Research Reports from the Communications Research Laboratory
at Ilmenau University of Technology

Design and Deployment Options for sub-6 GHz FDD Massive MIMO Systems

Muhammad Bilal Amin

TECHNISCHE UNIVERSITÄT ILMENAU

Fakultät für Elektrotechnik und Informationstechnik

Design and Deployment Options for sub-6 GHz FDD Massive MIMO Systems

Muhammad Bilal Amin

Dissertation zur Erlangung des
akademischen Grades Doktor-Ingenieur (Dr.-Ing)

Anfertigung im: Fachgebiet Nachrichtentechnik
Institut für Informationstechnik
Fakultät für Elektrotechnik und Informationstechnik

Gutachter: Univ.-Prof. Dr.-Ing. Martin Haardt
Univ.-Prof. Dr.-Ing. Giovanni Del Galdo
Dr. Simone Redana

Vorgelegt am: 09.07.2019
Verteidigt am: 25.02.2021

urn:nbn:de:gbv:ilm1-2021000064

Erklärung

Ich versichere, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet.

Weitere Personen waren an der inhaltlich-materiellen Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich hierfür nicht die entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten (Promotionsberater oder anderer Personen) in Anspruch genommen. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalte der vorgelegten Dissertation stehen.

Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer Prüfungsbehörde vorgelegt.

Ich bin darauf hingewiesen worden, dass die Unrichtigkeit der vorstehenden Erklärung als Täuschungsversuch bewertet wird und gemäß § 7 Abs. 10 der Promotionsordnung den Abbruch des Promotionsverfahrens zur Folge hat.

Linz, March 15, 2021 Muhammad Bilal Amin

To my parents, Bushra and Muhammad Amin

Acknowledgements

First and foremost, I would like to thank the Almighty Allah (*God*) for always steering me in the right direction, making things easier for me and granting me the wisdom and strength to write this thesis. Also, in general, for His guidance and protection throughout my life.

This thesis has come to fruition with the support of many people. I would like to acknowledge the assistance and guidance of Prof. Martin Haardt, who is my main supervisor. He took a lot of time out of his busy schedule to constantly guide me throughout this work. His comments, suggestions and questions were also very helpful in ensuring the quality of this thesis.

I would also like to express my deepest gratitude to Mr. Wolfgang Zirwas, who was my adviser and supervisor at Nokia Bell Labs in Munich. He offered me with invaluable guidance and support during my research. It was a privilege to work together with him on his wonderful ideas and innovations. Without his knowledge and assistance, this study would not have been successful.

This work was mostly carried under the umbrella of several European projects, for example, METIS and FANTASTIC5G in which Nokia Bell Labs was/is a technical partner. I would also like to thank Dr. Simone Redana for his support and assistance regarding administrative issues at Nokia Bells Labs. I am also grateful to Dr. Simone Redana and, in addition to Prof. Giovanni Del Galdo, for kindly accepting the role of an examiner of this dissertation and providing me with valuable review comments.

Last but not the least, I would like to thank my family as this endeavor would not have been possible without their tremendous support. I am deeply grateful to my parents for their infallible love and guidance. A big thanks to my mother for laying my academic foundation and instilling in me the desire to learn, and to my father for his continuous support and for nudging me to follow my passions. I would also like to express my deep appreciation to my beloved wife. This work could come to fruition only with her help, support, patience and inspiration. A special

thanks to my brother for being a constant source of inspiration and a role model, and to my sister for her unconditional love, affection, and encouragements. Finally, I would like to thank and compliment my kids, Ibrahim and Huda. They brought along a joy to life with them that I didn't think was possible.

Abstract

To satiate the rapidly growing demand for capacity by wireless devices, upcoming Fifth Generation (5G) mobile radio systems are targeting diverse and in some measure, extremely demanding concepts, techniques, and scenarios with respect to the number of served users, data rates, and latency. Significant performance benefits are expected from massive MIMO, especially in combination with tight inter-cell cooperation including Joint Transmission Coordinated MultiPoint (JT-CoMP), User Equipment (UE)-sided interference cancellation and ultra dense deployment of small cells.

