 + R_1^2 = x_i^2 + y_i^2 - 2x_i x - 2y_i y + x^2 + y^2 \tag{2.4}$$

Taking condition $i = 1$ in (2.1) and subtracting it from (2.4) gives

$$R_{i,1}^2 + 2R_{i,1}R_1 = x_i^2 + y_i^2 - 2x_{i,1}x - 2y_{i,1}y - x_1^2 + y_1^2 \tag{2.5}$$

where $x_{i,1}$ and $y_{i,1}$ are respectively defined as x_i–x_1 and y_i–y_1. If k_i is defined such that $k_i = x_i^2 + y_i^2$ for $i = 1,2,3$, then (2.5) can be written as

$$R_{i,1}^2 + 2R_{i,1}R_1 = -2x_{i,1}x - 2y_{i,1}y + k_i - k_1 \tag{2.6}$$

The unknowns are the location of the mobile (x, y) and the distance from the first base station (reference BS) to the mobile (R_1), which can be easily manipulated to produce a set of quadratic equations which, for the classical case of 3 base stations (which is utilised throughout this thesis), can be written in matrix form as

$$\begin{bmatrix} x \\ y \end{bmatrix} = -\begin{bmatrix} x_{2,1} & y_{2,1} \\ x_{3,1} & y_{3,1} \end{bmatrix}^{-1} \times \left\{ \begin{bmatrix} R_{2,1} \\ R_{3,1} \end{bmatrix} R_1 + \frac{1}{2} \begin{bmatrix} R_{2,1}^2 - k_2 + k_1 \\ R_{3,1}^2 - k_3 + k_1 \end{bmatrix} \right\} \tag{2.7}$$

This solution produces the mobile coordinates (x, y) in terms of R_1. If this solution (i .e. x and y in terms of R_1) is substituted into (2.1), with $i = 1$, a single equation with R_1 as an unknown is produced. Taking the positive root of the solution of that equation, and substituting it back into (2.7) gives the estimated mobile/UE coordinates. In cases where the 2 roots of the quadratic equation are both positive, select the smaller root. If both roots are negative, select the bigger root. It was observed through ray-tracing simulations discussed in Chapter 3 section 3.2, that selection of the smaller root when they are both positive does not always lead to a closer approximation. Using the bigger positive root produced a closer approximation in some cases, so position estimation can be done using both roots, and subsequent outlier rejection can be used to exclude a solution which places the target outside the zone/area of interest.

There are other TDOA techniques which may not be based on solving the hyperbolic equations or requiring synchronised BSs as discussed above. Xu et al. [70], demonstrated a system called Whistle, which is capable of doing TDOA localisation without the need for the receivers to be time synchronised. In their scheme, they made use of the target signal and another successive artificially generated secondary signal, together with sample counting techniques, to achieve a high resolution for the time differences. Their so called 'two-signal sensing' approach utilise sample counting in the sense that the receivers count the number of samples between the two signals, and then derive the corresponding time delay (between the two signals). For the rest of their TDOA position estimation, they adopted Chan's hyperbolic estimator [69]. Their experiments were conducted with acoustic signals, but the approach can be applied to RF signals.

2.2.3 TDOA CRLB and GDOP

Equation (2.7) can be arranged in matrix form as [71]

$$\boldsymbol{\theta} = R_1 H^{-1}\boldsymbol{a} + H^{-1}\boldsymbol{b} \tag{2.8}$$

where $\boldsymbol{\theta} = \begin{bmatrix} x \\ y \end{bmatrix}$, $\boldsymbol{H}^{-1} = \begin{bmatrix} x_{2,1} & y_{2,3} \\ x_{3,1} & y_{3,3} \end{bmatrix}$, $\boldsymbol{a} = \begin{bmatrix} R_{2,1} \\ R_{3,1} \end{bmatrix}$ and $\boldsymbol{b} = \frac{1}{2}\begin{bmatrix} R_{2,1}^2 - k_2 + k_1 \\ R_{3,1}^2 - k_3 + k_1 \end{bmatrix}$

The Cramer-Rao Lower Bound (CRLB), which can be used to evaluate the performance of TDOA, specifies the lower bound on the variance of the mobile position estimates

$$E[(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^2] \geq \boldsymbol{I}(\boldsymbol{\theta}) \tag{2.9}$$

where $E[\cdot]$ denotes the expectation function and $\boldsymbol{I}(\boldsymbol{\theta})$ denotes the Fisher Information Matrix (FIM). If the vector $\boldsymbol{R} = [R_1, R_2 ... R_B]^T$ is used to denote the actual ranges of the mobile from each BS, and $\hat{\boldsymbol{R}}$ is a vector that contains all the estimated ranges R_i for $i = 1...B$, then the estimated ranges can be modeled as $\hat{\boldsymbol{R}} = \boldsymbol{R} + \boldsymbol{\epsilon}$ where $\boldsymbol{\epsilon}$ represents the zero-mean Gaussian noise. The FIM can then be expressed as [72]

$$\boldsymbol{I}(\boldsymbol{\theta}) \triangleq E\left[\left(\frac{\partial}{\partial \boldsymbol{\theta}} \ln f(\hat{\boldsymbol{R}}|\boldsymbol{\theta})\right)^2\right] = E\left[\frac{\partial}{\partial \boldsymbol{\theta}} \ln f(\hat{\boldsymbol{R}}|\boldsymbol{\theta}) \cdot \left(\frac{\partial}{\partial \boldsymbol{\theta}} \ln f(\hat{\boldsymbol{R}}|\boldsymbol{\theta})\right)^T\right] \tag{2.10}$$

where $f(\hat{\boldsymbol{R}})|\boldsymbol{\theta}$ denotes the joint PDF of $\hat{\boldsymbol{R}}$ condition on the unknown mobile coordinates. From (2.10), it is shown in [73] that the CRLB can be expressed as

$$\left[\boldsymbol{I}(\boldsymbol{\theta})^{-1}\right]_{2\times 2} = \left(\boldsymbol{J}\boldsymbol{I_d}\boldsymbol{J^T}\right)^{-1} \tag{2.11}$$

where $\boldsymbol{I_d} = diag(\sigma_1^{-2}, \sigma_2^{-2}, ...\sigma_B^{-2})$ with σ_i^{-2} being the ith measurement noise variance, and $\boldsymbol{J}$ is the Jacobian matrix which is expressed as [71]

$\boldsymbol{J} = \begin{pmatrix} cos\phi_1 & cos\phi_2 & ... & cos\phi_B \\ sin\phi_1 & sin\phi_2 & ... & sin\phi_B \\ 1 & 1 & ... & 1 \end{pmatrix}$ where ϕ_i is the angle between the mobile and the ith BS.

For a mobile position with variances σ_x^2 and σ_y^2 for the 2D position estimates, the Geometric Dilution Of Precision (GDOP) can be expressed as

$$GDOP = \sqrt{\frac{\sigma_x^2 + \sigma_y^2}{\sigma_r}} \tag{2.12}$$

where σ_r is the range measurement error standard deviation [56]. Spirito [74] also derived an alternative expression which emphasises the geometric relationship between the mobile and the BSs.

2.2.4 AOA localisation model

The model for AOA localisation used in this thesis is based on the Three-object Triangulation Algorithm (ToTal) [75], which has been shown to be simpler and more efficient than other traditional triangulation algorithms. Although the algorithm was designed for robot positioning using beacons, its design perfectly matches the use of base stations as receivers, and the robot substituted by a mobile device or UE. The ToTal algorithm is chosen because it matches or corresponds to the use of 3 BS receivers in the selected TDOA model. This allows positioning via both TDOA and AOA to be performed using the same base stations and same measurements at the BSs which in-turn allows for easier and controlled comparison of the two localisation algorithms. Although the ToTal algorithm is designed to have the beacons transmit to the robot, with the robot calculating its position, in the implementation in this thesis, the reciprocal of this process is used, with the mobile transmitting to BSs, and the angle measurements done at the 3 BSs. The algorithm uses a reference beacon (similar to the use of reference BS in the TDOA model [69]) and then solves the robot coordinates relative to this beacon. The algorithm is presented below for completeness and to show how it has been adapted for the purposes of this research. Figure 2.3 shows the triangulation setup.

Given the three BSs (B_1,B_2 and B_3) with coordinates $\{x_1, y_1\}, \{x_2, y_2\}, \{x_3, y_3\}$, and the three angles of arrival ϕ_1, ϕ_2, ϕ_3, respectively, the aim is to calculate the coordinates of the mobile station $\{x_R, y_R\}$. Any BS can arbitrarily be taken to be the reference and placed at the origin (meaning its coordinates become zeros). In this case B_2 is taken to be the origin. The modified coordinates of the BSs relative to the origin become $B_2' = \{0,0\}$, $B_1' = \{x_1 - x_2, y_1 - y_2\} = \{x_1', y_1'\}$ and $B_3' = \{x_3 - x_2, y_3 - y_2\} = \{x_3', y_3'\}$.

Following the ToTal algorithm, the modified coordinates can be calculated as

$$x_1' = x_1 - x_2, \quad y_1' = y_1 - y_2, \quad x_3' = x_3 - x_2, \quad y_3' = y_3 - y_2, \tag{2.13}$$

The cotangents of the bearing angles between the base stations are calculated, with T_{ij} being the cotangent between BSs B_i and B_j. The bearing angle between B_1 and B_2 is given by $\phi_1 - \phi_2$, so, $T_{12} = cot(\phi_2 - \phi_1)$, $T_{23} = cot(\phi_3 - \phi_2)$, and because all the three angles are linked, any

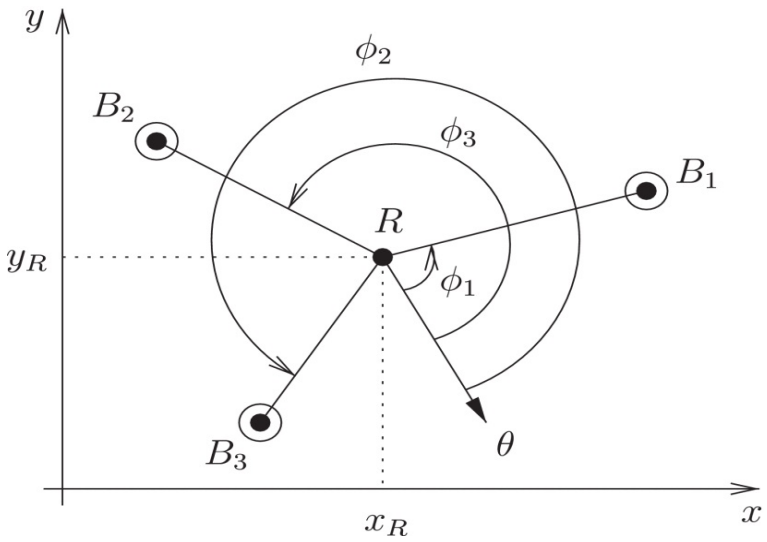


Figure 2.3: Triangulation setup. [75].

cotangent can be calculated from the values of the other two cotangents, so in this case, the third cotangent should be calculates as

$$T_{31} = \frac{(1 - T_{12}T_{23})}{(T_{12} + T_{23})}. \tag{2.14}$$

If C_{ij} is a circle that passes through B_i, B_j, and R, its center c_{ij} with coordinates $\{x_{ij}, y_{ij}\}$ can be calculated as

$$x_{ij} = \frac{(x_i + x_j) + T_{ij}(y_i - y_j)}{2}, \quad y_{ij} = \frac{(y_i + y_j) - T_{ij}(x_i - x_j)}{2} \tag{2.15}$$

which when the modified coordinates from (2.13) are used, the three sets of centre coordinates become

$$\begin{aligned} x'_{12} &= x'_1 + T_{12}y'_1, \quad y'_{12} = y'_1 - T_{12}x'_1, \\ x'_{23} &= x'_3 - T_{23}y'_3, \quad y'_{23} = y'_3 + T_{23}x'_3, \\ x'_{31} &= (x'_3 + x'_1) + T_{31}(y'_3 - y'_1), \quad y'_{31} = (y'_3 + y'_1) - T_{31}(x'_3 - x'_1). \end{aligned} \tag{2.16}$$

The parameter k_{ij} represents the components of the power line between 2 circles [76]. The power line of circles C_{12} and C_{23} can be defined as $k_{12} - k_{23}$, and for the modified coordinates, the k'_{ij} parameters are calculated as

$$k'_{ij} = x'_i x'_j + y'_i y'_j + T_{ij}(x'_j y'_i - x'_i y'_j), \tag{2.17}$$

so k'_{31} is calculated as

$$k'_{31} = x'_1 x'_3 + y'_1 y'_3 + T_{31}(x'_1 y'_3 - x'_3 y'_1). \tag{2.18}$$

It is only necessary to calculate k'_{31} since BS2 (B_2) is at the origin, $x'_2 = 0$ and $y'_2 = 0$, and consequently $k'_{12} = 0$ and $k'_{23} = 0$. The triangulation solution is obtained by intersecting the three power lines, which is effectively solving the linear system,

$$\begin{aligned} x(x'_{12} - x'_{23}) + y(y'_{12} - y'_{23}) &= k'_{12} - k'_{23} \\ x(x'_{23} - x'_{31}) + y(y'_{23} - y'_{31}) &= k'_{23} - k'_{31} \\ x(x'_{31} - x'_{12}) + y(y'_{31} - y'_{12}) &= k'_{31} - k'_{12.}. \end{aligned} \tag{2.19}$$

The three power lines meet at the power centre whose coordinates are the mobile station position $\{x_R, y_R\}$ which can be obtained from solving (2.19) [76]

$$\begin{aligned} x_R &= \frac{\begin{vmatrix} k'_{12} - k'_{23} & y'_{12} - y'_{23} \\ k'_{23} - k'_{31} & y'_{23} - y'_{31} \end{vmatrix}}{D} \\ y_R &= \frac{\begin{vmatrix} x'_{12} - x'_{23} & k'_{12} - k'_{23} \\ x'_{23} - x'_{31} & k'_{23} - k'_{31} \end{vmatrix}}{D} \end{aligned} \tag{2.20}$$

with the denominator D, being equal to

$$D = \begin{vmatrix} x'_{12} - x'_{23} & y'_{12} - y'_{23} \\ x'_{23} - x'_{31} & y'_{23} - y'_{31} \end{vmatrix} = \begin{vmatrix} x'_{12} & y'_{12} & 1 \\ x'_{23} & y'_{23} & 1 \\ x'_{31} & y'_{31} & 1 \end{vmatrix} \tag{2.21}$$

D can be simplified as

$$D = (x'_{12} - x'_{23})(y'_{23} - y'_{31}) - (y'_{12} - y'_{23})(x'_{23} - x'_{31}). \tag{2.22}$$

Because $k'_{12} = 0$ and $k'_{23} = 0$, the UE/MS position can be calculated as [75]

$$\begin{aligned} x_R &= x_2 + \frac{k'_{31}(y'_{12} - y'_{23})}{D}, \\ y_R &= y_2 + \frac{k'_{31}(x'_{23} - x'_{12})}{D}. \end{aligned} \tag{2.23}$$

2.2.5 Hybrid TOA-AOA

The simplest, yet powerful method to determine the position of a mobile in LOS scenarios, using just a single base-station, is the hybrid TOA-AOA technique [77] [78]. In this scheme, it is assumed that the base station is capable of determining the DOAs of a mobile device. This is especially the case when antenna arrays are used at the BS as is the case in Massive MIMO systems. DOA estimation is discussed in section 2.4.2 and also in Chapter 5. TOA in this case uses a single or multiple measurements to estimate the range R, to the mobile. The estimated UE/MS range is used together with an AOA measurement, to compute the position of the mobile, in 2D as shown in Figure 2.4 below.

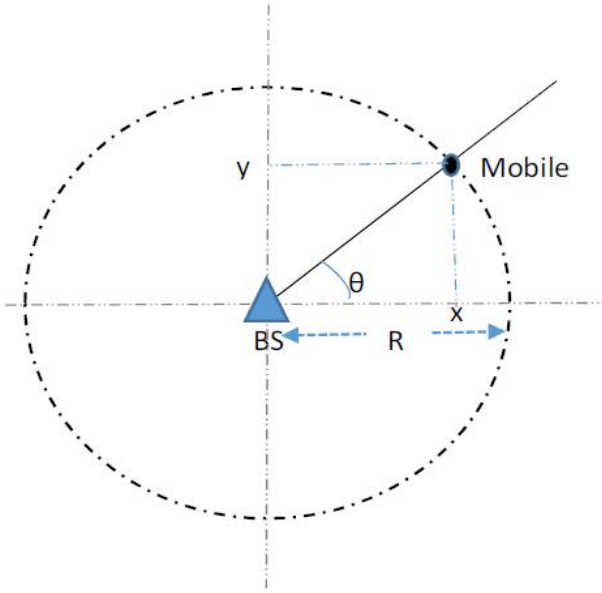


Figure 2.4: Single BS TOA-AOA.

From 2.4, knowing the range R, and the angle of arrival θ, the position of the mobile device can be calculated as

$$x = x_{BS} + R.\cos\theta, \qquad y = y_{BS} + R.\sin\theta. \tag{2.24}$$

where $\{x_{BS}, y_{BS}\}$ are the BS coordinates.

Range detection using TDOA (just to obtain R) is also possible using time delay measurements from adjacent BSs and DOA measurement from the local BS. This means when there are no LOS rays, single bounce ground reflected rays may be used together with the range obtained via TOA or 3 BS TDOA, for 2D localization. The Hybrid TOA-AOA algorithm used for simulations in this thesis selects a ray that is LOS for localization, and if not present, it considers a single bounce

ground reflected ray or rooftop diffracted ray (appears like a LOS link in 2D). This was possible because the ray-tracing software is capable of providing such information and this was used to evaluate these schemes for an urban environment. Advantages of using the range obtained by TDOA in this hybrid algorithm are that TOA requires two-way communication between BS and UE, which may not be convenient or possible in some applications; and also since TDOA utilizes at least 3 BSs, other surrounding BS may have a better link, and the estimated range may become better. However, the TDOA range estimate may become very poor if all the 3 links are NLOS. Other related TOA-AOA hybrid schemes for NLOS scenarios in cellular systems, are discussed in [79].

2.2.6 Sources of errors and their effects in TDOA and AOA

Errors in time-delay estimates may arise as a result of various phenomena. One of the most common sources of error are measurement noise, which may be inherent in the measuring equipment like amplifiers and A/D converters, which is really the receiver noise floor. Whilst this may cause some small variations in the time delay estimates, it does not affect the AOA estimate except where the signal capture process has significant defects. AOA accuracy is fundamentally affected by noise and also depends on gain/phase calibration of receivers and the accuracy of the antenna array manifold. The other source of errors is the malfunctioning of the receiver. This can lead to fatal errors that arise from erroneous time-delay estimates from the failed receiver. This means the data from the receivers have to be monitored to ascertain when a particular receiver has failed. Also, multipath propagation and correlated noise makes estimation difficult because they introduce an extra peak to the cross correlation functions, and so may be very difficult to detect. If that extra peak is higher than the target/source's, then an erroneous time-delay estimate is used. TDOA requires the receivers/base stations to be synchronised and the positions or coordinates of the base stations have to be accurate. Errors in the position estimates of the receivers will translate into errors in the unknown position of the transmitter, although such type of errors are generally not fatal, but reduce the accuracy for the estimation. Synchronisation errors also lead to erroneous results. Tighter synchronisation is difficult when the receivers are widely separated as is the case for distributed Massive MIMO system. It is also difficult when the receivers/sensors are numerous. Wang and Ho [80] demonstrated a method to sequentially obtain algebraic solutions for the target position, the receiver positions and also the synchronisation offsets. The approach was shown to meet the CRLB in terms of accuracy, although the estimates are not jointly obtained.

2.3 Hybrid data fusion for localisation

The concept of fusion of different estimates was described in [81], where it is noted that targets like mobile stations or other mobile user-equipment often move about following particular patterns,

which can often be predicted with no sharp jumps between positions. This type of user behaviour can be exploited to provide supplementary information for position tracking. Hybrid data fusion (HDF) can provide seamless positioning or navigation through widely differing environments. The above paper investigated HDF based on particle filter (PF), Position Kalman filter (PKF) and extended Kalman filter (EKF) [72] [82]. These filters were used to combine 3GPP-LTE TDOA and GNSS measurements. The paper made use of ray tracing simulations and a hidden Markov mobility model which is based on gas diffusion models. In [83] HDF is applied to the problem of Ultra-Wide Band (UWB) localisation. RSS and TOA measurements are combined to produce improved positioning algorithms. The approach is based on analysing the effect of adding TOA measurements to an RSS based system. Ranging is first done separately through RSS and TOA and then a fusion of the estimated ranges is done. The results show improved positioning in environments with reasonable noise levels. Ouyang et.al. [84] suggested a data fusion framework based on the weighted least squares estimator (WLSE). The three presented schemes for fusion are measurement, estimate and mixed fusion. Their results show that when the raw measurement vectors are correlated, the best result is achieved by measurement fusion with estimate fusion being the worst. In case of uncorrelated measurement vectors, all three types of fusion achieve same performance. His work with Zhang and others [85] proposed a TOA/AOA based algorithm for CDMA networks. Their algorithm is an extension of Taylor series least squares (TS-LS) TOA systems to include AOA measurements and their results show a better performance for the hybrid scheme. They also incorporated EKF and Unscented Kalman Filter (UKF) in their tracking algorithms and UKF was found to be better because it truly captures the noises' statistical mean and variance. Prieto et.al. [86] also presented ALPA (adaptive likelihood particle filter) which is a particle filter devised to specially address the issue of non-linear and non-Gaussian behaviours of measurements over time. Though their work considered measurements collected in a wireless LAN (WLAN) environment, their framework is based on the determination of the likelihood functions which represent the relationship between measurements and distance, of which such likelihoods can be dynamically adapted to other propagation conditions. The relevance of HDF in this work, is in the realisation that the estimates presented in this thesis can be further improved via HDF, thereby improving the location accuracy. This is especially the case for positioning of mobile UE where inertial measurements on the UEs can be used to improve the location accuracy.

2.4 Time delay and DOA Estimation

This section takes a brief look at the basic theory of time delay and DOA (or AOA-DF) estimation. This is relevant as limitations and advances in these topics, affect the localisation mechanisms that are discussed in this thesis. Time delay estimation is relevant for TDOA localisation techniques and DOA estimation is relevant for AOA localisation techniques. The estimation methods discussed are those that apply to multi-antenna systems, in line with the theme and body of work

presented in this thesis.

2.4.1 Time delay estimation

Time Delay Estimation (TDE) is a topic of interest in many scientific and technological fields. Time delay is a basic parameter in various applications. In wireless communications, a common application is its use in localisation, where either, two or more spatially separated and tightly synchronised transmitters can have their signals' time delays, at a receiver, converted to geolocation, such as the case with Global Satellite Navigations Systems (GNSS), or a case where two or more receivers/sensors receive a signal from a target, and the time delays from the receivers are converted into location via ranging. The methods for time delay estimation may be classified into, approximation model-based methods [87] [88] [89], explicit time-delay parameter methods [90] [91] [92] and higher-order statistics methods [93]. Other non-traditional TDE techniques can also be found in literature such as the use of neural networks [94] and wavelets [95].

Approximation model methods, in the form of either time domain approximation or frequency domain approximation methods, are popular and common for wireless and sensor networks. This thesis therefore provides a brief review on approximation methods. It is particularly concerned with those methods that may be suitable for array systems. Mobility is not a factor in these as it is separately tackled by use of tracking algorithms and hybrid data fusion. The most popular technique involves generalised cross-correlation [96] as in [65] for TDOA arrays. The received signals at the array are correlated and are assumed to be stationary Gaussian processes that contain non-cross-correlating noise. Performing cross-correlation produces a curve with a peak at the position of the time delay estimate. Time delay estimation accuracy depends on the SNR. Some may say bandwidth is critical for the performance of estimators, but it can be argued that increasing bandwidth only reduces the SNR required to achieve a given time resolution. Indeed, a Cramer-Rao Bound for the Maximum Likelihood (ML) time-delay estimator in [96], includes both bandwidth and SNR which supports the above argument. For next-generation systems, greater bandwidths can be assumed, as compared to current systems, with GHz bandwidths possible at millimetre-wave frequencies. Cellular densification in urban environments is also expected to improve SNR. These assumptions make TDE and TDOA localisation, together with the rest of the work presented in this thesis, relevant for next-generation wireless systems.

For the work presented in Chapter 3 section 3.2 and in Chapter 4, time delay estimates from the ray-tracing tool are used, so in this thesis, no work or simulations on time delay estimation has been conducted. As shall become clear in Chapter 5, the practical experimentation focused on AOA. This was necessitated by the limitations of the testbed with regard to time delay estimation, as outlined in Appendix C and more details about this can be found at the beginning of Chapter 5.

2.4.2 DOA Estimation

Localisation from super resolution DOA estimation requires multiple antennas at each base station. The following are DOA methodologies applicable for use with antenna arrays.

2.4.2.1 Classification

Spectral based solutions These techniques have super-resolution capability. Subspace-based approaches offer lower complexity compared to ML or Least-Squares (LS) based approaches. They are popular because of their computational simplicity. The two most popular subspace based approaches are MUSIC and Estimation Of Signal Parameters Via Rotational Invariance Techniques (ESPRIT) [51] (discussed in detail below). Subspace methods exploit the fact that the signal subspace is orthogonal to the noise subspace within the source signal. They are commonly referred to as super-resolution techniques.

- Noise subspace methods
  These methods take advantage of the fact that the manifold vectors lie within the signal subspace, with the noise subspace being orthogonal to them. MUSIC uses the noise subspace. The MUSIC algorithm is very sensitive to phase and gain errors, and sensor position. It requires careful calibration for it work well. The exhaustive search through all angles is computationally expensive [97].

- Signal subspace methods
  The noise subspace is discarded in signal subspace methods, to remain with only the signal subspace. This effectively increases the SNR. ESPRIT [98] uses the signal subspace. The ESPRIT algorithm itself is based on pairs of identical sensors although the pattern between sets of pairs could be different. The positions of each pair are being arbitrary, which makes calibration easier. Although ESPRIT somewhat relaxes the calibration overhead and computation, it however requires twice as many antennas.

Parametric Solutions Spectral-based solutions above are very popular because they are computationally attractive, but their performance may be insufficient for cases where the signal sources are highly correlated. Parametric methods discussed here, exploit the underlying model for data. The most common parametric solution is the ML technique. Stochastic ML and Deterministic ML are the main methods used in this case [99]. The major drawback of these methods is that they are computationally intensive since they require an exhaustive search, to find the estimates.

2.5 Localisation approaches: From DOA to positioning

Localisation that employs DOA estimation can be applied in a number of different ways, when antenna arrays are utilised. One way is to perform DOA on individual, separate base stations, and then use tri-angulation to get the location of the UE. This approach assumes that the DOA estimates that are used for tri-angulation are LOS estimates. Two or more DOA estimates from the BSs that are close the mobile station, can then be used to calculate the position of the mobile. To determine whether the DOA estimates are indeed LOS or not, the NLOS identification mechanisms discussed in Chapter 4, can be employed, so that only LOS DOAs are used for position fixing. Figure 2.5 shows the setup for an AOA localisation approach that utilizes DOA estimation on separate BSs.

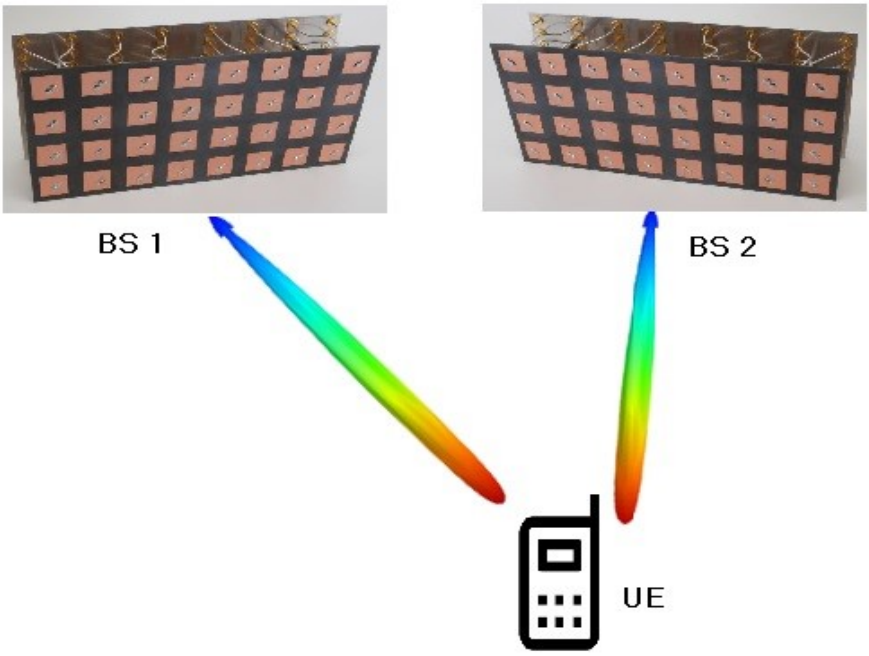


Figure 2.5: AOA localisation with 2 or more base stations.

For massive MIMO systems, this approach will be more suitable in distributed deployments [100], where the large number of antennas are actually distributed in the environment, yet still being synchronised. A discussion on how the Bristol Massive MIMO testbed can be distributed is provided in Appendix C. Utilising the testbed and the components already available for the system, i.e. the BS components and the UE components, the distributed deployments and positioning setup is shown in Figure 2.6 below.

Chapter 5 demonstrates and evaluates a novel alternative approach that avoids the challenges of distributing the antennas and takes advantage of the large geometries of linear, rectangular and possibly circular arrays. The approach entails performing DOA estimation using different

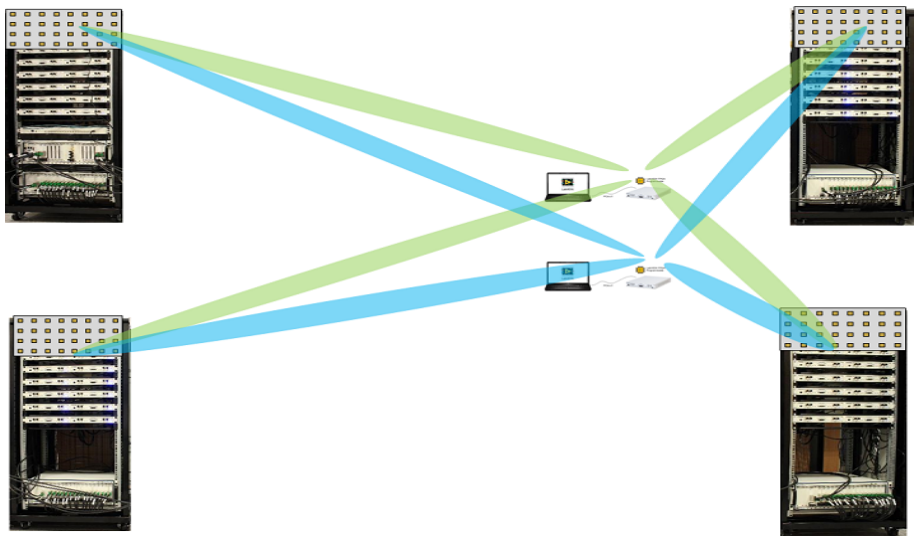


Figure 2.6: Massive MIMO Distributed deployment: Each cabinet and the sub-array on it act as base station and 2 USRP + laptop host sets, act as user equipment.

subsets of antennas on the collocated array. This allows DOA estimation, and subsequently triangulation, to be performed using just a single BS. This approach, however, depends on array size and geometry. It is suitable for LOS scenarios with smaller coverage areas, such as might frequently occur indoors. Linear or rectangular arrays will suffer from the Geometric Dilution of Precision (GDOP) problem [101] which is discussed in section 2.2.3. The key benefit from the use of sub-arrays for DOA estimation is that it allows localisation algorithms to have a choice over the selection of best DOA estimates, based on predetermined confidence criterion. This approach is presented in Chapter 5 sections 5.4 and 5.5. Algorithms that utilize arbitrary numbers of antennas selected from a massive array, have been proposed in literature [102] and this allows for a dynamic configuration of the DOA systems, that depends on the SNR for a given UE, and can provide a way for optimising energy vs location accuracy for large arrays. Figure 2.7 shows the single AOA approach using a rectangular array.

The approaches to AOA localisation from DOA estimates that have been discussed above so far assume that the DOA estimates are in line of sight. However, in many realistic outdoor situations, the estimated angles are those that are most likely to be a result of NLOS propagation. In cases where both NLOS and LOS DOAs are estimated or are expected, NLOS identification schemes can be used to identify those angles that are a result of LOS propagation, and then use them for localisation. Where no such LOS signals are received at the BS, there is a need to process these angles before they can be used for localisation. One approach is to use a scattering model to estimate the LOS angle from multiple angle measurements as seen on the BS [103]. The performance of this approach in terms of the accuracy of the estimated LOS angle, depends on the model and also the environment considered. A scattering model that is not suitable for the environment can lead to significant errors in the resultant localisation algorithm.

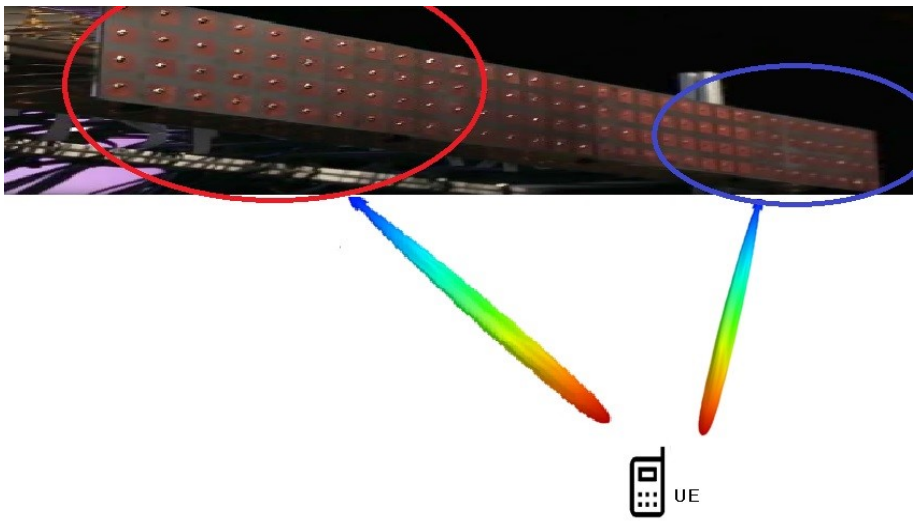


Figure 2.7: AOA localisation with a single base station.

Xinning Wei et. al. [104] proposed a technique that is capable of utilising multipath DOAs to obtain a 3D position of a MS. The technique utilises a MIMO geometric model, taking in departure angle measurements at the UE and arrival angle measurements at the BS. They assumed a single bounce reflection in every NLOS path, to produce a 3D Least-Squares (LS) solution to the mobile station position. Their approach requires the mobile station or UE to be able to determine AOD/AOA for its signals. This may be possible with future mobiles that may incorporate multiple antennas. If reliable AOD/AOA estimates can be obtained both on the UE (via multiple antennas on mobile device) and on the BS (via massive MIMO for example), then the position of the UE can be obtained. Figure 2.8. below shows the geometric model for the 3D NLOS scheme.

The setup of this model demonstrates a downlink propagation from the base station B to a mobile station M. Reciprocity is assumed in its application to an uplink scenario where the definitions of AOD and AOA become converse. The BS position is denoted by $\boldsymbol{p}_B = (x_B, y_B, z_B)^T$. The red arrows show a multipath ray that is reflected off a scatterer at position $\boldsymbol{p}_{S,n} = (x_{S,n}, y_{S,n}, z_{S,n})^T$, and is received at the mobile stations (MS) at an unknown position $\boldsymbol{p}_M = (x_M, y_M, z_M)^T$. $(\varphi_{D,n}, \vartheta_{D,n})$ denotes the angle of depature (AOD) for the nth NLOS propagation path. $(\varphi_{A,n}, \vartheta_{A,n})$ denotes the AOA for the nth NLOS path. d_n denotes the propagation path length which consists of the path length from the base station to the scatterer r_n and secondly, from the scatterer to the mobile station $(d_n - r_n)$. The solution to the mobile position is given in [104] as

$$\boldsymbol{p}_M = (\boldsymbol{M}^T \boldsymbol{M})^{-1} \boldsymbol{M}^T . \boldsymbol{b} + \boldsymbol{p}_B \tag{2.25}$$

where $\boldsymbol{M}$ denotes the matrix formed from the NLOS propagation paths. It can be noted that this approach can actually solve the position of both the UE and the scatterers at the same time, utilising a single BS. Their proposal and evaluation makes use of indoor measurements,

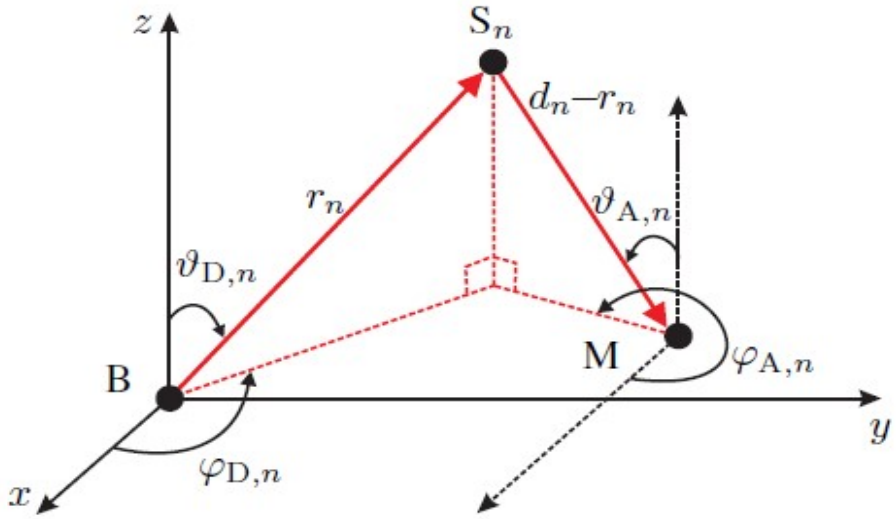


Figure 2.8: System model with one NLOS propagation path, the UE being denoted by "M", the BS being denoted by "B", and the scatterer being denoted by "S_n" [104].

but because the technique benefits from multipath it may also be suitable for application to outdoor environments like dense urban environments. Densification in such environments make it possible to still apply the single bounce reflection assumption. Base stations for next-generation systems like massive MIMO BSs may make this approach very lucrative and reliable when extended to outdoor urban environments.

2.6 Summary

This chapter has discussed key localisation techniques that are applicable to next-generation systems. A detailed comparison of TDOA and AOA is provided. A discussion of the potential benefits of localisation to next-generation systems is presented in Chapter 1 section 1.2. That discussion provided new motivations to the research on localisation and it highlighted key potential system benefits that can be derived from localisation. However, it should be noted that the foreground work in this thesis, apart from Chapter 4 which presents a direct localisation technique, discusses techniques that are intermediate steps towards localisation. Section 2.5 therefore provides a link between the work in Chapter 5 to localisation. This chapter also provides detail about the mathematical models that are considered in this thesis. Further literature relevant for LTE localisation is provided in Appendix A.

CHAPTER 3

Simulations and Preliminary Experiments

This chapter presents the tools and algorithms that are used throughout the thesis to evaluate the performance of proposed techniques. Each section in this chapter separately presents its results and the discussion emanating therefrom. First the Bristol University ray tracing tool is introduced and an outline of how this tool is used in preliminary experiments is presented. The preliminary ray-tracing based experiments seek to evaluate TDOA and AOA techniques in a common data framework. The results from these preliminary experiments is presented including explanations of how these results inform the decisions made in succeeding chapters. The signal model that is used in DOA estimation simulations is presented and example simulated MUSIC plots are used to demonstrate the benefit of having a large number of antennas, on DOA estimation super-resolution. Finally the proposed DOA estimation technique that can be used with the Bristol Massive MIMO testbed is presented. Within this section, a comparison of the proposed technique against the MUSIC algorithm is presented, using the same signal model. A summary of how the results in this chapter are relevant for the succeeding chapters is provided at the end.

3.1 Bristol University Ray Tracing Tool

The ray-tracing software that was developed at the University of Bristol, is used in throughout this thesis to either evaluate algorithms or validate experimental results. The ray-tracing tool, which uses a ray-launching engine, is based on a validated, realistic 3D ray-traced channel model as used in [105] [106]. The same ray-tracing model was used to generate most of the statistics that are now specified in the 3D extension of the 3GPP/ITU channel model [107]. The tool has also been extensively used in many other peer-reviewed publications [108] [109] [110]. It

incorporates a real-world environmental database of the City of Bristol (UK), which is a 3D Laser Illuminated Detection And Ranging (LIDAR) database of the city. Figure 3.1 shows an example of a point-to-point BS-MS link with all the determined multipath rays between the BS and the MS. Figure 3.2 shows the power-delay-profile (PDP) for a BS-MS point link. The time delay shown on the PDP are the absolute time delays for each ray between the BS and MS.

Figure 3.1: BS-MS point-to-point line with all determined rays.

Ray-tracing for the purposes of this thesis, is conducted at a carrier frequency of 3.51GHz to match the carrier frequency used on the Bristol massive MIMO testbed [8]. Although the testbed uses 3.51GHz, separate discussions of mmWave applications elsewhere in this thesis make assumptions of the mmWave frequencies and associated bandwidths. Also any localisation (AOA) techniques at this frequency can be exploited to perform downlink beamforming for mmWave. The transmit power is set to 32dBm and the receiver sensitivity is configured to be -120dBm. Isotropic antennas are used at both the BS and the MS at runtime, but any required antenna patterns on the receive and/or transmit end can be applied during post-processing. Table 3.1 shows the parameters that are produced for each ray.

3.2 Preliminary assessment of TDOA and AOA in an urban environment

This section looks at the issue of urban positioning with mobile radio networks utilising the traditional techniques of TDOA and AOA that are discussed in Chapter 2 section 2.2. The ray tracing tool, discussed above in section 3.1, is used to evaluate the two techniques in an outdoor area of the city centre of Bristol City. A dense picocell network is simulated covering a 1km^2 area of the city centre. Data from the ray tracing software is used for estimating the location of the

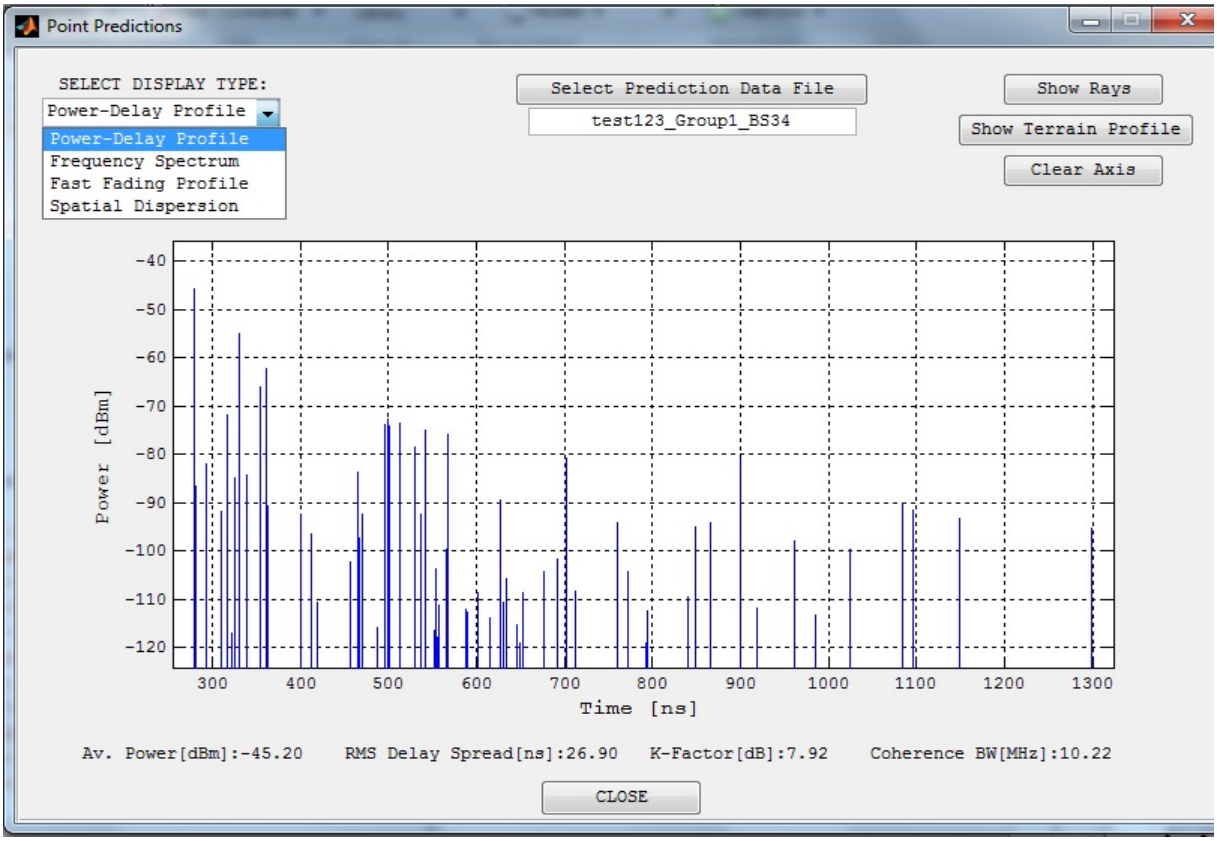


Figure 3.2: Power delay profile for a BS-MS Link.

Table 3.1: List of ray-tracer outputs

1. Easting coordinate of BS (x coordinate)
2. Northing coordinate of BS (y coordinate)
3. Height of BS (z coordinate)
4. Easting coordinate of MS
5. Northing coordinate of MS
6. Height of MS
7. Frequency
8. Transmit power
9. Time delay
10. Received power
11. Phase
12. Elevation AOD
13. Azimuth AOD
14. Elevation AOA
15. Azimuth AOA

mobile stations at hundreds of different random positions. Two dimensional (2D) TDOA, AOA and a hybrid TOA-AOA schemes are used, which only consider the position of the MS in the

azimuth domain. In this case, the considered scenarios are those where the base station - mobile station (BS-MS) links are known to be either line-of-sight (LOS) or non-line-of-sight (NLOS).

3.2.1 Ray-tracing setup

36 Base stations were placed in a 6x6 grid that covered a 1km^2 area of Central Bristol as shown in Figure 3.3. Since the study was considering a small cell deployment, the BS-BS distance of approximately 200m was considered. 1000 MS positions were then placed in a random uniform distribution in both axes of the grid area, and ray tracing was run against each BS-MS link.