The millimeter Wave bands for 5G will be allocated in 2019, so for now, all the performance targets of 5G have to be supported by the below 6 GHz Radio Frequency (RF) bands. This means that for paired and unpaired spectrum, Frequency Division Duplex (FDD) as well as Time Division Duplex (TDD) have to be supported. Massive MIMO, typically conceptualized as a strong over provisioning of antennas versus served users, asserts to solve various complicated issues like inter-cell interference, Multi-User MIMO (MU-MIMO) scheduling, coverage holes and capacity limits. However, it is deemed to work only for TDD systems as massive MIMO downlink transmission for FDD systems is particularly challenging due to two requirements. First, the Channel State Information (CSI) has to be obtained from a large number of antennas without an unreasonable overhead due to the transmission of orthogonal downlink reference signals from these antennas. Second, the relevant channel estimates have to be made available at the network side without an unrealistic uplink control signaling overhead. An exploding overhead for orthogonal reference signals or limited CSI accuracy, leads to pilot contamination, which has been shown to upper bound performance in the literature and is detrimental especially for sensitive interference cancellation schemes. Hence, to reduce the number of tranceivers in the FDD massive MIMO system, we have devised a framework based on the Grid of Beams (GoB) concept. It transforms the large number of antennas into effective antenna ports which can then even use the existing reference signals from Long Term Evolution - Advanced (LTE-A). Moreover,

to effectively combat inaccuracy issues in the CSI fed back to the base station, we need channel prediction algorithms which provide reliable performance over a large prediction horizon. To this end, we have devised an enhanced hybrid algorithm based on two state-of-the-art channel prediction algorithms. It provides strong prediction performance even for channel components received with low Signal-to-Noise Ratios.

Moreover, quintessential massive MIMO schemes concentrate on the base station side with the goal to achieve high spectral efficiency by MU-MIMO or large coverage by strong beamforming gains. The UE-sided analysis is typically limited to four or mostly eight antenna elements per UE, which can be justified by the corresponding UE complexity and the limited space to place more antenna elements for the below 6 GHz case. However, UE-sided beamforming would provide many benefits, ranging from improved channel estimation and prediction accuracy, effective interference suppression up to coverage and spectral efficiency gains on the system level. For this reason, we have proposed the novel concept of 'virtual beamforming' which allows us to form virtual massive MIMO arrays at the UE, even if it has a single physical antenna element. The UE is assumed to be moving at a constant speed and the received signals for the adjacent time slots are stored. These measurements are then utilized to form beamformers as if they were received by separate antenna elements. We only capture a small spatial sub-section of the channel and thus lose energy but this loss is partially compensated by the beamforming gain. Virtual beamforming leads to a reduction in the number of relevant multipath components per channel component as well as in the number of channel components themselves and gives a prediction accuracy gain of approximately 10 dB lower normalized mean square error compared to the case of a single antenna UE. On the other hand, virtual beamforming applied directly to the user data can be very inefficient as it requires the re-transmission of data symbols. To overcome this challenge we have devised a parallel transmission scheme over a set of coded virtual beams. This allows us to reduce the effect of the inherent re-transmission penalty of virtual beamforming and we have shown that reliable communication is possible. In addition, we have also devised a framework to combine varying number of virtual and physical antenna elements into a virtual massive MIMO array.

Zusammenfassung

Um den schnell wachsenden Kapazitätsbedarf drahtloser Geräte zu befriedigen, zielen die kommenden Mobilfunksysteme der fünften Generation (5G) auf unterschiedliche und teilweise extrem anspruchsvolle Konzepte, Techniken und Szenarien in Bezug auf die Anzahl der versorgten Nutzer, Datenraten und Latenzzeiten ab. Von massive MIMO werden erhebliche Leistungsvorteile erwartet, insbesondere in Kombination mit einer engen Zusammenarbeit zwischen den Zellen, einschließlich Joint Transmission Coordinated MultiPoint (JT-CoMP), Interferenzunterdrückung auf der Seite des Endgeräts (UE) und einem extrem dichten Einsatz von kleinen Zellen.

Die Millimeterwellenbänder für 5G werden 2019 zugewiesen, daher müssen vorerst alle Leistungsziele von 5G von den Funkfrequenzbändern unter 6 GHz unterstützt werden. Das bedeutet, dass für gepaartes und ungepaartes Spektrum sowohl Frequency Division Duplex (FDD) als auch Time Division Duplex (TDD) unterstützt werden müssen. Massive MIMO, typischerweise als starke Überprovisionierung von Antennen gegenüber den bedienten Nutzern konzipiert, verspricht, verschiedene herausfordernde Probleme wie Interzell-Interferenzen, Multi-User-MIMO (MU-MIMO) Scheduling, Abdeckungslücken und Kapazitätsgrenzen zu lösen. Es wird jedoch davon ausgegangen, dass es nur für TDD-Systeme funktioniert, da eine massive MIMO-Downlink-Übertragung für FDD-Systeme aufgrund von zwei Anforderungen eine besondere Herausforderung darstellt. Erstens muss die Kanalzustandsinformation (CSI) von einer großen Anzahl von Antennen erhalten werden, ohne dass ein unangemessener Overhead durch die Übertragung orthogonaler Downlink-Referenzsignale von diesen Antennen entsteht. Zweitens müssen die relevanten Kanalschätzungen auf der Netzwerkseite ohne einen unrealistischen Overhead für die Uplink-Steuersignalisierung zur Verfügung gestellt werden. Ein explodierender Overhead für orthogonale Referenzsignale oder eine begrenzte CSI-Genauigkeit führt zu einer Pilotverschmutzung, die in der Literatur als leistungsbegrenzend beschrieben wurde und insbesondere für empfindliche Interferenzunterdrückungssys-

teme nachteilig ist. Um die Anzahl der Transceiver in einem FDD-Massive-MIMO-System zu reduzieren, haben wir daher ein Framework entwickelt, das auf dem Grid of Beams (GoB)-Konzept basiert. Es wandelt die große Anzahl von Antennen in effektive Antennenports um, die dann sogar die vorhandenen Empfangssignale von Long Term Evolution - Advanced (LTE-A) nutzen können. Um Ungenauigkeiten in der CSI, die zur Basisstation zurückgeführt wird, effektiv zu bekämpfen, benötigen wir außerdem Algorithmen zur Kanalvorhersage, die eine zuverlässige Vorhersage über einen möglichst großen Zeithorizont liefern. Zu diesem Zweck haben wir einen verbesserten hybriden Algorithmus entwickelt, der auf zwei modernen Kanalvorhersagealgorithmen basiert. Er liefert eine zuverlässige Vorhersage auch für Kanalkomponenten, die mit niedrigen Signal-Rausch-Verhältnissen empfangen werden.

Darüber hinaus konzentrieren sich die wesentlichen Massive-MIMO-Schemata auf die Basisstationsseite mit dem Ziel, eine hohe spektrale Effizienz durch MU-MIMO oder eine große Abdeckung durch starke Beamforming-Gewinne zu erreichen. Die UE-seitige Betrachtung ist typischerweise auf vier oder meist acht Antennenelemente pro UE beschränkt, was durch die entsprechende UE-Komplexität und den begrenzten Platz zur Platzierung weiterer Antennenelemente für den Fall unter 6 GHz gerechtfertigt werden kann. UE-seitiges Beamforming würde jedoch viele Vorteile bieten, die von einer verbesserten Kanalschätzung und Vorhersagegenauigkeit über eine effektive Störungsunterdrückung bis hin zu einer besseren Abdeckung und spektralen Effizienz auf Systemebene reichen. Aus diesem Grund haben wir das neuartige Konzept des "Virtual Beamforming" vorgeschlagen, das es uns ermöglicht, virtuelle Massive-MIMO-Arrays am UE zu bilden, auch wenn es nur ein einziges physisches Antennenelement hat. Es wird angenommen, dass sich das UE mit einer konstanten Geschwindigkeit bewegt und die Empfangssignale für die benachbarten Zeitschlitze gespeichert werden. Diese Messungen werden dann zur Bildung von Strahlformern verwendet, als ob sie von separaten Antennenelementen empfangen würden. Wir erfassen nur einen kleinen räumlichen Teilbereich des Kanals und verlieren daher Energie, aber dieser Verlust wird teilweise durch den Gewinn der Strahlformung kompensiert. Virtuelles Beamforming führt zu einer Reduzierung der Anzahl relevanter Mehrwegekomponenten pro Kanalkomponente sowie der Anzahl der Kanalkomponenten selbst und ergibt einen Gewinn an Vorhersagegenauigkeit von ca. 10 dB geringerem normalisierten mittleren quadratischen Fehler im Vergleich zum Fall eines UE

mit einer einzelnen Antenne. Auf der anderen Seite kann virtuelles Beamforming, das direkt auf die Nutzdaten angewendet wird, sehr ineffizient sein, da es die erneute Übertragung von Datensymbolen erfordert. Um diese Herausforderung zu überwinden, haben wir ein paralleles Übertragungsschema über einen Satz kodierter virtueller Strahlen entwickelt. Dadurch können wir den Effekt der inhärenten Wiederübertragungsstrafe des virtuellen Beamforming reduzieren und wir haben gezeigt, dass eine zuverlässige Kommunikation möglich ist. Darüber hinaus haben wir ein Framework entwickelt, um eine unterschiedliche Anzahl von virtuellen und physischen Antennenelementen zu einem virtuellen Massive-MIMO-Array zu kombinieren.