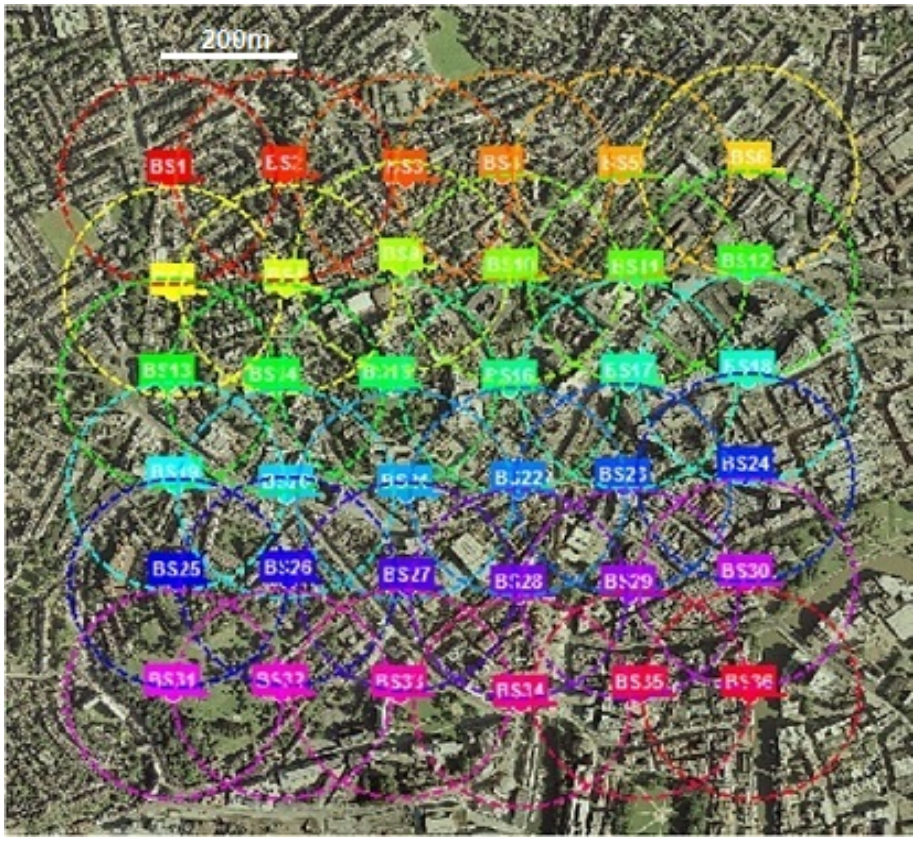


Figure 3.3: BS deployment.

3.2.2 Ray tracing parameters and pre-processing

Table 3.2 summarises the ray tracing parameters that were used for this experiment. The parameters are the used in all the succeeding ray-tracing experiments.

Pre-processing of the rays is carried out so that the data passed on to the localisation algorithms, improves the accuracy. All data/rays for each BS-MS link are sorted according to time delay. This was based on empirical observations that indicated that choosing rays with least delay produced better localization performance than selection based on received power level. This

Table 3.2: Ray tracing parameters

Parameter	Value
Environment	$1km^2$ area of central Bristol
Frequency	3.51GHz
BS transmit Power	32dBm
BS height	15m above clutter
MS height	1.5m
Receiver sensitivity	-120dBm
Antennas	Isotropic

effectively means that rays with the least delay are used for localisation. For the case where the NLOS/LOS knowledge is used to choose rays, a prioritisation scheme was devised, which gives priority to LOS rays. In practical systems, NLOS identification techniques would be key to such a prioritisation scheme. For a given BS-MS link, if multiple rays are all either LOS or NLOS, then the ray with the least delay is selected first, and so on. For 3 BS TDOA algorithm, each mobile station selects 3 BSs within its proximity, selects rays with the least time delays, and then use those rays for localization. Following the ray prioritization scheme, LOS rays are selected first, and if not present, or not enough for the localization scheme desired, then ground-reflected (GR) rays are selected, before pure NLOS rays are used. Ground reflected rays are given preference over other NLOS paths because, usually, ground reflected paths interfere with direct LOS paths. This means that the ground reflected multipath component may be irresolvable from the LOS path since they generally exceed the temporal and spatial resolution capabilities of most measurement systems. The severity of this issue depends on the antenna patterns, the BS/MS heights and how far the mobile station is from the base station [111]. Range error produced by ground reflected (also rooftop diffracted) paths may be smaller than other NLOS scenarios, in most cases. Figure 3.4 below demonstrates why ground reflected rays may have time delays comparable to LOS rays. From the ray-tracing data, ground reflections or rooftop diffractions are determined by rays that exhibit a LOS matching azimuth AOA and AOD but different elevation angles.

The concept of using individual rays and their characteristic data, is rooted in the assumption that next-generation systems will have both the adequate bandwidth, and SNR to provide the inherent ability to resolve individual multipath rays. These and additional assumptions for this setup, are discussed in detail in Chapter 4.

3.2.3 Localisation performance

The location/positioning error estimate is calculated as the distance between the estimated position and the actual position of the mobile station as obtained from the ray-tracer setup.

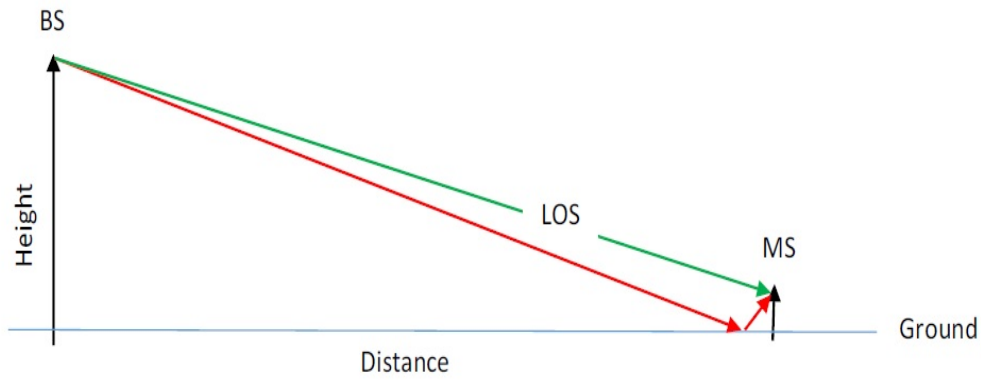


Figure 3.4: Ground reflected multipath (red).

$$e_i = \sqrt{[(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2]} \tag{3.1}$$

where (x_i, y_i) are actual coordinates for the ith MS taken from the ray-tracing tool, and $(\hat{x}_i, \hat{y}_i)$ are the corresponding estimated coordinates.

Considering reported smartphone GPS mean accuracy of 4.9m under open skies [112] and of 7-13m observed with recreation-grade GPS receivers under potential multipath [113], The accuracy target in this research is set at 10m. It should also be noted that these smartphones employ algorithms that use the phone's sensors to improve the position estimates. The results of the preliminary experiments in this chapter show that this accuracy level is achieved with 80% probability when NLOS identification is employed. It is therefore decided based on these factors that the standard accuracy evaluation for the rest of the work in this thesis is done at 80th percentile level, which is a probability level of 0.8 for all cumulative distribution function (CDF) curves presented.

3.2.4 Results and discussion

Performance of TDOA and AOA for an urban environment are presented in the location error CDF plots in Figure 3.5 below. Figure 3.6 compares TDOA vs AOA in both cases where NLOS/LOS knowledge is disregarded or used for prioritisation.

For a considered 80th percentile level, the results show that TDOA performs better than AOA when LOS rays are preferred according to the prioritisation scheme described above. This is a consequence of the fact that AOA errors are more severe than time delay errors, for localisation. The AOA algorithm assumes LOS propagation for the 3 BS-MS links that are selected for each

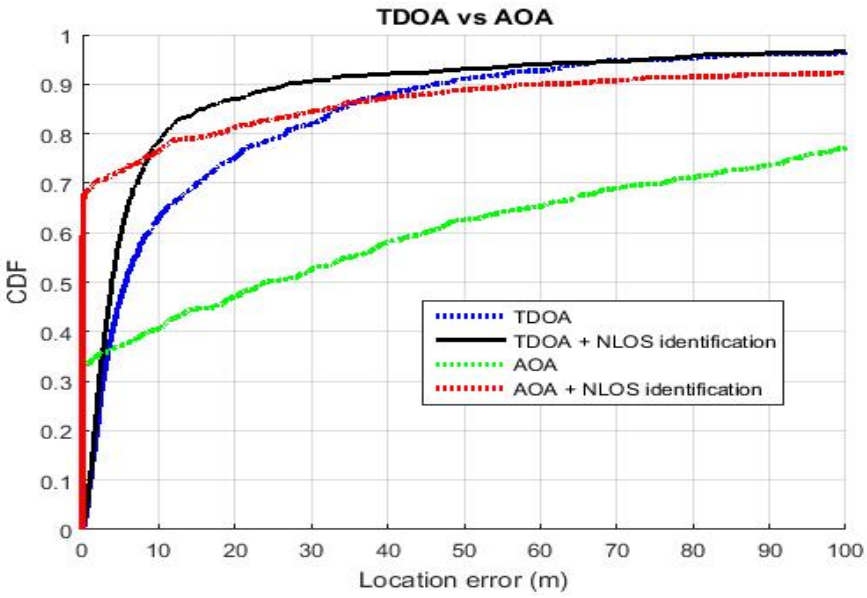


Figure 3.5: Performance of TDOA and AOA algorithms.

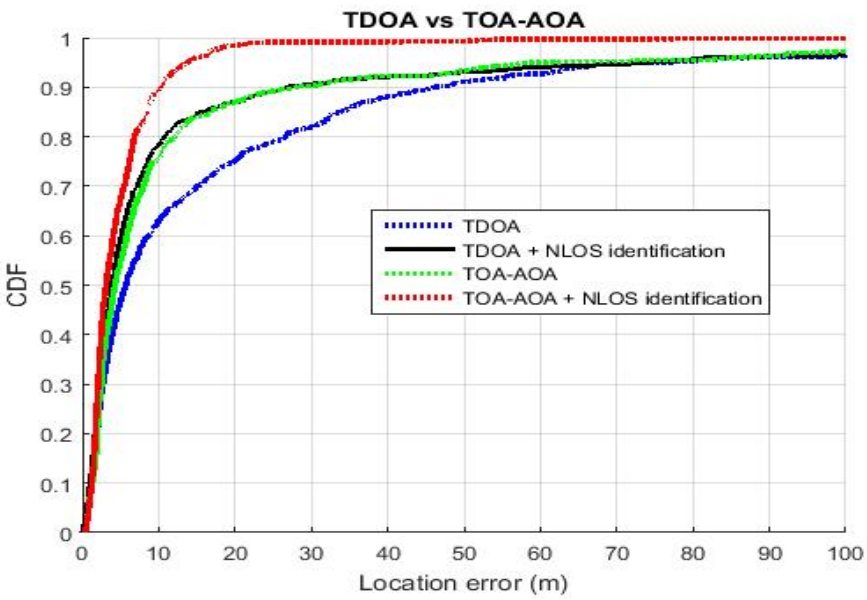


Figure 3.6: Effect of NLOS knowledge for TDOA and TOA schemes.

mobile position. The LOS angles may have errors, but if two or more of them are NLOS, then the algorithm is expected to produce location errors. The ray prioritization improves performance if there are LOS and ground reflected rays. This ray prioritization uses knowledge of the LOS/NLOS status of each ray as obtained from the ray-tracing tool. Here the usage of this knowledge, and the prioritisation scheme, is termed "NLOS Identification". This simply means that the algorithms

are prioritising the rays that have been identified to be LOS, GR and finally least delay NLOS, in that order. Without this NLOS identification, it can be seen that AOA performs much worse. Another observation when comparing TDOA vs AOA with NLOS identification, is that the performance of AOA is superior than TDOA for any given probability up to 0.75. This is as a result of the availability of LOS rays. AOA benefits from the LOS rays which produce perfect positioning, but once some of the 3, or if all of the rays used for the algorithm become NLOS, the localisation performance degrades. It should be noted however that in situations where all signals are LOS, AOA should produce better localisation accuracy that TDOA.

Because AOA results depend on LOS knowledge, obtaining 3 rays which are in LOS may be a challenge in dense environments, so the hybrid TOA/AOA improves the localisation performance by requiring only one LOS ray from only one BS. The probability of getting a single LOS or ground reflected ray between a BS and MS is much higher than that of getting 3 LOS rays between a MS and 3 different BSs. Figure 3.6 compares the hybrid TOA-AOA versus TDOA using the same set of data.

The results presented demonstrates the need for NLOS identification. NLOS identification in the above results is only achieved using apriori knowledge of the point-to-point rays generated by the ray-tracing tool, and because this information is not available in real deployments, this means NLOS identification in real systems could be a critical technique in improving localisation accuracy. This conclusion leads to the research and studies presented in Chapter 4.

3.3 Spectral-based DOA estimation using MUSIC

3.3.1 Signal model

Considering an antenna array with N elements that is receiving M narrowband signals ($M < N$) that are transmitted from the far-field source at the same carrier frequency F_c, the received signal can be modelled as

$$\boldsymbol{y}(t) = \boldsymbol{s}\boldsymbol{x}(t) + \boldsymbol{n}(t) \tag{3.2}$$

where $\boldsymbol{x}(t) = [x_1(t), x_2(t), ..., x_M(t)]^T$ is the transmitted signal envelop and $\boldsymbol{s} = [\boldsymbol{s_1}(\theta_1), \boldsymbol{s_2}(\theta_2), ..., \boldsymbol{s_M}(\theta_M)]$ is the manifold (steering vector) and $\boldsymbol{n}(t) = [n_1(t), n_2(t), ..., n_N(t)]^T$ is AWGN of power σ^2. The DOAs, $\theta_i : 1 \leq i \leq M$ can be estimated by considering L snapshots (block size) of the received signal $\boldsymbol{y}(t)$. The number of signals, M is not known at the beginning and this is considered as a detection problem. So in practice $\boldsymbol{y}(t)$ is taken over the L snapshots to produce $\boldsymbol{Y} = [y(t_1), y(t_2), \ldots, y(t_L)]$, yielding

$$\boldsymbol{Y} = \boldsymbol{S}\boldsymbol{X} + \boldsymbol{N} \tag{3.3}$$

The covariance matrix estimate $\hat{\boldsymbol{R}}_{\boldsymbol{yy}}$ is obtained by multi-dimensional correlation [114] as

$$\boldsymbol{R} = \hat{\boldsymbol{R}}_{\boldsymbol{yy}} = \frac{1}{L}\boldsymbol{Y}\boldsymbol{Y}^{\boldsymbol{H}} \tag{3.4}$$

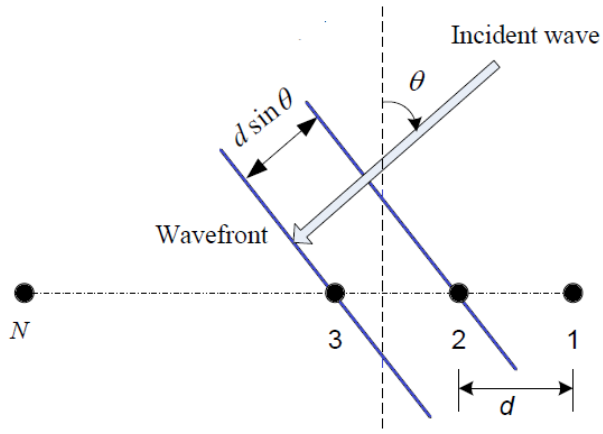


Figure 3.7: Plane wave impinging on a Uniform Linear Array.

For a uniform linear array (ULA) of N antenna elements with element-element spacing of d, the manifold vector for a signal with angle of arrival $\boldsymbol{\theta}$, as shown in Figure 3.7, can be written as

$$\boldsymbol{S}(\theta) = \left[1, e^{j2\pi(\frac{d}{\lambda})\sin(\theta)}, e^{j4\pi(\frac{d}{\lambda})\sin(\theta)}, \ldots, e^{j2\pi N(\frac{d}{\lambda})\sin(\theta)}\right]^T \tag{3.5}$$

where $\lambda = \frac{c}{F_c}$ is the wavelength and c being the propagation speed. For a Uniform Rectangular Array (URA), the manifold vector can be derived from the general case [115] [116]

$$\boldsymbol{S}(\theta,\phi) = \tilde{g}(\theta,\phi) \odot exp(-j\boldsymbol{\Delta}^T\boldsymbol{k}(\theta,\phi)) \tag{3.6}$$

where $\boldsymbol{\Delta} = \left[[x_1,y_1,z_1]^T,[x_2,y_2,z_2]^T,\ldots[x_N,y_N,z_N]^T\right] \in \mathbb{R}^{3\times N}$ is the antenna location matrix and $\boldsymbol{k}(\theta,\phi) \triangleq \frac{2\pi}{\lambda}[x,y,z]^T$ is the wave-number vector which, in the spherical coordinate system shown in Figure 3.8, can be expressed as

$$\boldsymbol{k}(\theta,\phi) \triangleq \frac{2\pi}{\lambda}\left[sin\theta cos\phi, sin\theta sin\phi, cos\theta\right]^T \tag{3.7}$$

θ and ϕ are elevation and azimuth angles respectively. $\tilde{g}(\theta,\phi) \in \mathbb{C}^N$ denotes the phase and directional gain of each element and $\odot$ denotes the Hadamard product. For the purposes of theoretical analysis, isotropic antennas are assumed, hence the manifold can be expressed as

$$\boldsymbol{S}(\theta\,\phi) = \frac{1}{\sqrt{N}}exp(-j\boldsymbol{\Delta}^T\boldsymbol{k}(\theta,\phi)) \tag{3.8}$$

and $\sqrt{N}$ is used in this case to normalise $\boldsymbol{S}(\theta,\phi)$. Now specifically for a URA, the source signal impinges on a 2-D array whose antenna location matrix can be expressed as $\boldsymbol{\Delta}^T = [d_x\mathbf{x}, 0, d_z\mathbf{z}]$,

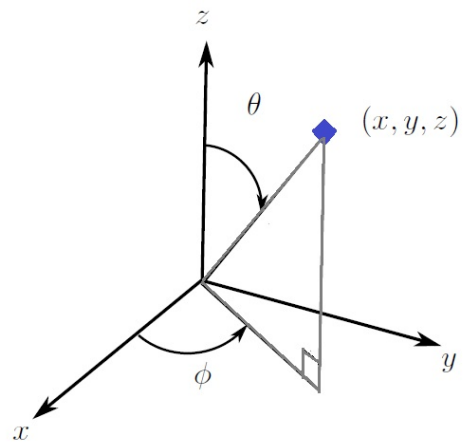


Figure 3.8: Spherical coordinate system

.

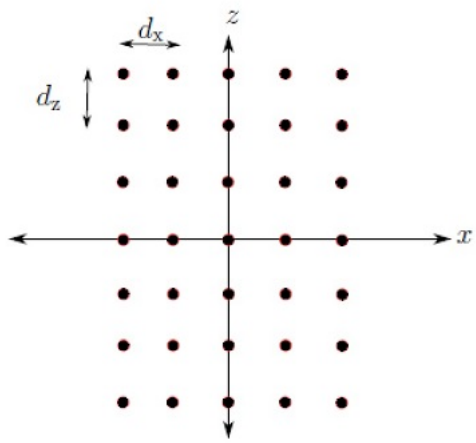


Figure 3.9: Uniform rectangular array

.

where d_x and d_z denote the element spacing in the x and z directions respectively as shown in Figure 3.9. The total number of antennas, N can be broken into number of antennas in x direction N_x and number of antennas in z direction N_z, such that the vectors $\mathbf{x}$ and $\mathbf{z}$ can be expressed as $\mathbf{x} = \mathbf{1}_N \otimes \tilde{\mathbf{x}}$ and $\mathbf{z} = \tilde{\mathbf{z}} \otimes \mathbf{1}_N$ respectively, where $\otimes$ denotes the Kronecker product, while $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{z}}$ are defined as

$$\begin{aligned} \tilde{\mathbf{x}} &\triangleq \left[-\tfrac{(N_x-1)}{2}, -\tfrac{(N_x-1)}{2}+1, \ldots, \tfrac{N_x-1}{2}\right]^T, \\ \tilde{\mathbf{z}} &\triangleq \left[-\tfrac{(N_z-1)}{2}, -\tfrac{(N_z-1)}{2}+1, \ldots, \tfrac{N_z-1}{2}\right]^T. \end{aligned} \tag{3.9}$$

This means that the manifold for the URA can be expressed as [116]

$$\boldsymbol{S}(\theta,\phi) = \frac{1}{\sqrt{N}} exp\left(-j\frac{2\pi}{\lambda}(d_x sin\theta \mathbf{x} + d_z cos\phi \mathbf{z})\right) \tag{3.10}$$

3.3.2 Eigen Decomposition

To estimate the power in noise and signal, and also the number of sources or transmitters, eigen decomposition is used. The received signal covariance matrix $\boldsymbol{R}$ can be written as [117]

$$\boldsymbol{R_{yy}} = \boldsymbol{S}\boldsymbol{R_{MM}}\boldsymbol{S^H} + \boldsymbol{R_{nn}}. \tag{3.11}$$

Assuming that the M sources are uncorrelated, $\boldsymbol{R_{MM}}$ is a diagonal matrix which represents the power in signal, which can be denoted by $\boldsymbol{P}$.

Applying eigenvalue decomposition (EVD) [118] to $\boldsymbol{R}$ yields eigenvalues that satisfy

$$\underbrace{\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_M}_{M} > \underbrace{\sigma^2{}_n = ... = \sigma^2{}_n}_{N-M} \tag{3.12}$$

where $\lambda_i : i = 1...M$ are signal eigenvalues and σ_n is the value for all noise eigenvalues. Consider $\boldsymbol{q_i} : i = 1...M$ to be the signal eigenvectors while $\boldsymbol{q_i} : i = M+1,...,N$ are noise eigenvectors. $\boldsymbol{R}$ can be represented as

$$\boldsymbol{R} = \sum_{i=1}^{M}(\lambda_i - \sigma_n^2)\boldsymbol{q_i}\boldsymbol{q_i^H} + \sigma_n^2 I \tag{3.13}$$

$\boldsymbol{R}$ can also be written as

$$\boldsymbol{R} = \sum_{i=1}^{M}|p_i|^2\boldsymbol{s}(\theta_i)\boldsymbol{s^H}(\theta_i) + \sigma_n^2 I \tag{3.14}$$

where $|p_i|^2$ is the average power of the ith source. So the expression $\boldsymbol{SPS^H}$ can be expressed as

$$\boldsymbol{SPS^H} = \sum_{i=1}^{M}(\lambda_i - \sigma_n^2)\boldsymbol{q_i}\boldsymbol{q^H}_{\boldsymbol{i}} = \sum_{i=1}^{M}|p_i|^2\boldsymbol{s}(\theta_i)\boldsymbol{s^H}(\theta_i). \tag{3.15}$$

It can be noticed that the signal eigenvectors $\boldsymbol{q_i} : i = 1...M$ are linear combinations of the steering vectors $\boldsymbol{s}(\theta)$ associated with M sources, or equivalently, the signal eigenvectors and corresponding steering vectors, span the signal subspace. The remaining eigenvectors $\boldsymbol{q_i}$: $i = M+1,...,N$, span the noise subspace, which happens to be orthogonal to the signal subspace. The number of emitting sources M may be calculated as N minus the multiplicity of the smallest eigenvalues of $\boldsymbol{R}$.

3.3.3 Detection

The number of sources M is usually assumed to be known in most DOA algorithms. The estimation of M is the job of detection algorithms (strictly, order determination). The Akaike Information Criterion (AIC) [119] and the Minimum Description Length (MDL) algorithms [120] are the most commonly used techniques. Figure 3.10 below shows the error performance of the two algorithms as a function of SNR.

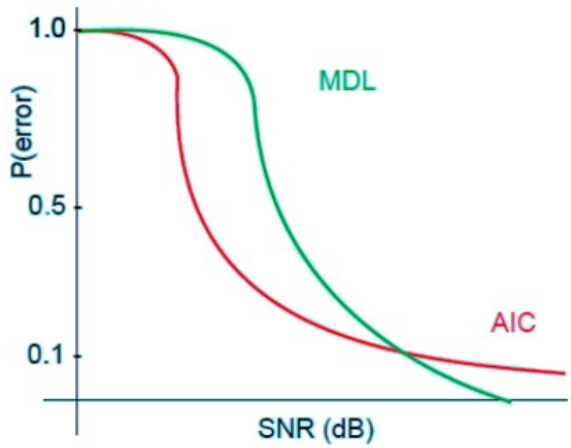


Figure 3.10: Performance of MDL and AIC algorithms. [120].

The signal subspace and noise subspace eigenvectors can be respectively separated as

$$\begin{aligned} \boldsymbol{U_s} &= [\boldsymbol{q_1} \quad \boldsymbol{q_2} \quad \dots \quad \boldsymbol{q_M}], \\ \boldsymbol{U_n} &= [\boldsymbol{q_{M+1}} \quad \boldsymbol{q_{M+2}} \quad \dots \quad \boldsymbol{q_N}] \end{aligned} \tag{3.16}$$

and the MUSIC algorithm can be expressed, using the spatial spectrum function $P(\theta)$, which is commonly denoted $P_{MU}(\theta)$, as [121]

$$P_{MU}(\theta) = \frac{1}{\boldsymbol{s^H}(\theta)\boldsymbol{U_n}\boldsymbol{U_n^H}\boldsymbol{s}(\theta)} \tag{3.17}$$

using the noise subspace, but it is also possible to use the signal subspace

$$P_{MU}(\theta) = \boldsymbol{s^H}(\theta)\boldsymbol{U_s}\boldsymbol{U_s^H}\boldsymbol{s}(\theta) \tag{3.18}$$

and search for peaks, but it is common practice to use the equation that produces lower dimension size. Whats used in practice, are the estimates $\hat{\boldsymbol{U}}_{\boldsymbol{s}}$ and $\hat{\boldsymbol{U}}_{\boldsymbol{n}}$ which are computed from the eigen decomposition of the estimate/sample covariance matrix $\hat{\boldsymbol{R}}_{\boldsymbol{yy}}$. Figure 3.11 demonstrates a MUSIC spectrum for simulated two co-channel signals received at angles of 20 and 40 degrees.

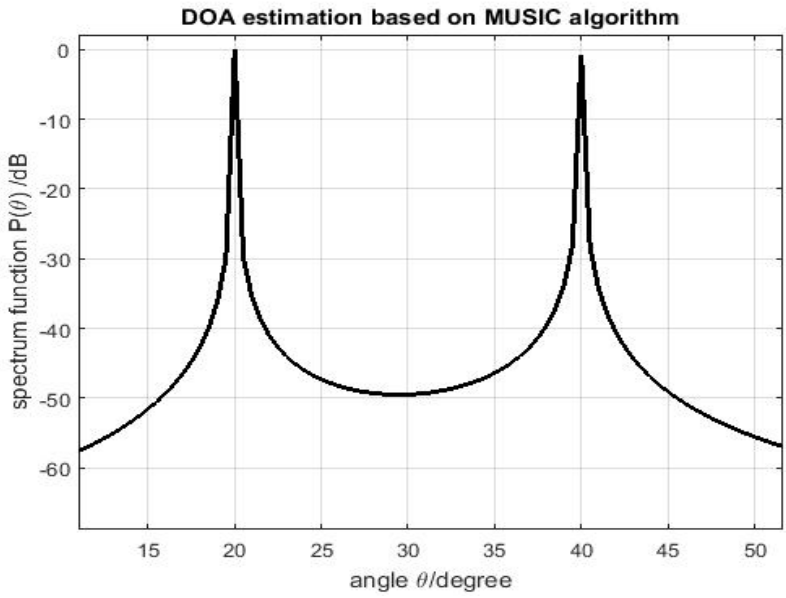


Figure 3.11: Simulated MUSIC spectrum.

MUSIC is a super resolution technique that is able to resolve DOAs that are closely spaced. The minimum separation distance (Δ) for the sources depends on number of array elements N, sample block size L and the SNR, according to the relationship in (3.19) [122].

$$\Delta = \left\{ \Delta : \frac{2880(N-2)}{LN^4\Delta^4} \left[1 + \sqrt{1 + \frac{LN^2\Delta^2}{60(N-1)}} \right] = SNR \right\}. \tag{3.19}$$

Figure 3.12 shows a simulation of the performance of the MUSIC algorithm with increase in number of antennas. The simulated signals are being received at 67 and 72 degrees. A massive MIMO case (e.g. 100 antennas) produces very narrow peaks at the angles where the signals are received. It can be noted that with 4 antennas, the algorithm failed to produce 2 distinct peaks, which demonstrates the problem of resolution when the angles are close together. The super resolution performance of the 100 antennas case suggest that massive MIMO arrays should offer greater DOA estimation accuracy.

When sources are correlated/coherent, preprocessing may be required in such scenarios, and there are various flavours of MUSIC algorithms to deal with such cases.

3.4 Proposed DOA estimation solution - APML

This section presents a parametric DOA estimation algorithm based on Maximum Likelihood (ML) technique. An ML algorithm is proposed by the Author's industrial sponsor, Roke Manor Research, with an interest to evaluate how it can be applied to next-generation systems, and

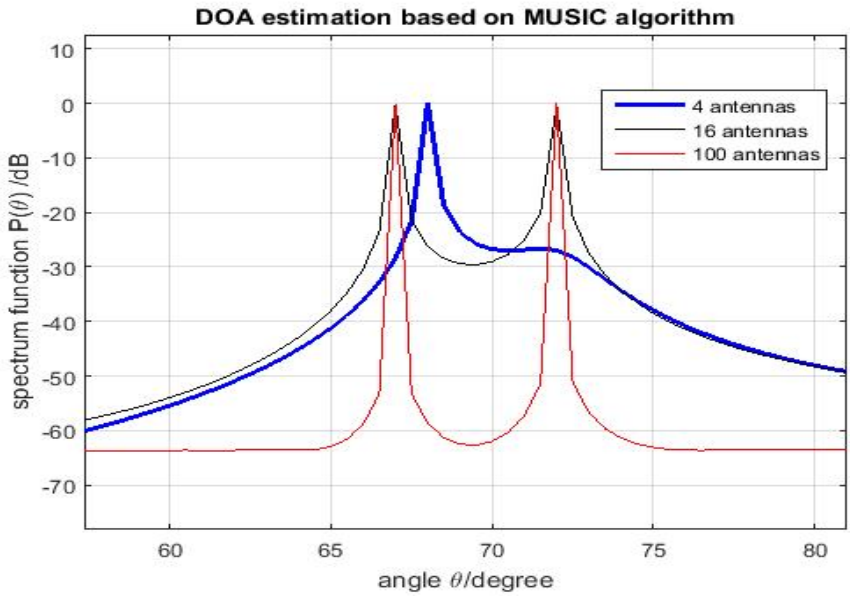


Figure 3.12: Simulated effect of number of antennas on super-resolution.

is modified by the author for use with the antenna array in experiments discussed in Chapter 5. The algorithm is presented and evaluated here as a precursor to Chapter 5 which uses real measurements utilising a massive MIMO testbed. The algorithm is based on the theory of vector space alternating projection [123] which has its origins in image processing [124] and later in Code Division Multiple Access (CDMA) communications [125]. Deterministic ML solutions consider the signals as deterministic and unknown quantities that should be estimated at the same time as the direction of arrival. Here the Alternating Projection ML (APML) algorithm is proposed. A variation of the APML is discussed and evaluated in a University College London (UCL) thesis by David Brandwood [126] who was also supported by Roke Manor Research. The APML algorithm works by creating an orthogonal projection to the received co-variance data, using the manifold vector of the initial determined signal alternately until all signal DOAs are recovered. For multi-antenna systems, the same signal model as used for MUSIC in section 3.3 is considered. Figure 3.13 below shows a block diagram of a generalised DOA system.

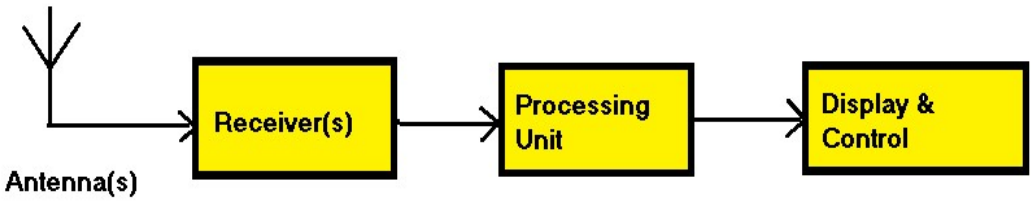


Figure 3.13: Generalised DOA system.

First the following assumptions for the system are made;

Assumptions

(i) The number of signals is known and is smaller than the number of antennas, M<N.

(ii) The set of any M steering vectors is linearly independent.

(iii) Isotropic and non-dispersive medium is assumed, i.e. uniform propagation in all directions.

(iv) Zero mean white noise and the signal(s) are uncorrelated.

(v) Far-Field sources

a) Radius of propagation is much greater than the size of the antenna array.

b) Plane wave propagation is assumed.

From the signal model in section 3.3, with the signal covariance matrix denoted by $\boldsymbol{R}$, and the measured manifold matrix denoted by $\boldsymbol{S}$, the beamscan result denoted by XX_i for $i = 1...M$ can be obtained and that beamscan is used to determine the signal direction of arrival. A conventional beam scan is performed to initialise the first signal. The first DOA is taken to be the angle at which the beamscan produces the highest peak.

$$\begin{aligned} XX_1 &= \left|\boldsymbol{S}^{\boldsymbol{H}}\boldsymbol{R}\boldsymbol{S}\right|, \\ \Rightarrow DOA_1 &= \underset{\theta}{\operatorname{argmax}}\ XX_1. \end{aligned} \tag{3.20}$$

$\boldsymbol{R}$ is the covariance matrix as defined in (3.4). Other signals are then initialised by taking the new manifold matrix $\boldsymbol{A}$, which contains only of the manifold vector corresponding to the first DOA determined in (3.20) above, so $\boldsymbol{A} = \boldsymbol{S}(DOA_1)$, and create the following projection which is orthogonal to $\boldsymbol{A}$

$$\boldsymbol{P_A} = \boldsymbol{I_N} - (\frac{\boldsymbol{A}}{\boldsymbol{A}^{\boldsymbol{H}}\boldsymbol{A}})\boldsymbol{A}^{\boldsymbol{H}}, \tag{3.21}$$

where $\boldsymbol{I_N}$ is an NxN identity matrix. A search is performed for the next signal by beamscan after applying the projection

$$\begin{aligned} XX_2 &= Re(\boldsymbol{B}^{\boldsymbol{H}}\boldsymbol{R}\boldsymbol{B}), \\ \Rightarrow DOA_2 &= \underset{\theta}{\operatorname{argmax}}\ XX_2. \end{aligned} \tag{3.22}$$

where $\boldsymbol{B} = (\boldsymbol{P_A}\boldsymbol{S})/\|\boldsymbol{P_A}\boldsymbol{S}\|$. This is repeated $\forall i : 1 \leq i \leq M$, each time updating $\boldsymbol{A}$ to add the manifold vector of the "found" signal, so the relevant matrices will contain

$$XX_i = \begin{cases} |\boldsymbol{S^H R S}| & \text{if} \quad i = 1 \\ Re(\boldsymbol{B^H R B}) & \text{otherwise} \end{cases} \quad \text{and} \quad DOA_i = \underset{\theta}{\text{argmax}} \; XX_i. \tag{3.23}$$

All recovered DOAs are stored in $\boldsymbol{\theta} = (DOA_1, \ldots, DOA_M)$.

The AP-ML full set of the estimates are refined by performing an iteration as follows:

Initialization;
set value of M;
set vector of initial DOAs, $\boldsymbol{\theta}$ (obtained from initialization, (3.20) to (3.23));
set iteration limit;
while *iteration limit not reached* **do**
  for *i = 1, ..., M* **do**
    Delete the column of $\boldsymbol{A}$ corresponding to the current DOA;
    Delete the current initial angle in the vector theta of initialised DOAs;
    For all manifold points, perform beamscan: $xx = Re(\boldsymbol{B^H R B})$;
    Estimate the $i^{th} DOA : DOA_i = \text{argmax}_\theta \; xx$;
    Store the DOA scans for each signal, for later cubic interpolation: $XX_i = xx$;
    Update A to contain the manifold vector for the new signal: $\boldsymbol{A} = [\boldsymbol{A}, \boldsymbol{S}(DOA_i)]$;
  end
  Record the M, DOAs estimated into new $\boldsymbol{\theta}$;
  Compare set of estimated DOAs in new $\boldsymbol{\theta}$ with old $\boldsymbol{\theta}$ to check for convergence;
  if *convergence not achieved?* **then**
    continue;
  else
    exit;
  end
end

Algorithm 1: APML algorithm estimate refinement

Cubic spline interpolation is then performed to refine the peaks in the resultant beamscans, $\boldsymbol{XX}$ and after that, plot the DOA scans for each signal $XX_i, \forall i : 1 \leq i \leq M$

The key to recovering all DOAs is the knowledge of the total number of co-channel signals expected, M. This can be estimated using MDL which is described in section 3.3. For simulation purposes, the value of M is known from the simulated signal that is synthetically generated. For the purposes of the DOA estimation presented in Chapter 5, an arbitrary number was chosen, which satisfy the requirement of the AOA localisation techniques considered. Since LOS signals are assumed to be dominant in the scenarios that were experimented on, one may argue that a value of $M = 1$ is sufficient to recover the LOS DOA in most cases, but in some cases, NLOS

signals may arrive with higher power. A value of $M = 3$ improves the chance of obtaining a LOS DOA. Some analysis on the beamscans of the top DOAs can be performed to determine the LOS DOA which can then be used for localisation.

To simulate and evaluate the performance of the algorithm, a synthetic signal is generated in MATLAB, constituting the DOAs to be recovered, and passed through a simulated Rayleigh channel with the following parameters.

- SNR = 20dB
- Path loss index = 4 (highly shadowed environment)
- Distances considered: 80m and 100m from the base station
- Antenna array sizes considered: 32 and 64 antennas

The choice of simulation parameters above is influenced by the hardware configuration of the testbed used in later chapters, in addition to the envisaged experiments, as discussed in Chapter 5 and also in Appendix C. 32 antennas represent the number of antennas on the Bristol Massive MIMO testbed sub-panel, which can be separated in a distributed deployment. Also 64 antennas were used in a single BS localisation experiment described in 5. A distance of 100m from the BS may represent a possible MS position in a small cell deployment for an urban environment.

3.4.1 Simulation results

Figure 3.14 shows an example beamscan plot for a signal with a DOA of -67 degrees, using 32 antennas, at a distance of 80m from the BS.

Figure 3.15 below shows how the algorithm performs in a simulated urban channel, as the number of antennas on the BS is increased. It demonstrates, as expected, that increasing the number antennas on the BS, improves estimation accuracy. The simulation data at 100m from the BS, with 64 antennas was also used to evaluate the sensitivity of the algorithm to gain and phase errors. Fig 3.16 shows the effect of gain and phase errors in the data. Without error performance data for any particular real system, it may be hard to produce a result of the effect of measurement errors, that is applicable for general cases. However, the result in Figure 3.16 provides a general idea of the kind of errors that are more impactful to the performance of the algorithm. In this case, it can be concluded that phase errors are more severe than gain errors.

The APML algorithm was chosen for DOA estimation presented in Chapter 5, mainly because of its ability for super resolution even with a loosely calibrated antenna array. Given the large number of antennas on the massive MIMO tested used, APML performance is expected to be comparable to the simulated results. To compare the APML algorithm with MUSIC (both using a similar signal model described in section 3.3, a synthetic signal was used as described in the preceding simulations, and DOA estimations was performed with different number of BS antennas. The average DOA error for all UE positions(angles) for a given BS configuration, was

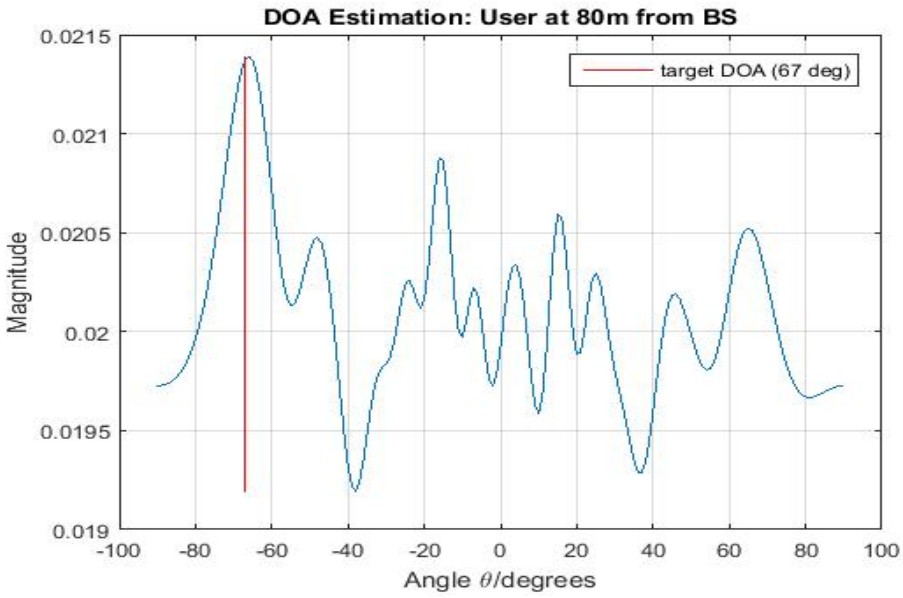


Figure 3.14: Example APML beamscan plot.

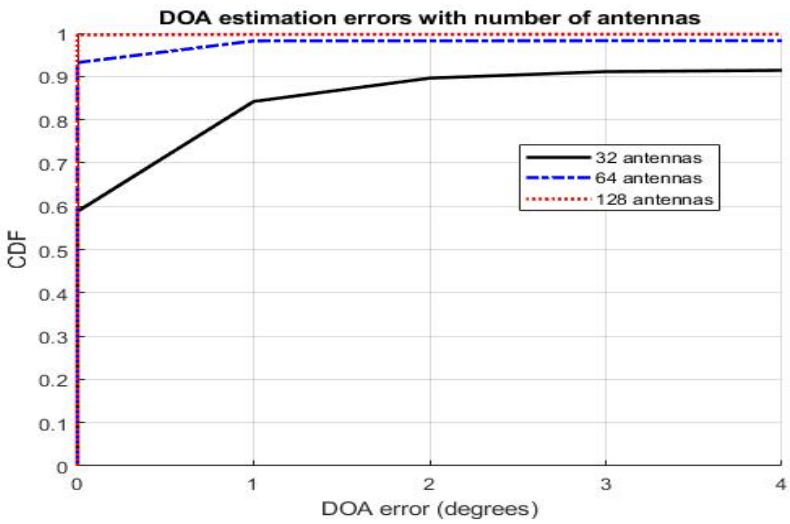


Figure 3.15: Effect of increasing number of antennas, on estimation accuracy.

calculated and rounded-up to the nearest degree. The results are plotted in Figure 3.17. The simulation results show that for the MUSIC algorithm is superior to the APML algorithm. This is notable for low number of BS antennas but as the number of antennas increase, the performance of APML algorithm improves to match that of MUSIC, for example in the case of 128 antennas. For a practical system, however, MUSIC has disadvantages as discussed in Chapter 2 section 2.4.2.1, in that it is very sensitive to gain and phase errors as well as BS antenna position errors. This means that it requires superior calibration schemes in order to get the performance that

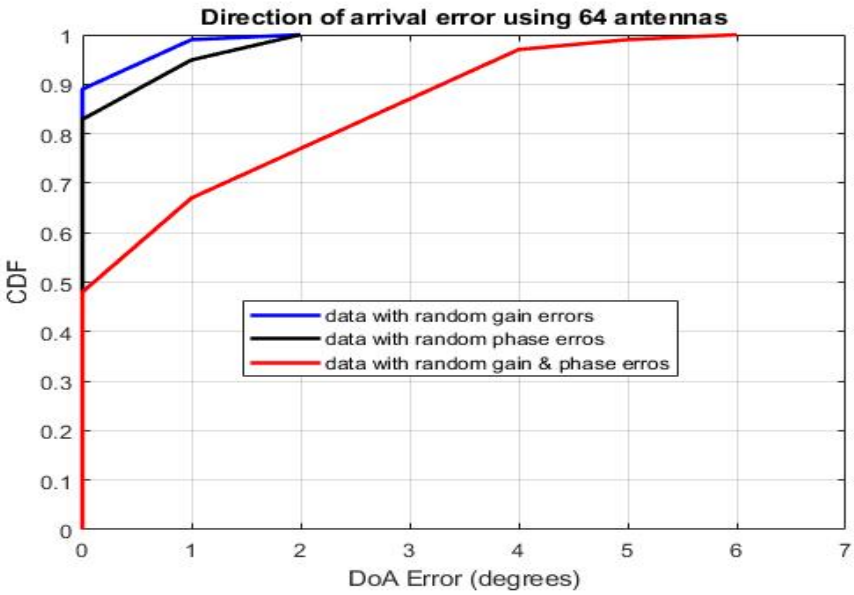


Figure 3.16: Effect of random gain and phase errors.

is demonstrated in simulations. The APML algorithm was adapted for use with the massive MIMO testbed presented in Appendix C and the experiments discussed in Chapter 5. The work presented in that chapter demonstrates the practicality of DOA estimation utilising the available testbed, with only minor software-based modifications.

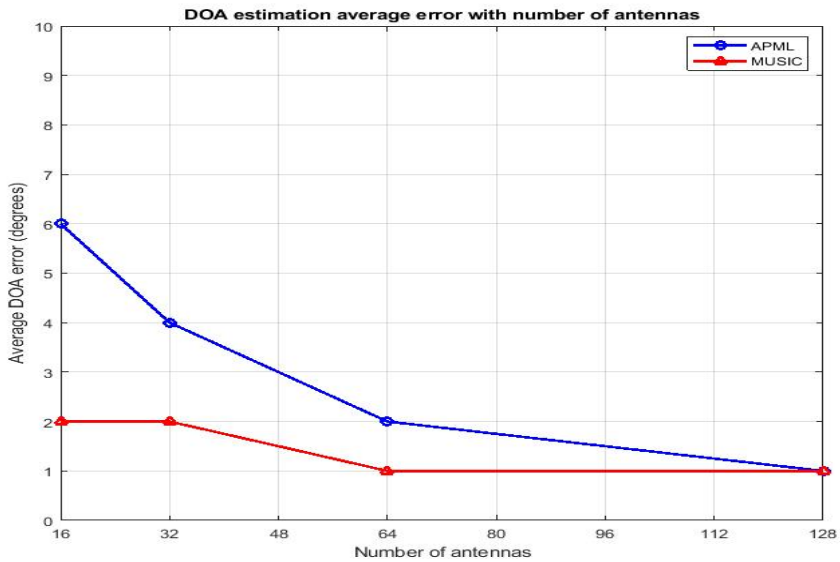


Figure 3.17: APML vs MUSIC simulation.

3.5 Summary

This chapter has discussed has presented the Bristol Ray-tracing tool which is key for the work that is presented in the next chapter. Preliminary experiments demonstrate how the data is processed and TDOA and AOA localisation techniques are evaluated. The key outcome in that section is the observation from Figure 3.5 that AOA benefits more from NLOS identification than TDOA. Conversely, it means AOA would be more suitable for LOS environments and TDOA would be more suitable for NLOS environments. Considering the proportion on mobile positions that achieve 10m accuracy, AOA improvement is approximately 90% as compared to about 25% for TDOA. Overall, the results demonstrate that NLOS identification is key to localisation. A proposed ML algorithm to be used with multi-antenna systems is presented, together with its simulated results. The APML algorithm performance improves with the number of antennas on the base station and its performance matches that of MUSIC when 128 antennas are used. From the results shown in Figures 3.16 and 3.6, the desired accuracy sought in the succeeding chapters of this research is set at 2 degrees for DOA estimation and 10m for outdoor location accuracy, both for an the established 80th percentile level as explained in section 3.2.3. To summarise how the results in this chapter are relevant to the work that is presented in succeeding chapters, section 3.2 introduces the ray-tracing software that is used in Chapter 4. The results in that section buttress the need for research into NLOS identification as contained in Chapter 4. Section 3.4 presents simulations which are relevant to the work presented in Chapter 5.

CHAPTER 4

LOCALISATION IN NLOS ENVIRONMENTS: MACHINE LEARNING APPROACH

This chapter discusses and examines the problem of localisation in Non-Line of Sight (NLOS) environments. It examines novel frameworks or schemes for NLOS identification and mitigation, as well as direct localisation, based on Least-Squares Support Vector Machines (LSSVM). Ray tracing data is used to evaluate these schemes and the traditional localisation algorithms, as discussed in Chapter 2, are employed to assess the performance of these schemes. The proposed frameworks for localisation using LSSVMs takes 2 approaches, first, the NLOS identification and mitigation approach, and second, the direct localisation approach.

Work presented in Chapter 3, section 3.2 demonstrates that NLOS identification is key to improving localisation performance in multipath environments. A method for achieving such NLOS identification using the same ray-tracing approaches, as described in section 3.2, allows a concise evaluation of the benefit of NLOS identification. This chapter also presents a technique for direct localisation using ray-trace data in section 4.4. This approach eliminates the need for NLOS identification and mitigation and it is shown to work best in multipath environments.

TDOA uses time delay measurements and AOA uses DOA measurements at the BSs. Usage of these measurements usually assumes a direct signal path between the base station and the mobile station. In reality, there may be no such direct path, especially in urban areas and other highly shadowed environments such as in urban areas and other environments like those for Internet of Things (IoT) devices. There is a need to take into account the multipath effect in the algorithms and techniques that are used for positioning. For TDOA, a very common approach is to identify signals that are a result of line-of-sight (LOS) propagation and those that are non-line of sight (NLOS). For non-line-of-sight signals, a mitigatory technique can then be applied to compensate for the positive range bias that result when the signal travels longer distances, as a

consequence of multiple reflections and refraction. Where enough line-of-sight signals are present, a preference of such signals can be applied in the selection of signals to use for positioning. The study in Chapter 3 section 3.2 demonstrates that even where a mixture of LOS and NLOS signals exist, simply applying a prioritisation of the LOS signals, followed by ground reflections before incorporating NLOS signals, can greatly improve the location accuracy. The AOA algorithm demonstrates that when all paths between the MS and BSs are in LOS, perfect positioning is possible, but the localisation performance deteriorates once NLOS paths get used. NLOS identification can therefore help improve localisation accuracy by selecting BSs that have a LOS path to the MS, and where these are not enough, single bounce ground reflected paths can be used, in case of 2D positioning since their azimuth angles of arrival are the same as LOS paths.

4.1 Chapter structure

This section provides descriptions of the chapter sections in-order to help understand the logical flow and organisation of content in this chapter since the content in all the sections present very related information. First is a brief recap on the NLOS problem and the common techniques that are employed for NLOS identification. This is important because it provides a picture of where the schemes proposed in this thesis fit in, with existing methods.

Second, as pointed out in the chapter introduction above, this chapter considers 2 approaches which are, NLOS identification plus mitigation approach, and direct localisation approach. Both use LSSVMs, so the next section (section 4.3) presents the proposed framework for NLOS identification and mitigation. The LSSVM model used is the same for both approaches, so it is outlined once in the first subsection of the first approach, and only the variations in the methodology and data processing are described for the second approach. The first approach itself has got 2 variations, which are the location specific approach (section 4.3.3), which is the flagship of the NLOS identification technique in this thesis, and the location independent approach (section 4.3.4).

Third is the direct approach in section 4.4, which as outlined above, used the same mathematical model, but with totally different methodology and data processing. This approach uses additional ray-tracing data from different areas of the greater City of Bristol. The data for the city centre area overlaps for both the first approach and this direct approach, which sets the basis for comparison of the approaches.

Lastly the comparison between the first approach (identification plus mitigation) and the direct localisation approach is presented in section 4.5.

4.2 NLOS identification for localisation systems

Non-line-of-sight identification can be achieved in a number of ways. The three main mechanisms involve considering the geometry of the channel to estimate the distance travelled by multipath

rays, or considering polarisation of the signal where every polarisation change is considered to be a reflection, or also using some statistical characterisation [127–132]. Most of the current NLOS identification and mitigation techniques involve some statistical approaches that require determination of the joint probability distributions of the underlying features, the outcome becomes very heuristic. Geometric model based approaches make assumptions of a maximum of just one reflection in the path between the BS and the MS [127], which may not be ideal in most cases. A comprehensive survey of NLOS identification and mitigation techniques is provided in [132]. A taxonomy of existing NLOS identification and mitigation methods is shown in Figure 4.1 below.

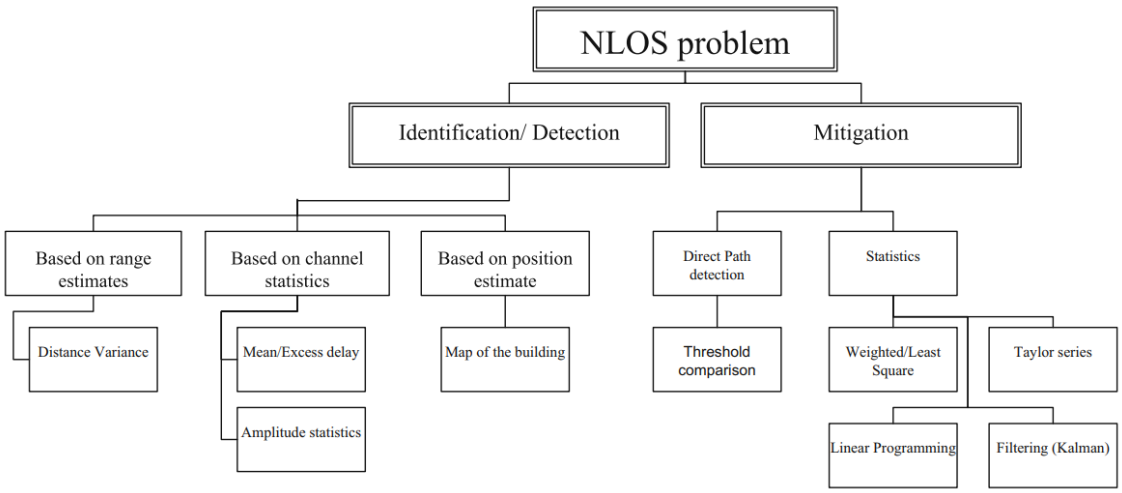


Figure 4.1: NLOS identification and mitigation methods [132].

The most common methods are those that use channel statistics. Parameters of the channel are computed at the receiver and hypothesis testing or set thresholds are used to determine if the received signal experienced an NLOS or LOS channel. Common parameters that are used are the amplitude, SNR, mean delay or excess delay and delay spread. The mitigation of the NLOS error in these methods is achieved via least-squares, weighted least-squares or Taylor series approaches. Machine learning techniques, in the form of support vector machines, is an optimisation-based approach that has been demonstrated to be effective for NLOS classification in an indoor environment [133] [134]. This section presents 2 techniques, one which is similar in approach, to [133] (discussion of the differences is provided below in section 4.3, comparison of results is provided in Table 4.5), but for outdoor urban environments, and the other which is location specific, both using ray-tracing data. The novel location specific approach presents a new framework for both NLOS identification and mitigation, that makes use of AOA measurements in addition to received power and time delay for each ray independently. Such a framework had not been proposed or used in this manner before. Usage of ray-traced data makes it easy to evaluate the proposed techniques since the raw ray tracing data includes the actual mobile positions, and

information on whether MS-BS links are NLOS or LOS, is already contained within the data for ready comparison with the estimated classification. This also makes it easier to quantify the effectiveness of the schemes, when the same data and framework as that used in the preliminary assessment in Chapter 3 section 3.2 is maintained. Use of support vector machines in these contexts may be thought of as applying artificial intelligence or machine learning concepts to the problem of localisation.

4.3 NLOS identification and mitigation using Least-Squares Support Vector Machines (LSSVM)

This study considers an optimisation-based machine learning technique, specifically the Least-squares Support Vector Machine (LSSVM) to perform both NLOS identification and mitigation. This approach does not require any statistical modelling of LOS and NLOS channels, which allows it to performs both tasks under a common framework. The approach in this study follows to some extent, a similar treatment as applied by Stefano Marano et. al. [133] where they used the LSSVM classifier and regressor for NLOS identification and mitigation respectively. The main differences of their work to this study are four-fold. Firstly, the type of data they used is different. They considered measured data in an indoor environment whereas this study utilises data obtain via ray tracing, in an outdoor urban environment. Secondly they used Channel Impulse Responses (CIR) and considered the whole CIR data for each mobile-base-station (MS-BS) link. In this study, a single tap from the Power Delay Profile (PDP) of each MS-BS link is used for localisation in the same way its applied in a separate scheme for NLOS identification presented in the location specific approach. In the location specific scheme, individual rays are classified separately rather than classifying the MS-BS link as a whole. Thirdly, because individual rays are used for localisation, mitigation is therefore applied to those individual rays in the form of range or time delay adjustments. Lastly the localisation algorithms considered in this study are those already proposed or identified in Chapter 2, for mobile networks, which are TDOA and TOA-AOA, which are different to those considered in their study . These differences mean that the location independent approach results will differ from those presented in [133] even where similar parameters are used. Such a comparison is provided in Table 4.5.

Support Vector Machines is a robust and effective technique for solving non-linear classification and function or density estimation problems. Support Vector Machines were originally introduced in statistical learning theory and structural risk minimisation, where convex optimisation problems, typically, the quadratic program, are solved. Least-Squares Support Vector Machines (LSSVM) is a reformulation of the standard SVMs in order to solve linear kernel-based systems. The solution is found by solving a set of linear equations instead of quadratic programming (QP) as in standard SVMs. LSSVMs were first proposed for classification by Suykens and Vendewalle [135] in 1999. They are classified under kernel-based learning methods. With

LSSVMs, you find links to regularisation networks, Gaussian processes, classical pattern recognition algorithms and extensions to unsupervised learning. They additionally put emphasis on, and exploit primal-dual interpretations. Where or whenever needed, it is also possible to impose sparseness, robustness and weightings to LSSVMs and there is a Bayesian framework that has been developed, with 3 levels of inference [136].

4.3.1 LSSVM NLOS Identification Model

The SVM methodology [135] seeks to construct a classifier of the form

$$(LOS/NLOS) = sign[Y(X)] \tag{4.1}$$

from

$$Y(X) = w^T \phi(X) + b \tag{4.2}$$

where $\phi(X)$ is the mapper that translates the input space, X, to a higher-dimensional space, which is explained further later in this section. w represents a vector which is normal to the separating hyper-plane and b is a parameter of the classifier that has to be determined empirically. Figure 4.2 shows the various components that make up a binary class support vector machine (SVM). Samples from the two classes that lie on either margin are called support vectors. The maximum margin $2/\|w\|$ and the margin for any given SVM, $b/\|w\|$ are shown.

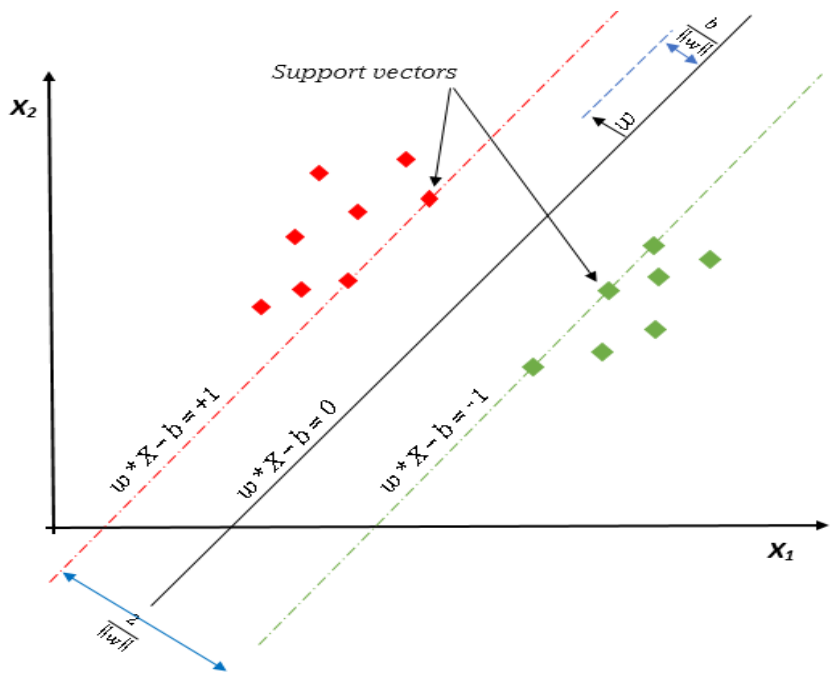


Figure 4.2: Support vector machine components

Given a training set of N data points $\{X_k, Y_k\}_{k=1}^{N}$ where the kth input is $X_k \in \mathbb{R}^n$ and $Y_k \in \mathbb{R}$ being the corresponding k^{th} output, an LSSVM classifier can be modeled as a function $\mathbb{R}^n \rightarrow \{+1, -1\}$, meaning the output sequence will have the binary class labels in the form which becomes the label or LOS status (i.e. "yes" for LOS and "no" for NLOS) corresponding to the input. The input space is a matrix with columns representing the chosen features, and rows representing the datapoints.

The classifier, according to the original formulation [137], should satisfy the following conditions

$$\begin{cases} w^T\phi(X_k) + b & \geq +1, \quad if Y_k = +1, \\ w^T\phi(Y_k) + b & \leq -1, \quad if Y_k = -1, \end{cases} \tag{4.3}$$

which is equivalent to,

$$Y_k[w^T\phi(X_k) + b] \geq +1, \quad k = 1, ..., N. \tag{4.4}$$

Since the above function is not explicitly constructed, if in the higher dimensional space, the separating hyper-plane is nonexistent, then the above inequality has to be violated, therefore a slack variable e_k is introduced to be able to tolerate miscalculations.

$$\begin{cases} Y_k[w^T\phi(X_k) + b] \geq 1 - e_k, & k = 1, ..., N \\ e_k \geq 0, & k = 1, ..., N. \end{cases} \tag{4.5}$$

Structural Risk Minimisation (SRM) principle is then used to formulate the minimisation problem for the classifier. Complete theoretical resources on structural risk minimization can be found in [138]. The LSSVM classifier minimisation problem [137] is formulated as

$$min J_1(w, b, e) = \frac{\mu}{2} w^T w + \frac{\lambda}{2} \sum_{k=1}^{N} e_k^2, \tag{4.6}$$

subject to

$$Y_k[w^T\phi(X_k) + b] = 1 - e_k, \quad k = 1, ..., N \tag{4.7}$$

which corresponds to regression, with its binary targets being, $Y_k = \pm 1$. The solution depends only on the ratio $\gamma = \lambda/\mu$, so the formulation can simply be written as,

$$min J_2(w, b, e) = \frac{1}{2} w^T w + \gamma \frac{1}{2} \sum_{k=1}^{N} e_k^2 \tag{4.8}$$

to form a single hyper-parameter γ which is then used to tune the trade-off between model complexity and level of tolerable training errors [135]. In order to solve this, one constructs a Lagrangian function [139]

$$L_2(w, b, e, \alpha) = \frac{1}{2} w^T w + \gamma \frac{1}{2} \sum_{k=1}^{N} e_k^2 - \sum_{k=1}^{N} \alpha_k \{[w^T\phi(X_k) + b] + e_k - Y_k\} \tag{4.9}$$

where $\alpha_k \in \mathbb{R}$, are Lagrange multipliers. The optimal point is obtained by setting,

$$\begin{cases} \frac{\partial L_2}{\partial w} = 0 & \xrightarrow{\text{yields}} & w = \sum_{k=1}^{N} \alpha_k \phi(X_k), \\ \frac{\partial L_2}{\partial b} = 0 & \xrightarrow{\text{yields}} & \sum_{k=1}^{N} \alpha_k = 0, \\ \frac{\partial L_2}{\partial e_k} = 0 & \xrightarrow{\text{yields}} & \alpha_k = \gamma e_k, k = 1,...,N, \\ \frac{\partial L_2}{\partial \alpha_k} = 0 & \xrightarrow{\text{yields}} & w^T \phi(X_k) + b + e_k, k = 1,...,N, \end{cases} \tag{4.10}$$

and it can be noted that these resultant equations that the elimination of w and e will yield a linear system, that can be written in matrix form as,

$$\begin{bmatrix} 0 & 1_N^T \\ 1_N & \Omega + \gamma^{-1} I_N \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ Y \end{bmatrix} \tag{4.11}$$

with $Y = [Y_1,...,Y_N]^T$, $\alpha = [\alpha_1,...,\alpha_N]^T$ and $1_N = [1,...,1]^T$ I_N is an $N \times N$ identity matrix and Ω is the kernel matrix. The parameters of the classifier α and b can then easily be obtained by solving the above linear system. For the training data set with N input space, b and α_k obtained after training, using (4.11), can then be used in the classifier which can be rewritten from (4.2), in a discrete form [140] as

$$Y(X) = \sum_{k=1}^{N} \alpha_k Y_k e_k^2 \psi(X, X_k) + b \tag{4.12}$$

The function $\psi(X, X_k)$ is the Kernel which is typically taken to be $X_k^T X$ for linear SVMs or $\psi(X, X_k) = \exp\left\{-\frac{\|X - X_k\|^2}{\sigma^2}\right\}$ for Radial Basis Function (RBF) SVMs [141], with σ, being a constant.

The NLOS mitigating LSSVM regressor is obtained by solving the optimisation problem [142]

$$\arg min \frac{1}{2}\|w\|^2 + \gamma \frac{1}{2}\|e\|^2 \tag{4.13}$$

subject to $Y_k = Y(X_k) + e_k, \forall k$, and in a similar way to the classifier, one obtains a linear regressor function of the form

$$Y(X) = \sum_{k=1}^{N} \alpha_k \psi(X, X_k) + b \tag{4.14}$$

4.3.2 Methodology

The three components of the NLOS classification scheme are, ray-tracing to generate training data for the geographical area to be covered, Ray-tracing to generate new MS positions that will be the test data (BS configuration is maintained), and lastly LSSVM classification for each BS-MS link (in case of the location independent approach) or LSSVM classification for each ray. The LSSVM engine used for these studies is extracted from the open source LSSVM Lab MATLAB tool box provided by the the pioneers of Least-Squares Support Vector Machines, Suykens et. al. and is outlined in [140].

Ray-tracing was setup in the same way as described in Chapter 3.

4.3.2.1 Assumptions

Channel reciprocity

The ray tracing software is designed to have the BSs acting as signal sources and the MSs as receivers, so channel reciprocity is assumed for any purpose that require uplink transmission.

Network deployment and capability

Network deployments are vastly different between operators depending on the type and use patterns. This study assumes a network with base stations that are capable of obtaining reliable AOA information, possibly through the use of antenna arrays as is the case with massive MIMO. Also, the network is assumed to be capable of resolving individual rays or multipath components. Next-generation wireless systems will utilize greater bandwidth than current generation systems, with GHz bandwidths possible at millimetre wave frequencies. Furthermore, bandwidth is not the limiting factor in time resolution as the Cramer-Rao bound on the maximum likelihood timing estimator [143] indicates that SNR is the limiting factor, and so super resolution algorithms can be used to improve timing resolution.

Initial/a priori NLOS knowledge

The a priori NLOS/LOS knowledge for each BS-MS link or ray (which is used to compare and evaluate the proposed technique), is based on the experimental setup knowledge from the ray-tracing software. A comparison of the AOD and AOA in both azimuth and elevation is done, and if they match (i.e. correspond to each other to form a straight line), then the link is considered to be LOS. If only the azimuth angles match, but the elevation angles differ, the range is checked to see if its comparable to expected range and if that is the case, then the ray is considered to be a single bounce ground-reflected path (GR path). The phrase “ground reflected” (GR) for paths or rays is used loosely in this case to refer to rays that are single bounce, ground reflected or rooftop diffracted rays. If there is no match in either azimuth and elevation, and the range (from the time delay) is much more than expected (the difference is greater than the expected range) then the link is considered to be “pure NLOS”. The designations “GR” and “pure NLOS” are simply used to distinguish these scenarios for the purposes of building the appropriate training data as described in section 4.3.

Noise

Noise in the measured values as presented in the ray-tracer outputs, is neglected. Received power, AOA/AOD information and time delay estimation done in the ray tracing software is considered to be accurate enough for purposes of this study. No noise modelling is built into the algorithms used. In a practical system, noise will have an effect on the available SNR. If all recorded values are affected uniformly, then it will have a scaling factor, which does not affect the classification results.

4.3.2.2 Data pre-processing and localisation

Similar pre-processing to that described in Chapter 3 sections 3.2 is used to select rays for localisation. A major addition to the preprocessing concerns the processing for the location agnostic approach which uses CIRs between BS and MS, for classification. Since this takes all the rays for each link, ray-prioritisation is only required on the mitigation phase, during localisation. The two localization algorithms used to compare the performance of the proposed Direct method for localization in LOS and NLOS environments, are TDOA and a hybrid TOA-AOA scheme. TDOA uses 3 Base stations, and one of them should be the same BS that is used by the TOA-AOA scheme which only require one BS. Detailed information on these algorithms, is provided in Chapter 2, and these are used to evaluate the performance of the classification and mitigation schemes proposed in this chapter. The AOA localisation algorithm is not considered in this chapter because of its reliance on LOS propagation. This chapter focuses on dense urban environments and the AOA algorithm was demonstrated in Chapter 3 to perform worst in such an environment, hence why it is not being considered here. However, it is worth noting that if NLOS identification results in a scenario where an MS sees at least 3 BSs that are in LOS, then the AOA algorithm provides the best localisation accuracy.

4.3.3 Location specific approach

Data from the ray tracer includes, for each ray, parameters like the BS position, time delay, received power and Angle of Arrival (AOA). This information is what forms the input data points $X_k \in \mathbb{R}^n$. The number of columns of X depends on the total parameters that are being used for classification. Using a priori NLOS knowledge, the output Y_i for each input point indicates the NLOS/LOS label. If a particular k^{th} path/ray is found to be LOS, then $Y_k = +1$, otherwise $Y_k = -1$ for an NLOS path. The output sequences used for training, $Y_k \in \{-1, +1\}$ forms a column vector whose size is equal to the training data size. Tens of thousands of data points were generated from ray tracing by placing thousands of MSs uniform-randomly within the area of interest. This forms the training set $\{X_k, Y_k\}_{k=1}^{N}$ from which various sizes of N can be extracted, making sure that it contains 50% of NLOS and 50% LOS rays. The LSSVM was trained using 10-fold cross-validation. In a 10-fold cross-validation, the training dataset is partitioned into 10 approximately equal-sized blocks, each containing 50% LOS and 50% NLOS paths. The LSSVM is then trained on 9 blocks and evaluation of performance is performed using the remaining block. This is carried-out for a total of 10 times, utilising each of the 10 blocks only once for evaluation and 9 times for training. The 10-fold cross-validation produces the tuning parameters γ and σ for the LSSVM. These parameters are then used for the training phase, which produces the parameters α and b. After the training phase, the new ray tracing data, which represents the data to be used for positioning, is then run against the trained LSSVM to determine if individual rays are either LOS or NLOS, in the classification phase. The RBF kernel was used in all cases because it yields the best validation and test set performance [141]. The classification result for each path is compared

against the a priori LOS status. This approach provides a robust and fast verification method, and thus helps to evaluate the performance of the LSSVM classifier for different features without the need for conducting a measurement campaign.

4.3.3.1 Input space and feature selection

The base input space comprises of the BS location's eastings and northings coordinates in a 2D space. The proposition is that, with the knowledge of a particular BS's location and each ray's parameters/measurements, the LSSVM should be able to determine if that ray is LOS or NLOS. The base input space is extended by including features of interest. In this location specific approach, the time delay and received power measurements for each path/ray are considered (separately, and combined). The dynamic range of the features is reduced, by taking their logarithms. This was done after determining the best training dataset type (TD2). Also after determining the optimum configuration, i.e. training data set size and combination of features, the base input was extended by considering the AOA on the BS. For both classification and mitigation, different sets of training data were constructed with differing content for the NLOS data points portion, where a mixture of ground reflected paths plus other NLOS scenarios produced training data 1 (TD1), all ground reflected paths for training data 2 (TD2), all other NLOS scenarios excluding ground reflected (GR) paths incorporated in training data 3 (TD3). The goal is to evaluate the constitution of training data that gives the best performance.

4.3.3.2 Cross-validation size and training size determination

The main mechanism for identification and mitigation requires training the LSSVM with a set of data, to obtain both the tuning parameters (via cross-validation) and the simulated result. Choosing an optimal cross-validation and training data size means that one makes an appropriate compromise on the computing and memory requirements against the accuracy or performance of the classifier or regressor. A study to determine the optimal training size that produces acceptable tuning parameters was conducted. This is effectively determining the dataset size for cross-validation. For this, the Receiver Operating Characteristic (ROC) curves are used. A ROC curve gives information about the quality of the classifier. It illustrates the separation abilities of a given binary classifier. If the area under the curve is 1 on given test data, then a perfect separating classifier is achieved. The goal therefore is to achieve a classifier with a ROC curve that has area under it of 1. This is a curve that seeks to pass through the top left corner and the closer the curve passes through that corner, the better the classifier. Both sensitivity and specificity of the classifier range from 0 to 1, and an ideal classifier will have both values at 1. The x-axis in Figure 4.3 is 1-specificity, which means a perfect classifier will have its x-values at zero. Further details and experiments on how ROC curves are used to evaluate performance of SVMs are also found in [144]. Different data set sizes were used in cross-validation to produce

tuning parameters that were then used to classify a separate data set and ROC curves were plotted. Figure 4.3 shows the ROC curves for different data sizes.

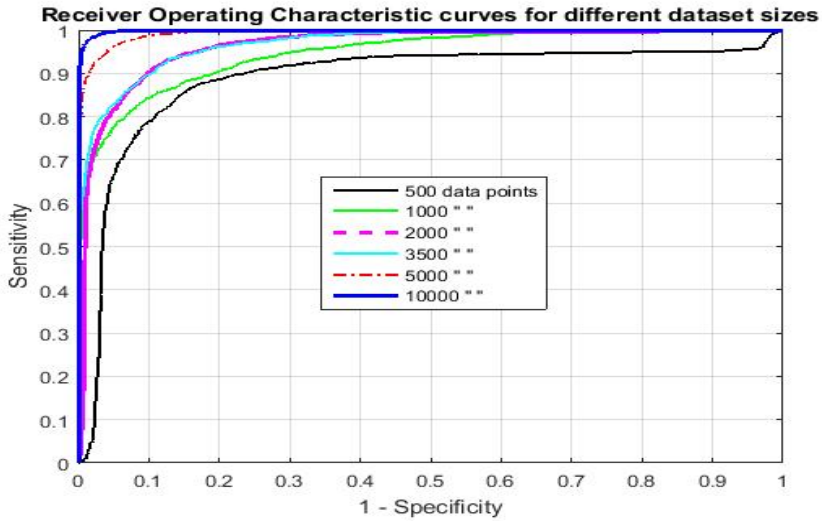


Figure 4.3: ROC curves for different dataset sizes.

It is evident from Figure 4.3 that the bigger the dataset size used for cross validation, the better the specificity of the classifier. A classifier that produces 95% accuracy with 5% or less false positives is a very good classifier for our purposes. It is evident that a dataset sizes of 5000 and 10 000 satisfies this. It was decided that a dataset size of 10 000 be used for cross-validation phase to maximise the sensitivity, and in line with other empirical evaluations on data sizes as outline in this section.

For the training phase, a study was conducted to determine an optimal training dataset size. Tuning parameters from the 10 000 cross-validation size, were used, to train the LSSVM with different dataset sizes and the classification performances, in terms of the probabilities, were recorded. Table 4.1 below shows the results.

Table 4.1: Training Size Determination

Training Dataset size	**Classification error probability**
50	0.8405
100	0.7385
500	0.4854
1000	0.2660
2500	0.1482
5000	0.0534
10000	0.0430
22384	0.0353

It can be noted that the classifier starts to show very good performance with a data set size of

5000. Doubling the dataset size from that point only improves the error performance by roughly 1%. Also, using the maximum available dataset size of 22 384 only provided an improvement of less that 1%. The data points were limited to 22 384 as a consequence of the total number of BS-MS links that were generated during ray-tracing. Inserting more MS positions within the grid of concern results in more data points. Given these results and the fact that a cross-validation dataset size of 10 000 was chosen, a similar training size of 10 000 was considered suitable for uniformity with cross validation, while providing a near-best error performance. The new data that is used to test the classifier is generated within the same area of the city by uniform-randomly placing a thousand MS positions. Running ray-tracing with a thousand MS positions within the 36 BS grid produced at least 8000 datapoints on each run, and these were analysed and everytime it was noted that these datapoints (rays) contain at least about 25% LOS rays and at most, just below 50% LOS rays, after 20 ray-tracing runs.

4.3.3.3 NLOS mitigation

Localisation strategies considered in this study are based on the time delay between the MS and the BS. NLOS propagation leads to positively biased range estimates. Mitigation of the NLOS positioning error is achieved by estimating the ranging error, and in this case the LSSVM function estimation is used to estimate the error in the time delay measurement (or alternatively, the error in the range estimate) that is recorded in the ray-tracing simulation. Training data is constructed through calculations of the time delay error, by comparing the expected LOS propagation time (given the knowledge of the actual BS and MS locations from the ray tracer setup) and the measured time delay. Where range error is used as the output parameter, it is obtained from the relationship ($range\_error = c \times time\_delay\_error$) where c is the propagation speed. This is done for all the datapoints. The input space X comprises of the base input (base station position) and selected features or combinations of them. The output Y is the time delay error or range error, obtained as described above. The LSSVM is then trained and the obtained regressor parameters are used to estimate errors using separate localisation data. Regressor performance evaluation is done by subtracting the regressor output, in its form as range error, from the actual ranges as obtained by considering the actual BS and MS positions from the ray-tracer. Cumulative Distribution Functions (CDF) of the residual range error after mitigation, are plotted against the original measured range error, to evaluate the improvement after mitigation. NLOS mitigation is achieved by subtracting the output of the LSSVM for each input, from the measured time-delays (or ranges). In all the above cases, mitigation was applied regardless of whether the path or link was NLOS or LOS.

4.3.3.4 Results and discussion

The results presented in this section demonstrate the performance of the LSSVM classifier and regressor as an NLOS identification and mitigation technique, respectively. Table 4.2 defines the

3 different sets of training data (TDs) and their different mixes as discussed above in section 4.3.3.1.

Table 4.2: Training Data Sets Composition

Training Dataset Label	**Composition**
TD1	50% LOS, 25% GR and 25% Pure NLOS
TD2	50% LOS and 50% GR
TD3	50% LOS and 50% Pure NLOS

Table 4.3 shows that best performance is achieved when training data 2 (TD2) is used, i.e. when the training data points consist of half LOS rays and the other half being GR rays.

Table 4.3: LSSVM NLOS Identification Performance

Features	**Probabilities**		
	False LOS identification	**Missed LOS identification**	**Identification error**
Using training data 1 (TD1)			
Time delay (τ)	0.196	0.015	0.206
Received power (α)	0.058	0.008	0.065
τ & α	0.042	0.001	0.042
Using training data 2 (TD2)			
τ	0.189	0.014	0.202
α	0.048	0.012	0.061
τ & α	0.036	0.001	0.037
log(τ) & log(α)	0.033	0.002	0.035
log(τ), log(α) & AOA (θ)	0.019	0.0001	0.019
Using training data 3 (TD3)			
τ	0.237	0.008	0.244
α	0.107	0.007	0.113
τ & α	0.103	0.002	0.104

GR rays are good approximations to LOS when the mobile target is far away from the base station as compared to the heights of both BS and MS ($R >> h_{BS} > h_{MS}$) where R is the range or distance between BS and MS, h_{BS} and h_{MS} are the BS and MS heights respectively. Training with such data therefore provides a very fine separating hyper-plane which reduces classification errors. In the test data considered, no purely NLOS paths (NLOS excluding GR) were misclassified. It is also evident that incorporating AOA and combining the two features, delay and received power, produces the best performance. Figure 4.4 shows the CDFs of the ranging error when the LSSVM is trained with different sets of data and features. When training data 2 is used, it can be observed that mitigation performs well for components that originally had small range errors although as the original error grew, mitigation could not offer significant correction. This is mainly because training data 2 contains half LOS and half GR paths, which

produce smaller range errors. Training data 1 provides better mitigation for large range errors and TD3 perform better at very large range errors. One can therefore choose the training data to use depending on the network setup, bearing in mind the expected range errors. Overall, large range errors are hard to mitigate effectively because they also introduce a larger dynamic range for the regressor.

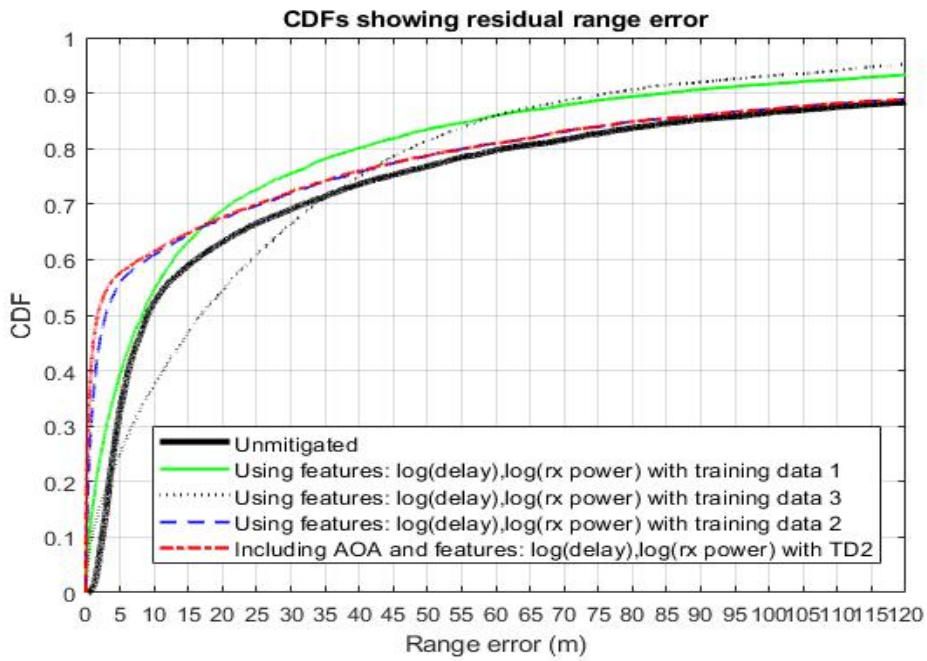


Figure 4.4: CDFs of residual ranging error.

When range error mitigation is only applied to NLOS rays, it was observed that no improvement results as compared to the case when mitigation was applied to all rays, which suggests that the performance of the LSSVM regressor is mainly affected by NLOS components. The TOA-AOA scheme is the best scheme to demonstrates the effect of ranging errors on localisation since it directly utilises a single range estimate for each ray. The algorithm used selects a GR ray whenever there are no LOS rays and reverts to 3 BS TDOA estimated range when all rays are purely NLOS. This allows mitigation to be applied to mainly GR rays which have smaller original range errors. The effect of applying mitigation to this scheme is shown in Figure 4.5. It can be noted that for the 80th percentile level, NLOS mitigation improves localisation performance by approximately 4m and produces a maximum location error of 10m.

4.3.4 Location independent approach

The location specific approach outlined above classifies and mitigates individual rays, between a BS and MS. With the location independent approach, features that are extracted from the channel's impulse response, like delay spread, can be used to determine if the link is LOS or

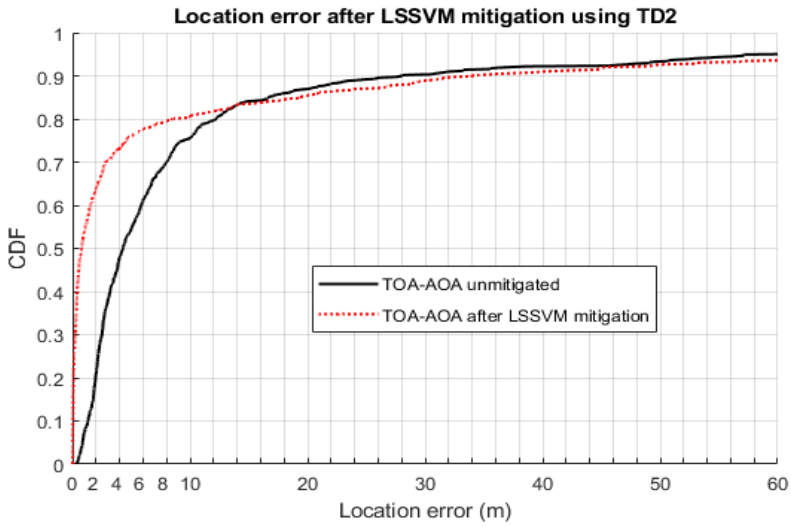


Figure 4.5: CDF of location error after mitigation.

NLOS. Once identification is performed, an assumption is then made that if the link is LOS, it is the first arriving ray that is actually the line of sight one. This ray is then used for localization. For this scheme, the position of the BS is not required. This means that it is possible to re-use parameters from one urban environment, in another similar environment without the need for re-training. The output sequence used in the training data Y_k is taken to be $Y_k = +1$ if the BS-MS link contains a LOS ray, and $Y_k = -1$ if the link does not contain a LOS path. The RBF kernel was used in all the cases for the same reasons outlined for the location specific case. The fundamental methodology in terms of ray-tracing is similar to what is presented under location-specific approach in section 4.3.3. Training data 2 (TD2) is used, which in this case takes BS-MS links with that include GR rays as best case NLOS rays. A similar dataset size determination as presented for the location specific approach is done, which indicated a size of 10 000 data points to be optimum. Mitigation follows similar methodology as conducted under location-specific mitigation and it similarly applies to the first arriving ray, whose parameters are used for localisation. Since mitigation is applied per individual ray, just as in the location-specific approach, it was not necessary to repeat the analysis, even after performing identification using this location independent approach.

4.3.4.1 Features

The approach in the location-specific case has been to classify and mitigate individual multipaths between a BS and a MS since it is the data for these individual rays that is used for localisation as described in Chapter 3. This section extends this approach by considering the overall channel between a BS and MS. The channel impulse response (CIR) is used to calculate the features that are used for NLOS classification. These features could be values or channel statistics derived

from the CIR. The input space is therefore taken to be the different combinations of the following features of interest. The following features were considered;

- Minimum delay for the link (τ_{min})
  This corresponds to the path with shortest delay, and the assumption that this would be the delay of the first arriving ray was used.

- Maximum received power for the link (α_{max}) The maximum amplitude of the received signal between the BS and MS. This is taken as the amplitude of the highest tap/path from the ray tracer. No assumption on this was used, so the rays with highest amplitude were used regardless of their delay.

- Half-power delay of the channel (τ_a)

$$\tau_a = \frac{\sum_{k=1}^{M} \tau_k^2 \alpha_k^2}{\sum_{k=1}^{M} \alpha_k^2} \tag{4.15}$$

  τ_a represents the time for half the power to arrive, α_k and τ_k represent the received amplitude and time delay of the kth ray, for a total of M rays.

- RMS Delay spread (τ_{rms})
  Root-mean-square delay spread for each link is calculated as,

$$\tau_{rms} = \sqrt{\frac{\sum_{k=1}^{M} (\tau_k - \tau_a)^2 \alpha_k^2}{\sum_{k=1}^{M} \alpha_k^2}} \tag{4.16}$$

  with the parameters as defined in the equation for τ_a

- Azimuth spread (θ_{rms})
  Root-mean-square azimuth spread for each link as seen on the BS, is calculated as,

$$\theta_{rms} = \sqrt{\frac{\sum_{k=1}^{M} (\theta_k - \theta_a)^2 \alpha_k^2}{\sum_{k=1}^{M} \alpha_k^2}} \tag{4.17}$$

  with the parameters as defined in the equation for θ_a being the half power azimuth spread, calculated in a similar way as the half-power channel delay in equation 4.15.

Other features that were considered include the total energy in a signal, which is the summation of received power amplitudes for all rays but empirical evaluation indicated that there was no correlation with LOS status of the link, hence this was not used. Also, because of the stationary nature of the channel data that was being used, other statistical features like the kurtosis and the K-factor, could not be confidently derived to aid in the classification. However, in a real implementation, this author believes these features will add further improvement to the classifier using the location independent approach. Power delay profiles with very few rays usually do not provide enough information in order to extract reliable channel parameters. Links that produced less than ten rays, from the ray-tracer simulation, were excluded from the analysis.

4.3.4.2 Results and discussion

It can be noted from Table 4.4 that most combinations of 3 or more of these features produce an error probability within 1% of each other. The best error probability performance of this approach

Table 4.4: Location Independent LSSVM NLOS Identification Performance

Features	Error probability
τ_{min}	0.3933
τ_a	0.2910
θ_{rms}	0.2495
τ_{rms}	0.2143
α_{max}	0.1754
$\tau_a, \tau_{rms}, \alpha_{max}$	0.1299
$\tau_{min}, \tau_{rms}, \alpha_{max}$	0.1320
$\tau_{min}, \tau_a, \tau_{rms}, \alpha_{max}$	0.1279
$\tau_{min}, \tau_a, \tau_{rms}, \alpha_{max}, \theta_{rms}$	0.1211

is shown to be approximately 0.12 . The same tuning parameters were used for classification in a different part of the city using all 5 features, and achieved an error probability of 0.1465 without re-training which indicated that indeed the approach is not location specific. Also a similar or comparable environment would be expected to produce similar classification performance. However, where there are significant differences in the density of the urban area, including differences in the urban canyon and foliage, the performance of the classifier in this approach can vary. These results are comparable to those summarised in [133] using similar features, although their study was for an indoor environment. This means the use of channel statistics as features for NLOS classification with LSSVM is a robust approach which can indeed be extended to any environment. Table 4.5 shows a comparison of the 2 different approaches presented

Table 4.5: Comparison of Identification Performance

Features	Error Probability		
	Location independent	Location specific[1]	Result from [133][2]
Using maximum received power	0.175	0.061	0.130
Using maximum received power and delay time based measurements	0.127	0.035	0.100

[1]Best case scenario (training data 2)

[2]Second row result is based on received power and signal rise time as compared to time delay measurements/statistics used in this thesis

in this thesis, against the results presented in [133]. Their second result (second row) is based on received power and signal rise time, but has been compared in this case to the result in this study, where time delay statistics together with received power are used. These results demonstrates that the location independent approach in this study is comparable to their results (within 5% error probability) because they both use a similar approach with differences outlined at the beginning of section 4.3. The location specific approach, which uses a completely different framework as described in section 4.3.3, is superior to both. Comparisons have only been provided where comparable features are used.

4.3.5 Summary

This study has demonstrated that LSSVM can be used for NLOS identification and mitigation in outdoor urban environments. Approximately 40% improvement to location accuracy at 80th percentile level, has been demonstrated after NLOS mitigation. The location-specific approach means if the LSSVM is trained with data from specific area, it cannot be used for NLOS identification in another area. One may have to train for a very large area, covering the whole city or even the whole country where possible. The system should then allow training once, with unlimited use of the data for the area as required. This means there will be a need for availability of ray-tracing data for each city/region considered and also periodic or continual updating of that data as the built environment changes. NLOS positions that are a result of dynamic environmental changes, like mobile scatters, are not captured in the design of the location specific approach. However there are ways to deal with such dynamic cases as discussed in section 4.4.4.

4.4 Direct Localisation using Ray-tracing and LSSVMs

In this section, the concept of LSSVM function estimation, as applied to NLOS mitigation described in section 4.3.3.3 is modified, to directly estimate position of the MS by estimating the x-coordinates and y-coordinates of the mobile station MS separately. The approach in this section, which will be referred to as the “Direct method" or "Direct approach”, achieves localisation of the mobile without first having to go through NLOS identification or mitigation. The ideas discussed in this section are completely independent from the previous sections, and it presents a novel framework for direct localisation that avoids the need to perform NLOS mitigation. Related approaches like fingerprinting [145] involve matching the received signal quantities, commonly the received signal strength, to the values that are pre-recorded in the fingerprinting database, for any particular environment. A fingerprinting database is built by taking or collecting measurements per grid, of the area of interest. Positioning accuracy in this scheme, therefore, depends on the grid size used. Fingerprinting requires cell matching before correlation with grids around that cell, whereas the LSSVM method discussed here, handles BS matching and location estimation within the same framework. Also, fingerprinting employs either probabilistic algorithms, such

as maximum likelihood, to estimate the position or deterministic algorithms that calculate the similarity between the UE measurement and the database grid-based measurements. Because of the challenges and limitations of fingerprinting, it is commonly considered as an augmenting scheme to other approaches to improve accuracy. Ray-tracing has also been demonstrated to be an effective tool for localisation, including when used together with finger-printing [146–148].