Contents

List of Figures

Chapter 1

Introduction

1.1 Motivation

1.1.1 A Brief Introduction to 5G

The pioneering scientist and inventor, Nikola Tesla predicted mobile phones back in 1909 and smart phones in 1926 when he said in an interview with Collier's magazine, "*When wireless is perfectly applied the whole earth will be converted into a huge brain, which in fact it is, all things being particles of a real and rhythmic whole. We shall be able to communicate with one another instantly, irrespective of distance. Not only this, but through television and telephony we shall see and hear one another as perfectly as though we were face to face, despite intervening distances of thousands of miles; and the instruments through which we shall be able to do this will be amazingly simple compared with our present telephone. A man will be able to carry one in his vest pocket*" [Ber15].

It took us many decades to reach Tesla's vision but we are definitely there and ready for the next step. Smart phones have become a necessity for daily life by providing the functionality of several devices like a camera, music player, radio, navigation tool, entertainment and social media device, and even as a digital wallet. To enable these functions, cellular networks have been concurrently going through a dramatic technological evolution resulting in, the so called, *generations* of cellular networks which are typically abbreviated as the number of the generation followed by a capital G. For a detailed description of the evolution of the cellular networks, please see [RTMR06]. Here, we provide a short summary, which is also tabulated in Figure 1.1. The 1G networks, introduced in the 1980s, were analog systems that focused on the provisioning of voice services. The introduction of 2G networks in the 1990s was enabled by the advancement of digital processing techniques. The foundation of 2G was the Global System for Mobile (GSM) communications, which is one of the most famous communication systems and is still in use in most parts of the world. GSM uses circuit switching technology to enable voice services and also introduced messaging services. Circuit switching is suited to voice traffic but could not provide high data rates for bursty data traffic and thus, GSM was upgraded through General Packet Radio Service (GPRS) and later by Enhanced Data Rates for GSM Evolution (EDGE), both of which allow the additional provisioning of data services via packet switching. The GSM, GPRS, and EDGE networks rely on Time Division Multiple Access (TDMA). Another contemporary class of 2G networks was also introduced that instead relies on Code Division Multiple Access (CDMA), known as IS-95A and IS-95B, popularized mainly by Qualcomm. The TDMA based technologies evolved into 3G cellular networks in the form of High Speed Packet Access (HSPA), whereas the CDMA based IS-95 systems evolved

into the CDMA2000 series [HM00, BDHM01]. The Fourth Generation (4G) is comprised of the well-known Long Term Evolution (LTE) and Long Term Evolution - Advanced (LTE-A) technologies which focus on technologies like Multiple-Input and Multiple-Output (MIMO) and Orthogonal Frequency Division Multiplexing (OFDM) to increase throughput [HT09, SBT11].

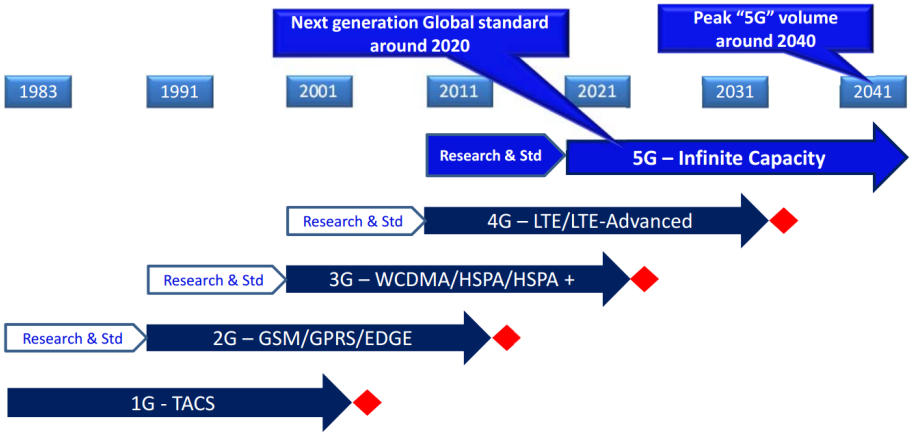


Figure 1.1: A short summary of the evolution of cellular networks [Taf13].

As stated by Tom Marzetta, it is a timeless truth that *the demand for wireless throughput will always grow* [Mar15]. It is evident in future trends as there will be devices other than just smart phones connected to the network including (connected) cars resulting in the envisioned Internet of Things (IoT). Additionally, the introduction of new applications like virtual or augmented reality, and services like instantaneous language translation impose new challenges and influence the requirements for a future Fifth Generation (5G) mobile network.

As the name suggests, 5G is the 5^{th} major rung in the evolutionary ladder of cellular mobile networks. There is no official concrete definition of 5G yet but according to a 5G-Public Private Partnership (5G-PPP) white paper, "*5G will not only be an evolution of mobile broadband networks. It will integrate networking, computing and storage resources into one programmable and unified infrastructure. The impact of 5G will go far beyond existing wireless access networks with the aim for communication services, reachable everywhere, all the time, and faster.*" [GP15]. Hence, it is generally believed that 5G will combine evolving current systems like LTE-A and WiFi, with revolutionary novel technologies to meet the demanding new performance requirements. Nokia foresees that the users will expect a *Zero latency Gigabit experience* from a 5G system [Nok16]. *Zero latency* in this context means that 5G needs to bring down latency low enough, so that whatever the use case, the radio interface does not become the bottleneck. *Gigabit experience* translates to data reception and transmission speeds of Gigabits per second to users. The other

general requirements for 5G have been summarized in Figure 1.2.

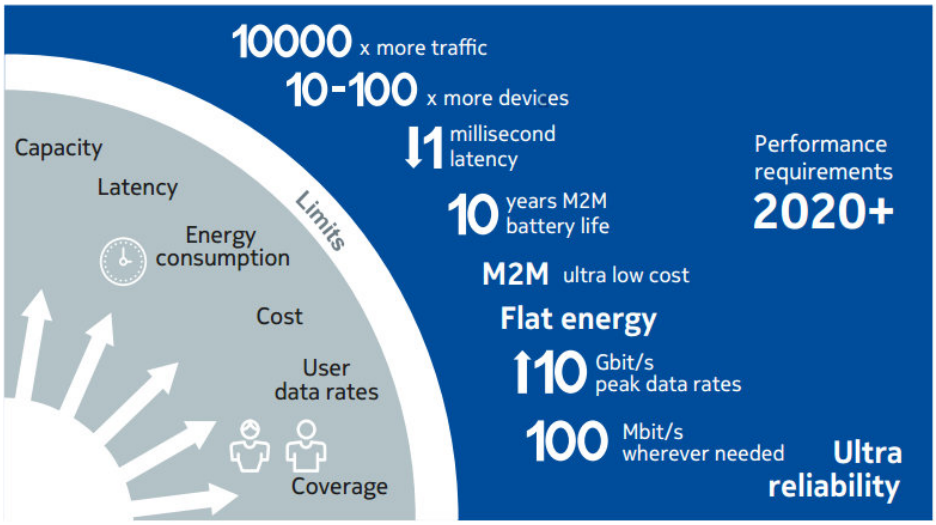


Figure 1.2: The summary of key requirements for 5G [Nok16].

The research and development work on 5G, carried out by both industry and academia, was started at the start of this decade and has progressed significantly. The overall network requirements, implementation challenges, candidate technologies and their enablers have been studied and expressed in various publications like, for example, in [ABC+14, DGK+13, BHL+14, OBB+14, HRTA14, CLRH+14]. There are also several dedicated European Union projects working on 5G like the recently concluded METIS I [Que15], its successor METIS II and FANTASTIC5G. Five core services have been envisioned in the FANTASTIC5G project to be the driving force of 5G (which are similar to the requirements found in [ITU15] and [3GP17]), namely [Sch16]:

1. **Mobile BroadBand (MBB)**
2. Massive Machine Communications (MMC)
3. Mission Critical Communications (MCC)
4. Broadcast/Multicast Services (BMS)
5. Vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2X) communications

A common key element that all these core services share is ubiquitous coverage.