4.4.1 Experimental Setup and Methodology

A ray-tracing setup similar to the one described in section 4.3.2 was used to generate data for 3 different areas of the greater City of Bristol, UK. These areas were chosen to represent 3 different environments which are; a dense urban area, an urban peripheral area and farm land. Same BS and MS heights were used across the different environments for uniformity and also because the coverage areas considered where of the same size. MS positions are randomly placed within the base-stations' coverage area and ray-tracing is run for each BS-MS link. MS positions falling onto obscure areas like court yards do not produce any ray data because no signal could reach any of the BSs which are located outside of those areas, so they are excluded from the study. Assumptions outlined in section 4.3.2.1 are applied in this case. Chapter 3 section 3.2 carries the detailed account of the ray-tracing setup and data processing. Section 4.3.3.3 in this chapter provides detail on how ray-tracing data is processed for LSSVM function estimation or regression. Localisation performance is evaluated in the same way as described in previous experiments and the same localisation algorithms are used.

4.4.1.1 Training and regression

The key parameters from the ray-tracer that are used in this study are; the BS and MS locations (x and y coordinates), the azimuth AOA at the BS, the received power and the time delay, for each ray. The inputs of the regressor form a $(N \times 5)$ matrix whose 5 columns are: the BS x-coordinate, BS y-coordinate, the signal/ray's AOA at the BS, the logarithm of its time delay, and the logarithm of its received power. The output sequences used for training form a column vector with the x-coordinates or the y-coordinates of the MS depending on the coordinates being estimated at that point. Data-points (N) in this case therefore, refers to the total number of MPCs that are used, each having the above parameters.

Training data points are created by randomly placing MSs within the coverage area of interest. For each BS-MS link, the first arriving paths are chosen. From these, those that are determined to be LOS are grouped separately to those determined to be NLOS. The training dataset size is determined in a similar way as described in section 4.3.2, and a training data size of at least 3000 data points was shown to be sufficient to produce high performing tuning parameters, but to keep with the standard from previous experiments, a dataset size of $N = 10000$ is used for model and code uniformity. These datapoints consist of half LOS and half NLOS MPCs. Training

data was generated per each considered area and it is that training dataset that is used for the LSSVM location estimation within that area.

Training yields the regressor tuning parameters and constants which are then used to estimate the coordinates of the MS for any new given data set. Training is done separately for the x and y coordinates using the appropriate output sequences. This approach means estimation of the MS position is O(2) as compared to the traditional regression for NLOS mitigation. It is however possible to just estimate, say, the y-coordinate and use it together with LOS information where available (via NLOS identification or otherwise), to calculate the x-coordinates for those positions that are determined to be in LOS as shown in Figure 4.6 below.

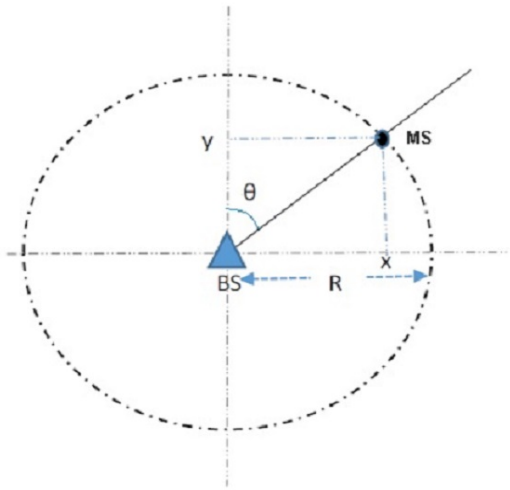


Figure 4.6: Obtaining the second coordinate for LOS scenarios.

Once the estimate of the y-coordinate is obtained, the x-coordinate can then be calculated as follows,

$$x_i = y_i.tan(\theta_i), \tag{4.18}$$

where x_i is the x-coordinate corresponding to the y-coordinate y_i, and θ_i is the AOA, for each ith MS position. This approach is only suitable for LOS positions. After obtaining the estimates for both the x and y coordinates of the MS, the location error is calculated as in Equation 3.1. The methodology can be extended to 3D by incorporating the estimation of the elevation coordinate, z, in a similar way. As illustrated in Chapter 3 section 3.2, the ray-tracer output data includes the elevation, or height of the MS. This data can be used for training and estimation. Training

will have to be done separately for the z coordinate and the consequence will be increased computation.

4.4.1.2 Post-processing and outlier removal

The ray-tracing setup has BS-BS distance of 300m so a coverage radius for each BS of 150m is considered for determining outliers. The BS deployment seeks to approximate envisaged 5G deployments, where a dense deployment of small cells is expected. The process of determining and excluding outliers involves calculating the distance R_i between the known BS position and the estimated MS location, i.e. the range of the MS, which is calculated as follows,

$$R_i = \sqrt{(BSx_i - M\hat{S}x_i)^2 + (BSy_i - M\hat{S}y_i)^2} \tag{4.19}$$

where BSx_i is the ith BS's x-coordinate and $M\hat{S}x_i$ is the estimated i^{th} MS x-coordinate. The other symbols' meaning follow, for the y-coordinate. It also follows that for multiple MS positions using the same BS, BSx_i and BSy_i are constant, for all i. A BS receives multiple rays from an MS and each ray is used to estimate the MS position. For a single MS position, some rays will estimate the position better than others, so those rays that result in the BS-MS distance greater than 150m are discarded. Empirical tests for an urban environment (see Appendix B Figure B.5) show that more regressor errors start increasing for MS positions beyond 100m from the BS. Outlier removal criteria may be tightened to any distance but that will create more coverage black spots, hence 150m range was selected. On average, the total number of data points that were excluded because of this criterion were around 10%. Figure 4.7, below, shows the effect of excluding those rays that are resulting in outlying MS positions. This result was produced following ray-tracing in the same city centre area as that considered in section 4.3.

4.4.2 Environments considered

The three environments considered are the city centre, city peripheral area, and open area (park). A zoomed-in smaller portion of these areas are respectively shown in Figures 4.8, 4.9 and 4.10 below. These figures are used to explain the results in section 4.4.4. The city centre area considered in this case is the same city centre area that was considered in the NLOS identification and mitigation schemes in section 4.3. Ray tracing is run against each of these areas to generate both the training data and the data used for performance evaluation. Sampled MS positions are indicated on the figures and they are color-coded according to quantized received signal power.

4.4.3 Multipath characterisation in environments considered

To characterise and be able to compare the extent of multipath in a given environment, the average RMS Azimuth Spread of Arrival (ASA) was used. Each ray arriving at the BS from an MS represents a BM-MS link, and for each BS-MS link, the RMS azimuth spread is calculated

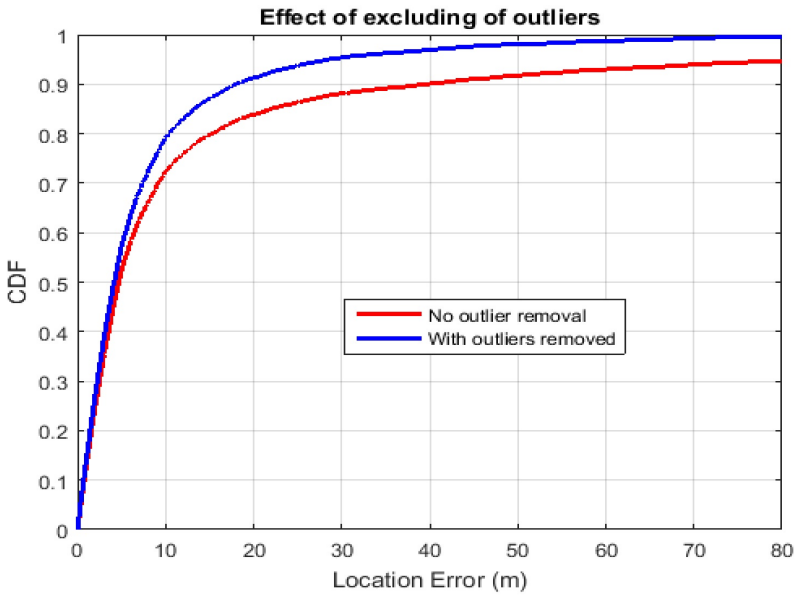


Figure 4.7: Outlier removal.

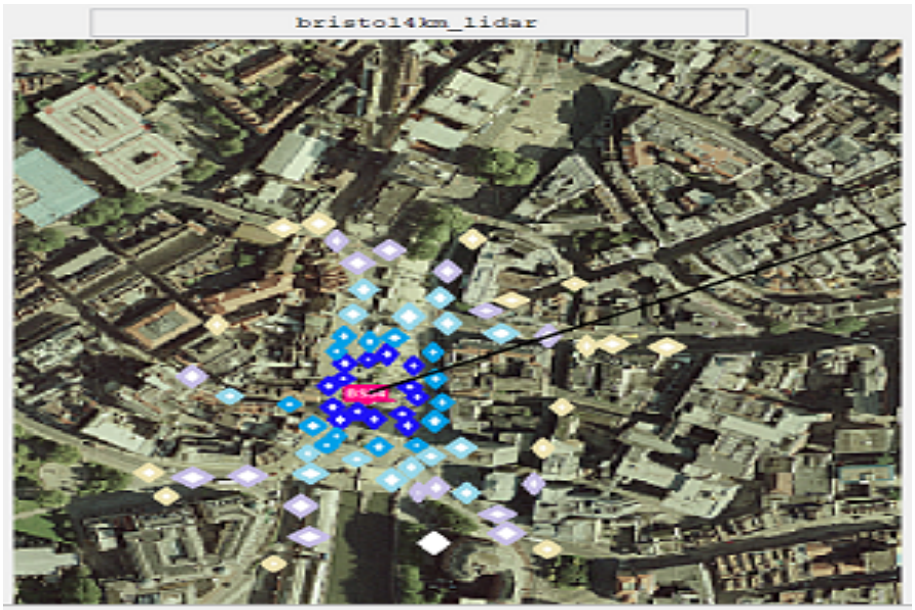


Figure 4.8: Dense urban area / City centre (sampled color-coded positions: same color means positions with same received signal power).

as shown in (4.17). An average is then calculated from the values obtained for all the MSs seen by the BS within a given environment. Table 4.6 shows the differences in the average azimuth spread values for the 3 environments considered.

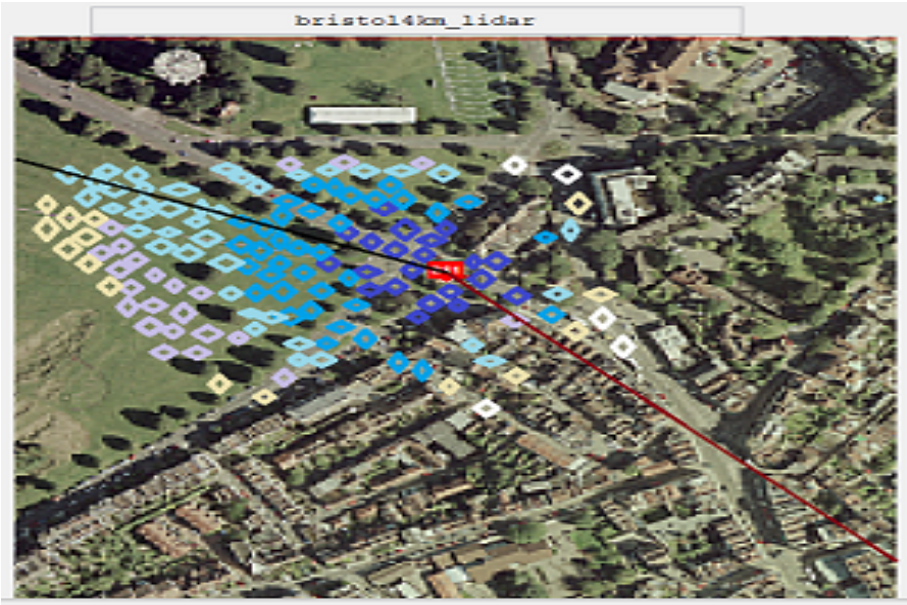


Figure 4.9: Urban peripheral area.

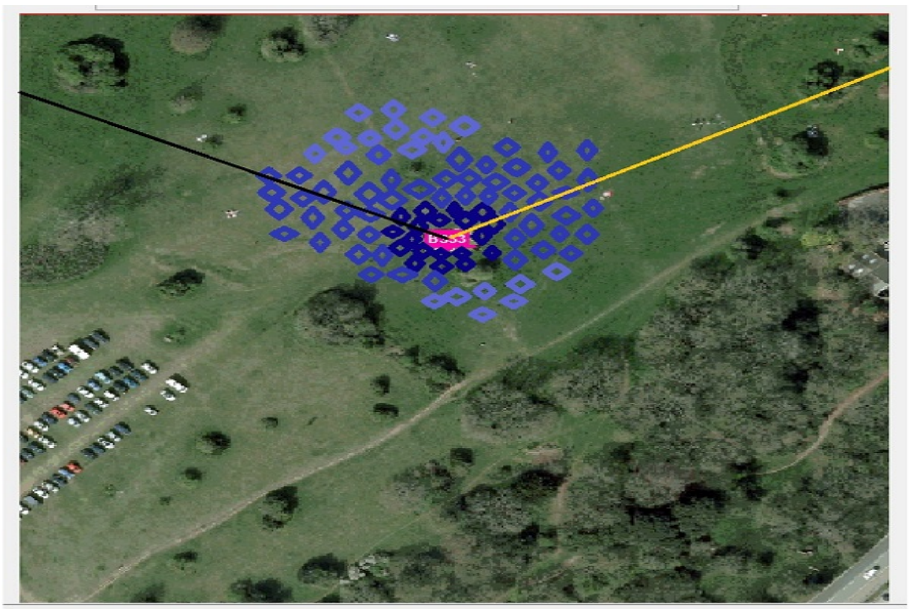


Figure 4.10: Park / Farmland, showing trees and open areas.

4.4.4 Results and discussions

Figure 4.11. below shows that the Direct localisation approach performs best in a dense urban environment, achieving location accuracy of 10m or less for all probabilities up to the 80th percentile. This is mainly because it benefits from the uniqueness of the multipath generated in such environments (unique spatial signature for each BS). The uniqueness or level of multipath is

Table 4.6: Mean azimuth spread

Environment	Mean azimuth spread
Dense urban area / City centre	29°
Urban peripheral area	15°
Park / Farmland or open area	7°

determined by the average azimuth spread as shown in Table 4.6. This can also be demonstrated by the fact that, given a set of measurements for received power, time delay and angle of arrival, the probability of getting multiple MS positions that can record similar measurements, from the BS, is small in such multipath environments. The LSSVM regressor works by maximising the margin of the separating hyper-plane. When the data points for a key parameter are homogeneous, as is the case for AOA measurements in LOS environments, the results for both classification and regression deteriorates. The Direct approach produced best location accuracy in urban

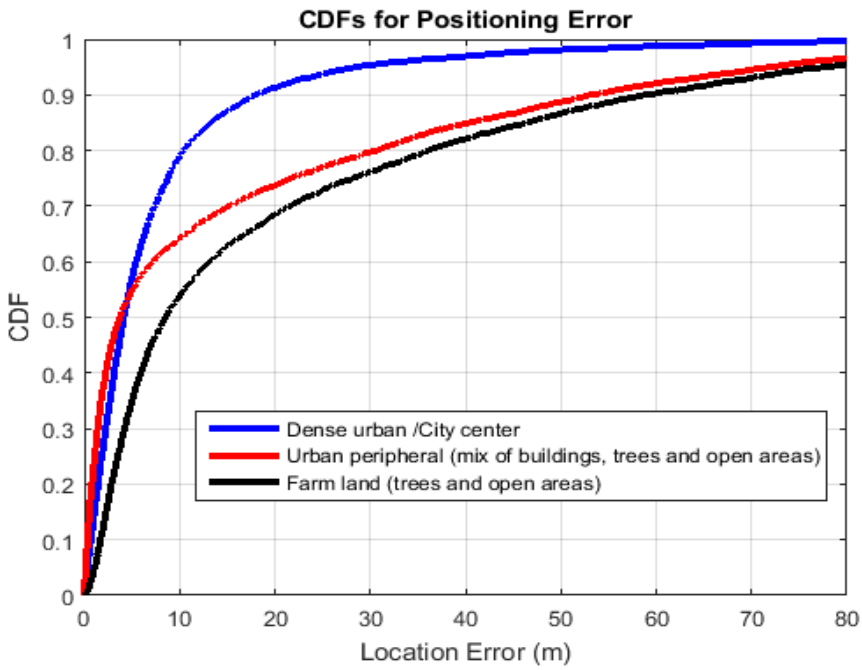


Figure 4.11: Localisation performance for the 3 environments.

peripheral environment up to the 50th percentile because the BS was closer to the buildings, so a corresponding portion of MS positions that were within 50m of the BS, were within the built-up area and hence benefited from the uniqueness of the multipath parameters as shown by its average azimuth spread value in Table 4.6, which is the second largest for the 3 environments considered. It can also be noted that the average azimuth spread is in second place to that of the dense urban environment, but it is still more than double that of the open area environment. The performance degradation after 0.5 corresponds to the portion of the MS positions that are

in non-built-up areas. That is why the curves for urban peripheral and open areas appear to converge as location error increase. Also it clearly shows that the performance in open areas is the worst for any considered percentile level, because the scheme does not benefit from multipath in such an environment as shown by its smallest azimuth spread, so the probability of recording similar parameters at completely different and far apart positions, increases as the homogeneity of the measurements increase in LOS environments. To demonstrate the multipath benefit for the Direct approach, one may consider a straight line extending from the BS (line of bearing) outward in any direction. It is easy to see that the line will cross more MS positions of the same given received power, in farmland/open-area type (LOS) environments than in dense urban (NLOS) areas. This also explains why AOA is a key metric, as demonstrated in Figure 4.12 under section 4.4.5 below, because it resolves ambiguities where multiple measurements have same estimated range and received power.

4.4.5 Sensitivity to measurement errors

In practical systems, both measurement errors and model errors (discrepancies between ray-tracing model and the actual environment) would be present. The sensitivity of the localisation estimate to measurement errors in the three quantities (AOA, Time delay and received signal strength) used to determine location, has been investigated by introducing a Gaussian distributed error in each. This is done in order to provide an indication of the parameters/quantities that are more sensitive to errors, hence more critical for the performance of the LSSVM regressor. A 5% standard error was introduced in each individual parameter in turn and also in all parameters at once and location estimation was performed for each case. The same ray-tracing data from the above results was used, following the methodology described in this section. Figure 4.12 demonstrates that errors in the data affect the localisation accuracy. These errors could be introduced by electromagnetic noise or equipment errors. Considering the established 80th percentile level, errors in all 3 parameters worsens the location accuracy by 40m. The results in Figure 4.12 also show that the Direct approach is most sensitive to AOA errors. An AOA error for a multipath component in a dense environment corresponds to a totally different path between the BS and the MS with totally different channel characteristics. It can be expected therefore, that an increase in such errors can significantly impact the performance of this Direct scheme. However, this is good news because in practice, next-generation wireless systems that employ antenna arrays at the BS, like Massive MIMO [8], will likely provide very accurate AOA estimates, so this approach sits well with envisaged fifth generation (5G) systems. It should however be noted that the challenges of accurate calibration in real array systems remain, which affects the angular resolution.

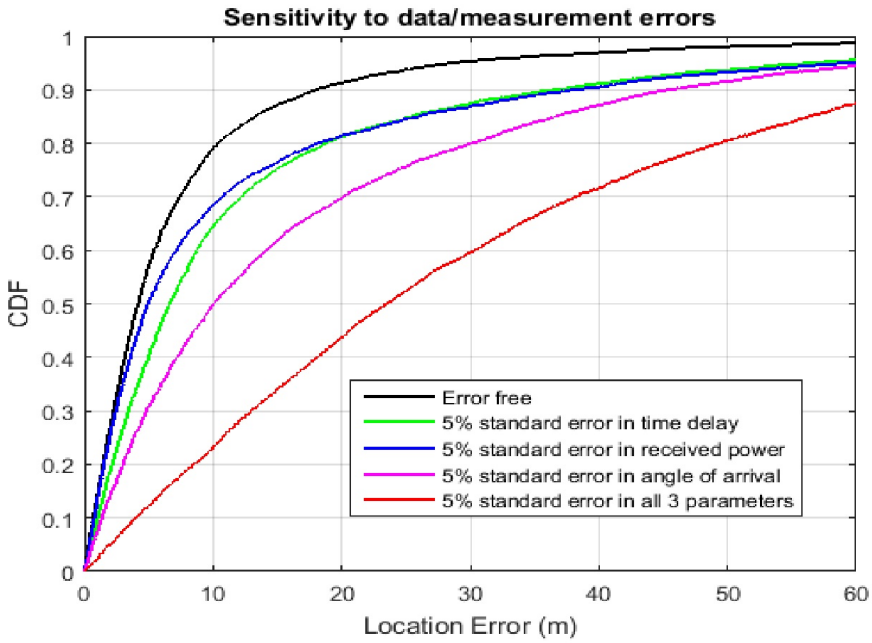


Figure 4.12: Sensitivity to measurement errors.

4.5 Comparison between NLOS mitigation and the Direct approach

The Direct method is a location specific scheme in the sense that its performance will depend on the environment. However, similar environments, in terms of their multipath, are expected to produce similar or comparable localisation accuracy performance. Training data must be generated for each new area the same was as in the location-specific approach for NLOS identification.

It is demonstrated in section 4.3 that the TDOA and TOA-AOA localization methods benefit from NLOS mitigation. The process of NLOS mitigation in these cases involve running the LSSVM twice, once for classification and then secondly for regression during mitigation. The Direct method also involves running the LSSVM twice, once to obtain the x-coordinate, and second to obtain the y-coordinate, meaning these two schemes' complexity, as far as the LSSVM processing is concerned, is comparable. However, it should be noted that the NLOS identification and mitigation scheme additionally involve running the TOA-AOA and TDOA localization algorithms themselves, after mitigation is applied. It is of interest to compare the performance of these approaches in different scenarios.

Comparison of the NLOS identification and mitigation scheme with the Direct method is presented in this section. TOA-AOA and TDOA localisation algorithms are used to compare localisation performance in the city center area and a predominantly LOS area. Comparisons are performed using the same dataset in-order to evaluate the schemes under the same framework.

TDOA requires each MS position to be able to see at least 3 BSs. MS positions that could not see at least 3 BSs were excluded. TOA-AOA and the Direct method uses only one BS per MS position, so for each MS, TDOA selects 3 BSs using prioritization described in Chapter 3 section 3.2 while TOA-AOA and the Direct method choose the closest of the 3 TDOA BSs (shortest delay) for localisation.

4.5.1 Direct method vs TOA-AOA after mitigation

The comparison of localisation performance between the TOA-AOA and the Direct method is shown in Figure 4.13, and it shows that the TOA-AOA method performs better within the 80th percentile level. TOA-AOA performs better below the 80th percentile because by its design, it

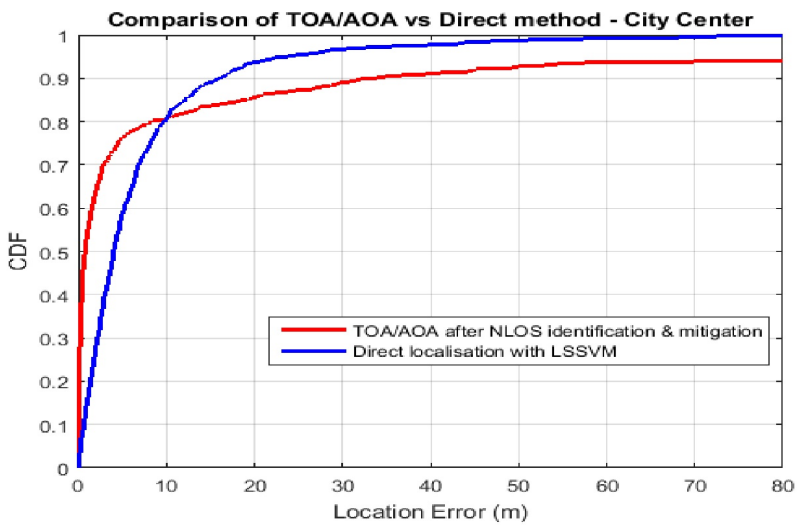


Figure 4.13: Direct method vs TOA-AOA.

uses LOS paths and only requires a single LOS path per MS position. Where there is no LOS MPCs between the BS and the MS, the algorithm selects ground reflected (GR) paths as outlined in the prioritisation scheme discussed in Chapter 3 section 3.2. It can be noted that TOA-AOA performance deteriorates if probability above the 80th percentile is considered. This is mainly due to the percentage of MS positions that do not have LOS paths to the BS. Use of ground reflected paths or any other mitigated NLOS paths results in growing position errors. When the correct AOA is obtained in a LOS link, the source of error then becomes, mainly, the time delay error, which produces the range between the BS and MS. Measurement errors and the positive bias of this delay can be reduced by using the mitigation technique discussed in section 4.3 and thus the localisation performance can be improved. In a LOS environment it was determined that TOA-AOA performs better than the Direct method at any percentile level.

4.5.2 Direct method vs TDOA after mitigation

The LSSVM direct localisation method performs better than TDOA in the dense urban environment as shown in Figure 4.14, below. This is because it is benefiting from multipath in such an environment as discussed in section 4.4.4. Also, TDOA depends on the performance of NLOS identification and mitigation, so in a case where there are insufficient LOS paths, the time delays used may result in a significantly over-estimated ranges.

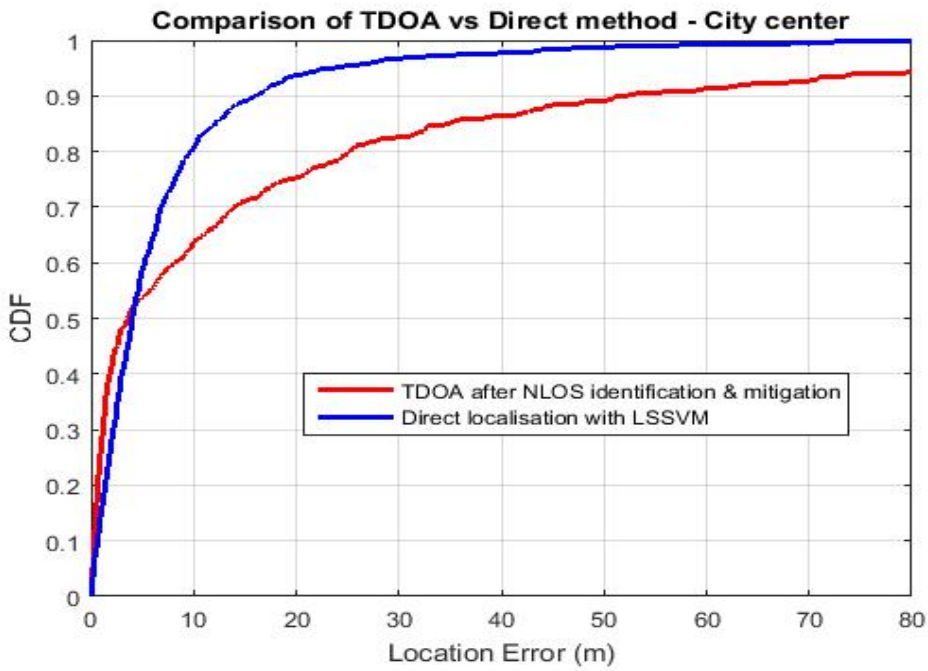


Figure 4.14: TDOA vs Direct method in an urban environment.

TDOA performance is comparable or better to the Direct method below the 50th percentile, which correspond to the proportion on LOS and mitigated GR rays. This demonstrates that TDOA mitigation using LSSVMs is more effective for those rays that originally presented small range errors but as the range error get large (pure NLOS cases), the LSSVM mitigation fails to improve localisation. This is also evident when one analyses Figure 4.4 where mitigation with training data 2 (TD2) improves range errors that were originally small.

An additional experiment was conducted for the open areas, which represents a largely LOS environment to compare how TDOA performs against the Direct method using the data that was generated for the open area considered in section 4.4.2. Figure 4.15, below, shows that in open areas, TDOA produces better location accuracy considering the established 80th percentile level. For this established standard, TDOA performs better over the Direct method, by approximately 5m. This is because of the availability and quality of LOS paths. TDOA performance becomes comparable to the Direct method as some NLOS paths start to be included in the algorithm due

to unavailability of sufficient LOS paths. Also, some MS positions may suffer from geometric dilution of precision (GDOP) when the 3 BSs chosen are in an undesirable geometry, such as in a straight line.

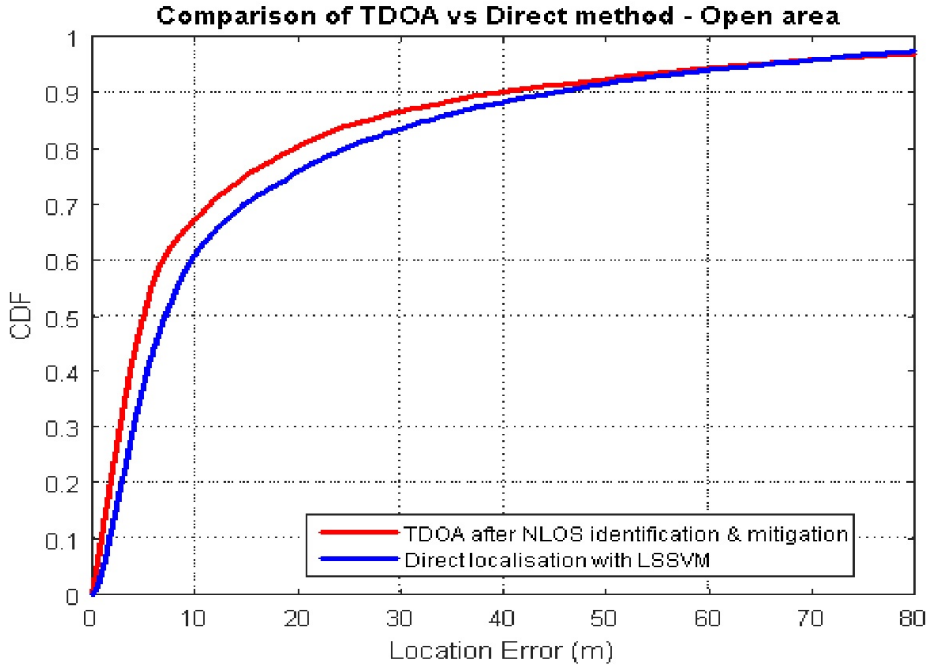


Figure 4.15: TDOA vs Direct method in a LOS environment.

4.5.3 Discussion and conclusions

The results shown above are location specific. Different performance could be expected with a change in location, but the MS positions considered in this study are in the order of thousands, so it could be expected to produce similar results in similar environments. To that end, an experiment was also conducted using data from a separate area of the city, for the NLOS environment, and a different open area, for the LOS environment, to evaluate the performance of the Direct method in similar environments, and similar CDF curves were observed. These are shown in Appendix B. This section has demonstrated an approach to urban localisation using ray-traced data and LSSVMs. It demonstrates that the direct localisation approach provides better localisation accuracy under NLOS conditions, compared to the process of NLOS identification and mitigation and then exploiting the traditional localisation algorithms like TDOA and TOA. Because the direct approach is essentially a single BS localisation scheme, it has been demonstrated that AOA errors can greatly affect the location accuracy. In such circumstances, multi BS schemes like TDOA may be able to handle errors in one BS better. Granularity and performance of the Direct method can be further controlled by the size of training data. A larger training data size improves the tuning parameters. Because training can be done per each BS with the tuning parameters

referenced per each BS. More BSs can be used to obtain estimates for an MS position, and if each estimate can be assigned some confidence value, the use of multiple BSs can improve accuracy. However, running training for each BS separately, and then running estimation for just one MS, takes time, considering the total of 36 BS and thousands of mobile positions considered. The sensitivity of the approach with mobile scatterers is not currently known and requires further study. There are possibilities to mitigate the effects of mobile scatters by employing Doppler Shift processing techniques [149] [150] [151]. These and similar techniques are used in radar processing to remove clutter, and can be used to remove multipath components that are resulting from moving scatterers. If the main mobile scatterers are assumed to be vehicles, increasing the height of BSs may improve the availability of LOS components. Even when the MS is mobile, the multipath components will be as a result of fixed scatterers like buildings and other street infrastructure, which are taken care of within the ray-tracing based frameworks discussed, or via other map-based techniques [145]. Densification through deployment of numerous micro BSs on street lamp posts may actually mean the probability of getting a LOS component increases. Further accuracy for mobility or tracking scenarios can be improved by data fusion with the MS's inertial measurements. This is useful even for MS's with GPS but which are in areas where no GPS signal can reach as discussed in Chapter 1.

4.6 Summary

This chapter has demonstrated that optimisation-based machine leaning approaches, specifically, the LSSVM can be used for NLOS identification and mitigation in outdoor urban environments. The chapter has also demonstrated that ray-tracing data can be exploited together with LSSVMs to provide direct localisation in radio cellular systems. These approaches can contribute immensely to mobile network-based localisation strategies, which in-turn, can be critical to 5G systems, where geolocation information can be exploited to benefit various subsystems. The use of ray-tracing tools, together with machine leaning algorithms for geolocation purposes is first demonstrated in this research. A key outcome from the first section is the performance of the location specific approach for NLOS identification. A demonstrated over 98% accuracy means for any similar city with a validated ray-traced database, NLOS identification to that level of accuracy can be expected. This level of accuracy has not been seen in literature for a city outdoor environment. Section 4.4 has presented a localisation scheme that does not depend on NLOS identification. The results have shown that the scheme performs best in dense multipath environments. The scheme is dependent on the availability of DOA measurements. The technique is more sensitive to DOA errors than received power level or time delay errors. In environments where the average azimuth spread is small (see Table 4.6), such as in LOS environments, there could be only one or a few dominant rays and if these rays have similar DOA (or small difference between their azimuth AOA measurements), then the impact of AOA as a parameter offering

an additional degree of freedom is lessened. The direct scheme estimates the two coordinates of the mobile in two separate processes which matches the other discussed schemes (Location independent and location specific approaches) that run NLOS identification first, followed by range mitigation, in terms of computation. Although the location specific approach and the direct approach make use of existing algorithms, their application frameworks together with the use of ray-traced data is a new approach to tackling challenges related to device localisation. LSSVM runtime for NLOS identification (classification) or for mitigation (regression) depends on the number of base stations and mobile stations used for training (the number of data points used). In the study presented in Table 4.1, it is evident that a datapoint size of 5000 is adequate to produce a greater that 90% classification accuracy accuracy. This corresponds to a training run of approximately 5 minutes using a dual core computer. The author notes that with advances in edge computing as discussed in Chapter 6, the computational hurdle will not be an issue for future real time processing. Also if a train once approach is considered, where training parameters are reused until the environment is deemed to have changed, then the computational hurdle for training becomes only a periodical issue which can be scheduled.

CHAPTER 5

EXPERIMENTATION WITH A MASSIVE MIMO TESTBED

It has already been discussed and illustrated from Chapter 1 through to Chapter 3 that localisation in next-generation wireless systems benefits from use of multi-antennas at the base station. The case for Massive MIMO is demonstrated by showing how direction of arrival (DOA) accuracy increases with the number of antennas for both subspace-based and parametric estimation algorithms. DOA estimation is key to geometric localisation approaches like triangulation in AOA. This chapter takes further the proposed DOA estimation algorithm introduced in Chapter 3, and applies this to the massive MIMO testbed described in Appendix C. The localisation approaches that utilise the DOA estimates are presented in Chapter 2 section 2.5, which outline the use cases of the estimates for the ultimate localisation of the mobile station or user equipment. The assumption considered when one moves from DOA estimate to position fixing, is also outlined. This is followed by a discussion of how the Alternating Projection Maximum-Likelihood (APML) algorithm (Discussed in Chapter 3, section 3.4) is implemented with the massive MIMO testbed, showing how the parameters of the components of the system are utilised in the algorithm. Relevant assumptions that simplify the solution are also presented. It is important to note that Chapters 3 and 4 outcomes highlight the importance of angular information, both for the AOA based algorithms and for NLOS identification plus mitigation. It was also noted that the LSSVM direct localisation scheme is sensitive to AOA errors. These observations tie in with the focus of this chapter which is DOA estimation. DOA estimation is therefore an intermediate step towards localisation in both cases, i.e. either using traditional techniques like AOA , TOA-AOA, or machine learning approaches as discussed in Chapter 4.

The rest of the chapter is then devoted to the practical experimentation which carries details on the 4 measurement campaigns and their DOA estimation results. Each campaign carries detail on experimental setup and methodology, together with challenges and limitations, as well

as the results and conclusions. The key outcomes are summarised at the end of the chapter. The terms MS and UE may be used interchangeably depending on the technologies being discussed in this thesis, but the main convention in this thesis is to use the term MS in the context of ray-tracing based experiments mainly used in Chapters 3 and 4, and the term UE, which is synonymous in multi-user communications like LTE and multi-user MIMO, is used in Chapter 5.

5.1 Array manifold measurement and calibration

Systems that use antenna arrays for direction finding or beamforming rely heavily on the array manifold [152] [153]. As discussed in Chapter 2 section 2.4 and also 2.2.6, the parametric array processing achieves super-resolution because the techniques are not constrained by the Rayleigh resolution limit, as seen with non-parametric techniques. But achieving super-resolution requires accurate knowledge of the array manifold. A small difference between the true manifold and the assumed manifold, as used in the processing, can significantly reduce the super-resolution capability and thus degrading the performance of the direction of arrival estimation.

The array manifold [154] gives the general response of the array, which is dependent on the array geometry. An approach, to build the array manifold from anechoic chamber measurements was devised as detailed below. The array manifold determined in this way will be relevant for LOS scenarios, which most of the measurements in Chapter 5 considered.

To get the response of each and every patch antenna on the array, in the anechoic chamber, was not feasible, because the array with all the 128 elements/patches, is bulky, and could not be mounted in the chamber. Also taking vertical and horizontal polarised measurements for all the antennas becomes a lengthy process. Since the 32 element sub-arrays that make up the full array are separable, it was possible to measure the patches on just one 32-element sub-array. The measurements could then be applied for the rest of corresponding elements, by employing antenna pattern transformations and rotations. This approach assumes that these sub arrays, and the individual patches were built with a degree of uniformity and precision.

In the anechoic chamber, the 32-element array was set at a distance of 8m from a single transmitter, using 3.51GHz to match the carrier frequency used on the testbed. A single patch antenna was connected to the receiver while the rest had 50Ω terminations for both H and V ports. Also for each patch, both horizontal and vertical polarisation measurements were taken. Figure 5.1 shows the 32 element sub-array mounted on the rotating stand in the anechoic chamber.

In all the experiments described in this thesis, the full array was comprised of the 4 sub arrays joined together to form a 4x32 single array. The response of the antennas at the joints is also critical as it is expected to be different from either any edge or centre elements. To capture this effect, it was deemed necessary to measure the antennas at the joints while 2 sub panels are joined together. These measurements can then be applied for the remaining 2 joints using the mentioned uniformity assumption. Figure 5.2 shows the back of the joint of 2 sub-panels during

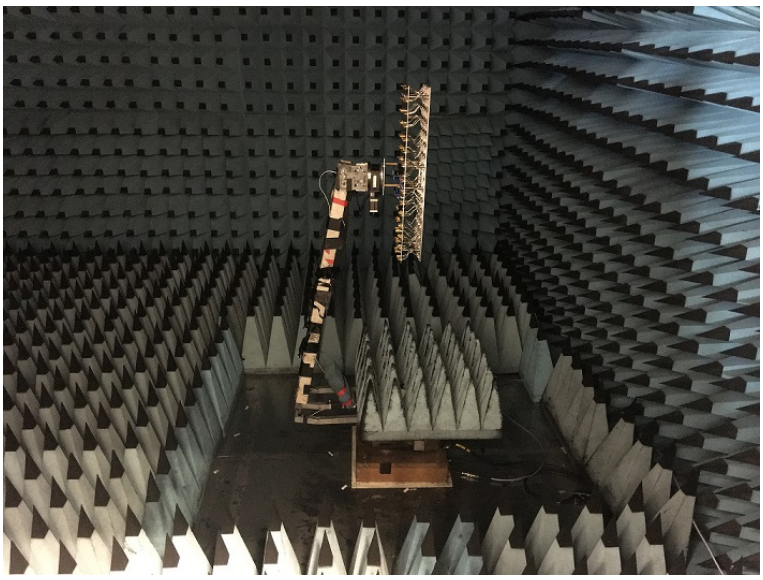


Figure 5.1: 32-element sub-array in the anechoic chamber

anechoic measurements.

Figure 5.2: Backview of the joint of 2 sub-panels

This uniformity assumption was tested for patches on the same sub-panel. Corner patches and edge patches were compared after relevant pattern transformations, and the correlation was calculated. Figure 5.3 shows the comparison between antenna patches 8 and 32 (top left corner and bottom left corner patches) in the vertical polarisation. The correlation coefficient was calculated to be 0.96. This supported the assumption that measurements for certain patches can be used for corresponding patches within the full array. Additional patterns and measurements that characterise the array are included in Appendix D.

The experiments described in this thesis were conducted in a vertical polarisation configura-

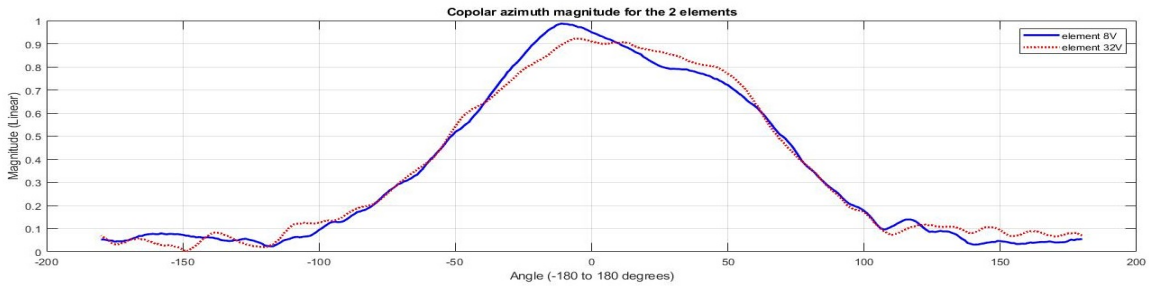


Figure 5.3: Azimuth pattern for top left and bottom left corner patches

tion. The manifold was therefore synthesized with vertical polarisation measurements at a fixed elevation angle of 90 degrees. This is the angle when the array is directly facing the transmitter. This was done to simplify the processing because the direction of arrival algorithm considered is 2D and only considers the azimuth directions. Although this can be extended for 3D scenarios, it was important to prove the system using simple configurations before complex scenarios can be considered. In building the manifold matrix, only the azimuth angles from -90°to +90°were considered because of the rectangular array. From Figure 5.3 above, it is clear that beyond this region, the array does not receive any useful signal.

From the signal model discussed in Chapter 3, Section 3.3.1, it can be noted that for a single signal source, and taking L snapshots, the estimated correlation matrix would be as shown in equation (3.4) with vector $\boldsymbol{Y_i}$ being a 128x1 snapshot vector. By analysing equations (3.11) to (3.16) it can be noted that only the first eigenvector $\boldsymbol{q_1}$, span the signal space. This becomes the principal eigenvector and in this case there is a direct relation between $\boldsymbol{q_1}$ and $\boldsymbol{S}(\theta)$. $\boldsymbol{S}(\theta)$ can therefore be estimated by $\boldsymbol{q_1}$ for calibration, and this can be repeated for every degree step capture obtained from the anechoic chamber measurement, from $-90\,\mathrm{deg}\,to + 90\,\mathrm{deg}$. This approach assumes uncorrelated noise, but because in reality, the noise may be correlated, this adds to the errors in accurately measuring the array manifold. After noting that a constant gain and phase response across the array is not possible due to imperfect manufacturing, impedance matching and also the cables connecting the antenna array to the transceivers, it can also be noted that another effect that introduces errors in the array response, is antenna coupling [155]. This is due to the radiation between adjacent or nearby antennas as well as coupling with a common feed network. There could also be interactions between the antennas and its surrounding objects. For an antenna array, individual elements also re-radiates the received energy [156] and the amount of re-radiation scattering is dependent on the impedance matching.

For an arbitrary planar array with M elements, the array error matrix can be defined as [157]

$$\Gamma = diag(\alpha_1 exp(jv_1), \alpha_2 exp(jv_2) ... \alpha_M exp(jv_M)) \tag{5.1}$$

which represents the array gain/phase error matrix with $\Gamma \in \mathbb{C}^{M \times M}$, where α_m is the normalised gain factor and v_m is the phase shift. If the effects of the array support structures are neglected,

the antennae coupling of the array can be defined as the energy interchange between any adjacent antenna elements. This coupling is defined by Eberhardt [157] as $C \in \mathbb{C}^{M \times M}$ with

$$[C]_{ij} = \begin{cases} 1 & i = j \\ c_{ij} & i \neq j \end{cases} \quad with c_{ij} = c_{ji} \tag{5.2}$$

The data model used in [158] which takes into consideration, the sensor error matrix and the antennae coupling can be expressed as

$$\boldsymbol{Y} = \boldsymbol{\Gamma}.\boldsymbol{C}.\boldsymbol{S}.\boldsymbol{X} + \boldsymbol{N} \tag{5.3}$$

The error matrix includes the gain and the phase errors for the array. In other calibration approaches, antenna position errors are considered, but in this work, they are neglected because they are small as compared to other errors [157] but such approaches are discussed in [159], [160] and [161].

To simplify the problem, the antenna coupling and the array error matrix are combined together into a single matrix Ξ which is used in the standard signal model described in Chapter 3, and which leads to a simplified model that can be defined as

$$\boldsymbol{Y} = \boldsymbol{\Xi}.\boldsymbol{S}.\boldsymbol{X} + \boldsymbol{N} \tag{5.4}$$

In addition to the differences in the response of the physical antennas, the differences in the response of the radio receivers will impact AOA accuracy. Calibration schemes have to compensate for the hardware responses in both the UE and BS radio chains. An evaluation of these effects within the testbed radio chains is discussed in Appendix D. Measurement errors caused by the receivers can be addressed by periodic calibration as discussed by Eberhardt in his works [157] [162]. Similar scheme which uses reciprocity calibration is discussed below in section 5.1.1 and having compounded the error matrix from coupling and other phenomenon into an error matrix Ξ as shown in (5.4), reciprocity calibration can be used to address such errors.

5.1.1 Reciprocity calibration

A Massive MIMO TDD system assumes a reciprocal channel between the user and the BS, in order to transmit the users' data via the same channel that is seen during uplink pilot [15]. Equipment imperfections disrupt that reciprocity, so reciprocity calibration seeks to factor out the different responses of these receive and transmit chains. Whilst the electromagnetic propagation channel can be considered to be reciprocal [163], the RF chains on the UE and the BS are generally not reciprocal [164], so reciprocity calibration seeks to calculate the downlink precoding coefficients that can factor out the non-reciprocal transceiver responses, thereby allowing the use of the reciprocity assumption for the testbed. The procedure for reciprocity calibration involves estimating the calibration coefficients first, then applying the coefficients to the channel estimates for the uplink.

Research on reciprocity calibration has been gaining traction in recent years, owing to the possibility of using TDD based designs in 5G. Common approaches for calibration, involve the use of dedicated circuitry [165], but can be very complex for a massive MIMO system with hundreds of antennas. Other over-the-air techniques make use of some of the UEs in the environment for calibration purposes [164]. The challenges with these approaches gave rise to research on new techniques that utilise the antennas at the BS only [166] [167] [168] [169]. These techniques make use of channel measurements between selected reference antennas and the rest of the other antennas on the BS. [169] specifically presents a calibration mechanism based on mutual coupling between the antennas. Figure 5.4 below shows how the reciprocity calibration's estimation step, is conducted.

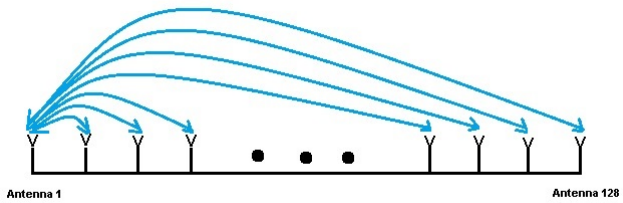


Figure 5.4: Reciprocity calibration using the first antenna

The following section presents the estimation for reciprocity calibration coefficients, and how they are applied for direction of arrival estimation.

An N antenna base station serving K users has to acquire the uplink channel state information (CSI), $\hat{h}_{k\to n}$, for all $n = 1,2,\ldots,N$ and $k = 1,2,\ldots,K$, from uplink pilot signals. Because of the above mentioned non-reciprocal nature of the user and BS transceiver RF chains, the main challenge is the estimation of the downlink reciprocal CSI $\hat{h}_{n\to k}$ from that of the uplink. The non-reciprocity which may be a result of random phase and/or amplitude errors from the RF hardware, may also be due to a combination of dynamic effects that may emanate from internal clocking hardware like multipliers, dividers and phase locked loops (PLLs). These effects are also in addition to static effects caused by manufacturing defects. This means that the channel between two transceivers will be a product of the frequency response of the transmitting chain (tx), the actual wireless channel (h) and the frequency response of the receiving chain (rx). For a given pair of radios, i and j, the estimated channel between them can be written as,

$$\hat{h}_{i\to j} = tx_i \cdot h_{i\to j} \cdot rx_j \tag{5.5}$$

Estimation of the reciprocal channel $\hat{h}_{j\to i}$, is done by first defining the reciprocity coefficients, $b_{i\to j}$, as,

$$b_{i\to j} = \frac{\hat{h}_{i\to j}}{\hat{h}_{j\to i}} = \frac{tx_i \cdot h_{i\to j} \cdot rx_j}{rx_i \cdot h_{j\to i} \cdot tx_j} = \frac{tx_i \cdot rx_j}{rx_i \cdot tx_j} = \frac{1}{b_{j\to i}} \tag{5.6}$$

Due to the generally accepted physical channel reciprocity, if both uplink and downlink channels are captured within the channel coherence time, then $h_{i\to j} = h_{j\to i}$. If the calibration coefficients are known, the downlink channel can be estimated from the uplink channel by, $\hat{h}_{j\to i} = \hat{h}_{i\to j}/b_{i\to j}$ Figure 5.5 shows a representation of the channel components between two radios.

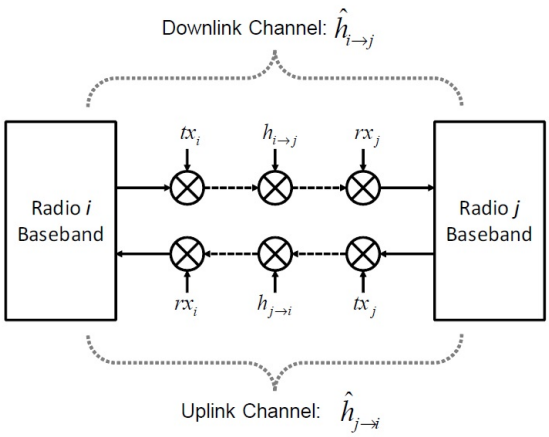


Figure 5.5: Uplink and downlink channel components, $\hat{h}$ represents estimated channel. Dashed lines depict part of the channel that is wireless and $h_{i\to j} = h_{j\to i}$ [167]

For a user k transmitting an uplink pilot to the BS with N antennas, $\hat{h}_{k\to n}$ represents the estimated uplink channel, while $\hat{h}_{n\to k}$, represents the downlink channel to estimated using N reciprocity calibration coefficients, which are $b_{n\to k}$, for $n = 1,2,\ldots,N$. But obtaining $b_{n\to k}$ coefficients requires transmission between every BS antenna n and every user k, together with the feedback from both. Furthermore, because the UEs do not share clocks with the BS, there will be drift over time so frequent calibration will be needed. This is what drives the research into internal calibration schemes, those that use the BS antennas only.

As demonstrated in Figure 5.4 above, this internal calibration schemes involves a reference antenna, say antenna 1, transmitting to all the other antennas. The desired calibration coefficients in this case become $b_{n\to 1}$, for $n = 2,3,\ldots,N$. It can be noted that these coefficients remain stable over long period of time because all the antennas on the BS share the same clock. The calibration coefficients between any two radios, can be obtain if the coefficients between each of them, and a reference radio, is known, as demonstrated by,

$$\frac{b_{i\to j}}{b_{i\to y}} = \frac{\frac{tx_i \cdot rx_j}{rx_i \cdot tx_j}}{\frac{tx_i \cdot rx_y}{rx_i \cdot tx_y}} = \frac{tx_y \cdot rx_j}{rx_y \cdot tx_j} = b_{y\to j} \tag{5.7}$$

So the down link channel $\hat{h}_{n\to k}$ can be found following Equation 5.6 by, $\hat{h}_{n\to k} = b_{n\to k} \cdot \hat{h}_{k\to n}$, where $b_{n\to k}$ can be obtained following Equation 5.7 above as, $b_{n\to k} = b_{1\to k}/b_{1\to n}$, which suggests that the full channel state information can be obtained simply by transmitting a single pilot from each UE, and also just a single pilot from the BS's reference antenna. But this requires sending

feedback for the channel estimate, $\hat{h}_{1\to k}$, of the downlink from the reference antenna, for each of the k UEs, a situations that decreases the channel capacity. Simplification of this approach sets the coefficients of the reference antenna, $b_{1\to k}$, to 1, for all k, the key observation being that if all the BS antennas are affected by the same phase error, then the resultant spatial beam-pattern, still remains unchanged. The CSI estimation for each BS antenna deviates from its actual CSI, by a scaling factor common to all BS antennas, hence the multi-user beamforming for massive MIMO should produce the same spatial beam-pattern. For the purposes of direction of arrival estimation, the beamscan technique used here, should also find the same beam-pattern. The resulting relative channel, which will be used for downlink, is calculates as,

$$\hat{h}'_{n\to k} = \frac{\hat{h}_{n\to k}}{b_{1\to n}} \tag{5.8}$$

The complete process of acquiring and calibrating the channel data for direction of arrival estimation involves the 3 steps of: (i) determining all the internal calibration coefficients, $b_{1\to n}$, by sending pilots between the BS reference antennas and every other BS antenna, (ii) estimating the uplink channel $\hat{h}_{k\to n}$, by sending orthogonal pilots from the k UEs to the BS antennas, and (iii) determining $\hat{h}'_{n\to k}$ by using (i) and (ii) and using it for beamscan in the APML algorithm. This calibration scheme was shown to produce average angle deviation from true angle, of less than 2.6% using real-world measurements as shown in Figure 5.6 [167].

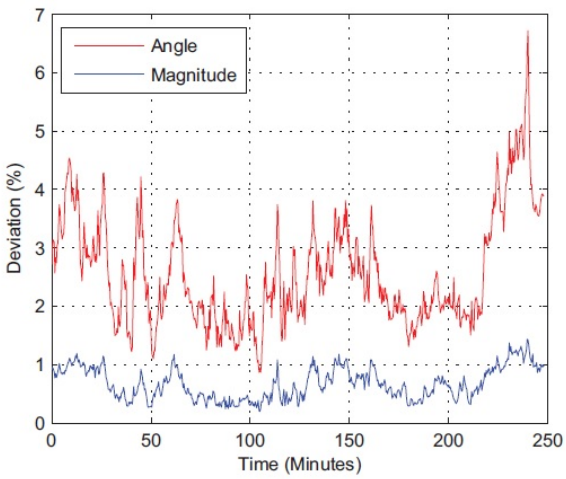


Figure 5.6: Reciprocity calibration scheme exhibits an average angle deviation of less than 2.6% (maximum 6.7%), and mean amplitude deviation of less than 0.7% (maximum 1.4%) [167]

The process of obtaining the calibration coefficients $b_{1\to n}$, involved running the testbed's BS for at least 50 minutes. This was done to allow the components to reach their operating temperatures, so that the temperature effect could be minimised. Experiments were conducted to determine the time it takes for the coefficients to remain stable, by checking these coefficients over time from the start of the BS. It was determined that these remain fairly stable after 40 minutes.

The calibration coefficients used were recorded in the middle of the measurements and also at the end of the measurements. Any of these two measurements can be used, or an average can be calculated. One antenna was used as a reference each time and a total of 34 antennas were used as references for a total of 68 data captures for reciprocity calibration.The reciprocity calibration is applied to the system as shown in Figure 5.7 below. Because the direction of arrival estimation processing was conducted offline, the Massive MIMO framework's base station LabVIEW code was modified to extract the calibration data.

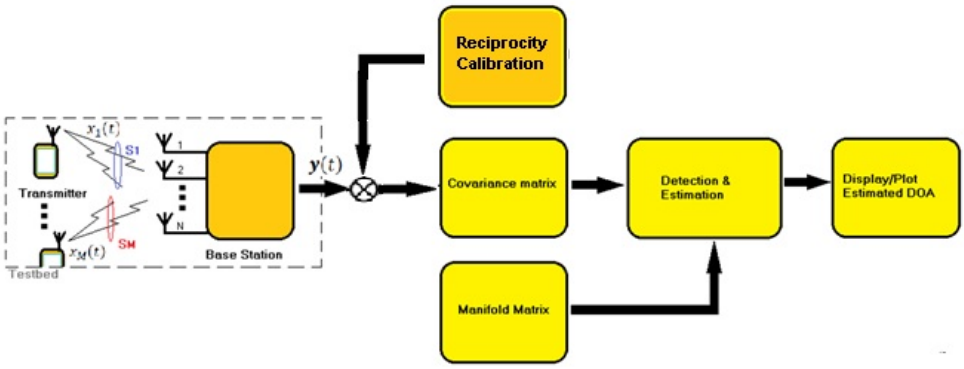


Figure 5.7: Complete system with reciprocity calibration

5.2 APML algorithm implementation

The techniques discussed above all describe how DOA estimates can be used to achieve localisation of the mobile station. The challenge of determining the DOA for signals arriving at the BS is discussed in Chapter 3, together with the proposed APML algorithm in section 3.4. This section provides a description of how the APML and the massive MIMO testbed outlined in Appendix C, come together to provide a practical DOA estimation system. The setup of the equipment is described in Appendix C section C.1. The experiments are designed in a way that allows channel captures to be collected during a measurement campaign, with the processing being done offline. This approach is acceptable for 2 reasons. Firstly, it allows multiple channel captures for a large number of scenarios and positions without requiring any equipment reconfiguration, meaning multiple experiments could be performed at once. Secondly, this removed the challenge of real-time processing for the ML algorithm used. It is highlighted in Chapter 2 that a major drawback of the ML algorithms is the complexity because of the exhaustive search, so offline processing removes the limitation that could otherwise be present when real-time tracking is considered. The issue of complexity is discussed in Chapter 2 and it can be noted that there are other algorithms that could be used that have lower complexity [31] [170] [51] but have their own practical constraints as discussed in section 2.4.2. APML algorithm is compared to

the MUSIC algorithm in a simulation in Chapter 3 Figure 3.17. The design and limitations of the testbed, together with the calibration error sensitivity of subspace methods like MUSIC, both prescribed the decision on the algorithm and parameters that were considered for the experiments in this chapter. The author was responsible for calibration measurements, pattern measurements in the anechoic chamber and processing of the raw data into relevant parameters using MATLAB as well as taking into account the particular configuration on the testbed on each given measurement campaign. He was also responsible for the offline implementation of the algorithm and the subsequent analysis.

A discussion of the initial estimation of the number of co-channel sources is provided in Chapter 3. It is important to note that in the application of the APML algorithm with the testbed, an assumption is made that the highest peak in the initial DOA scan corresponds to the LOS angle and subsequent scans are not utilised, although in all cases an initial value of the number of sources (M) is set to 4 for the purposes of analysing the spread of the first 4 angles, which may provide an indication of the multipath environment. It is also worth noting that while this assumption simplifies the estimation process, it also leads to significant errors if the environment has considerable multipath. The experiments presented in the succeeding sections were all done in presumed LOS environments, so it is believed that the assumption holds for these environments.

Each measurement campaign involved moving the equipment on to the site, site logistics, cabling and testing out the system. Initial experimentation for DOA estimation with the testbed involved DOA processing of channel captures without any calibration to check if the algorithm could benefit from the large number of antennas to provide some DOA accuracy without any calibration. This was conducted during early Massive MIMO trials that are described in [171]. Another preliminary DOA experimentation was conducted during a Massive MIMO trial in collaboration with British Telecom, as discussed in Appendix E. These preliminary experiments did not produce useful accuracy. Subsequent experiments that are described in this chapter utilised a measured manifold obtained as described in section 5.1. In total, 3 measurement campaigns are presented in this chapter.

In all the experiments from each measurement campaign, the positions of the base station and every position where the UE (s) was placed, were recorded. This provided a way of generating the dataset of the actual position and angles of the UE(s) to the BS, assuming a LOS propagation. All experiments assumed LOS propagation, so the calculation of the DOA error was based on the deviation from the recorded LOS angle of the UE to the centre of the BS array. Where subarrays where used, the angle considered was that of the UE to the centre of the group of antenna elements that were used. The considered spatial aperture for the rectangular array was approximately 3dB beamwidth of each patch in order to exclude 30 degrees on each side of the antenna patterns, because in these regions, the main beam broadens and sensitivity and control is lost. These regions can be noticed in the amplitude patterns in Figure 5.3 and also in

the patterns presented in Appendix D.

5.3 Wills Memorial Hall measurement campaign

An DOA experiment was setup in the Wills Memorial Hall of the University of Bristol. This was a LOS propagation environment. The floor is a hard concrete one and the ceiling is approximately 15m high. The walls are wood panelled half-height and the rest is plastered. The BS was setup on the floor near the stage. The BS setup process involved meticulously cabling the antenna array, which stood on an adjustable 2-leg aluminium stand, to the transceivers on the collocated racks. Figure 5.8 and 5.9 show the BS setup and the schematic of the environment respectively.

Figure 5.8: Picture of the BS setup in the Wills Memorial Hall.

OTA synchronisation was used, and a separate antenna was used to transmit the synchronisation signal. A bespoke calibration scheme was employed, which involved the use of 2 UEs with one UE acting as the target for DOA estimation, while the other remained fixed on a particular known position (denoted with C in Figure 5.11, to act as the calibration source. The 2 UEs were set at a height of 1.2m, which was sufficient to be above the chairs.

Initially, the 2 UEs are setup at the same position, which was about 10m directly in front of the centre of the BS array. The 2 UEs transmit uplink pilots and the channel impulse response were captured at the BS. The angle of this initial position was recorded. The target UE was then moved to predetermined and recorded positions within the hall, and each time on a specific position, the channel captures were recorded both for the target UE and the calibration UE. The gain and phase changes that were seen on the calibrating UE were exploited as calibration data for the target UE. Figure 5.10 shows the target UE in the hall. Figure 5.11 shows the schematic of the positions of the target UE and the calibrating UE. The target UE was moved to positions

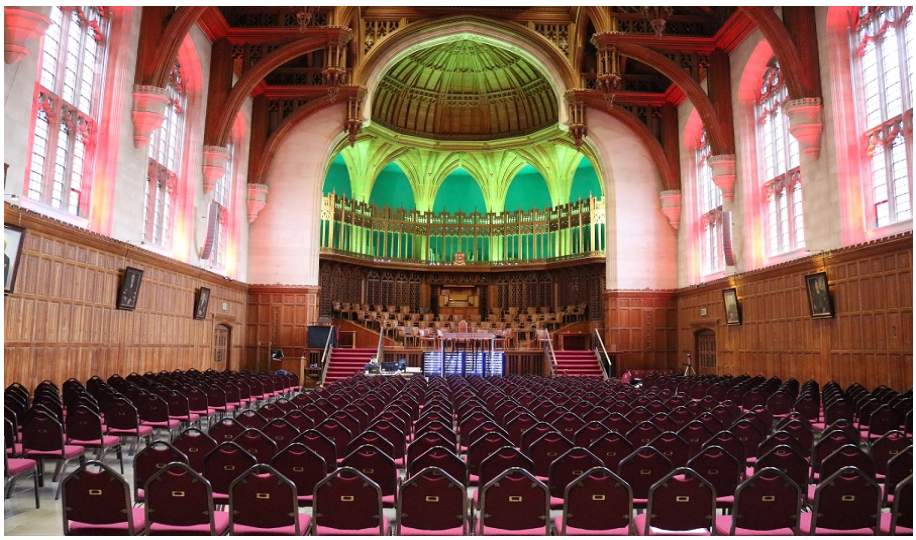


Figure 5.9: Picture of the Wills Memorial Hall environment.

Figure 5.10: User Equipment on a trolleys.

shown in Figure 5.11. The spread of the positions generally covered the bottom half of the hall as shown in Figure 5.11. Channel captures for the 2 UEs were post-processed in MATLAB and the calibration was applied to the target UE data before the APML algorithm was run and DOA estimates for each position were produced. The DOA error for each position was produced by comparing the actual angle to the estimated angle. The APML algorithm was employed in a way that assumes LOS propagation as discussed in Chapter 2 section 2.5. The DOA error was evaluated using data from all antennas (case 1), with data from 64 antennas that make up the 2 middle rows (case 2) and with 32 antennas that make the upper row of case 2 (case 3). The above configurations for case 2 and case 3 were chosen to exclude antenna patches at the top and

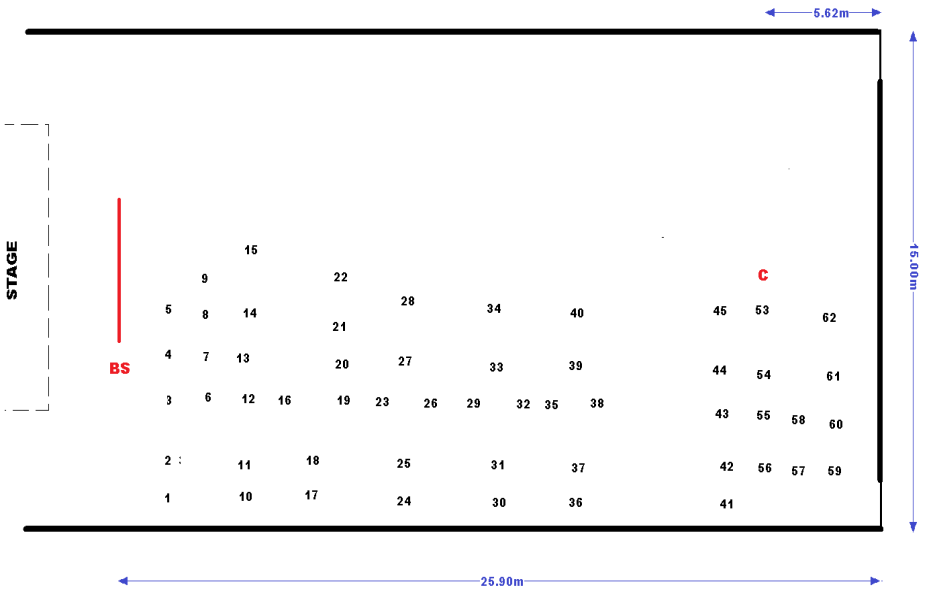
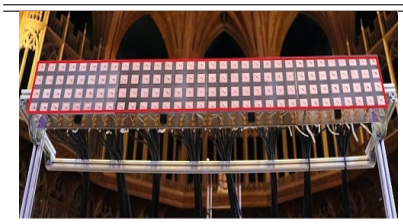
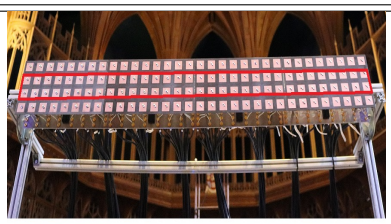
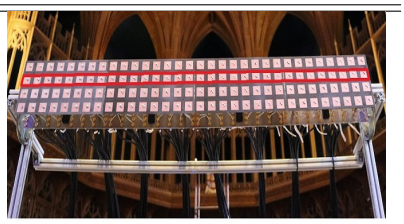


Figure 5.11: Schematic of the Wills Memorial Hall setup.

bottom edges of the array, which utilising a wider azimuth aperture by considering the antenna patches that span the whole array length. Table 5.1 shows the array configurations considered. The red lines show the bounds of the antenna elements that are selected.

Table 5.1: Array configurations considered (selected elements bounded by red lines)

Case 1 - 128 antennas	**Case 2 - 64 antennas**	**Case 3 - 32 antennas**
		

5.3.1 Results and discussion

The results for the positions considered in the study demonstrate as expected that increasing the number of antennas improves DOA estimation accuracy. Figure 5.12 shows the CDF curves of the DOA error for 32, 64 and 128 antennas for all the positions considered. If an arbitrary accuracy threshold of 10 degree is considered, it can be noted that case 1 (128 antennas) produced approximately 10% more positions that meet that threshold than case 2 (64 antennas). Similarly case 2 produces approximately 10% more positions that meet that threshold, than case 3 (32

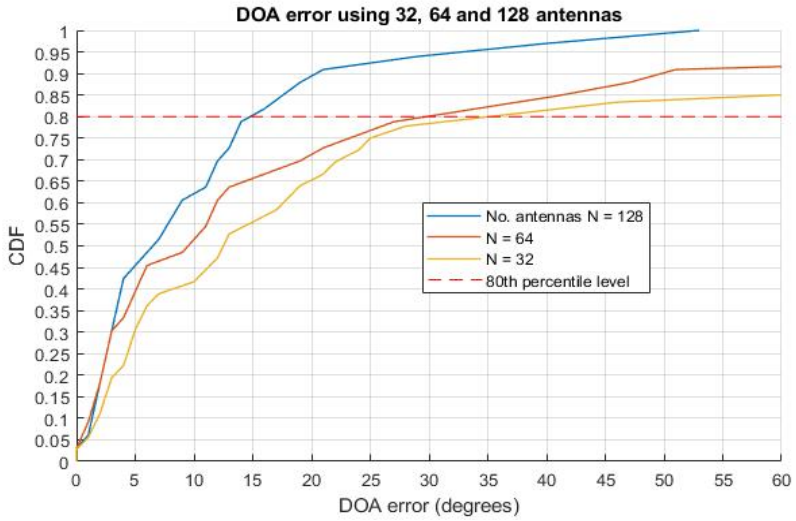


Figure 5.12: CDF of LOS angle error (degrees).

antennas). This shows that increasing the number of antennas improves the estimation accuracy, but that improvement is limited as a consequence of the calibration scheme used. Considering

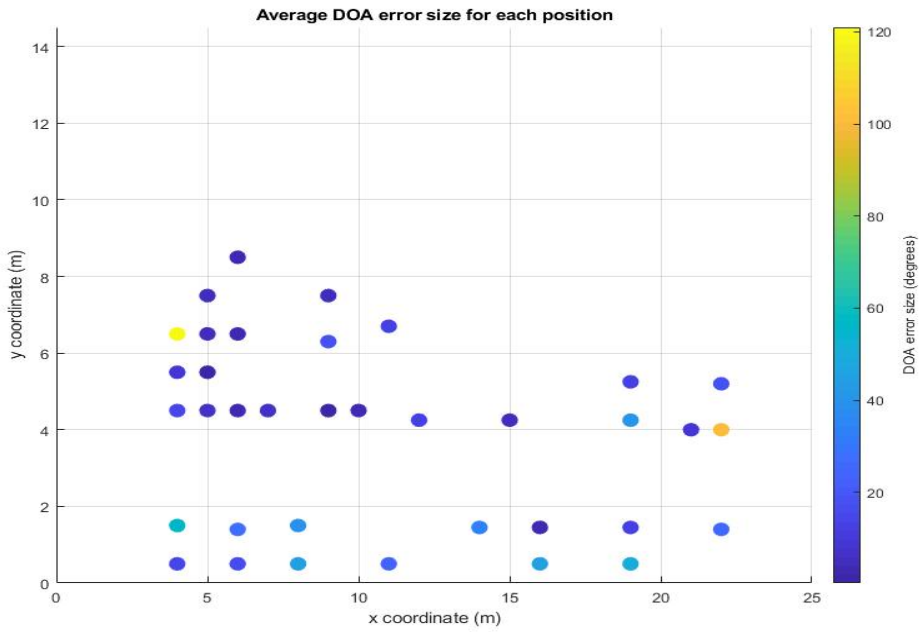


Figure 5.13: Scatter plot.

the 80th percentile level as the established standard in this thesis in Chapters 3 and 4, the best accuracy is obtained under case 1, which is 15 degrees. Cases 2 and 3 achieve approximately 30 and 35 degrees respectively for the prescribed percentile level. All the curves start to flatten after

DOA error bigger that 20 degrees. This suggests that there are some positions which resulted in large errors for all cases. Figure 5.13 shows a scatter plot of the average DOA error sizes corresponding to the studied positions in Figure 5.11. The average error for the 3 cases is plotted on each position in the 2D plane of the hall. It can be noted that the positions close to the bottom perimeter (left wall) of the hall recorded most of the largest DOA errors. This could be attributed to reflections off the wall. Position 5 shows an anomaly in all cases and it registered the largest DOA error. This suggests that the LOS propagation assumption did not hold for this position. Overall, the results could be improved by employing a different and better calibration scheme as becomes apparent in the succeeding measurement campaigns.

5.4 Millennium square measurement campaign

An outdoor experiment was conducted at the Millennium square in Bristol city centre. This experiment was organised to coincide with the first UK 5G Showcase [172] that was held over the same period at the same venue. Videos of the event can be viewed at [173] and [173]. This meant the equipment was being used for other 5G massive MIMO demonstration during the same period which ultimately led to a limitation in terms of UE placement for this particular DOA estimation experiment.

Figure 5.14: BS setup at Millennium Square. Left image shows the side view of the antenna array and right image shows the array on the balcony, from the ground.

The BS was setup on the balcony of the "We The Curious" building [174] on the northern side of the Millennium Square and it was set at a height of 12.7m from the ground. The antenna array was tilted down at an angle of approximately 15 degrees to the vertical, so that it could face the UEs whose antennas were mounted on the face of a temporary marquee at a height of 2m from the ground, and at a horizontal distance of about 40m from the BS. The UE height of

Figure 5.15: UE antennas on the side of the marquee.

2m was chosen so that the bodies of the dipole antennas can stand slightly above the marquee. The antennas were affixed to the plastic cover of the marquee but it should be noted that there was a thin metal tubing inside the marquee and although the antennas slightly stood above this inside tubing, some back-scattering effects would be expected in these circumstances. Figures 5.14 and 5.15 show the setup of the BS antenna array and the UE antennas in the Millennium Square environment.

Figure 5.16: UEs setup inside the marquee.

Figure 5.16 shows the UEs inside the marque with connections via sma cables to the antennas that are outside the marquee. Also Figure 5.17 below shows the relative positions of the BS and

Figure 5.17: BS and UEs setup on the Millennium Square.

the UEs in the Millennium Square environment. The environment was static at the time the channel captures were taken and LOS propagation was assumed although there was a lamp-post between the BS and UEs as indicated in the aerial photograph of the environment shown in Figure 5.18.

Figure 5.18: Aerial photograph of the Millennium Square.

The UEs were placed close together with an antenna separation of approximately 43cm, in a straight line parallel to the BS array. As described in Appendix C, the UEs are made up of a USRP with 2 transceiver chains and a Windows Laptop host, so this meant that each set of laptop and USRP (the UE package as shown in Figure C.7) was used as 2 UEs. A total of 6 UE sets

created a total of 12 users, which are effectively 12 separate mobile stations. The UE sets were placed in side the marque and connected to the antennas mounted outside the marquee using 3m SMA RF coaxial cables. The antenna separation was approximately 43cm as the antennas were spread evenly on the side of the marquee that faces the BS. The schematic that shows the relative positions of UEs and base stations is shown in Figure 5.19 below.

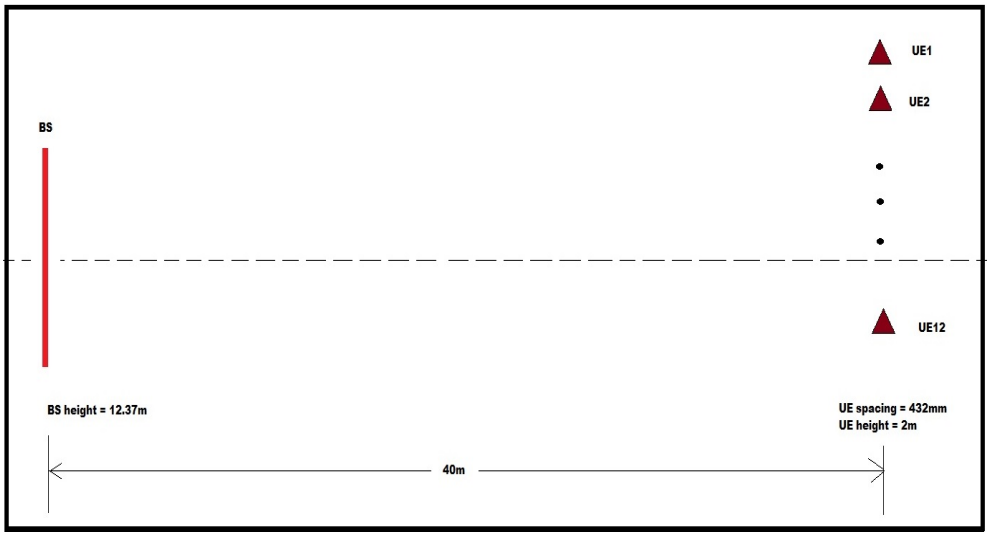


Figure 5.19: 2D schematic layout of the Millennium Square.

Reciprocity calibration data was acquired during this experiment. Reciprocity calibration and how its implemented, is discussed in section 5.1.1. The reciprocity calibration coefficients were generated after the BS was left running for over 4 hours to allows the equipment to reach its continuous operation temperatures. Over-the-air synchronisation was used. Because all the 12 UEs were synchronised, the 1st UE, the 10th and the 12th UEs were used as calibration sources in some of the cases considered. As discussed in Appendix C, the uplink uses a 20MHz carrier which is sampled at 30.72MS/s. DOA estimation was carried out using instantaneous channel captures and 100 channel captures were taken for each UE and 3 takes were generated separated in time by 15 minutes, meaning that at the end of the 3 takes, there were 300 instantaneous channel captures for each user. The availability of 100 captures per user per take allowed the DOA errors to be averaged, and the Root Mean Square Error (RMSE) of DOA error was calculated for each fixed UE position.

This study consisted of varying array configurations and also calibration employed. Table 5.2 below shows the different array configurations considered. The array configurations in Table 5.2 were used in cases that are described in Table 5.3. The red lines show the bounds of the antenna elements used.

Case 1 evaluates the estimation results when different UEs are used as calibrating sources. This case does not employ reciprocity calibration, so it is comparable to the Wills Memorial Hall

Table 5.2: Array configurations considered (selected elements bound by red lines)

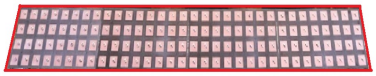
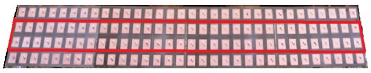
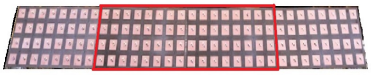
Config 1 - 128 antennas	**Config 2 - 64 antennas**	**Config 3 - 64 antennas**
		
Config 4 - 64 antennas	**Config 5 - 64 antennas**	**Config 6 - 32 antennas**
		
Config 7 - 64 antennas	**Config 8 - 64 antennas**	**Config 9 -64 antennas**
		

Table 5.3: Cases and array configurations considered - Millennium Square

Case	**Description (Array configuration + Calibration)**
1	Comparing calibrating sources - Uses all 128 antennas in Config 1 and UEs 1, 10 and 12
2	Comparing in-field UE calibration vs Reciprocity calibration or both - Uses 128 antennas in Config 1
3	Comparing different contiguous antenna selections - Uses 64 antennas in different configurations (Config 2 - 6)
4	Effect of antenna number on DOA estimation accuracy - Uses 32 antennas in Config 6, 64 in Config 3 and 128 antennas
5	Arbitrary antennas selections - Uses 64 antennas selected arbitrarily from the whole array (Configs 7-9)

experiment described in section 5.3. An arbitrary choice of the first, the last and the directly facing UEs was used to explore the effect of using different calibrating UEs. Case 2 seeks to compare the result when different calibrating data is applied. It takes the best calibrating UE from Case 1, and compares the results to when either reciprocity calibration is employed or both schemes are employed. Case 3 uses 64 antennas in different configurations and seeks to evaluate the effect of these array configurations. The best calibration mechanism from Case 2 is applied. Case 4 evaluates estimation accuracy when using 32, 64 and 128 antennas in a uniform array configuration. Case 5 evaluates estimation accuracy with arbitrary antenna selections. It uses 64 antennas selected in configurations shown in Table 5.2. In all cases the RMSE for each UE was calculated to the nearest degree and the results are presented below.

5.4.1 Results and discussion

The results in Figure 5.20 for Case 1, show that using UE number 10 as the calibrating source achieves the best DOA accuracy. The three UEs used for calibration are not identical, hence it

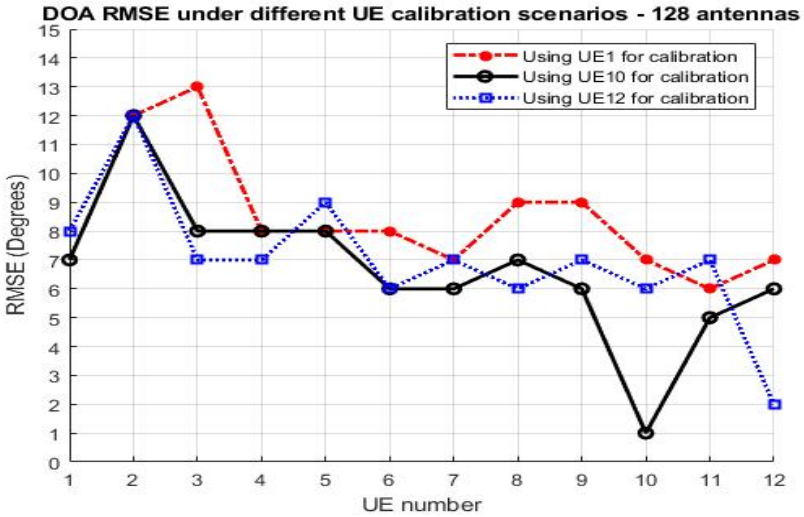


Figure 5.20: Case 1 - RMSE of DOA error (degrees).

is expected that the estimation performance with each UE will be different. If all the UEs were truly identical, only the gain and phase errors in the 128 receive chains would affect accuracy and calibration would reduce those errors. UE10 was chosen as a calibrating source for Case 2 because it produced the best estimation accuracy when all the other 11 UE positions are considered. The

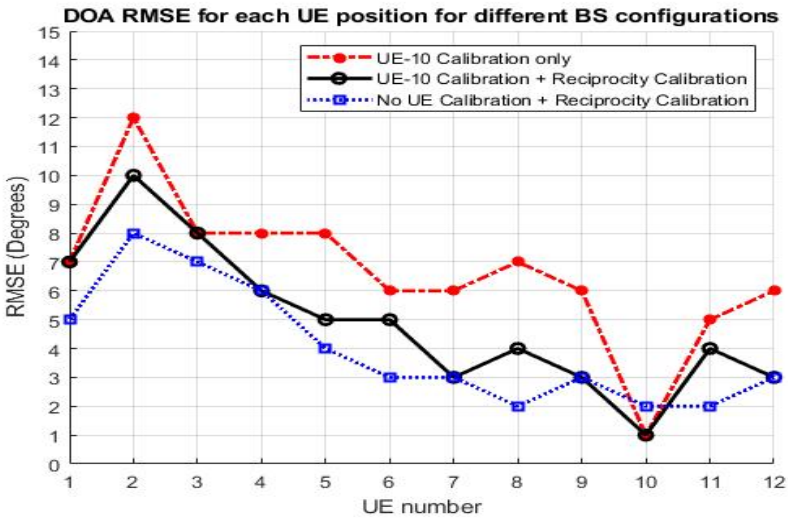


Figure 5.21: Case 2 - RMSE of DOA error (degrees).

next set of results shown in Figure 5.21 for Case 2 shows that there is no significant accuracy gain

from employing both UE calibration and reciprocity calibration. Reciprocity calibration alone, when compared to UE calibration used in section 5.3 shows that applying reciprocity calibration adequately reduces estimation errors. This suggests that the bulk of the errors are incurred during channel estimation, of which reciprocity deals with, quite effectively. This observation is in line with the theory discussed in section 5.1.1, which explains how reciprocity calibration reduces the hardware effects introduced by the transmit and receive radios. Figure 5.22 for Case

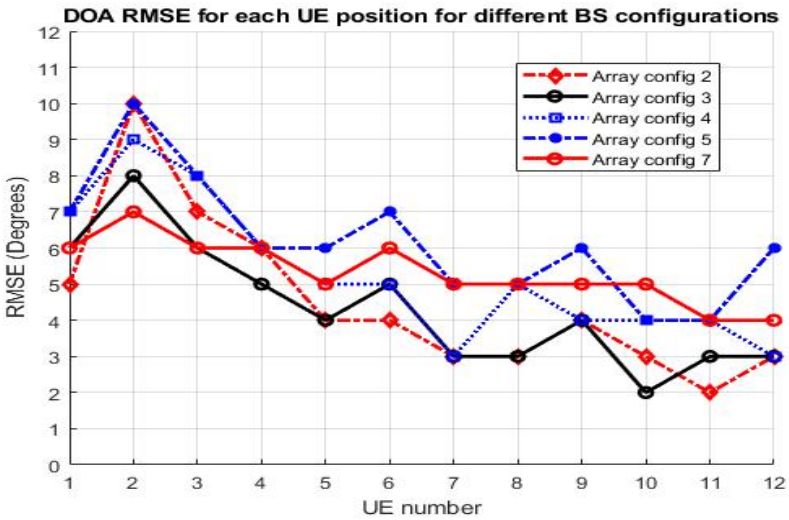


Figure 5.22: Case 3 - RMSE of DOA error (degrees).

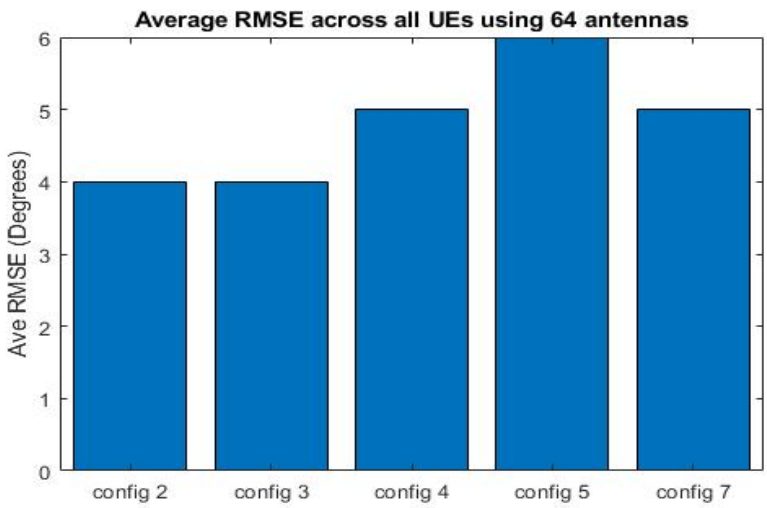


Figure 5.23: Case 3 - Average RMSE across all UEs (degrees).

3 shows that the best estimation accuracy is achieved when 64 antennas that are in the 2 rows that span the whole array, are selected. Figure 5.23 shows the calculated average RMSE across

all UEs for each array configuration. Config 5 shows the worst performance. Since the DOA estimation considered in this thesis is the azimuth DOA, this result suggests that azimuth DOA estimation benefits from a larger azimuth aperture. If the elevation angle was being considered, this author believes that a configuration that maximised the elevation aperture will also improve the elevation angle estimation. Figure 5.24 for Case 4 shows, as expected, that an increase

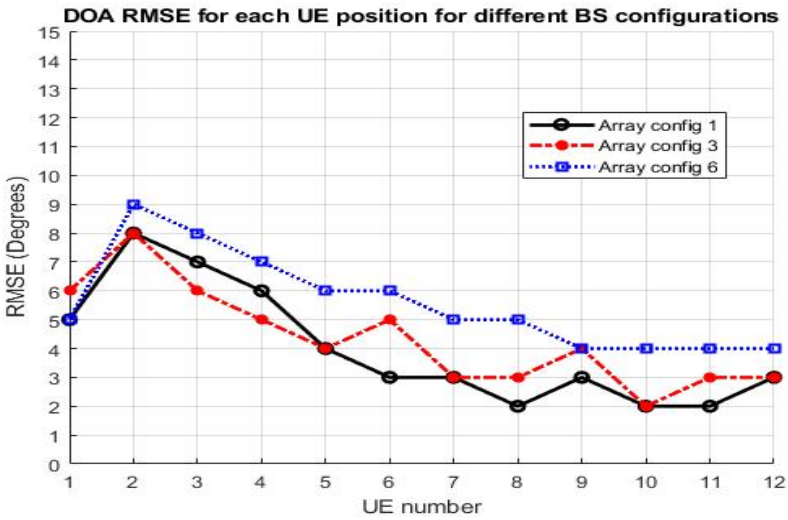


Figure 5.24: Case 4 - RMSE of DOA error (degrees).

antenna number improves estimation accuracy. The number of positions considered (12 in this case) is limited because of the reasons discussed in the experimental setup. Still the results suggest that for a large number of positions, a CDF which shows an estimation improvement with increase in number of antennas should be expected as shown in Chapter 3 simulations. Figure 5.25 for Case 5 shows varying estimation accuracy depending on the array elements and receive chains that are involved. For the array configurations considered, there is no single configuration that improves accuracy across all UEs. However, by establishing a criterion that is able to select the lowest error from each config, an overall estimation improvement can be achieved. There is a general trend of decreasing DOA RMSE, with increase in UE number and an inspection of the UE positions suggests this is because in the presence of small BS position error. When there is a BS position error, the UEs that are further away from the BS perpendicular are affected more in terms of LOS angular error, than those that are close to the perpendicular. Also, the LOS propagation assumption may not have been true in some cases due to objects like the lamppost as shown in Figure 5.18.

5.4.2 Millennium Square environment ray-tracing

Ray-tracing was run for the Millennium Square environment, with the configuration outlined in Chapter 3, in terms of the frequency and transmit power. The purpose of the ray tracing was

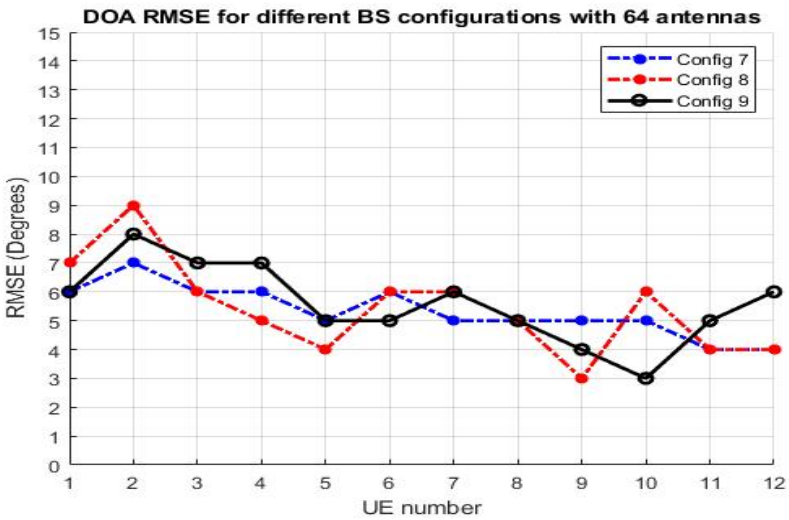


Figure 5.25: Case 5 - RMSE of DOA error (degrees).

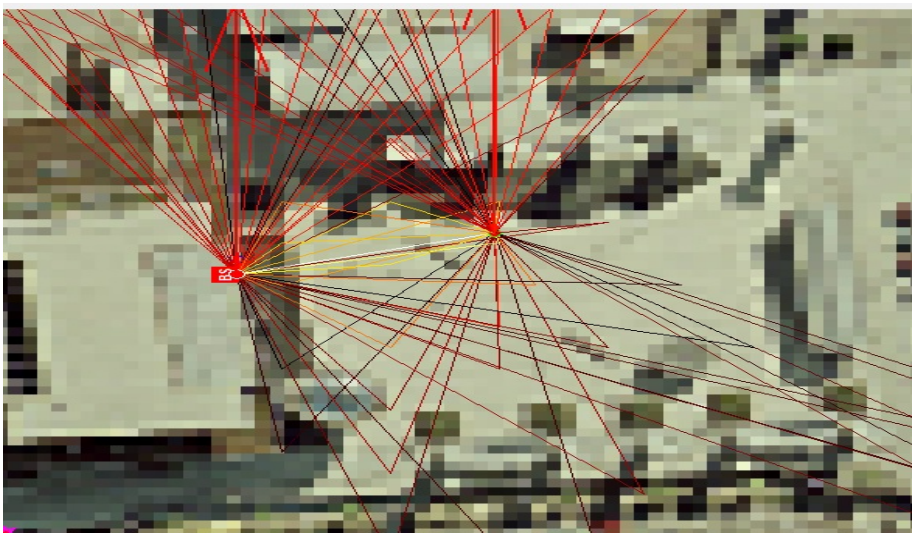


Figure 5.26: Ray-tracing for the Millennium square environment.

to show all the expected rays between a UE and the BS, and determine if these agree with the DOA estimation results from the channel measurements. The rays produced by ray-tracing also help to visualise the geometry of the multipath caused by the perimeter buildings and structures. In this case a single base station and a single mobile station placed approximately at the UE number 10 position. The multipath rays between the BS and the first UE are shown in Figure 5.26 below. The white line shows the LOS ray and any coloured lines are multipaths.