In this thesis, we will focus on the MBB core service for below 6 GHz frequency bands. The motivation behind this is that all the newly allocated Radio Frequency (RF) bands will be below 6 GHz as discussed in the World Radiocommunication Conference 2015 (WRC2015) [IR15]. The availability of centimeter and millimeter Wave (mmWave) bands will be decided in the upcoming WRC2019. However, many

companies like Qualcomm and Nokia, plan to introduce their 5G products as soon as the 2018 Winter Olympics in South Korea. This means that for now all the expected 5G traffic growth must be supported by the below 6 GHz frequency range as shown in Figure 1.3.

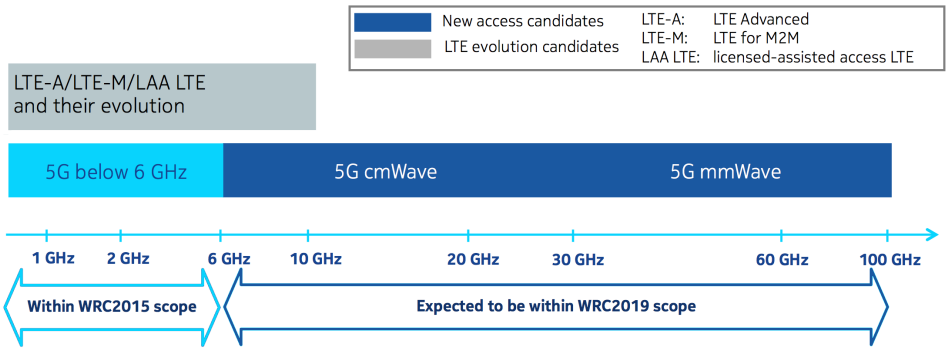


Figure 1.3: Timeline of spectrum allocation for 5G.

1.1.2 Mobile BroadBand: Massive MIMO and CoMP

Mobile BroadBand, which is the marketing name for high-speed wireless internet, has been the impetus behind the exponential growth of mobile networks as it enabled humans to satisfy an essential need, that is, to be able do more while on the go instead of being anchored to a physical location. For 5G, the focus of MBB will again be to address the human-centric demands for access to advanced multimedia content like streaming 8K video and 3-Dimensional (3D) or holographic videos, and other very demanding applications and services like virtual and augmented reality. The question that now arises is, *how to meet this high demand for wireless?* The potential solutions fall into either of the three categories:

1. deploy more base stations with each serving a proportionately smaller area,
2. squeeze the most out of all the spectrum bands by exploiting the currently underutilized or unused spectrum, and
3. efficiently use multiple antennas at the base stations and/or at the User Equipments (UEs).

The deployment of more base stations as epitomized by *small cells* is a promising technology but it is covered extensively in literature, for example, in [JMZ+14, BLM+14, WHG+14, GPA+15, MPT+13], and not the focus of this thesis. Also, gains based on the provisioning of new spectrum bands and bandwidth will soon reach a ceiling as the amount of available electromagnetic spectrum cannot increase [Mar15]. Thus, the solution of interest then comes in the form of the most important innovation in wireless communications known as *MIMO* which relies on multiple-antenna

systems at both ends of the transmission link. MIMO systems offer additional Degrees of Freedom (DoFs) in the spatial dimension because in typical urban environments, radio waves experience multipath propagation resulting in independent channels between the multiple transmit and receive antennas. These additional DoFs can be exploited in four ways for data transmission: the achievable data rate can be increased by the transmission of multiple data streams, the inherent array gain can provide robustness against noise and signal degradation, interference can be avoided through appropriate spatial processing, and the quality of the connection can be improved by the multiple transmission of a single data stream.

MIMO technology which formed the backbone of 4G systems like LTE-A, had been limited to a maximum of 8 antennas [Wan13, 3GP09]. To meet the demands of 5G, we need to go bigger and this has exactly happened in the form of massive MIMO which employs the use of a colossal number of coherently and adaptively operated service antennas [LTEM13, Mar10]. This results in immense improvement in throughput, coverage and energy efficiency by focusing the signal energy into tiny regions of space during transmission and reception.

Massive MIMO is typically defined as a system in which the number of service antennas M_T, is at least an order of magnitude higher than the number of users N_U; that is, $M_T/N_U > 10$ [JMZ+14]. This definition is controversial as the real design parameter is the number of users N_U because the number of transmit antennas M_T is generally fixed in a deployment and not a variable. It has been shown in [BLM15] and [DAZ+14] that massive MIMO performs quite well in the range $M_T/N_U < 10$ and it is even possible to let $M_T/N_U = 1$, which leads to a small per terminal rate but the sum spectral efficiency is still very high. Hence, a simple definition of a massive MIMO system is, "*It is a system with unconventionally many active antenna elements, M_T, that can serve an unconventionally large number of terminals, N_U. One should avoid specifying a certain ratio M_T/N_U, since it depends on a variety of conditions; for example, the system performance metric, propagation environment, and coherence block length*" [BLM15].