It can be seen from Figure 5.26 above that there was indeed some multipath propagation between the UE and the BS but a LOS component existed. The first 4 angles estimated for this UE all are matched a ray from the ray-tracer output to within 3 degree tolerance. Training data

was generated within the area of the city that includes the millennium square with one BS places at the position and height which is the same as was used in the above experiment. After training with the LSSVM as described in chapter 3, NLOS identification was run for the above ray-traced UE, against the BS, using the rays produced in the ray tracing shown in Figure 5.26 above and the result confirmed that the UE was indeed in a LOS position relative to the massive MIMO BS.

5.5 Merchant Ventures Building indoor measurement campaign

This measurement campaign was conducted in the Merchant Ventures Building (MVB) laboratory, which is the laboratory for the Communications Systems and Networks (CSN) research group. The BS array was mounted above the lab benches, supported by wooden boards as shown in Figure 5.27 below. Figure 5.28 also shows the lab environment, which includes benches, lab equipment and building pillars within the lab space.

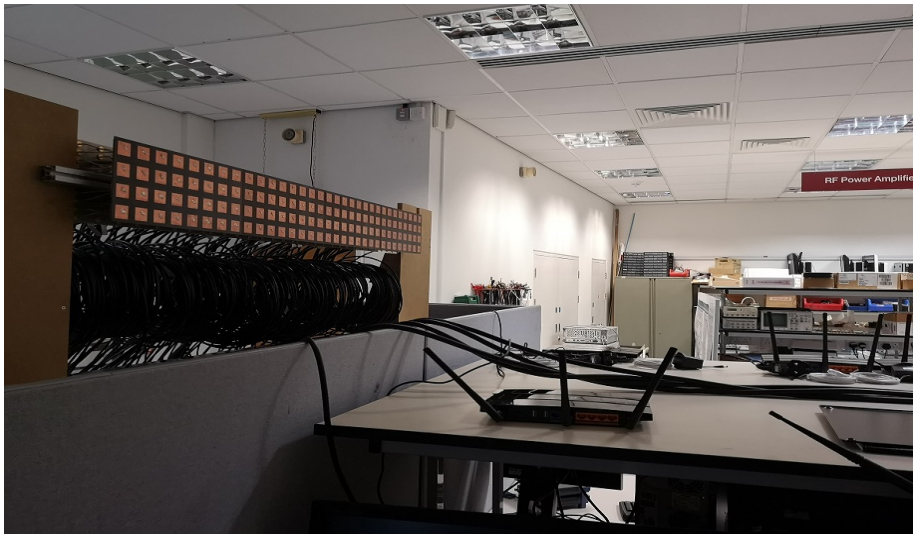


Figure 5.27: BS Antenna array setup in MVB.

In this experiment, a single UE was used, which was moved around the environment to different predetermined positions, each time taking a capture of the channel in a similar methodology as described for the Wills Memorial experiment described in section 5.3. These positions were recorded for later translation into actual LOS angles that are then compared with the estimated angles to produce the DOA estimation errors. The major difference with the Wills Memorial experiment is that in this experiment, reciprocity calibration was used. The calibration coefficients were generated after the BS equipment was run for at least an hour. The UE was setup on a trolley at a height of about 1.5m. Cable synchronisation was used in this case, and this created a restriction on the positions where the UE could reach although it was sufficient to

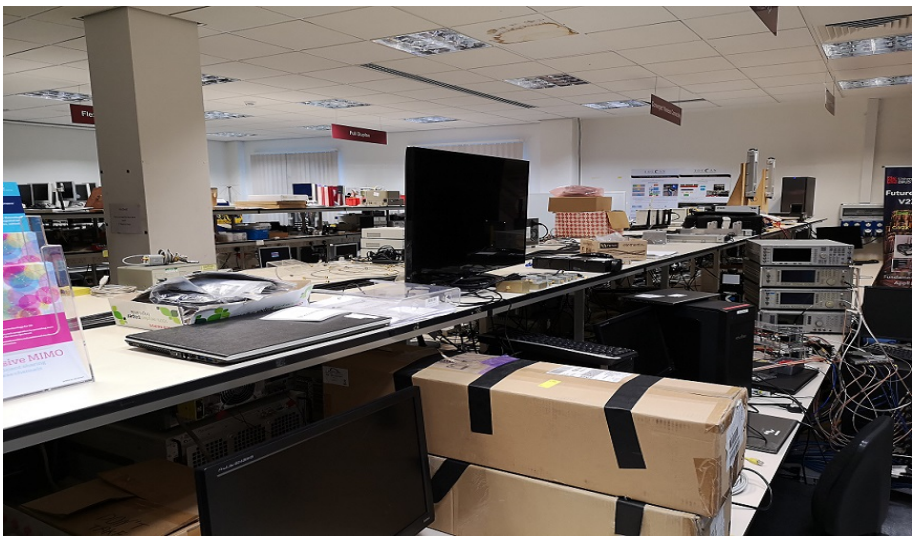


Figure 5.28: Picture of the MVB environment.

cover most of the areas within the lab. Figure 5.29 shows the schematic of the setup in the lab with numbered positions indicating the positions at which the UE channel captures were taken. Figure 5.30 shows the UE on a trolley.

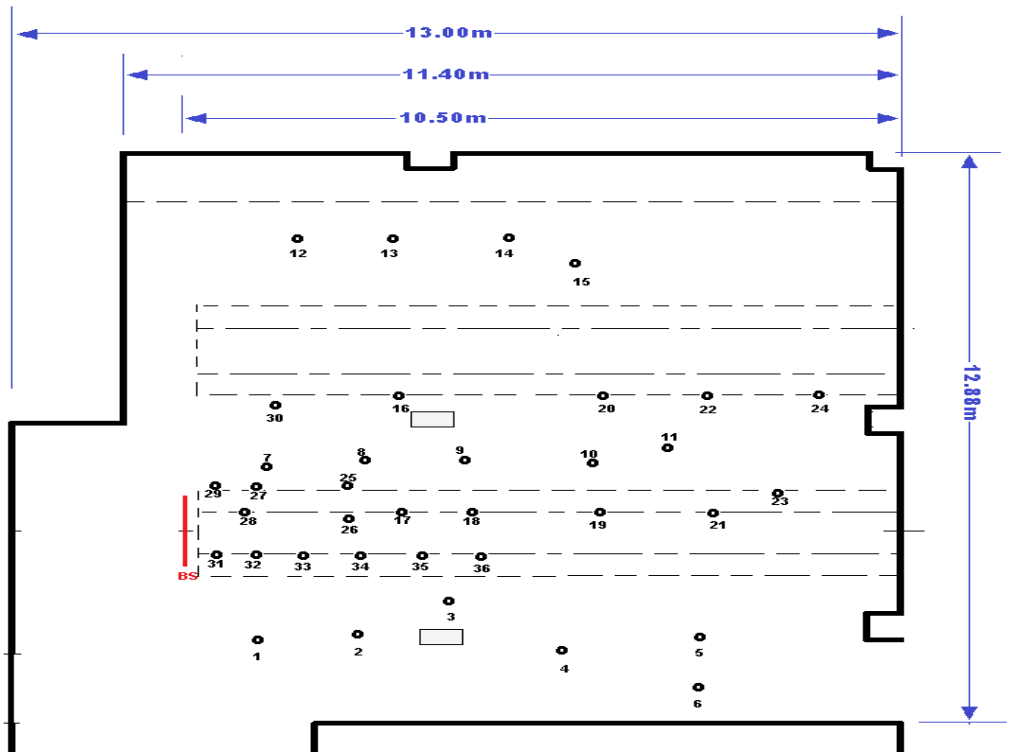
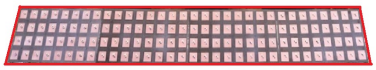
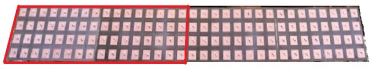
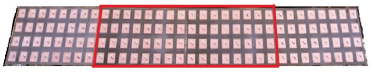
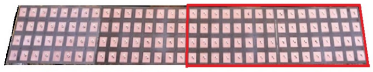


Figure 5.29: Schematic of the MVB environment.

Figure 5.30: UE setup in the MVB.

Table 5.4 shows the antenna configurations that were considered. Triangulation with a single base-station is also evaluated in this section, so DOA estimation was run using 64 antennas in configurations 3 to 5. These configurations also allow evaluation of the performance with same number of antennas but spanning the entire length of the array as shown in configuration 2.

Table 5.4: Array configurations considered

Config 1 - 128 antennas	**Config 2 - 64 antennas**	**Config 3 - 64 antennas**
		
Config 4 - 64 antennas	**Config 5 - 64 antennas**	
		

5.5.1 Results and discussion

The environment was static with no objects moving during captures. The UE positions considered clearly included a few pure NLOS positions from observation during the experiment and can also be identified by looking at the schematic in Figure 5.29 . Reciprocity calibration was used, as described in previous experiments. Two sets of reciprocity data were generated using the first and the last antenna elements, as the reference antenna respectively. There was no significant difference in the DOA accuracy between usage of either set of calibration data. The difference

between the estimated angles and the actual LOS angles between the centre of the array configuration used, and the UE position, are calculated as the DOA errors and a CDF curve of the errors for all the positions is produced for each configuration as shown in Figure 5.31 below.

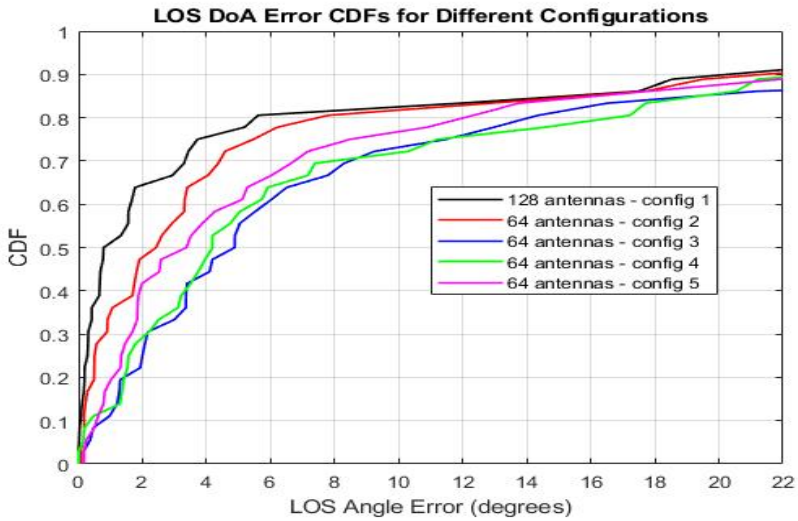


Figure 5.31: CDF of DOA error for all positions in the MVB lab.

From Figure 5.31, it can be noted that DOA estimation accuracy of less than 6 degrees DOA error, is achieved for the 80th percentile using all the antenna elements. While this may not be adequate for some localisation applications especially in indoor environments, it may be relevant for other 5G use cases like downlink beamforming where usage of comparable beamwidths might mean 5 or 6 degree errors do not impact on the downlink transmission. The best performance is obtained when all 128 antennas are used, in line with findings from previous experiments in preceding sections. It can be noted that for positions that produced an angle error grater that 8 degrees, the 128 antenna configuration (Config 1) and the 64 antenna configuration (Config 2) tend to converge. This is caused by positions that are in NLOS and their DOA estimation cannot be improved by employing more antennas. This can also be observed for all configurations, that as the error increases, the CDF curves appear to converge. For the configurations that use 64 antennas, the one with a larger azimuth aperture (Config 2) performs better with an accuracy of 8 degrees for the 80th percentile level. Of the configurations that use the 4x16 setup, the left configuration (Config 5) produced the best performance. The small differences in performance are a result of the differing radio effects in each set of considered radio chains. The overall performance for each case, as shown in Figure 5.31, was affected by cases where no LOS exists and these were determined to be positions 4, 6, 11, 12, 14, 15, 17, 20, 25, 27 and 36, from analysing the DOA estimation results together with the schematic of the environment. A scatter plot of the estimation results for each position is shown in Figure 5.32. As in previous experiments, calculation of DOA error assumed LOS paths between the BS and the UE at every position,

which clearly was not the case in some positions. The estimation of the total number of signals as discussed in Chapter 3, becomes very important for NLOS scenarios because it allows recovery of all multipaths and any possible mechanism of estimating the LOS angle from the multipaths such as described in Chapter 2 section 2.5, can then be used.

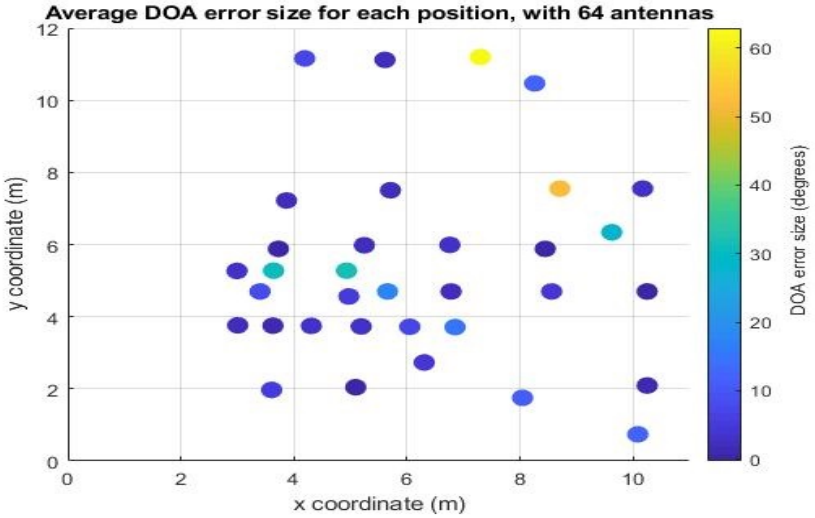


Figure 5.32: Scatter plot of DOA error in MVB lab.

DOA estimation was performed using 3 sets of configurations, which are the full array, left half sub-array and right half sub-array, shown in Table 5.4 as Configs 1, 3 and 5 respectively. This produced 3 DOAs for each position. The config that uses the full array was chosen in place of Config 4 because it produces better accuracy for the same boresight. The 3 angles were used in an AOA algorithm (ToTal) described in Chapter 2, to produce a position estimation for the UE. Whilst localisation of positions identified to be in NLOS above, would be of interest in a multiple BS scenario, these positions were excluded from the localisation evaluation because the setup is using a single BS and the resulting DOA errors led to location errors that were more that 50m outside of the area of consideration. The position error was calculated in a similar way as described in Chapter 3 section 3.2.3 and a CDF was produced as shown in Figure 5.33 below. The choice to use 64 antennas in Configs 3 and 5 was influenced by the need to use a large number of antennas possible in a way that produces 3 equidistant boresights on the antenna array, to allow AOA localisation with a single BS as originally suggested in Chapter 2 section 2.5. This will not be necessary in a multiple BS environment or in single BS setups that use multiple multipath components to estimate the UE position. The result show an accuracy of 5 metres for the 80th percentile level. This is not very useful for an indoor environment, considering the dimensions of the lab as shown in 5.29. The location accuracy was clearly affected by cases where no LOS DOAs exist. Additionally, because of the small separation distance between the boresight anchor points on the array, most of the UE positions had almost similar estimated DOAs on at least 2 boresights,

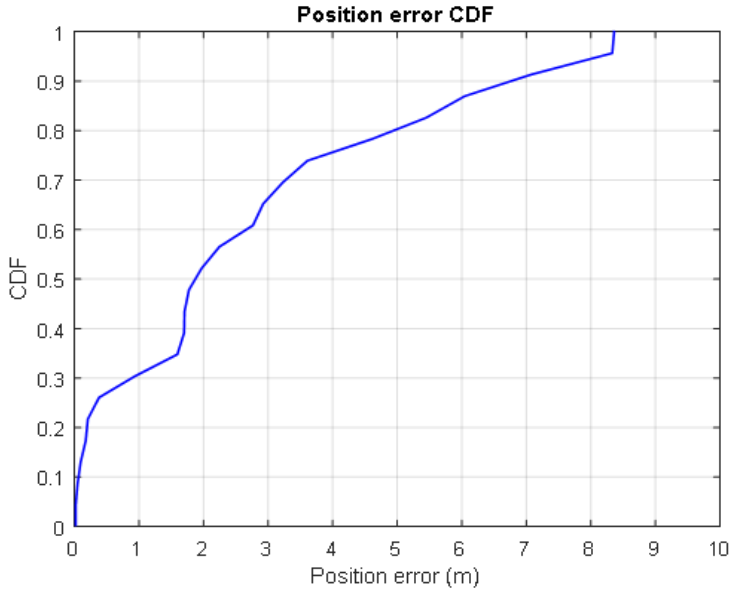


Figure 5.33: CDF of position error in the MVB.

thus positioning accuracy is greatly affected by Geometric Dilution of Precision (GDOP). This

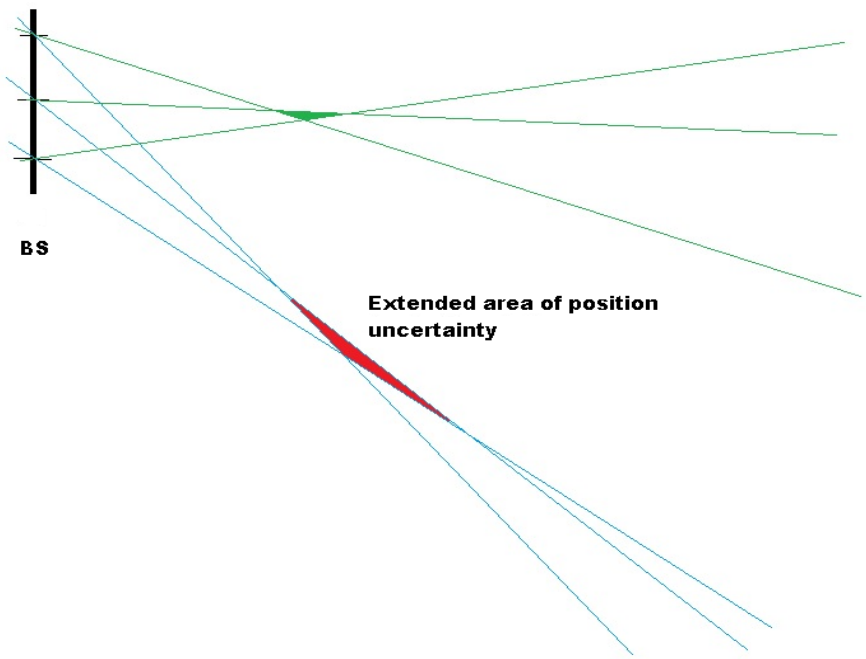


Figure 5.34: Effect of GPOP on AOA accuracy.

effect is very much pronounced for positions that are at either end of the array, meaning only positions aligning to the middle of the array can be estimated better as demonstrated in Figure

5.34. The shaded convergence areas for the 3 DOAs gets increasingly elongated, meaning more position uncertainty as the UE position moves further away from the BS array perpendicular.

5.6 Summary

This chapter has demonstrated DOA estimation using an existing Massive MIMO testbed. It has demonstrated simple DOA estimation using a loosely calibrated antenna array and the key benefit is that there are no modifications to the testbed hardware platform required in order to implement DOA estimation. The work in this chapter has also demonstrated that array manifold characterisation is key to improving the accuracy of estimation. It has been shown that reciprocity calibration can be exploited for DOA estimation in Massive MIMO systems. This is important because it allows reuse of the reciprocity calibration data for two purposes, which are estimation of the downlink channel and also DOA, at the same time.

A key outcome from this chapter is the real-world demonstration that multiple antennas on a BS allow arbitrary and contiguous antenna selections for DOA estimation. This offers more degrees of freedom to DOA estimation and AOA localisation where estimation for a single target can be done using different sets of antennas with the most accurate result being chosen according a set confidence criterion. Other outlier removal techniques can also be used to reject some results. Averaging can also be done, there by improving the accuracy of the system. The results presented in section 5.5 did not produce adequate 80th percentile level accuracy (centimeter level accuracy is generally envisaged for indoor environments [175]). There are improvements that can be made to improve on the accuracy, like employing more sophisticated calibration schemes and employing estimation algorithms that are robust to multipath propagation. The results presented in section 5.4 suggest that some applications like downlink beamforming in LOS environments may be fulfilled by the demonstrated accuracies, provided that the DOA RMSEs are within the required beam-widths. Even in cases where the accuracy is not within the beam width it could still be useful to reduce the number of search steps to find the right beam.

CHAPTER 6

Conclusions and Recommendations

This thesis has presented a rationale for why localisation is critical in next-generation wireless systems, with particular focus on 5G systems which can utilise location information to address some of the system challenges. This was outlined in the introduction chapter. The state of the art in localisation in mobile networks and a detailed analysis of the 2 key wireless network enabled localisation algorithms TDOA and AOA, were subsequently presented in Chapter 2 as well as in Appendix A. An evaluation of the TDOA and AOA methods using ray-tracing data was presented in Chapter 3 and the results demonstrated the need for NLOS identification in order to improve the accuracy of these algorithms to below 10m for outdoor urban environments. Because this was a key result, NLOS identification and mitigation schemes which utilise machine learning in the form of Least-squares Support Vector Machines were devised and evaluated using ray-tracing with a real-world database as discussed in Chapter 4. An extension of the mechanism which provides for direct localisation using ray-tracing data was also presented in Chapter 4. One of the two key localisation algorithms, i.e. AOA, required DOA estimation at the BS, so to facilitate the evaluation of the schemes and methodologies discussed in this thesis, DOA estimation utilising a physical massive MIMO testbed was conducted. The testbed, which is outlined in Appendix C, was used to carry out the experiments discussed in Chapter 5, demonstrating the advantage of having a large number of antennas at the BS. The observations and conclusions which can be drawn from the work that is presented in this thesis can contribute to design and implementation of localisation techniques for next-generation systems. They can also contribute to the design and implementation of 5G communication systems. Preliminary experiments in Chapter 3 inform on the choice of algorithm, between AOA and TDOA depending on the environment. Experiments in Chapter 3 sections 4.3.3 and 4.4 specifically show that if systems can resolve multipath components to individual rays, these can be used to either improve

NLOS identification and mitigation or for direct localisation. This should buttress the need for systems with adequate bandwidths to meet this requirement. Experiments in Chapter 5 inform on the best geometries for array DOA estimation. A comparison and ray-tracing based evaluation of TDOA against AOA informs on when to choose either algorithm. NLOS identification and mitigation using machine learning in a location specific framework has demonstrated performance that surpasses any current demonstrated or proposed technique. A direct localisation technique that uses machine learning has been shown to perform best in NLOS environments, which are otherwise difficult to achieve similar localisation performance with traditional techniques. Whilst a multiple BS setup would have provided a better localisation performance using the APML and AOA algorithms, Chapter 5 has demonstrated useful DOA estimation using a loosely calibrated antenna array.

These observations are discussed at the end of each Chapter and and the overall conclusions are expanded in this chapter. To provide an adequate treatment of the observations, the conclusions as well as the recommendations stemming from the work. The following discussion presents an itemised outline of the suggestions and recommendations that could be drawn from this work.

- A detailed discussion and comparison of AOA and TDOA is presented in section 2.2. Preliminary assessment of the TDOA and AOA techniques was presented in section 3.2.4. The work in this section demonstrated that the AOA localisation algorithm is more accurate than TDOA in LOS environments as demonstrated by Figure 3.5 which shows that for all positions where LOS angles of arrival were obtained, there was not any location error for the AOA technique. This was clearly not the case with TDOA. This section also highlights the importance of NLOS identification in localisation schemes. Figure 3.5 in this regard, demonstrates for TDOA that NLOS identification improves azimuth location accuracy by 20 metres considering the 80th percentile level. Most studies such as outlined in section 2.2.1 compare the two algorithms for a particular environment or scenario without considering the effect of treating NLOS and LOS measurements separately for that same environment. The work presented in section 3.2.4 suggests that a dynamic approach of utilising AOA for predicted LOS positions, and utilising TDOA for other positions can greatly improve the overall localisation accuracy by eliminating position errors for all LOS positions. This requires NLOS identification schemes, such as those discussed in Chapter 4, to be exploited to identify the LOS positions where the AOA algorithm can be used. The recommendation from this observation is to implement systems that are capable of dynamically choosing the localisation algorithm such as discussed in [176] for TDOA/AOA RFID localisation, and also hybrid fusion for tracking (discussed in section 2.3), in-order to leverage NLOS identification in such systems.

- Whilst the simulations in section 3.4.1 show as expected, that increasing the number of antennas improves DOA estimation with the Alternating Projection Maximum-Likelihood

algorithm, Figure 3.15 showed that a minimum of 32 antennas achieves a DOA error of just one degree for the established 80th percentile level. For 64 and 128 antennas, no error is produce at the prescribed percentile level. These results provide a clear demonstration of the benefit of multiple BS antennas to DOA (hence localisation). A simulation of the effect of gain and phase errors in Figure 3.16 demonstrated that the APML algorithm is more sensitive to phase errors than gain errors by approximately 8%. This result is useful when devising calibration schemes. It informs that phase calibration is key in improving the estimation accuracy.

- The results in Chapter 4 validate and demonstrate the effectiveness of Least-Squares Support Vector Machines as a technique to implement NLOS identification and mitigation. The results show that training the LSSVM with data that incorporates ground reflections (Training data 2) improves identification accuracy by 6.7% as compared to utilising NLOS measurements (Training data 3), and by half a percentage as compared to training data 1 which comprises of NLOS and LOS measurements. These results demonstrate that the choice of training data is critical in obtaining the best identification accuracy. A best case LOS identification error probability of 0.019 is achieved when time delay, received power and angle of arrival are used, as shown in Table 4.3. These results imply that it should be possible to identify all measured LOS channels in specific environments using ray-tracing data. There is a significant drop in the identification error probability, from the next best case result of 0.035 to the best case result of 0.019, a drop of 0.016 which represents nearly halving the number of errors. This is attributed to the inclusion of AOA as a feature in the best case scenario. This demonstrates that AOA is key to improving LOS identification for the location specific approach. Location independent classification results in Table 4.4 are in line with results obtained in [133] and in this case the best case identification error probability of 0.1211 was achieved. Section 4.4 evaluates the use of Least-squares Support Vector Machines for direct localisation. The results show that the direct scheme performs best in dense urban environments, with at least a 200% location accuracy improvement as compared to other environments, when considering the 80th percentile level. Again as observed in section 3.2.3, the results in this section validate the finding that the angle of arrival is key to improving the location accuracy. This is shown in Figure 4.12 where errors in the angle or arrival measurements worsen the location accuracy by at least 50% as compared to errors in other parameters, when considering the 80th percentile level. For 90th percentile, the direct method improves location accuracy by 50% over the TOA-AOA scheme in a dense urban area. The results in this section suggest that the direct method is most suitable for city center environments and can even be further improved by using tracking algorithms and fusion with inertial measurements.

- Ray-tracing has been demonstrated in Chapter 3 to be a useful tool in the evaluation of the schemes discussed in that chapter. Also in Chapter 4, ray-traced data was used to train

Least-Squares Support Vector Machines. This means that ray-traced data can be used to train similar Machine Learning algorithms. However, it should be noted that usage of these schemes which utilise ray-tracing is dependent on the availability of the ray-traced database for the environment of interest. This approach also requires the ray-traced data to be validated with real radio measurements. To that end, in the UK, Ordinance Survey (OS) [177] in their work which buttress the research presented in Chapter 4, has piloted 5G ray-traced maps with Bournemouth City. A video of their work is available online [178]. This signifies a move by local authorities in partnership with network operators, towards providing ray-traced network maps as the smart city market grows. The UK Department for Business, Innovation and Skills (BIS) released a smart city background paper in 2013 [179] as well as a 2013 report [180] which recognised location based services as a driver for smart cities. Specifically for 5G, the UK Department for Digital Culture Media and Sport in conjunction with Ordnance Survey released a planning guide in 2018 [181], targeted at planners and local authorities, which identifies ray-tracing as one of the key tools in the planning process. Fuschini et. al. in their ray-tracing review paper [182] argued that ray-tracing can be exploited for mmWave Massive MIMO downlink beamforming. Examples of commercial 5G ray-tracing models include Huawei's high precision 5G model [183] and Siradel's Volcano 5G ray-tracing model [184]. While the general availability of ray-tracing data for cities could be said to be improving as cited above, it is also recommended that telecommunications operators should help facilitate availability and validation of ray-traced data for cities, and also facilitate the periodic updating of the data especially in response to the changes in the built-up environment. These developments will mean that the possibility of exploiting machine learning techniques for localisation increases when computational cost decreases through allowing train-once approaches for environments that remain fairly static. This further makes the direct localisation technique that is presented in Chapter 4 more practical for urban environments. This approach is further supported by the thrust towards vehicular traffic free cities. In an article in the Independent online, titled "A car-free future? How UK cities are moving towards a pedestrian age" [185], a number of cities that include Birmingham Edinburgh, Leeds, Sheffield and York are cited to have proposals to ban vehicles in the city center. This development helps the radio channels between the numerous 5G base stations and users to be come more deterministic if the assumption that placing the BSs at a greater height (e.g. on buildings) is capable of availing LOS components to human traffic.

- Chapter 5 demonstrates the use of a Massive MIMO testbed for direction of arrival estimation. All the results obtained are consistent with the expectation that estimation accuracy improves as the number of antennas increases. Figures 5.20 to 5.25 and Figure 5.31 suggest that as the azimuth aperture of the array increases, estimation accuracy also improves. This can be demonstrated by considering Case 3 using array configurations in Table 5.2

together with Figure 5.23, config 5 which uses half the array in azimuth, performed worse by 50% as compared to the lowest average RMSE result of 4 degrees (configs 2 and 3). This result suggests that for azimuth DOA estimation, rectangular arrays that extend in the azimuth dimension are more suitable.

- Robust calibration methods are key in direction of arrival estimation techniques as discussed in Chapters 2 and 3. Gain and phase calibration of the transmit and receive antennas, calibration of the transmit and receive radio chains, calibration of the cable interconnects, manifold calibration and mutual coupling calibration, are all key to the performance of RF DOA estimation system. Also, whether the system is coherent over all the radios, matters. The work in Chapter 5 has demonstrated that reciprocity calibration can be exploited for DOA estimation in a massive MIMO system.

- The work presented in section 5.5 clearly demonstrated that a single base station AOA localisation using a uniform rectangular array will not produce an acceptable localisation accuracy even in indoor environments, probably because of the geometric dilution of precision problem. This author recommends other array geometries like circular arrays for such scenarios. Furthermore, approaches like the 3D TDOA schemes discussed in [104] and as presented in section 2.5, can be employed to offer improved localisation accuracy in NLOS environments. It is particularly recommended to investigate and evaluate the feasibility and accuracy of using this scheme, using ray-tracing data. This can then inform on the approach's feasibility with a DOA testbed like the Massive MIMO testbed used in this thesis.

6.1 Future work

While focus has been on usage of common DOA estimation algorithms to estimate the angle of arrival at the BS as discussed above, there have been other non-standard approaches to DOA estimation which may remove the challenge of calibration in array systems. One such technique is the usage of machine leaning for DOA estimation as presented in [186], [187] and [188]. It can be noted that these approaches sit well with the work presented in Chapter 4 of this thesis and the author believes ray-tracing again can be used to evaluate these schemes.

DOA estimation has been demonstrated using the Alternating Projection Maximum-Likelihood algorithm. It would be of interest to devise a mechanism to utilise the testbed to achieve DOA estimation using other algorithms. Relevant calibration techniques can be designed and appropriate changes to the testbed implemented in order to evaluate these algorithms with a real-world practical testbed. The Alternating Projection Maximum-Likelihood algorithm is suitable with offline analysis as in this case but the computational intensiveness may render it unsuitable for real-time systems. The advances in Mobile Edge computing [189] may address some of the computing challenges.

Overall, the research in this thesis demonstrated that accurate 2D localisation can be achieved by exploiting machine learning algorithms with a ray-tracing model. The significance of this is three-fold, first, the advances in usage of AI techniques in radio communications implies that this technique sits well with the wireless systems of the future and should accelerate research into related edge computing systems to support the required computational loads for real-time systems. Second, the exploitation of ray-traced data should accelerate development of very accurate and frequently updated map-based RF databases as already noted above in the case of the work by Ordinance Survey, in light of the performance of the proposed schemes. Third, the technique has been demonstrated to address the challenge of multipath well, to the point of providing accurate localisation in dense multipath environments. The research has also demonstrated the potential of large scale, commercial off the shelf MIMO systems for direction of arrival estimation, which in future may challenge the assumption that expensive, high-precision RF systems with well characterised antennas are required for AOA estimation and localisation.

The following pieces of work are proposed for studies that will further demonstrate 2D localisation in a dense urban environment utilising the Massive MIMO testbed.

- Distributed sub-array measurements will be key for a complete AOA localisation system with multiple BSs. An experimental design that makes use of a distributed deployment of the massive MIMO testbed, as discussed in Chapter 2 section 2.5, must be developed which will facilitate measurements with the four BS racks separated.

- Extending the MIMO testbed to estimate TOA, in addition to DOA and thus extending the localisation model on the testbed to use TOA, in addition to AOA, and running simulations against the ray-tracing model.

- Running outdoor trials, in the Bristol area, comparing the estimated location (TOA and TDOA) against the actual location.

APPENDIX A

TDOA AND LOCALISATION IN LTE

This appendix give an overview of the state of localisation in 4G LTE. It provides the key references to literature on specifications and standards, together with a background literature that builds into the choices made by this author, especially regarding uplink TDOA.

A.1 Introduction

The fourth generation (4G) cellular systems that use the Long-Term Evolution (LTE) standard have support for four main methods for user positioning [6]. It supports satellite-based positioning both in form of autonomous positioning using GNSS systems and also assisted GNSS. LTE also supports Mobile Radio Cellular (MRC) positioning using TDOA, Hybrid methods like Hybrid-GNSS where MRC technique like TDOA is combined with GNSS. Lastly LTE supports the Control Plane – User Plane session handling mechanism. These are described in succeeding sub-sections.

A.1.1 Satellite Based Positioning

The two main modes of satellite positioning are described, with highlight on their limitations and challenges.

A.1.1.1 Autonomous

Satellite positioning uses a minimum of 3 satellites (assuming the use of a very accurate clock), to estimate the position of a mobile device. Time difference of arrival of the satellite signals is used. While TDOA in GNSS and TDOA using MRC uses the same mechanism and theoretic constructs or models, in this thesis, TDOA will be used to refer to the MRC positioning algorithm,

and all the GNSS mechanisms are simply referred to as GNSS positioning. The drawbacks of satellite-based positioning can be described as follows;

- **LOS requirements** The mobile device would need to have a clear line of sight (LOS) with at least 3 satellites for the positioning system to work. This is a challenge in urban areas where tall buildings and other moving objects (e.g. buses) may block the line-of-sight. Also satellite positioning will not work indoors since indoor penetration of the signals is poor.
- **Weaker signals** The received positioning signal from the satellites is very weak and requires sensitive receivers to pick them up.

A.1.1.2 Assisted GNSS

A-GNSS is meant to improve the time to first fix (TTFF). It uses both the satellites and the cellular network. Positioning in this setup can be achieved in two way, Mobile-Assisted and Mobile-Based. Mobile-assisted scheme utilises the mobile device to measure the signal parameters and then send the results to a location server, whilst the Mobile-based has the mobile device measuring and then running the localisation/positioning code, before sending the location results to the network. Obviously, the drawbacks of GNSS as identified before, still affect A-GNSS. These limitations encourage us to look closely at MRC positioning methods and their hybrid variants.

A.2 Mobile Radio Cellular Positioning

In LTE MRC positioning, there are two main standard positioning mechanisms, which are Enhanced Cell ID (eCID) and TDOA.

A.2.1 Enhanced Cell ID (eCID)

Based on the cell of origin (COO), this basically gives location as the cell in which the UE is connected to. This technique is preferably and commonly used for devices that do not have a GNSS receiver. In its simplest form, location accuracy can only be a function of the cell size. The technique works by taking in the cell size, and some measurements on the radio signals in 3 ways.

(i) eCID + distance estimation from 1 Base Station Uses (RSRP) and can locate the device down to a circle (whose radius is determined by the received signal power at device)

(ii) eCID + distance estimation from at least 3 BSs Uses (RSRP, TDOA,TADV or RTT) to run some Trilateration algorithm and together with eCID and give a finer resolution to the position of the device.

(iii) eCID + AOA from at least 2BSs similar to (ii) but uses AOA technique instead of RSRP or other Time-based measurements.

A.2.2 TDOA

This is the preferred method, and is discussed in detail in the succeeding sections. TDOA may be implemented in two different ways which are; Observed TDOA (O-TDOA) or Uplink TDOA (U-TDOA).

- **OTDOA**
  O-TDOA utilises signal observations on the UE, from different BSs. In LTE, Observed Time Difference of Arrival (OTDOA) uses Reference Signal Time Difference (RSTD). Neighbour cells (eNBs) are used to derive an OTDOA to the serving cell. The reference signal is embedded into the overall downlink signal. The receiver/UE receives multiple reference signals from different eNBs and measures the TDOAs, each describing a hyperbola. At least 3 pairs are required. The location of the target is the intersection of the 3 hyperbolas.

- **U-TDOA**
  U-TDOA systems determine the target position through trilateration. The time difference that is measured from the received signal, is converted to a constant difference distance between the 2 receivers, which defines a hyperbolic curve as described for the generic TDOA system. U-TDOA was debuted in research phase since 3GPP Release 9. Further 3GPP activities surrounding U-TDOA from Release 11 to the current release, can be found in 3GPP documents on [190].

Comparison between OTDOA vs UTDOA

Potential issues of Downlink TDOA (OTDOA)

- Different RSTD measurement accuracy between UEs(Handsets).
- PRS Transmission has impact to system capacity although a long interval in its periodic transmission can reduce the impact, provided the handset has a more efficient algorithm for coherent integration.
- Indoor positioning remains a challenge with OTDOA.

Advantages of UTDOA

- UTDOA is becoming more interesting because it is transparent to the UE. Advanced signal processing is possible at the eNB to improve UTDOA performance in terms of Sounding Reference Signal (SRS) detection and hearability, eg interference randomisation and cancellation.
- UTDOA is easy to deploy and/or upgrade from either OTDOA or to future versions.

TDOA in succeeding sections and chapters will focus mainly on its implementation in the form of UTDOA.

A.2.3 Comparison between A-GNSS vs MRC approaches

Table A.1 summaries the differences, advantages and disadvantages between A-GNSS and Mobile Radio Cellular positioning approaches.

Table A.1: Comparison of GNSS and MRC Positioning systems [2]

A-GNSS	**Mobile Radio Systems**
Weak received signals	Comparatively stronger received signals
Low bandwidth	High bandwidth (e.g. 20MHz or higher for LTE)
Similar received power from all satellites	Stronger signal from serving BS and strong interference from adjacent BSs
Long synchronisation procedures	Short synchronisation procedures
Very accurate satellites synchronization using atomic clocks	Synchronization of the base station not apriori guaranteed
Signal known a-priori due to low data rates	Complete signal not known a-priori to support high-data rates, only certain pilots
Line of sight (LOS) access as normal case, not suitable for urban or indoor areas	Non-line of sight (NLOS) access as normal case, suitable for urban or indoor areas
3-dimensional positioning	mostly suitable for 2-dimensional positioning

A.2.4 LTE positioning architecture

This section describes the basic components within the location or positioning infrastructure for LTE.

A.2.4.1 The 3 Step positioning

Positioning in LTE can be considered as a 3 Step process;

- Providing the initial assistance information for position estimation
- Executing measurements and reporting the results
- Running position estimation code based on the results.

Figure A.1 shows the basic components for positioning in an LTE network and these components are described below.

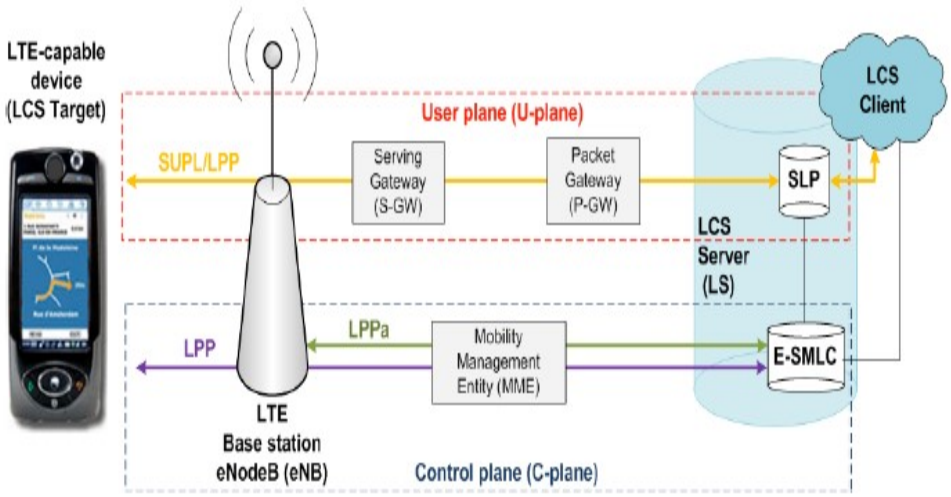


Figure A.1: Network location architecture in LTE [2].

LCS: The location service client requests service. This is usually installed on the target system, but all code would be executed on cloud-based servers
LS Server: The location service server (LS) is usually a logical or physical component which is responsible for collecting measurements and other information from the UE and eNB to assist in measurements and positioning estimation.
LCS target: The LCS Target is usually the User Equipment (UE)

A.2.4.2 U-Plane and C-plane Communications

The client device can communicate with the location server in two ways; over the user-plane (U-Plane) or over the control-place(C-Plane) Evolved Serving Mobile Location Centre (E-SMLC) is used in the C-Plane whilst, Secure User Plane Location (SUPL), an Open Mobile Alliance (OMA) defined general-purpose protocol for positioning, is used in the User-Plane. Both the E-SMLC and SUPL Location platform (SLP) can be located in the same server, physically, because they are simply logical components or entities.

A.2.4.3 LPP and LPP Annex Protocols

Exchange of information happens via the LTE Positioning Protocol (LPP) and LTE Positioning Protocol Annex (LPPa) protocols[6] [191]. Communication between the LS Server and eNB utilises LPPa. In OTDOA for example, the base station is the one responsible for the configuration of the signals that are used in positioning measurements like the positioning reference signals (PRS). The base station also provides information back to E-SMLC and allows it to take inter-frequency measurements where necessary.

A.2.5 Summary

Table A.2 summaries LTE supported positioning methods highlighting where the measurements and estimation algorithms are based.

Table A.2: Overview of supported positioning methods

Method	Measurement	Estimation
A-GNSS	UE	UE or Location Server (LS)
eCID	UE	LS
OTDOA	UE or eNB	LS
UTDOA	eNB	LS

The techniques discussed in this appendix describe the basic or core architecture of LTE positioning. Further enhancements to the protocols, and other enhancement proposals can be found in the technical specification documents within releases up-to the current release, all which can be found on the 3GPP website [190]. Also the study items for 5G positioning support have already been published [192].

APPENDIX B

LSSVM IN DIFFERENT AREAS OF THE CITY

This appendix describes and shows a separate area of the City of Bristol against which ray tracing was performed, as described in Chapter 4. The figures in this appendix show the locations of the base stations as well as the localisation results.

Figure B.1 shows a different area of the Bristol City center, which shows the extent of the built-up environment. This area is characterised by multipath propagation and is comparable to the city center environment used in Chapter 4. Figure B.2 shows the BS placement in the city center environment considered. The Direct method was applied using the ray-tracing data from this area and the result is shown in Figure B.4, which agrees with the result obtained in Chapter 4 section 4.4. Figure B.3 shows an open area within the Bristol ray-tracing database which was used to evaluate localisation performance. This area is comparable to the open area considered in Chapter 4 section 4.4 but the BS placements were different for each case. The result is also shown in B.4. Figure B.5 shows the variation of the estimated range errors for sampled positions, with distance from the BS.

Figure B.1: Ray traced area of the city center environment.

Figure B.2: BS placement in the city center environment.

Figure B.3: Ray-traced open area environment.

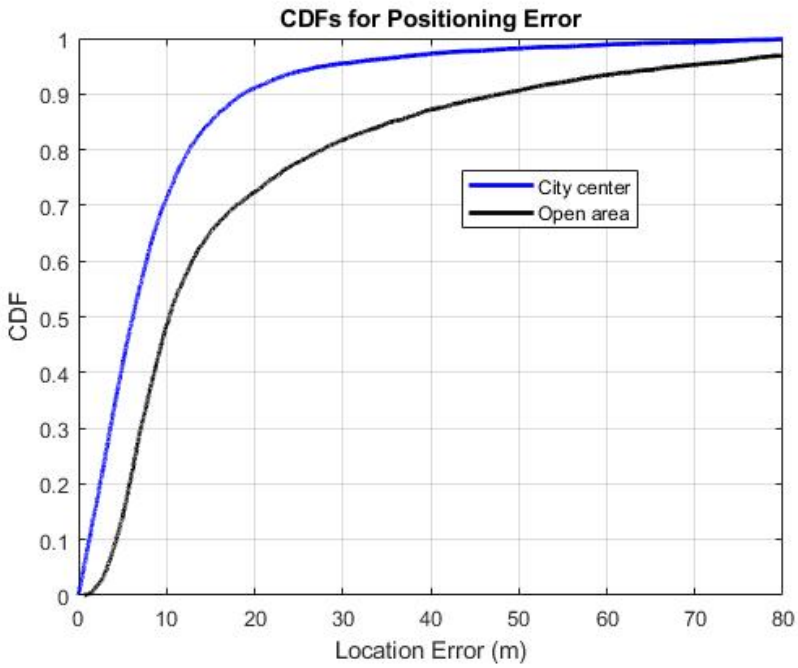


Figure B.4: Localisation performance for the 2 environments.

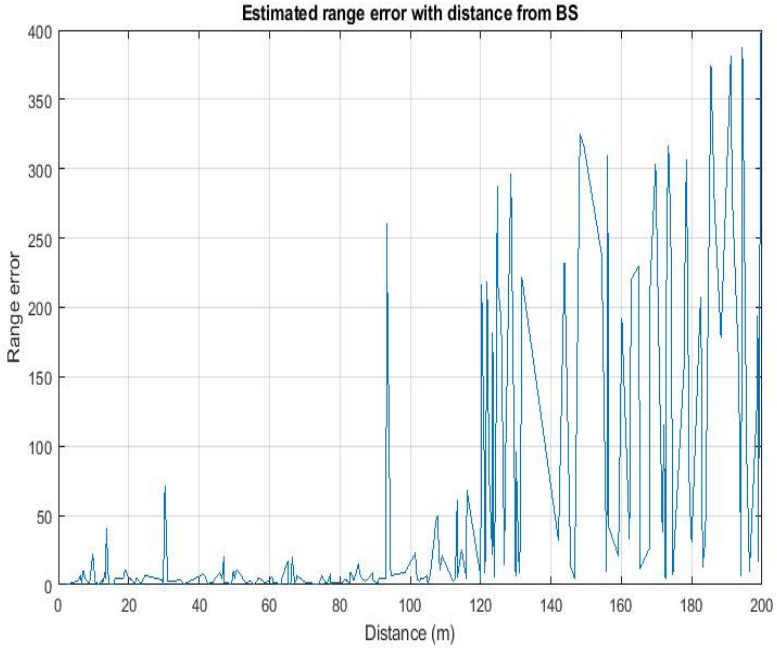


Figure B.5: Estimated range error, with distance from BS.

APPENDIX C

Massive MIMO testbed for DOA estimation: Hardware Design and Implementation

It was established in Chapter 2, that localisation in mobile radio systems exploit the variations of signal parameters like received power, time delay and angle of arrival, from the target or UE, to the base station. Exploiting these parameters for localisation, requires special consideration to the equipment that is being used. In LTE, Location Management Units (LMUs) are used. Special Positioning Reference Signals (PRS) were defined since LTE Release 9 [6], because the prior cell-specific reference signals in Release 8, were not adequate for positioning. Synchronisation between the BSs is also paramount, especially for schemes that perform localisation using time delay-based methods like TOA and TDOA. For AOA based methods, the antennas at the BS need to be calibrated. The subspace methods discussed in Chapter 2 require careful calibration as they are quite sensitive to gain and phase errors. It was also established that AOA at the BS is a key parameter for the localisation schemes discussed in Chapter 4. Firstly, it is established that AOA information, together with time delay and received power, are key to improving the NLOS identification schemes using LSSVMs. Secondly, accurate AOA was demonstrated to be critical for a direct localisation scheme that uses ray tracing data and LSSVMs, in an urban environment. Furthermore, it can be noted that AOA information is also critical for downlink beamforming. Chapter 2 discusses ideas on how AOA information can be exploited to address the problems of pilot contamination, power allocation and dynamic resource allocation, among other radio systems challenges. It can therefore be concluded that AOA information is critical for localisation and many other aspects of 5G and next generation wireless systems. This chapter provides details of the massive MIMO test bed at the University of Bristol that was used for DOA estimation to achieve the results presented in Chapter 5. It should be noted that this testbed was not designed with DoA extraction in mind and the spatial

multiplexing approach for multiuser communication avoids the need for array and front-end calibration, whereas it is key for DoA extraction. The chapter also details the design of the base station with respect to the hardware components, synchronisation and the antenna array. It also provides details on the design of the user equipment. Experiments conducted to characterise the BS radio chains, and the antenna array are presented in D. A basis for reciprocity calibration, its formulation and how it was conducted on the testbed, is also presented.

C.1 Bristol Massive MIMO testbed design

The Bristol Massive MIMO testbed was extensively used in experiments that resulted in world records for spectrum efficiency by Bristol University, UK, in collaboration with Lund University, Sweden with results covered by the press [193]. The author was involved in practical aspects of these experiments, setting up the base station, and recording the positions of UEs. Detailed accounts of the testbed's design and implementation can be found in the publications of the author's colleague and lead in the Massive MIMO Project at Bristol, Paul Harris, which are cited at the end of section C.1.1. This section will focus on the design aspects which are relevant to the use of the testbed for direction of arrival estimation.

C.1.1 Design parameters

The design of the Bristol massive MIMO testbed aligns with TDD LTE specifications. Table C.1. summarises the design parameters.

Table C.1: Testbed Design Specifications and Features

Number of antennas	128
Number of UEs	12
Max Single Tone Tx Power per chain	15dBm
Carrier Frequency Range	1.2–6GHz (Licence at 3.51 GHz)
Carrier bandwidth	20MHz
Baseband Sampling Rate	30.72MS/s
UL/DL MCS	QPSK, 16-QAM, 64-QAM, 256-QAM
MIMO Detection/Precoding	MMSE, ZF and MRC/MRT
Waveform	LTE Spec OFDM
Duplexing	TDD

The system parameters were chosen to align with LTE specifications because it is generally expected that initial 5G system prototypes will seek compatibility with existing LTE standards. The system, which is part of the Bristol Is Open (BIO) initiative [194], was designed using National Instruments (NI) commercial off-the-shelf products as listed in Table C.2.

NI provided a unified hardware abstraction through their LabVIEW platform, for rapid development and prototyping. Their LabVIEW for Communications Suite, which incorporates

Table C.2: Hardware components for the Bristol MMIMO system

Component	**Model**	**Description**
Host	PXIe8135	2.3GHz Quad-Core PXI Express Controller Up to 8 GB/s system and 4 GB/s slot bandwidth
SDR	USRP RIO 2943R / 2953R	2 RF Front Ends and 1 Xilinx Kintex-7 FPGA Centre frequency variable from 1.2–6GHz 830 MB/s bidirectional throughput on up to 15 DMA channels
Reference Clock Source	PXIe-6674T	10MHz reference clock source with < 5 ppb accuracy and 6 configurable I/O connections
Reference Clock Distribution	OctoClock	10MHz 8-channel clock and timing distribution network
Switch	PXIe-1085	Industrial form factor 18-slot chassis, 7 GB/s bidirectional throughput per slot, 2 switches per chassis with inter-switch traffic up to 3.2 GB/s. Links between chassis bound to 7 GB/s bidirectional
Expansion Module	PXIe-8374	PXI Express (x4) Chassis Expansion Module, Software-transparent link without programming, Has a star, tree or daisy-chain configuration

FPGA programming, allowed the development of this scalable and reconfigurable system, which was designed to be able to allow for any sub-6GHz system to be implemented. The picture of the BS is shown in Figure C.1 below.

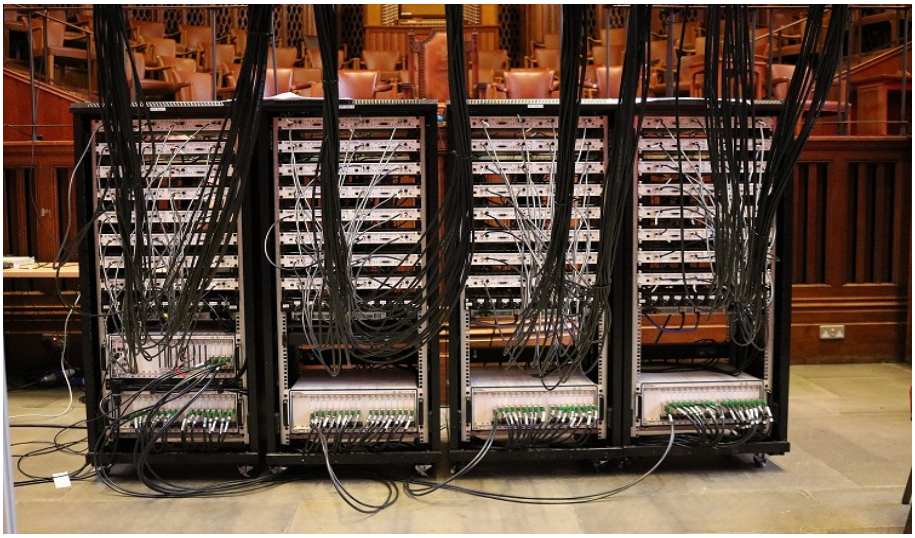


Figure C.1: Bristol Massive MIMO testbed Basestation showing the 4 racks.

The Bristol Massive MIMO testbed serves as a proof of concept and evaluation platform. This means issues like power efficiency and portability were not key aspects of the design. In a real-world commercial deployment, application specific integrated circuits are generally used in

the realisation of the actual commercial products, hence many of the challenges with this testbed are expected to be inherently addressed in a commercial product. The base station (BS) hardware uses 4 chassis of the National Instruments' PXIe-1085 (a rugged version of the PCIe), with 18 slots each. Each chassis is linked back to the central processing chassis, which is a PXIe-8135, which houses 4 FLEX-RIO 7976R co-processors. This central chassis also houses the PXIe-6674T timing card for the reference clock, and the Microsoft Windows controlling computer, which provides the interface to configure and program the system. The design of the Bristol Massive MIMO testbed also considered a distributed deployment, so each of the PXIe-8135 are housed in their own cabinet. Each cabinet should be able to be deployed separately from the others and optical fibre links used to connect the distributed cabinets to the central cabinet that houses the processing unit. This scheme is aligned to the overall goals of the Bristol Is Open (BIO) [194] framework, which has access to the City of Bristol fibre network as shown in Figure C.2 below.

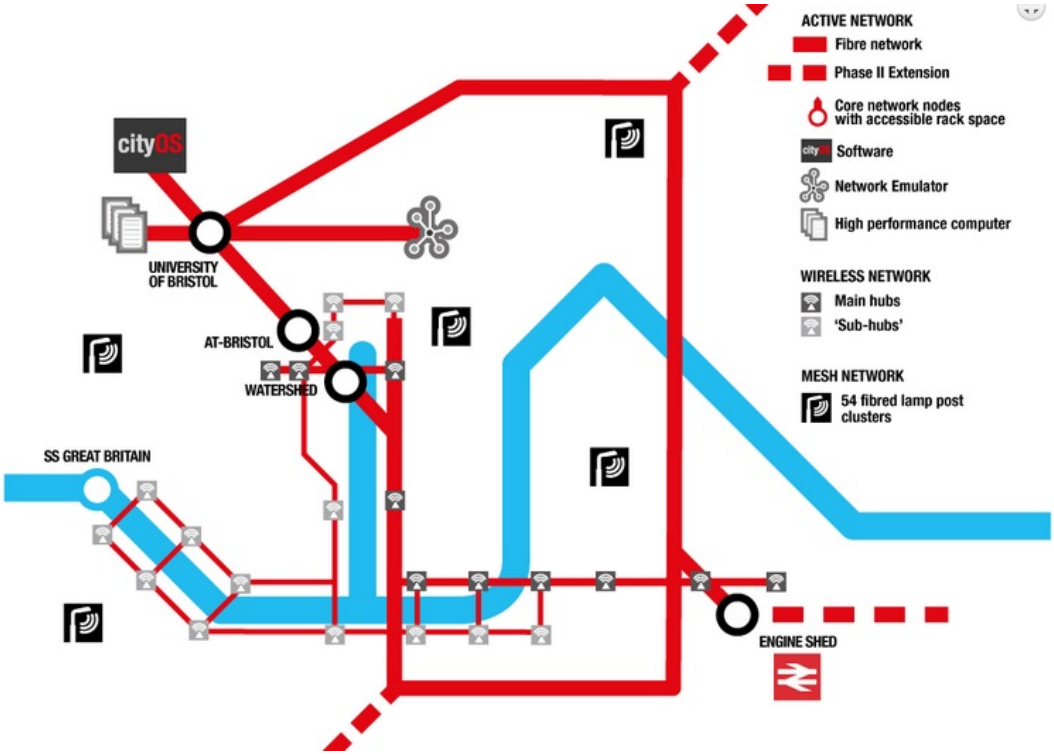


Figure C.2: Bristol Is Open (BIO) Infrastructure.

Each cabinet of the BS holds the 16 Remote Radio Heads (RRHs), which are NI's Universal Software Radio Peripherals with Reconfigurable Input/Out (USRP-RIOs). Each USRP-RIO has an associated programmable Xilinx Kintex-7 K7410T FPGA. Also each USRP-RIO within a cabinet, acts as the RRH for 2 RF chains. A subsystem is defined to consist of 8 URSP-RIOs. This means each cabinet holds 2 subsystems with one performing the function of uplink integration (bandwidth splitting and antenna combining) and another performing downlink

integration (bandwidth combining and antenna splitting). The Xilinx Kintex-7 K7410T FPGA in each USRP-RIO has direct access to the Analogue-to-Digital Converters (ADCs) and the Digital-to-Analogue Converters (DACs) of the device, and these are used to perform the OFDM modulation/demodulation at the front-end. CSI estimation and MIMO processing are performed on the Flex-RIO 7976Rs, which, each have a Digital Signal Processing (DSP) tailored Xilinx Kintex-7 FPGA, whilst visualisation and other functions are performed on the host controller. Figure C.3 shows an overview of the distribution of tasks in the subsystems.

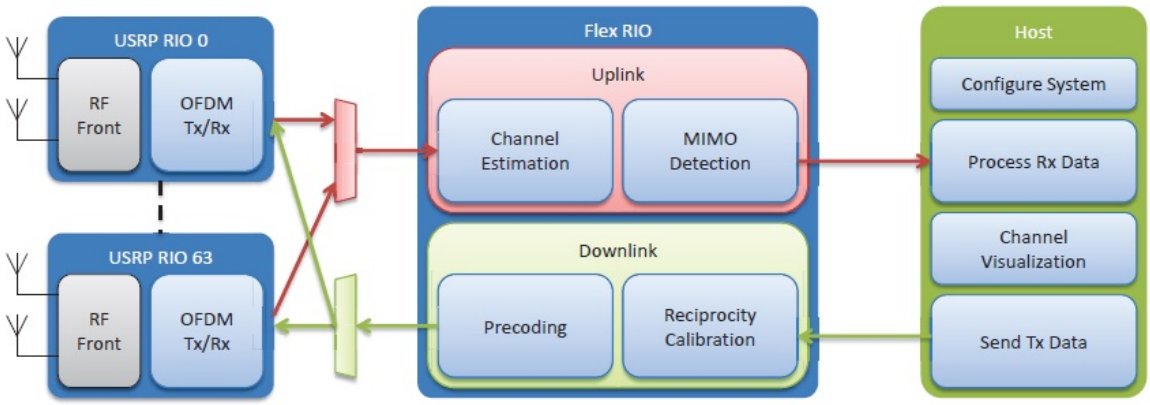


Figure C.3: Distribution of different functions in the subsystems.

The Massive MIMO testbed uses Time Division Duplex (TDD). TDD operation relies on reciprocity of the channel between the Uplink (UL) and the Downlink (DL) and it only requires orthogonal UL pilots from K users. Reciprocity calibration is discussed in section 4.3. All aspects of the air-interface such as the waveform, Modulation Coding Scheme (MCS) options, frame structure and frame timings were made to align as closely as possible with the LTE TDD standard whilst retaining enough flexibility for research purposes. Both UL and DL operations use OFDM on a single 20 MHz carrier sampled at a rate of 30.72 MS/s. It is critical for MIMO operation, that sampling synchronisation is maintained across all the RRHs. This is achieved by having each USRP's Phase- locked Loop (PLL), locked to a common 10 MHz reference clock. A digital trigger is then used to synchronously start each radio at the same point in the frame schedule. The accuracy of the 10 MHz clock is 80 parts-per-billion (ppb) and is produced by an Oven-controlled Crystal Oscillator (OCXO) which is present on the PXIe-6674T timing card. This is then fed to the Ettus Research Octoclock-G distribution module as shown in Figure C.4.

Further information on the design and configuration of the BS hardware can be found on the following publications by the author's colleague, who was the lead for the testbed [8] [100] [195] [196].

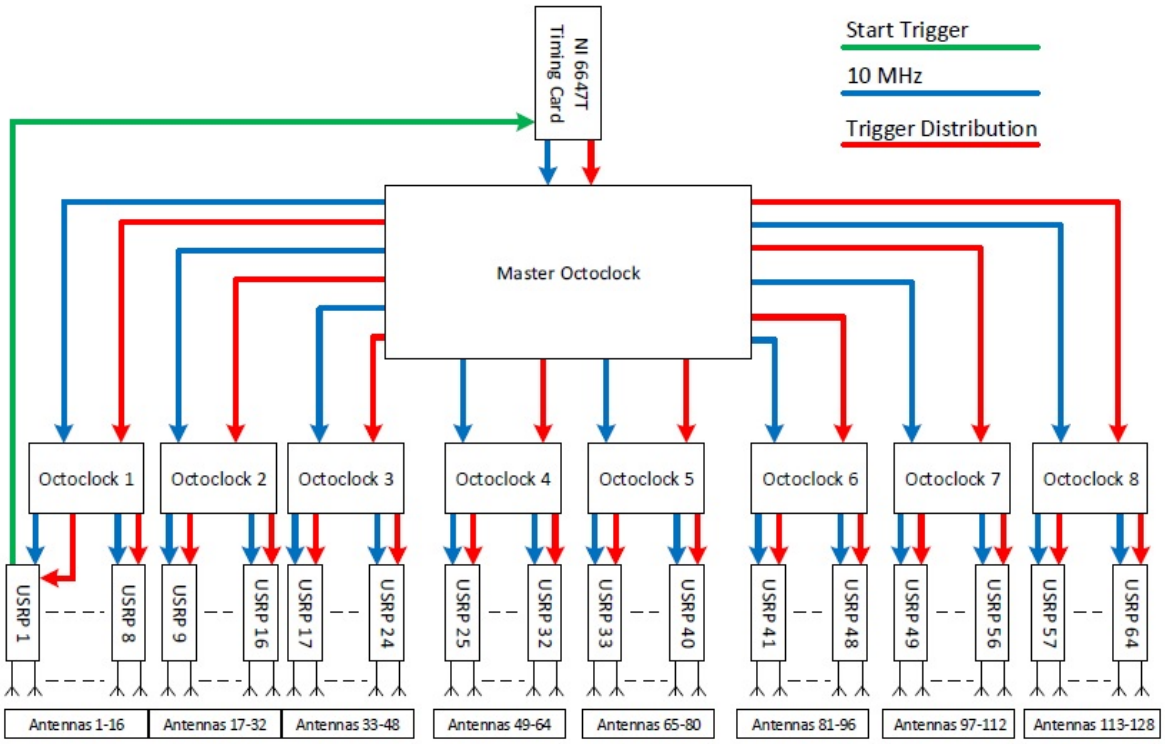


Figure C.4: Distribution of reference clocks and trigger.

C.1.2 User Equipment

The UE setup in all experiments made use of a sleeve dipole antenna. Figure C.5 shows the sleeve dipole antenna and its patterns for both horizontal and vertical polarisation. Table C.3 shows the radiated power and directivity for both polarisations. It can be noted that the sleeve dipole antenna is heavily vertically polarised.

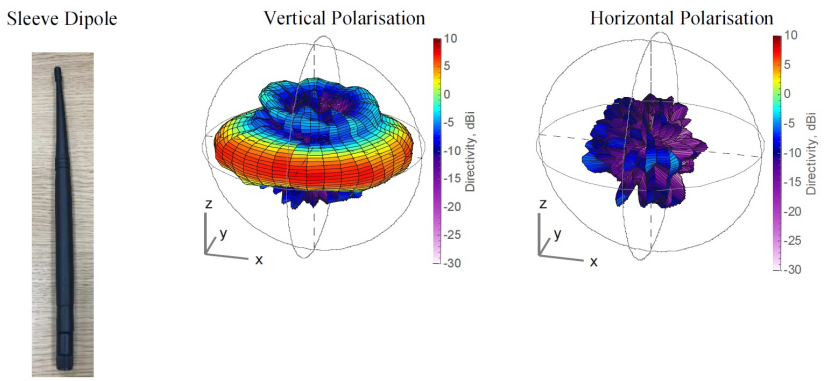


Figure C.5: UE Sleeve Dipole antenna.

Figure C.6: User Equipment showing the USRP and host laptop.

Table C.3: UE antenna statistics

Percentage Power		**Maximum Directivity (dBi)**	
Vertical	Horizontal	Vertical	Horizontal
92%	8%	7.87	-3.77

The mobile clients or UEs for the system are made up of the same model of USRPs that are used as the RRHs in the BS. A single USRP, that is connected to a laptop via a PCIe link, can be configured as either one dual-antenna UE or two single-antenna UEs, depending on specific experimental requirements. Each UE utilises a similar 2-channel OFDM chain as the one used on the RRHs, on its FPGA. Figure C.7 shows a picture of the USRP and Laptop, which form 2 single antenna UEs. For cases where the spatial separation between the 2 antennas, is not adequate, one UE chain can be disabled, and the system can be operated as just one single antenna UE, or a dual antenna UE when both antennas are connected to the single active chain. In most experiments discussed in Chapter 5, only one RF chain with a single antenna was active and it is this configuration which was considered as a single UE. Figure C.8 shows a screenshot of the UE LabVIEW software panel used to monitor and configure the UE, on the host laptop computer.

C.1.3 Synchronisation

As shall become apparent in Chapter 5, the approach to synchronisation, for the purposes of this research, is twofold. For outdoor experiments where the distances between the BS and the UEs is considerably large, Over-The-Air (OTA) synchronisation of the UE devices and the BS, is used. For indoor experiments, the UE devices could be tethered to the BS so that they can share a common clock and timing reference. The available cables are a maximum of 20m, so cables synchronisation could only be used for distances of 20m or less, from the BS. OTA synchronisation is also used in one of the indoor experiments. OTA synchronisation utilises an LTE primary

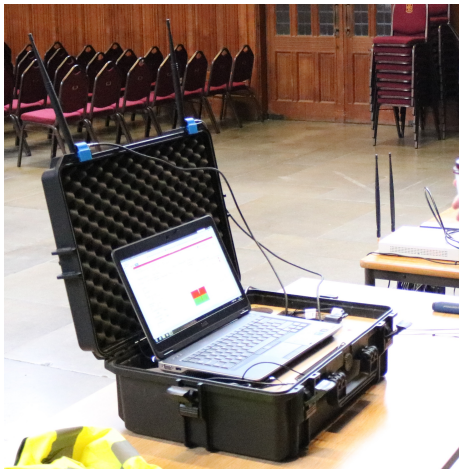


Figure C.7: User Equipment showing the USRP and host laptop.

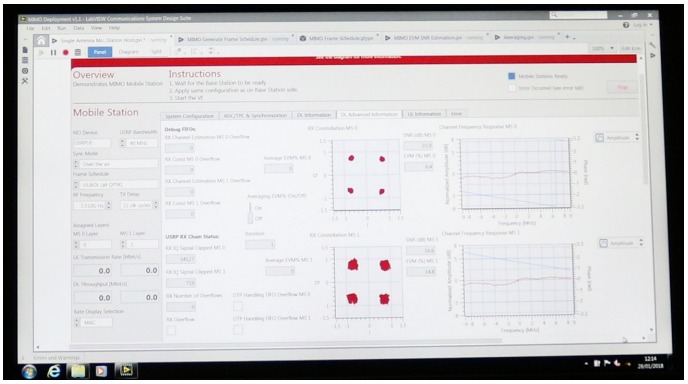


Figure C.8: Screenshot of the UE software on host laptop.

synchronisation signal (PSS), which is implemented using a frequency-shifted bank of replica filters, with carrier frequency offset (CFO) and timing adjustments being made on the FPGA [35].

C.1.4 Channel data capture

The approach followed in the experimentation considered the use of either a single UE, which was moved to different positions within the environment, each time capturing the channel impulse response, or use of multiple UEs fixed at different positions. Those positions are recorded, and the channel measurements are used for direction of arrival estimation at the BS. Each UE transmits an orthogonal UL pilot for CSI estimation. This is achieved by assigning each UE different frequency-orthogonal subcarriers within the 1 OFDM symbol. With a resource block (RB) size of 12 subcarriers and a subcarrier spacing of 15 kHz (in line with LTE), each UE can be assigned 1 frequency-orthogonal pilot per resource block. The QPSK pilot sequence is generated once for 300 subcarriers and concatenated to cover the full 1200 subcarriers that are used by the system. The

estimates for each UE can be interpolated across frequency if anything more than a zero-order hold is desired. Interpolation was applied for the results that are presented in Chapter 5. The fast TDD switching frequency means training more users by adding in more UL pilot symbols is not feasible, although more users can be accommodated by assuming a larger coherent bandwidth and increasing the resource block size. The channel impulse response (CIR) snapshots were captured on the BS host controller by writing 1 UL pilot symbol, the same one used for channel metrics and floating-point MIMO detection, to hard disk.

The massive MIMO software framework was developed by National Instruments and the original code was created using LabVIEW 2014. It was later ported to LabVIEW Communications Suite by NI, leading to the LabVIEW Communications System Design Suite, which is a new integrated development environment (IDE) that offers closer integration with software-defined radio (SDR) technology, in particular, the NI's USRPs, enabling rapid wireless prototyping and algorithm development. More information on the resultant MIMO Application Framework, can be found on NI's published extended white paper [197].

C.2 The 128-element Antenna array

The testbed was designed in such a way that the four cabinets should be able to be separated, each housing radioheads that make up 32 RF chains. The design of the antenna array follows this approach by having sub-panels that can be combined to make a 128-element array and could also be separated into four 32 element panels, each which can be connected to 2 subsystems housed in each cabinet. This allows for distributed deployments. When all the panels are co-located to form a single 128 element array, the joining-up of the 32-element subpanels can be done in any desired permutation of configuration, to explore the impact of azimuth or elevation dominance upon performance in a particular test scenario. Figures C.9 and C.10 shows a single 32-element sub-panel front and back, respectively. The array was designed and prototyped in-house at University of Bristol using Duroid 5880 and the final construction was contracted to Kinnier Dufort.

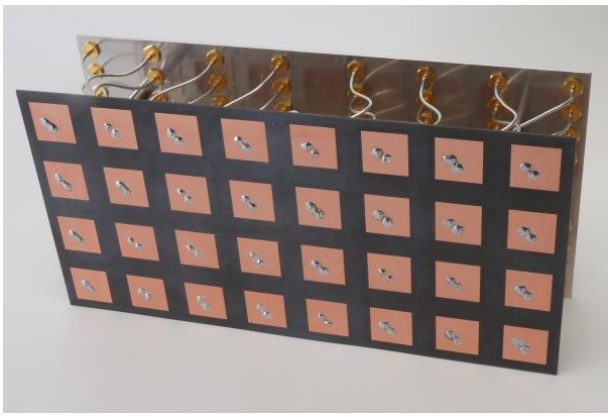


Figure C.9: 32 element sub-panel.

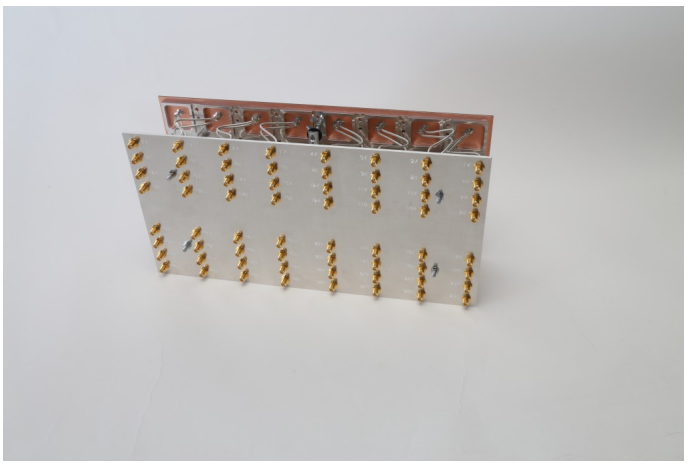


Figure C.10: Rear view of the 32 element sub-panel.

C.2.1 Array design

Duroid 5880 was chosen for the array design, because it offered a high-grade RF substrate which satisfies the requirements for improved dielectric consistency and design repeatability. The array was designed for 3.51GHz, with half wavelength spacing. Figure C.11 shows the S11 measurement for a single patch.

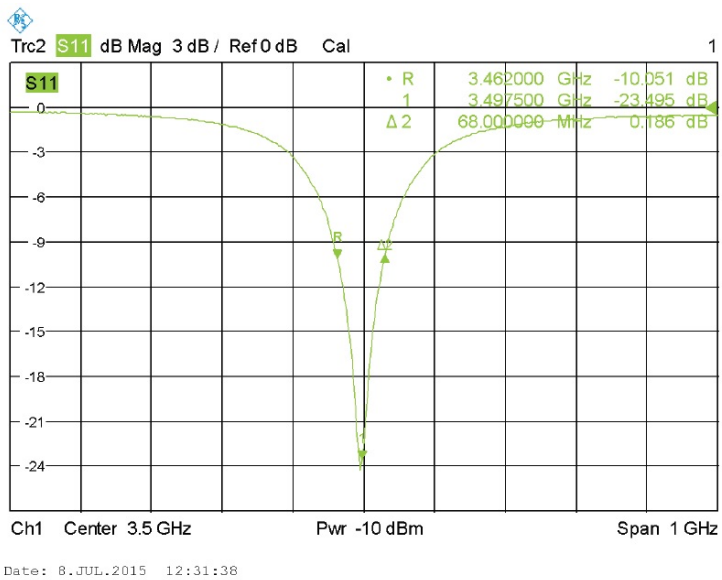


Figure C.11: S11 Measurement for Duroid 5880 patch

Each patch antenna has horizontal (H) and vertical (V) polarisation options to allow for connections using desired polarisation configurations. This was incorporated into the design, in order to reduce the wear and tear on SubMiniature version A (SMA) RF connectors by providing a

fast and flexible way for polarisation reconfiguration of the array. A two-panel sandwich solution, with rear-panel mounted female SMA connectors was chosen to reduce strain on the joints as shown in Figure C.12. To maintain the half wavelength spacing between the adjacent patches on the joint between neighbouring panels, metal spacers are used. On deployments, the sandwich for the whole array is attached to a metal stand using sliding bolts. The experiments relevant for this thesis considered an azimuth dominant configuration as shown in Figure C.13

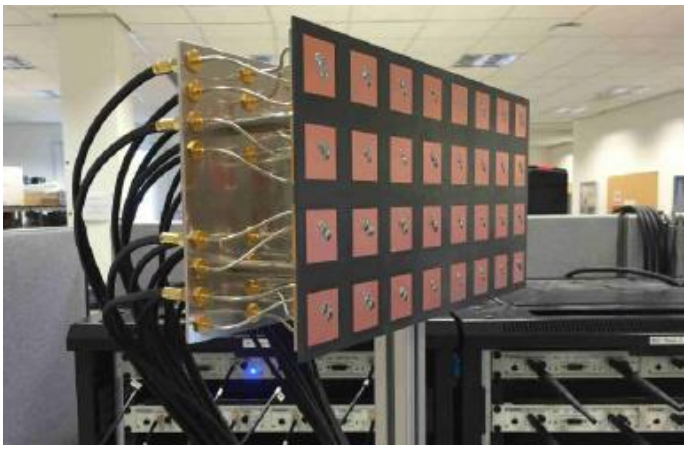


Figure C.12: 2 panel sandwich solution with H and V connections.

Figure C.13: 128 array in an azimuth dominant 4x32 configuration

3D pattern measurements of one 4x8 sub-panel were taken in the University of Bristol's anechoic chamber. Figure C.14 shows the pattern results for an edge and centre antenna elements. The maximum directivity gain was measured to be approximately 6.8 dBi for both above cases.

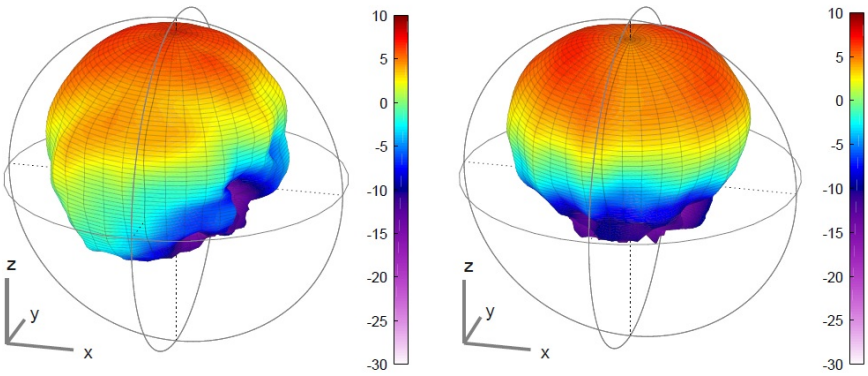


Figure C.14: 3D radiation pattern measurements (Left) left corner element, (Right) right center element

C.3 Limitations and challenges

The system has a bandwidth of 20MHz at a carrier frequency of 3.51GHz which does not provide enough resolution to apply time of arrival techniques. It is generally accepted that such a testbed is not exactly a channel sounding equipment, so trade-offs on accuracy and functionality are expected. This meant that direction of arrival techniques were quite key to exploiting the testbed for any localisation efforts.

The following issues are personal experiences and setup constraints which are not hardware limitations, so they did not affect the results presented but are outlined simply for providing the experiences on the logistical and setup issues experienced by the author. Any data as a result of the issues discussed here was not used, so did not affect the results. The OTA synchronisation was in continual development during the research. On some experiments, some of the modifications made to it were not working well, so the data acquired on those experiments consisted of captures where synchronisation was broken. Pre-processing was carried out in order to determine and exclude such data.

Conducting experiments using the testbed comprised of moving heavy equipment to the location, cabling of the antenna array to the BS radio chains and setting up of the UEs. For a team of about 5 researchers, each of whom was using the testbed differently according to their research, these experiments had to be coordinated. Because different experiments were conducted with colleagues at the same time, this meant that there were limitations to the setup, which applied to this particular testbed. Experiments contained here, had to accommodate colleagues' experiments, hence the setup and configuration had to take that into account. Every experiment required its own set of code for capturing the channel impulse response, which corresponded with particular wiring connections and configuration of the BS on the day.

Processing of data produced by the testbed has been done offline owing to the fact that the

LabVIEW framework was not designed with third-part applications in mind, so modifications of the code would be necessary to accommodate our application, and also how well the testbed design fits with the computational intensity required of the DOA estimation algorithm used in this thesis, was not evaluated. These pieces of work fell out of the scope of this research.

The testbed was designed with a distributed deployment capability. This deployment was not tested or used because it required activation and configuration of the equipment and network to support it. Distributed deployment would have allowed the BS to be split into 4 sub arrays, but with all radio chains still sharing the same clock. This would have allowed DOA estimation at each sub-system, regardless of how far these sub arrays were separated. If the sub-arrays covered the same location, then triangulation would be possible and the localisation of the UEs within that environment becomes possible. As shall be presented in Chapter 5, localisation via triangulation, using the collocated system, was tested in an indoor environment, and the results are presented.

C.4 Summary

This appendix has covered the design of the Bristol University's Massive MIMO testbed. The author was involved in the initial system testing. They conducted array measurements in Bristol University's anechoic chamber. They also conducted experiments to characterise the receive radio chains on the BS. Preliminary experimentation without any manifold and calibration, had not achieved DOA estimation because the system was not initially designed to do this, so it validated the need for the manifold and calibration to get the DOA estimation right. Chapter 5 looks at the specific experiments that were conducted to generate the data that was used for direction of arrival estimation.

APPENDIX D

ANTENNA ARRAY AND BS RADIO CHAINS CHARACTERISATION

This appendix provides additional details about the characterisation of the antenna array. It provides more data on the array measurements that were conducted in the anechoic chamber, and the validation of that data. It has been discussed in Chapter C that for each elements, the vertical polarisation was used. Also because the analysis in this thesis focuses only on the azimuth, the extracted data only uses a fixed elevation angle of 90 degrees. For each element the range measurements were validated by comparing the first set of data to the last set of data. In a linear scale of 0 degrees to 180 degrees elevation, the first measurement (at 0 degrees) is expected to be highly correlated with the last measurement at 180 degrees when inverted. All measured elements were checked to see if that was the case and indeed those measurements were correlated as shown in Figures D.1, D.2 and D.3.

Element measurements were also compared to check correlation between the corresponding elements as described in section 5.1. The results validated the approach of utilising corresponding measurements in place of those elements that were not measured. Figures D.4 and D.5 below demonstrate that correlation in addition to Figure 5.3 provided in section 5.1.

Samples of the 3D radiation patterns is provided below in Figures D.6 to D.11, with information such as the gain and the directivity of each element. The vertical polarisation port for each element was used for these measurements. All measured elements had an efficiency of 100%. To explore the uncertainties in the radio chains, an experiment was conducted to have one of the UEs transmitting to the receiver radios directly via cable. An 8-port splitter was used to connect to the BS radios. The first port was connected to the first radio, and the other 7 ports were connected to the rest of the radios, sequentially, each time recording the channel impulse response. The first receiver radio connection was maintained for normalisation. Over-The -Air synchronisation was used, both for simplicity and also to measure its effects. 100 CIR captures for each antenna were used to build the average magnitude response and also the

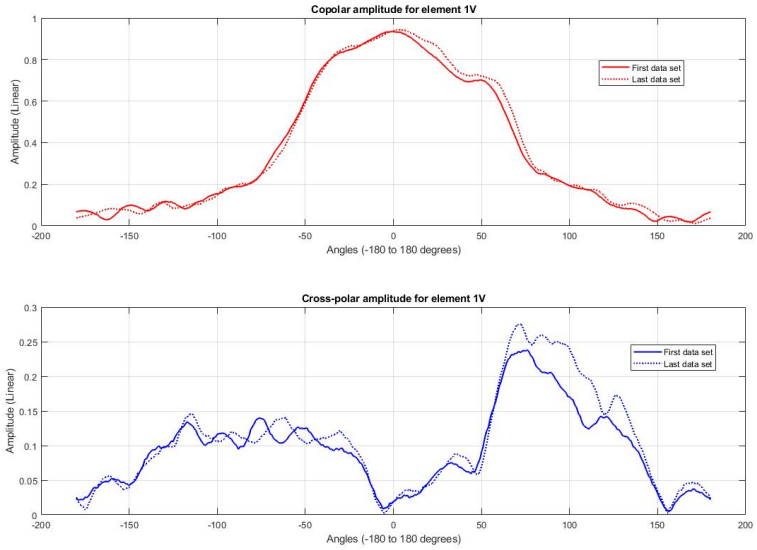


Figure D.1: Amplitude range measurements for element 1

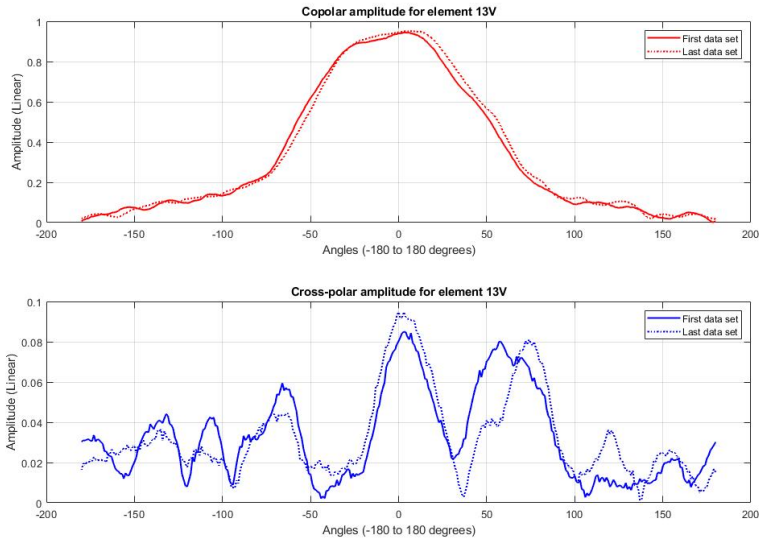


Figure D.2: Amplitude range measurements for element 13

power delay profiles. An average maximum magnitude difference of 3dB was recorded. Timing jitter was observed in the measurements which could be explained by errors in the over-the-air synchronisation scheme that was in use at the time. Figure D.12 shows the Power delay profiles for all 128 receivers.

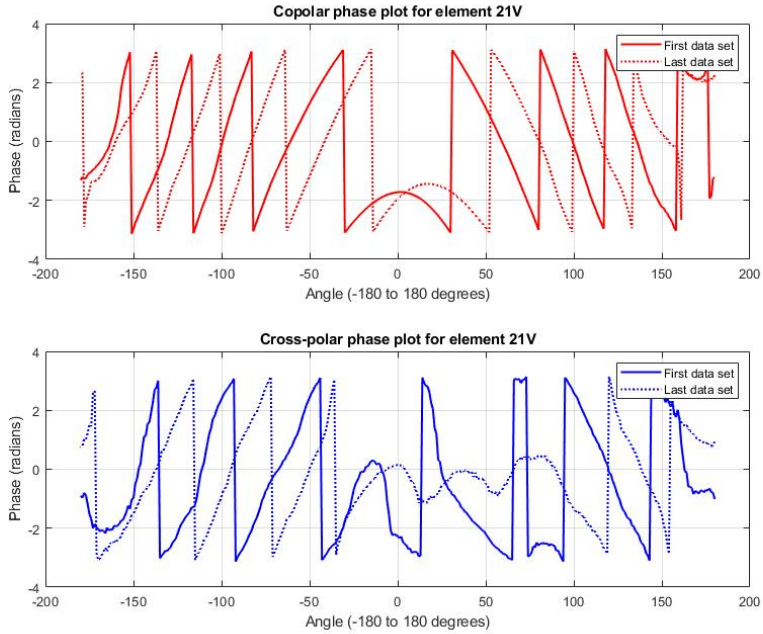


Figure D.3: Phase range measurements for element 21

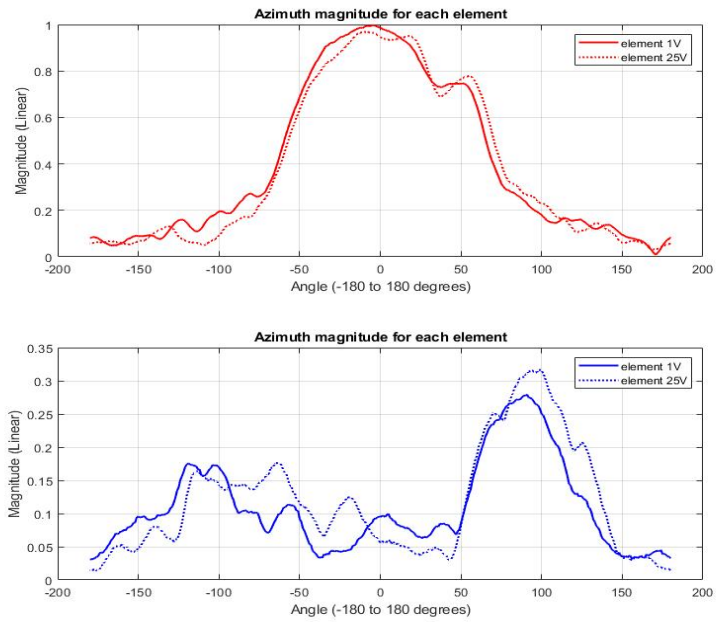


Figure D.4: Amplitude range measurements for elements 1 and 25

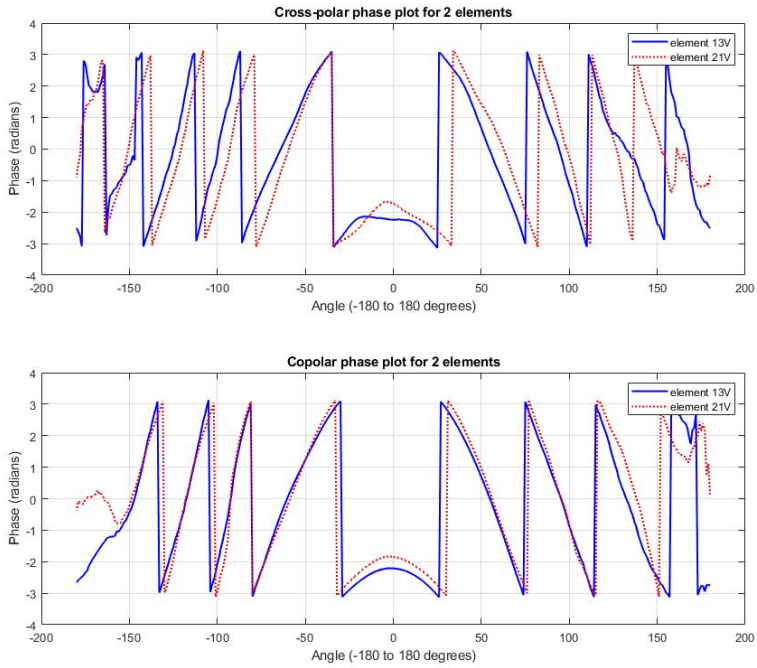


Figure D.5: Phase range measurements for elements 13 and 21

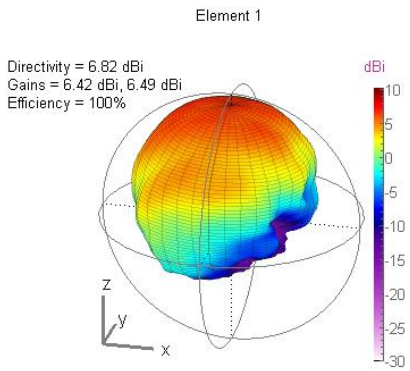


Figure D.6: 3D radiation pattern for element 1

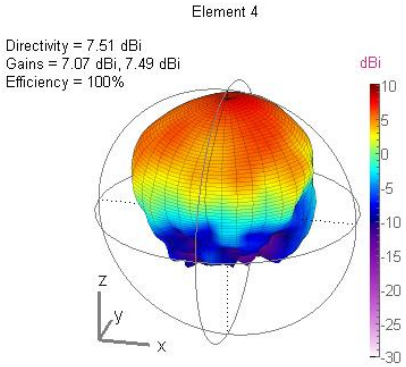


Figure D.7: 3D radiation pattern for element 4

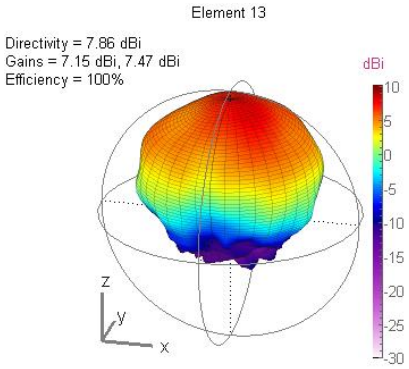


Figure D.8: 3D radiation pattern for element 13

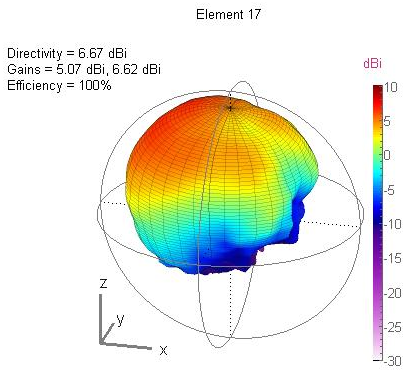


Figure D.9: 3D radiation pattern for element 17

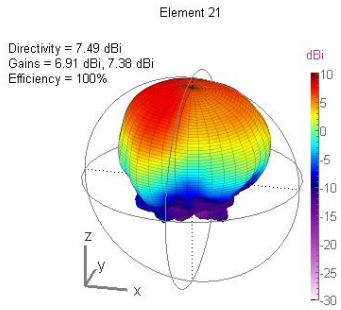


Figure D.10: 3D radiation pattern for element 21

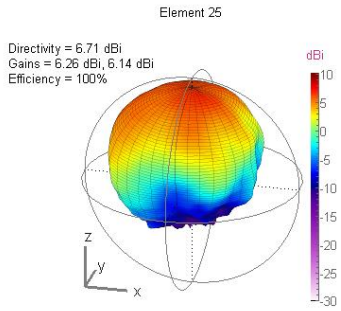


Figure D.11: 3D radiation pattern for element 25

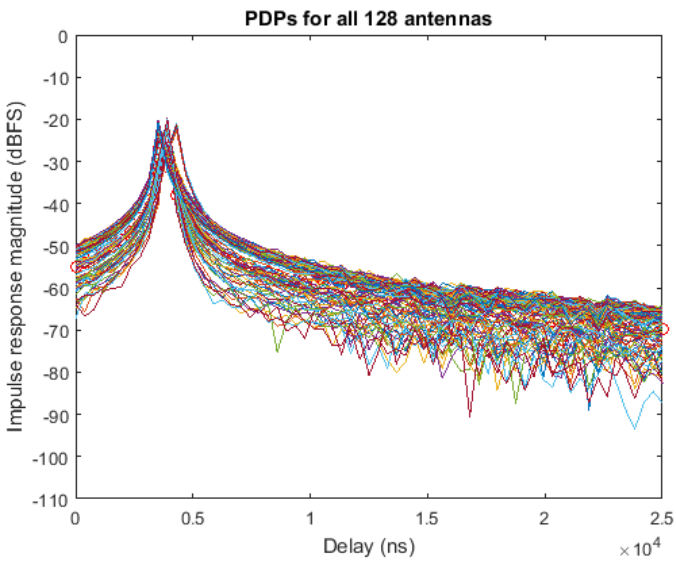


Figure D.12: Power delay profiles for all 128 receivers

APPENDIX E

BT TRIAL

This appendix give an overview of a preliminary DOA experimentation that was conducted during a Massive MIMO trial in collaboration with British Telecom (BT) [198]. Massive MIMO field trials were conducted at BT Labs in Adastral Park, Suffolk, UK. Press articles on the trials are available online [199] [200] [201]. A preliminary assessment of the DOA estimation capability was conducted. The BS was set up on the indoor stage of a large hall. A site-specific BS array manifold was measured using a single UE. The UE was moved from -60 degree position in increments of 10 degrees up to the +60 degree mark in an arc, 10 meters away from the center of the BS array as shown in Figure E.3. On each position, the UE calibration transmitted a pilot signal and the channel response is measured at the base station. This data was then processed to build the array manifold, which was then used in the DOA processing. Figures E.1 and E.2 show the BS setup and the UE during measurement respective. The same UE was used then moved two other position further away from the BS. Figure E.3 shows the complete setup. The average angle errors of 11 and and 14 degrees for the two positions was obtained. This suggested that a n improved manifold measurement should improve the accuracy. Further measurements were taken in the work that is described in Chapter C.