Another important component of 5G MBB is cooperation between base stations, also known as *Coordinated MultiPoint (CoMP)* [IDM+11, FKV06, SRH13]. In CoMP, either just the control-plane information or additionally user-plane information is shared between the base stations, depending on the CoMP scheme used. This means that CoMP can be considered as a geographically distributed MIMO system comprising the antennas of the multiple distributed base stations and the antennas of the multiple UEs served by them [BJ+13, SKM+10]. The aim of CoMP is to eliminate the interference between adjacent cells which is beneficial for the overall performance but especially helps the users close to the cell borders.

To exploit the spatial dimension through MIMO or CoMP and jointly serve

multiple users on the same time and frequency resource, base stations employ *spatial precoding* in the downlink [FG98]. Precoding is performed in such a way that at the geographic location of each user, the signals transmitted from the multiple antennas overlap constructively for each user's own data signal and deconstructively for all the other users' signals. *To perform precoding, knowledge of the full Channel State Information at the Transmitter (CSIT) is required.* The availability of accurate CSIT is paramount for massive MIMO and CoMP to perform properly but efficient acquisition of CSIT is one of the main challenges faced by these schemes.

1.2 State-of-the-Art

Let us first consider a point-to-point link between a transmitter and receiver. Ideally, both of the participants should be able to communicate with each other simultaneously, referred to as Full-Duplex (FD) transmission, but it is very difficult to achieve as pointed out by Andrea Goldsmith, "*It is generally not possible for radios to receive and transmit on the same frequency band because of the interference that results. Thus, bidirectional systems must separate the uplink and downlink channels into orthogonal signaling dimensions typically using time or frequency dimensions*" [Gol05]. In reality, most wireless communication systems like 4G LTE are half-duplex that emulate a full-duplex system through a separation of the directional communication channels either in frequency (known as Frequency Division Duplex (FDD)) or in time (known as Time Division Duplex (TDD)). However, it has been proven that full-duplex communication is indeed possible in [CJS+10, JCK+11, BMK13, SSG+14, DS10, TZH16] and for 5G systems, further research is being carried out as FD could potentially result in a doubling of the transmission speed. Nevertheless, FD systems are still in their infancy with a potential gain of 40-80% over half-duplex in small area systems [Rik15]. Thus, the first iteration of 5G systems would still use half-duplex FDD or TDD transmission.

Channel State Information (CSI) is usually required at both the base station and the UE. It is utilized to alter the Modulation and Coding Scheme (MCS), the power allocation and user scheduling among other things. Massive MIMO systems as envisioned by Tom Marzetta, rely on TDD operation because in such systems, uplink reference signals can provide the base station with not only the uplink CSI but also the downlink CSI because of channel reciprocity which results in a reduction in the overhead of CSI acquisition [Mar15]. Moreover, the time required to transmit uplink pilots is independent of the number of base station antennas in TDD mode [Mar10].

On the contrary, in FDD systems, CSI is estimated or predicted at the UE by making use of Channel State Information Reference Signals (CSI-RSs) or pilots, which is then quantized to reduce the overhead and fed back to the base station

through a low-rate feedback link. This means that FDD systems require considerably more time and resources for CSI acquisition as compared to a TDD system. Additionally, they quickly become unfeasible with an increasing number of antenna elements as the time required to transmit downlink reference signals or pilots is proportional to the number of base station antennas. To summarize, FDD massive MIMO systems suffer from the following practical issues:

- large number of RF chains are required
- reference signal and feedback overheads are huge
- global CSI sharing needs to be done in real-time
- the computational complexity for precoders at the base stations is huge.

In spite of that, FDD systems dominate over TDD in current cellular networks because they offer benefits continuous channel estimation, small latency, backward compatibility etc. Hence, it is important to develop solutions for FDD massive MIMO systems as most operators have paired spectrum, especially under 6 GHz, which they would like to migrate from older technologies to 5G for better performance.