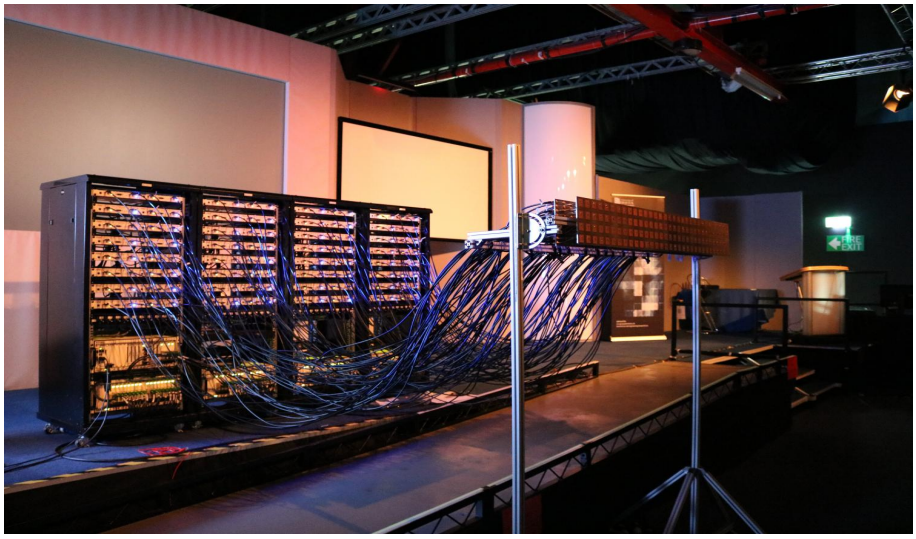


Figure E.1: BS setup in the hall.

Figure E.2: UE setup.

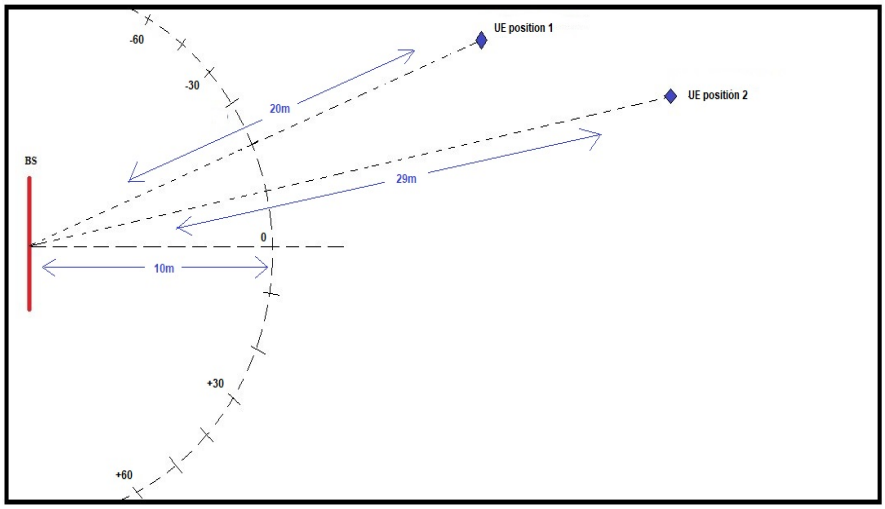


Figure E.3: DOA experiment with a custom manifold.

References

[1] Fortune Business Insights, "Report ID FBI101060 - https://www.fortunebusinessinsights.com/industry-reports/location-based-services-market-101060," December 2019.

[2] M. Thorpe, M. Kottkamp, A. Rössler, and J. Schütz, "LTE Location Based Services: Technology Introduction, White Paper," 2013.

[3] FCC, "https://www.fcc.gov/document/wireless-e911-location-accuracy-requirements-3," November 2018.

[4] P. Misra and P. Enge, *The Global Positioning System: Signals, measurements, and performance*. Ganga-Jamuna Press, 2011.

[5] GSA, "https://www.gsa.europa.eu/newsroom/news/results-are-galileo-increases-accuracy-\location-based-services," November 2018.

[6] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); LTE Positioning Protocol (LPP); (3GPP TS36.355 Release 9)," 2010.

[7] O. Edfors, L. Liu, F. Tufvesson, N. Kundargi, and K. Nieman, *Massive MIMO for 5G: Theory, Implementation and Prototyping*. John Wiley Sons, Ltd, 2016, ch. 9, pp. 189–230. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119116493.ch9

[8] P. Harris, W. B. Hasan, H. Brice, B. Chitambira, M. Beach, E. Mellios, A. Nix, S. Armour, and A. Doufexi, "An overview of massive mimo research at the university of bristol," in *Radio Propagation and Technologies for 5G (2016)*, Oct 2016, pp. 1–5.

[9] S. Malkowsky, J. Vieira, L. Liu, P. Harris, K. Nieman, N. Kundargi, I. C. Wong, F. Tufvesson, V. Öwall, and O. Edfors, "The world's first real-time testbed for massive mimo: Design, implementation, and validation," *IEEE Access*, vol. 5, pp. 9073–9088, 2017.

[10] HUAWEI, "Road to a Super Connected World; (Mobile World Congress Keynote)," December 2019.

[11] Samsung, "https://www.samsung.com/global/galaxy/what-is/5g-speeds/," November 2019.

[12] CISCO, "Mobile Traffic Approaching The Zettabute Milestone; (CISION PR Newswire)," February 2019.

[13] WORLD ECONOMIC FORUM, "https://www.weforum.org/agenda/2018/01/the-world-is-about-to-become-even-more-interconnected-here-s-how/," January 2018.

[14] INSTITUTE FOR DIGITAL COMMUNICATIONS, "http://www.idc.lnt.de/en/forschung/massive-mimo-systems/," November 2018.

[15] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, November 2010.

[16] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive mimo for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, February 2014.

[17] A. Kaye and D. George, "Transmission of multiplexed pam signals over multiple channel and diversity systems," *IEEE Transactions on Communication Technology*, vol. 18, pp. 520–526, October 1970.

[18] G. J. Foschini, "Transmission of multiplexed pam signals over multiple channel and diversity systems," *Bell Labs Technical Journal*, vol. 1, pp. 41–51, October 1996.

[19] G. D. Golden, G. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, "Transmission of multiplexed pam signals over multiple channel and diversity systems," *Electronics Letters*, vol. 35, pp. 14–16, January 1999.

[20] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, p. 311–335, March 1998. [Online]. Available: https://doi.org/10.1023/A:1008889222784

[21] E. Telatar, "Capacity of multiantenna gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, November 1999.

[22] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive mimo: Benefits and challenges," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 742–758, Oct 2014.

[23] D. Gesbert, M. Kountouris, R. W. H. Jr., C. b. Chae, and T. Salzer, "Shifting the mimo paradigm," *IEEE Signal Processing Magazine*, vol. 24, pp. 36–46, September 2007.

[24] T. L. Narasimhan and A. Chockalingam, "Channel hardening-exploiting message passing (chemp) receiver in large-scale mimo systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, pp. 847–860, October 2014.

[25] H. Halbauer, A. Weber, D. Wiegner, and T. Wild, "Energy efficient massive mimo array configurations," in *2018 IEEE Globecom Workshops (GC Wkshps)*, 2018, pp. 1–6.

[26] Z. Wang, C. Qian, L. Dai, J. Chen, C. Sun, and S. Chen, "Location-based channel estimation and pilot assignment for massive mimo systems," in *2015 IEEE International Conference on Communication Workshop (ICCW)*, June 2015, pp. 1264–1268.

[27] D. Fan, F. Gao, G. Wang, Z. Zhong, and A. Nallanathan, "Angle domain signal processing-aided channel estimation for indoor 60-ghz tdd/fdd massive mimo systems," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 9, pp. 1948–1961, Sept 2017.

[28] Z. Yuan, G. Yu, W. Li, Y. Yuan, X. Wang, and J. Xu, "Multi-user shared access for internet of things," in *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, May 2016, pp. 1–5.

[29] H. Nikopour and H. Baligh, "Sparse code multiple access," in *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Sep. 2013, pp. 332–336.

[30] R. Hoshyar, F. P. Wathan, and R. Tafazolli, "Novel low-density signature for synchronous cdma systems over awgn channel," *IEEE Transactions on Signal Processing*, vol. 56, no. 4, pp. 1616–1626, April 2008.

[31] S. Verdu, *Multiuser Detection*, 1st ed. Cambridge University Press, 1998.

[32] A. G. Burr, "Capacity bounds and estimates for the finite scatterers mimo wireless channel," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 812–818, June 2003.

[33] H. Xie, F. Gao, S. Zhang, and S. Jin, "A unified transmission strategy for tdd/fdd massive mimo systems with spatial basis expansion model," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 4, pp. 3170–3184, April 2017.

[34] X. Liu, J. Sha, H. Xie, F. Gao, S. Jin, Z. Zhang, X. You, and C. Zhang, "Efficient channel estimator with angle-division multiple access," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 66, no. 2, pp. 708–718, Feb 2019.

[35] W. B. Hasan, *On Evaluating the Practicality of Using Massive MIMO in Real-World Scenarios and Identifying Solutions to Operational Deployments*. University of Bristol, 2019.

[36] W. B. Hasan, P. Harris, H. Bromell, A. Doufexi, and M. Beach, "Adaptive user grouping based on evm prediction for efficient robust massive mimo in tdd," *IEEE Access*, vol. 7, pp. 162 683–162 696, 2019.

[37] B. Chitambira, S. Armour, S. Wales, and M. Beach, "Nlos identification and mitigation for geolocation using least-squares support vector machines," in *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, March 2017, pp. 1–6.

[38] H. Li, S. Habibi, and G. Ascheid, "Handover prediction for long-term window scheduling based on sinr maps," in *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2013, pp. 917–921.

[39] J. M. R. Jerez, M. R. Garcia, and A. D. Estrella, "Effects of power control errors and multipath fading on ber in a cellular cdma system," in *Vehicular Technology Conference Fall 2000. IEEE VTS Fall VTC2000. 52nd Vehicular Technology Conference (Cat. No.00CH37152)*, vol. 6, Sept 2000, pp. 2593–2598 vol.6.

[40] T. K. Sarkar, Z. Ji, K. Kim, A. Medouri, and M. Salazar-Palma, "A survey of various propagation models for mobile communication," *IEEE Antennas and Propagation Magazine*, vol. 45, no. 3, pp. 51–82, June 2003.

[41] C. Phillips, D. Sicker, and D. Grunwald, "A survey of wireless path loss prediction and coverage mapping methods," *IEEE Communications Surveys Tutorials*, vol. 15, no. 1, pp. 255–270, First 2013.

[42] F. Diggelen, *A-GPS: Assisted GPS, GNSS, and SBAS*, March 2009.

[43] Y. He and A. Bilgic, "Iterative least squares method for global positioning system," *Advances in Radio Science*, vol. 9, pp. 203–208, August 2011.

[44] A. Tian, D. Dong, D. Ning, and C. Fu, "Gps single point positioning algorithm based on least squares," in *2013 Sixth International Symposium on Computational Intelligence and Design*, vol. 2, 2013, pp. 16–19.

[45] ITU, "Comparison of Time-Difference-of-Arrival and Angle-of-Arrival Methods of signal geolocation, https://www.itu.int/pub/R-REP-SM.2211-2-2018," June 2018.

[46] K. J. Krizman, T. E. Biedka, and T. S. Rappaport, "Wireless position location: fundamentals, implementation strategies, and sources of error," in *1997 IEEE 47th Vehicular Technology Conference. Technology in Motion*, vol. 2, May 1997, pp. 919–923.

[47] G. Aruna and S. Bhattacharyya, "Capacity and snr analysis of a millimeter wave cellular network with phased antenna array," in *2018 International Conference on Computer Communication and Informatics (ICCCI)*, 2018, pp. 1–5.

[48] S. S. Moghaddam and F. Akbari, "Efficient narrowband direction of arrival estimation based on a combination of uniform linear/shirvani-akbari arrays," *International Journal of Antennas and Propagation*, vol. vol, 2012.

[49] Z. Matoušek, M. Babjak, and J. Ochodnický, "Correlation-based tdoa algorithm for the target localization at low signal to noise ratio," in *2016 New Trends in Signal Processing (NTSP)*, 2016, pp. 1–5.

[50] ITU-R, *ITU Handbook on Spectrum Monitoring*, 2011.

[51] O. A. Oumar, M. F. Siyau, and T. P. Sattar, "Comparison between music and esprit direction of arrival estimation algorithms for wireless communication systems," in *The First International Conference on Future Generation Communication Technologies*, 2012, pp. 99–103.

[52] M. Compagnoni, A. Pini, A. Canclini, P. Bestagini, F. Antonacci, S. Tubaro, and A. Sarti, "A geometrical–statistical approach to outlier removal for tdoa measurements," *IEEE Transactions on Signal Processing*, vol. 65, no. 15, pp. 3960–3975, 2017.

[53] G. Fokin, "Tdoa measurement processing for positioning in non-line-of-sight conditions," in *2018 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)*, 2018, pp. 1–5.

[54] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero, R. L. Moses, and N. S. Correal, "Locating the nodes: cooperative localization in wireless sensor networks," *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 54–69, July 2005.

[55] D. J. Torrieri, "Statistical theory of passive location systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-20, no. 2, pp. 183–198, March 1984.

[56] N. Patwari, A. O. Hero, M. Perkins, N. S. Correal, and R. J. O'Dea, "Relative location estimation in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 51, no. 8, pp. 2137–2148, Aug 2003.

[57] J. D. Moell and T. N. Curlee, *Transmitter Hunting: Radio Direction Finding Simplified*. McGraw Hill Professional, 1987, vol. 2701.

[58] P. Q. C. Ly, S. D. Elton, and D. A. Gray, "Aoa estimation of two narrowband signals using interferometry," in *2010 IEEE International Symposium on Phased Array Systems and Technology*, 2010, pp. 1004–1009.

[59] C. Mensing and S. Plass, "Positioning algorithms for cellular networks using tdoa," in *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, vol. 4, 2006, pp. IV–IV.

[60] M. A. Landolsi, H. R. Khan, A. S. Al-Ahmari, and A. H. Muqaibel, "On the performance of wireless sensor network time difference of arrival localization under realistic interference and timing estimation errors," *International Journal of Distributed Sensor Networks*, vol. 15, no. 2, p. 1550147719832537, 2019. [Online]. Available: https://doi.org/10.1177/1550147719832537

[61] H. Seuté, C. Enderli, J. Grandin, A. Khenchaf, and J. Cexus, "Experimental measurement of time difference of arrival," in *2016 17th International Radar Symposium (IRS)*, 2016, pp. 1–4.

[62] V. Kozmin, A. Pavlyuk, and A. Tokarev, "Comparison of spectrum monitoring coverage features of aoa and tdoa geolocation methods," *Electrosviaz*, no. 2, 2014.

[63] Y. R. Hamdy and S. A. Mawjoud, "Performance assessment of u-tdoa and a-gps positioning methods," in *2012 International Conference on Future Communication Networks*, April 2012, pp. 99–104.

[64] G. Carter, "Time delay estimation for passive sonar signal processing," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 29, no. 3, pp. 463–470, June 1981.

[65] W. Hahn and S. Tretter, "Optimum processing for delay-vector estimation in passive signal arrays," *IEEE Transactions on Information Theory*, vol. 19, no. 5, pp. 608–614, Sep. 1973.

[66] J. S. Abel and J. O. Smith, "Source range and depth estimation from multipath range difference measurements," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 8, pp. 1157–1165, Aug 1989.

[67] B. T. Fang, "Simple solutions for hyperbolic and related position fixes," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 26, no. 5, pp. 748–753, Sep. 1990.

[68] W. H. FOY, "Position-location solutions by taylor-series estimation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-12, no. 2, pp. 187–194, March 1976.

[69] Y. T. Chan and K. C. Ho, "A simple and efficient estimator for hyperbolic location," *IEEE Transactions on Signal Processing*, vol. 42, no. 8, pp. 1905–1915, Aug 1994.

[70] B. Xu, R. Yu, G. Sun, and Z. Yang, "Whistle: Synchronization-free tdoa for localization," in *2011 31st International Conference on Distributed Computing Systems*, June 2011, pp. 760–769.

[71] C. Gentile, N. Alsindi, R. Raulefs, and C. Teolis, *Ranging and Localization in Harsh Multipath Environments*, 11 2013.

[72] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1993.

[73] Yihong Qi, H. Kobayashi, and H. Suda, "Analysis of wireless geolocation in a non-line-of-sight environment," *IEEE Transactions on Wireless Communications*, vol. 5, no. 3, pp. 672–681, March 2006.

[74] M. A. Spirito, "On the accuracy of cellular mobile station location estimation," *IEEE Transactions on Vehicular Technology*, vol. 50, no. 3, pp. 674–685, May 2001.

[75] V. Pierlot and M. V. Droogenbroeck, "A new three object triangulation algorithm for mobile robot positioning," *IEEE Transactions on Robotics*, vol. 30, no. 3, pp. 566–577, June 2014.

[76] D. Pedoe. The Mathematical Association of America; 2nd Edition, 1997.

[77] C.-Y. Wen and F.-K. Chan, "Adaptive aoa-aided toa self-positioning for mobile wireless sensor networks," *Sensors*, vol. 10, no. 11, pp. 9742–9770, 2010. [Online]. Available: https://www.mdpi.com/1424-8220/10/11/9742

[78] B. Zhou, L. Ai, X. Dong, and L. Yang, "Doa-based rigid body localization adopting single base station," *IEEE Communications Letters*, vol. 23, no. 3, pp. 494–497, 2019.

[79] C.-S. Chen, Y.-J. Chiu, and J.-M. Lin, "Hybrid toa / aoa schemes for mobile location in cellular communication systems," vol. 4, no. 2, 2010, pp. 287–291.

[80] Y. Wang and K. C. Ho, "Tdoa source localization in the presence of synchronization clock bias and sensor position errors," *IEEE Transactions on Signal Processing*, vol. 61, no. 18, pp. 4532–4544, Sep. 2013.

[81] S. S. C. Mensing and A. Dammann, "Hybrid data fusion and tracking for positioning with gnss and 3gpp-lte," *International Journal of Navigation and Observation*, vol. 2010, no. 812945, June 2010.

[82] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, Feb 2002.

[83] M. Laaraiedh, S. Avrillon, and B. Uguen, "Hybrid data fusion techniques for localization in uwb networks," in *2009 6th Workshop on Positioning, Navigation and Communication*, March 2009, pp. 51–57.

[84] M. C. R. W. Ouyang, A. K.-S. Wong, "On data fusion for parametric-model-based wireless localization," *International Journal of Wireless Information Networks*, vol. 19, no. 1, pp. 22–37, March 2012.

[85] V. Y. Zhang, A. K. Wong, K. T. Woo, and R. W. Ouyang, "Hybrid toa/aoa-based mobile localization with and without tracking in cdma cellular networks," in *2010 IEEE Wireless Communication and Networking Conference*, April 2010, pp. 1–6.

[86] J. Prieto, S. Mazuelas, A. Bahillo, P. Fernandez, R. M. Lorenzo, and E. J. Abril, "Adaptive data fusion for wireless localization in harsh environments," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 1585–1596, April 2012.

[87] J. Falk, P. Händel, and M. Jansson, "Direction finding for electronic warfare systems using the phase of the cross spectral density," in *RadioVetenskap och Kommunikation (RVK) :*, 2002, pp. 264–268.

[88] A. Hero, H. Messer, J. Goldberg, D. J. Thomson, M. G. Amin, G. Giannakis, A. Swami, J. K. Tugnait, A. Nehorai, A. L. Swindlehurst, J. . Cardoso, L. Tong, and J. Krolik, "Highlights of statistical signal and array processing," *IEEE Signal Processing Magazine*, vol. 15, no. 5, pp. 21–64, Sep. 1998.

[89] C. Carlemalm, S. Halvarsson, T. Wigren, and B. Wahlberg, "Algorithms for time delay estimation using a low complexity exhaustive search," *IEEE Transactions on Automatic Control*, vol. 44, no. 5, pp. 1031–1037, May 1999.

[90] L. Ljung, "Identification for control: simple process models," in *Proceedings of the 41st IEEE Conference on Decision and Control, 2002.*, vol. 4, Dec 2002, pp. 4652–4657.

[91] A. Elnaggar, G. A. Dumont, and A. . Elshafei, "Recursive estimation for system of unknown delay," in *Proceedings of the 28th IEEE Conference on Decision and Control,*, Dec 1989, pp. 1809–1810 vol.2.

[92] G. Ferretti, C. Maffezzoni, and R. Scattolini, "Recursive estimation of time delay in sampled systems," *Automatica*, vol. 27, no. 4, pp. 653 – 661, 1991.

[93] C. L. Nikias and J. M. Mendel, "Signal processing with higher-order spectra," *IEEE Signal Processing Magazine*, vol. 10, no. 3, pp. 10–37, July 1993.

[94] Y. Kurita, T. Hashimoto, and Y. Ishida, "An application of time delay estimation by anns to frequency domain adaptive i-pd controller," in *IJCNN'99. International Joint Conference on Neural Networks. Proceedings (Cat. No.99CH36339)*, vol. 3, July 1999, pp. 2164–2167 vol.3.

[95] P. C. Ching, H. C. So, and S. Q. Wu, "On wavelet denoising and its applications to time delay estimation," *IEEE Transactions on Signal Processing*, vol. 47, no. 10, pp. 2879–2882, Oct 1999.

[96] C. Knapp and G. Carter, "The generalized correlation method for estimation of time delay," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 24, no. 4, pp. 320–327, August 1976.

[97] C. Stoeckle, J. Munir, A. Mezghani, and J. A. Nossek, "Doa estimation performance and computational complexity of subspace- and compressed sensing-based methods," in *WSA 2015; 19th International ITG Workshop on Smart Antennas*, 2015, pp. 1–6.

[98] V. Vasylyshyn, "Direction of arrival estimation using esprit with sparse arrays," in *2009 European Radar Conference (EuRAD)*, 2009, pp. 246–249.

[99] Z. Yang, J. Li, P. Stoica, and L. Xie, "Sparse methods for direction-of-arrival estimation," 2017.

[100] P. Harris, S. Zang, A. Nix, M. Beach, S. Armour, and A. Doufexi, "A distributed massive mimo testbed to assess real-world performance and feasibility," in *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*, May 2015, pp. 1–2.

[101] G. Dudek and M. Jenkin, *Computational Principles of Mobile Robotics*, 2nd ed. New York, NY, USA: Cambridge University Press, 2010.

[102] A. D. Lonkeng and J. Zhuang, "Two-dimensional doa estimation using arbitrary arrays for massive mimo system," *International Journal of Antennas and Propagation*, vol. 2017, December 2017.

[103] J. Kelner and C. Ziółkowski, *Multi-Elliptical Geometry of Scatterers in Modeling Propagation Effect at Receiver*, 09 2018, pp. 115–141.

[104] X. Wei, N. Palleit, and T. Weber, "Aod/aoa/toa-based 3d positioning in nlos multipath environments," in *2011 IEEE 22nd International Symposium on Personal, Indoor and Mobile Radio Communications*, Sep. 2011, pp. 1289–1293.

[105] K. H. Ng, E. K. Tameh, A. Doufexi, M. Hunukumbure, and A. R. Nix, "Efficient multielement ray tracing with site-specific comparisons using measured mimo channel data," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 3, pp. 1019–1032, May 2007.

[106] E. K. Tameh and A. R. Nix, "The use of measurement data to analyse the performance of rooftop diffraction and foliage loss algorithms in a 3-d integrated urban/rural propagation model," in *VTC '98. 48th IEEE Vehicular Technology Conference. Pathway to Global Wireless Revolution (Cat. No.98CH36151)*, vol. 1, May 1998, pp. 303–307.

[107] T. A. Thomas, F. W. Vook, E. Mellios, G. S. Hilton, A. R. Nix, and E. Visotsky, "3d extension of the 3gpp/itu channel model," in *2013 IEEE 77th Vehicular Technology Conference (VTC Spring)*, June 2013, pp. 1–5.

[108] E. Mellios, A. R. Nix, and G. S. Hilton, "Ray-tracing urban macrocell propagation statistics and comparison with winner ii/+ measurements and models," in *2012 Loughborough Antennas Propagation Conference (LAPC)*, Nov 2012, pp. 1–4.

[109] S. Zhang, D. Kong, E. Mellios, A. Doufexi, and A. Nix, "Comparing theoretic single-user and multi-user full-dimension mimo data throughputs in realistic city-wide lte-a deployments," in *2015 IEEE Global Communications Conference (GLOBECOM)*, Dec 2015, pp. 1–7.

[110] A. Navarro, D. Guevara, N. Cardona, and J. López, "Measurement-based ray-tracing models calibration in urban environments," in *Proceedings of the 2012 IEEE International Symposium on Antennas and Propagation*, July 2012, pp. 1–2.

[111] O. Landron, M. J. Feuerstein, and T. S. Rappaport, "A comparison of theoretical and empirical reflection coefficients for typical exterior wall surfaces in a mobile radio environment," *IEEE Transactions on Antennas and Propagation*, vol. 44, no. 3, pp. 341–351, March 1996.

[112] F. van Diggelen and P. Enge, "The world's first gps mooc and worldwide laboratory using smartphones," in *Proceedings of the 28th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2015)*, September 2015, pp. 361–369.

[113] K. Merry and P. Bettinger, "Smartphone gps accuracy study in an urban environment," *PLOS ONE*, vol. 14, no. 7, pp. 1–19, 07 2019. [Online]. Available: https://doi.org/10.1371/journal.pone.0219890

[114] Z. Fang and A. Manikas, "Doa and range estimation of multiple sources under the wideband assumption," in *2016 IEEE Global Communications Conference (GLOBECOM)*, 2016, pp. 1–6.

[115] H. Van Trees, *Optimum Array Processing - Part IV of Detection, Estimation, and Modulation Therory*. Wiley, May 2002.

[116] A. Manikas, *Differential Geometry in Array Processing*. Imperial College Press, 2004. [Online]. Available: https://www.worldscientific.com/doi/abs/10.1142/p309

[117] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, 1986.

[118] D. S. Watkins, ch. 6, pp. 233–263. [Online]. Available: https://epubs.siam.org/doi/abs/10.1137/1.9780898717808.ch6

[119] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716–723, 1974.

[120] S. Haykin, J. P. Reilly, V. Kezys, and E. Vertatschitsch, "Some aspects of array signal processing," *IEE Proceedings F - Radar and Signal Processing*, vol. 139, no. 1, pp. 1–26, Feb 1992.

[121] Wu Shilong, Luo Jingqing, and Liu Youjun, "A method for recognition of spectrum peaks in 2-d music algorithm," in *2005 IEEE International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications*, vol. 1, 2005, pp. 704–707 Vol. 1.

[122] L. Godara, *Handbook of Antennas in Wireless Communications*. CRC Press, 2002.

[123] J. V. NEUMANN, *Functional Operators (AM-22), Volume 2: The Geometry of Orthogonal Spaces*. Princeton University Press, 1950. [Online]. Available: http://www.jstor.org/stable/j.ctt1bc543b.1

[124] H. Stark and Y. Yang, "Vector space projections: A numerical approach to signal and image processing, neural nets, and optics," 1998.

[125] O. C. Park and J. F. Doherty, "Generalized projection algorithm for blind interference suppression in ds/cdma communications," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 44, no. 6, pp. 453–460, 1997.

[126] D. Brandwood, *Topics in the Accuracy and Resolution of Superresolution Systems*. University College London, 2009.

[127] X. Wei, N. Palleit, and T. Weber, "Aod/aoa/toa-based 3d positioning in nlos multipath environments," in *2011 IEEE 22nd International Symposium on Personal, Indoor and Mobile Radio Communications*, Sep. 2011, pp. 1289–1293.

[128] I. Guvenc and C. Chong, "A survey on toa based wireless localization and nlos mitigation techniques," *IEEE Communications Surveys Tutorials*, vol. 11, no. 3, pp. 107–124, August 2009.

[129] Z. Xiao, H. Wen, A. Markham, N. Trigoni, P. Blunsom, and J. Frolik, "Identification and mitigation of non-line-of-sight conditions using received signal strength," in *2013 IEEE 9th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, Oct. 2013, pp. 667–674.

[130] L. Cong and W. Zhuang, "Non-line-of-sight error mitigation in tdoa mobile location," in *GLOBECOM'01. IEEE Global Telecommunications Conference (Cat. No.01CH37270)*, vol. 1, Nov 2001, pp. 680–684.

[131] M. P. Wylie and J. Holtzman, "The non-line of sight problem in mobile location estimation," in *Proceedings of ICUPC - 5th International Conference on Universal Personal Communications*, vol. 2, Oct 1996, pp. 827–831.

[132] J. Khodjaev, Y. Park, and A. Saeed Malik, "Survey of nlos identification and error mitigation problems in uwb-based positioning algorithms for dense environments," *annals of telecommunications - annales des télécommunications*, vol. 65, no. 5, pp. 301–311, June 2010.

[133] S. Marano, W. M. Gifford, H. Wymeersch, and M. Z. Win, "Nlos identification and mitigation for localization based on uwb experimental data," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 7, pp. 1026–1035, Sep. 2010.

[134] S. Tian, L. Zhao, and G. Li, "A support vector data description approach to nlos identification in uwb positioning," *Mathematical Problems in Engineering*, vol. 2014, May. 2014.

[135] J. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural Processing Letters*, vol. 9, no. 3, pp. 293–300, June 1999.

[136] T. Van Gestel, J. A. K. Suykens, G. Lanckriet, A. Lambrechts, B. De Moor, and J. Vandewalle, "A bayesian framework for least squares support vector machine classifiers," *Neural Computation*, vol. 14, pp. 1115–1148, 2002.

[137] V. N. Vapnik, *The Nature of Statistical Learning Theory*. Berlin, Heidelberg: Springer-Verlag, 1995.

[138] SVMS.ORG, "http://www.svms.org/srm/," December 2018.

[139] B. Beavis and I. Dobbs, *Optimisation and Stability Theory for Economic Analysis*. Cambridge University Press, 1990.

[140] J. A. K. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, and J. Vandewalle, *Least Squares Support Vector Machines*. WORLD SCIENTIFIC, 2002. [Online]. Available: https://www.worldscientific.com/doi/abs/10.1142/5089

[141] T. van Gestel, J. A. Suykens, B. Baesens, S. Viaene, J. Vanthienen, G. Dedene, B. de Moor, and J. Vandewalle, "Benchmarking least squares support vector machine classifiers," *Machine Learning*, vol. 54, no. 1, pp. 5–32, Jan 2004.

[142] T. Hofmann, B. Schölkopf, and A. J. Smola, "Kernel methods in machine learning," *The Annals of Statistics*, vol. 36, no. 3, p. 1171–1220, Jun 2008. [Online]. Available: http://dx.doi.org/10.1214/009053607000000677

[143] D. Dardari, C. Chong, and M. Z. Win, "Improved lower bounds on time-of-arrival estimation error in realistic uwb channels," in *2006 IEEE International Conference on Ultra-Wideband*, Sep. 2006, pp. 531–537.

[144] W. Xu-hui, S. Ping, C. Li, and W. Ye, "A roc curve method for performance evaluation of support vector machine with optimization strategy," in *2009 International Forum on Computer Science-Technology and Applications*, vol. 2, Dec 2009, pp. 117–120.

[145] Y. hui Wang, H. Li, X. long Luo, Q. ming Sun, and J. nan Liu, "A 3d fingerprinting positioning method based on cellular networks," *International Journal of Distributed Sensor Networks*, vol. 10, no. 7, p. 248981, 2014.

[146] M. N. De Sousa and R. S. Thomä, "Mobile station localization emitter in urban nlos using multipath ray tracing fingerprints and machine learning," in *2018 8th International Conference on Localization and GNSS (ICL-GNSS)*, June 2018, pp. 1–6.

[147] J. Schmitz, F. Schröder, and R. Mathar, "Tdoa fingerprinting for localization in non-line-of-sight and multipath environments," in *2015 International Symposium on Antennas and Propagation (ISAP)*, Nov 2015, pp. 1–4.

[148] S. Kikuchi, A. Sano, and H. Tsuji, "Blind mobile positioning in urban environment based on ray-tracing analysis," *EURASIP Journal on Advances in Signal Processing*, vol. 2006, no. 1, p. 038989, Dec 2006.

[149] M. Johnson, L. Freitag, and M. Stojanovic, "Improved doppler tracking and correction for underwater acoustic communications," in *1997 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 1, 1997, pp. 575–578.

[150] L. Stankovic, T. Thayaparan, M. Dakovic, and P.-B. Vesna, "Micro-doppler removal in the radar imaging analysis," *IEEE Transactions on Aerospace Electronic Systems*, vol. 49, pp. 1234–1250, 04 2013.

[151] R. Uijlenhoet, A. Overeem, and H. Leijnse, "Dual-polarization radar clutter detection and removal," in *AGU Fall Meeting Abstracts*, December 2018, pp. A21B–07, provided by the SAO/NASA Astrophysics Data System. [Online]. Available: https://ui.adsabs.harvard.edu/abs/2018AGUFM.A21B..07U

[152] H. R. Karimi and A. Manikas, "Manifold of a planar array and its effects on the accuracy of direction-finding systems," *IEE Proceedings - Radar, Sonar and Navigation*, vol. 143, no. 6, pp. 349–357, Dec 1996.

[153] A. Manikas, A. Alexiou, and H. R. Karimi, "Comparison of the ultimate direction-finding capabilities of a number of planar array geometries," *IEE Proceedings - Radar, Sonar and Navigation*, vol. 144, no. 6, pp. 321–329, Dec 1997.

[154] T. E. Tuncer, T. K. Yasar, and B. Friedlander, *Classical and Modern Direction-of-Arrival Estimation*, T. E. Tuncer and B. Friedlander, Eds. Boston: Academic Press, 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/B9780123745248000027

[155] W. L. Stutzman and G. A. Thiele, *Antenna Theory and Design*, 3rd ed. John Wiley Sons, 2012.

[156] P. David, "Scattered and absorbed powers in receiving antennas," *Antennas and Propagation Magazine, IEEE*, vol. 46, pp. 144 – 145, 03 2004.

[157] M. Eberhardt, P. Eschlwech, and E. Biebl, "Investigations on antenna array calibration algorithms for direction-of-arrival estimation," *Advances in Radio Science*, vol. 14, pp. 181–190, 09 2016.

[158] B. Friedlander and A. J. Weiss, "Direction finding in the presence of mutual coupling," *IEEE Transactions on Antennas and Propagation*, vol. 39, no. 3, pp. 273–284, 1991.

[159] A. J. Weiss and B. Friedlander, "Array shape calibration using sources in unknown locations-a maximum likelihood approach," *ICASSP-88., International Conference on Acoustics, Speech, and Signal Processing*, pp. 2670–2673 vol.5, 1988.

[160] E. K. L. Hung, "Matrix-construction calibration method for antenna arrays," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 36, no. 3, pp. 819–828, 2000.

[161] Boon Chong Ng and Chong Meng Samson See, "Sensor-array calibration using a maximum-likelihood approach," *IEEE Transactions on Antennas and Propagation*, vol. 44, no. 6, pp. 827–835, 1996.

[162] M. Eberhardt, M. Lehner, A. Ascher, M. Allwang, and E. Biebl, "An active uhf rfid localization system for fawn saving," *Advances in Radio Science*, vol. 13, pp. 87–94, 11 2015.

[163] C. A. Balanis, *Antenna Theory: Analysis and Design, 4th Edition*. Wiley, 2016.

[164] F. Kaltenberger, H. Jiang, M. Guillaud, and R. Knopp, "Relative channel reciprocity calibration in mimo/tdd systems," in *2010 Future Network Mobile Summit*, June 2010, pp. 1–10.

[165] M. Petermann, M. Stefer, F. Ludwig, D. Wubben, M. Schneider, S. Paul, and K. Kammeyer, "Multi-user pre-processing in multi-antenna ofdm tdd systems with non-reciprocal transceivers," *IEEE Transactions on Communications*, vol. 61, no. 9, pp. 3781–3793, September 2013.

[166] R. Rogalin, O. Y. Bursalioglu, H. Papadopoulos, G. Caire, A. F. Molisch, A. Michaloliakos, V. Balan, and K. Psounis, "Scalable synchronization and reciprocity calibration for distributed multiuser mimo," *IEEE Transactions on Wireless Communications*, vol. 13, no. 4, pp. 1815–1831, April 2014.

[167] C. Shepard, H. Yu, N. Anand, E. Li, T. Marzetta, R. Yang, and L. Zhong, "Argos: Practical many-antenna base stations," in *Proceedings of the 18th Annual International Conference on Mobile Computing and Networking*, ser. Mobicom '12. New York, NY, USA: ACM, 2012, pp. 53–64. [Online]. Available: http://doi.acm.org/10.1145/2348543.2348553

[168] R. Rogalin, O. Y. Bursalioglu, H. C. Papadopoulos, G. Caire, and A. F. Molisch, "Hardware-impairment compensation for enabling distributed large-scale mimo," in *2013 Information Theory and Applications Workshop (ITA)*, Feb 2013, pp. 1–10.

[169] J. Vieira, F. Rusek, O. Edfors, S. Malkowsky, L. Liu, and F. Tufvesson, "Reciprocity calibration for massive mimo: Proposal, modeling, and validation," *IEEE Transactions on Wireless Communications*, vol. 16, no. 5, pp. 3042–3056, May 2017.

[170] J. Verhaevert, E. V. Lil, and A. V. [de Capelle], "Direction of arrival (doa) parameter estimation with the sage algorithm," *Signal Processing*, vol. 84, no. 3, pp. 619 – 629, 2004.

[171] P. Harris, S. Zhang, M. Beach, E. Mellios, A. Nix, S. Armour, A. Doufexi, K. Nieman, and N. Kundargi, "Los throughput measurements in real-time with a 128-antenna massive mimo testbed," in *2016 IEEE Global Communications Conference (GLOBECOM)*, 2016, pp. 1–7.

[172] UNIVERSITY OF BRISTOL, "https://www.bristol.ac.uk/engineering/research/smart/events/layered-realities-weekend/," March 2018.

[173] WATERSHED - LAYERED REALITIES: A 5G SHOWCASE, "https://www.watershed.co.uk/audio-video/layered-reality-5g-weekend-an-introduction," March 2018.

[174] WE THE CURIOUS, "https://www.wethecurious.org," March 2018.

[175] H. Koyuncu and S.-H. Yang, "A survey of indoor positioning and object locating systems," *International Journal of Computer Science and Network Security (IJCSNS)*, vol. 10, 01 2010.

[176] D. Kim, J. Ha, P. Kim, and K. You, "Tdoa/aoa localization in rfid system using dual indirect kalman filter," December 2011.

[177] ORDNANCE SURVEY, "https://www.ordnancesurvey.co.uk/business-government/innovation/connectivity," May 2020.

[178] ORDNANCE SURVEY VIDEO, "https://www.youtube.com/watch?v=tUmw0RSjzMU," September 2017.

[179] DEPARTMENT FOR BUSINESS, INNOVATION AND SKILLS, "Smart cities: Background paper - https://www.gov.uk/government/publications/smart-cities-background-paper," October 2013.

[180] UK DEPARTMENT FOR BUSINESS, INNOVATION AND SKILLS, "The smart city market: Opportunities for the uk - https://www.gov.uk/government/publications/smart-city-market-uk-opportunities," October 2013.

[181] UK DEPARTMENT FOR DIGITAL CULTURE MEDIA AND SPORT, "5g planning - https://www.gov.uk/government/publications/ordnance-survey-5g-planning-and-mmwave-environment-reports," February 2018.

[182] F. Fuschini, E. M. Vitucci, M. Barbiroli, G. Falciasecca, and V. Degli-Esposti, "Ray tracing propagation modeling for future small-cell and indoor applications: A review of current techniques," *Radio Science*, vol. 50, no. 6, pp. 469–485, 2015. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015RS005659

[183] HUAWEI, "Huawei 5g wireless network planning solution white paper - https://www.huawei.com/uk/industry-insights/outlook/mobile-broadband/insights-reports/5g-wireless-network-plan-solution-whitepaper," 2018.

[184] SIRADEL, "https://www.siradel.com/solutions/software/volcano-5g/," May 2020.

[185] THE INDEPENDENT, "https://www.independent.co.uk/environment/car-free-cities-pedestrianisation-cycling-driverless-vehicles-york-oslo-birmingham-a9299856.html," May 2020.

[186] J. A. Rohwer, C. T. Abdallah, and C. G. Christodoulou, "Least squares support vector machines for direction of arrival estimation," in *IEEE Antennas and Propagation Society International Symposium. Digest. Held in conjunction with: USNC/CNC/URSI North American Radio Sci. Meeting (Cat. No.03CH37450)*, vol. 1, June 2003, pp. 57–60.

[187] J. A. Rohwer, C. T. Abdallah, and C. G. Christodoulou, "Least squares support vector machines for direction of arrival estimation with error control and validation," in *GLOBECOM '03. IEEE Global Telecommunications Conference (IEEE Cat. No.03CH37489)*, vol. 4, Dec 2003, pp. 2172–2176.

[188] N. C. Raj, P. V. Aswathy, and K. V. Sagar, "Determination of angle of arrival using nonlinear support vector machine regressors," in *2007 International Conference on Signal Processing, Communications and Networking*, Feb 2007, pp. 512–515.

[189] "5G.CO.UK - https://5g.co.uk/guides/what-is-mobile-edge-computing/," 2020.

[190] 3GPP, "http://www.3gpp.org/DynaReport/FeatureListFrameSet.htm," December 2018.

[191] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); LTE Positioning Protocol A (LPPa); (3GPP TS36.455 RELEASE 9)," 2010.

[192] 3GPP, "Study on NR positioning support; (3GPP TS38.855 RELEASE 16)," 2018.

[193] P. Harris, S. Malkowsky, L. Liu, and M. Beach, "Building massive antenna arrays to handle the explosion of wireless signals," *IEEE Spectrum*, 2017.

[194] BRISTOL IS OPEN, "http://www.bristolisopen.com/," November 2018.

[195] P. Harris, W. B. Hasan, S. Malkowsky, J. Vieira, S. Zhang, M. Beach, L. Liu, E. Mellios, A. Nix, S. Armour, A. Doufexi, K. Nieman, and N. Kundargi, "Serving 22 users in real-time with a 128-antenna massive mimo testbed," in *2016 IEEE International Workshop on Signal Processing Systems (SiPS)*, Oct 2016, pp. 266–272.

[196] P. Harris, *A Real World Evaluation of Massive MIMO: Design, Implementation and Deployment*. University of Bristol, 2017.

[197] NATIONAL INSTRUMENTS, "MIMO Application Framework Extended White Paper; "http://www.ni.com/gate/gb/GB_INFOMIMOFRAMEWP/US"," November 2018.

[198] UNIVERSITY OF BRISTOL PRESS, "BT Massive MIMO Trials - http://www.bristol.ac.uk/news/2017/february/massive-mimo-trials.html," February 2017.

[199] [ONLINE], "https://www.evaluationengineering.com/industries/communications/wireless-5g-wlan-bluetooth-etc/article/13015180," March 2017.

[200] BRISTOL IS OPEN, "https://www.bristolisopen.com/bristol-and-bt-collaborate-on-\massive-mimo-trials-for-5g-wireless," February 2017.

[201] [ONLINE], "https://phys.org/news/2017-02-bristol-bt-collaborate-massive-mimo.html," February 2017.