The current body of research on FDD massive MIMO CSI acquisition typically falls into three categories based on: temporal correlation, spatial correlation and compressive sensing. All of these potential solutions, generally assume some form of *sparsity* in the channel that can be exploited. This sparsity can arise from a strong spatial correlation requiring the estimation of only a few strong eigen-directions or from the Channel Impulse Responses (CIRs) being sparse in time due to a limited number of scatterers around the base station.

Temporal channel correlation is leveraged by [CLK15, CCLM13] to reduce the CSIT overhead by utilizing trellis-code based quantization codebooks. In contrast, a memory-based channel training sequence design is presented in [CLB14]. They have proposed both open-loop and closed-loop training frameworks with the results favoring the closed-loop training framework as it provides a better performance especially in the low Signal-to-Noise Ratio (SNR) regime and for high number of transmit antennas. It reduces the overhead because the UE only needs to indicate to the base station, the best subsequent training signal to be sent out of the common set of training signals shared in advance between them, based on the previously received training signals and prior knowledge of the channel. In [RL14] and references therein, the authors have utilized a compressive sensing approach to exploit this sparsity. However, it is important to note that at mmWave frequencies, the CIRs may be sparse [AASS+14] but at lower frequencies, for example the below 6 GHz range which we are interested in, measurement results show that sparsity assumptions are rather dubious [GERT15].

Similarly, the spatial correlation of channel coefficients can also be exploited to reduce the CSIT overhead. The pioneering work in this category was presented in [ANAC13] and extended in [NAAC14, Nam14, JMCN15, AASS+14]. They have proposed the Joint Spatial Division and Multiplexing (JSDM) scheme which divides the users into groups with approximately the same channel covariance eigenspace. A two-stage precoding is performed, namely the pre-beamforming matrix and a Multi-User MIMO (MU-MIMO) precoding matrix for the effective channel, which includes the pre-beamforming as well. This provides the ability to counteract both inter-group interference and the interference inside each group. Other somewhat similar two-stage precoding and user grouping schemes have been proposed by [CL14, XYM14].

In addition to the CSI acquisition overhead, practical MIMO systems suffer from a phenomenon called *channel out-dating* or *channel aging* [TH13]. This means that the channel has already changed resulting in a mismatch in the fed-back CSI and the real channel before the CSIT is used at the base station. This phenomenon is more pronounced in high mobility environments and results in a severe performance degradation. The underlying cause of this is feedback delay, which is inevitable as channel estimation or prediction, processing, and feedback require a certain amount of time. A promising solution has been recognized in the form of *prediction of the channel information into the future* which can effectively combat the performance degradation due to feedback delays [DHHH00a, RHSK08]. For a detailed survey of the state-of-the-art of channel prediction schemes, please refer to Section 4.2.

1.3 Open Issues

Channel State Information overhead reduction and improved channel prediction for practical, below 6 GHz, FDD massive MIMO systems is a new topic. For this reason, several issues related to the prediction performance, the reference signal design, and the overall CoMP framework are still open. The most important open issues are summarized as follows:

1. Can massive MIMO work in FDD operation?
2. How can we effectively use massive MIMO in the FDD mode without unbearably increasing the reference signal overhead?
3. How does the overall framework of an FDD based massive MIMO system look like at the system level especially with the incorporation of CoMP?
4. How can we predict the weaker channels in the FDD massive MIMO case?
5. How can we reduce the CSI acquisition overhead in an effective manner without resorting to sparsity in the channel?

6. How can we avail the benefits of massive MIMO at the UE side as well?
7. How can we efficiently reduce the number of channel components and multipath components in the received signal of the UE?

1.4 Contributions and Outline of the Thesis

The main focus of this thesis is the development of massive MIMO techniques for the more challenging FDD case to reduce the CSI acquisition overhead and to enhance the performance of channel prediction schemes. Overall, the techniques developed in this thesis enable us to provide an optimistic answer to the first open issue: Yes, massive MIMO can work in FDD operation.

We start off with a description of the overall system or network level concept of a prospective FDD 5G cellular network in **Chapter 2**. As CoMP, especially in the form of Joint Transmission Coordinated MultiPoint (JT-CoMP), is an important candidate for achieving the required performance goals for 5G, an Interference Mitigation Framework - Advanced (IMF-A) has been developed in the EU project ARTIST4G [GS12]. We provide a detailed description of the core components of this framework. The main benefit of IMF-A is that it effectively mitigates inter-cooperation area interference, allowing for transmission schemes to be optimized for just one cooperation area. To cater for the requirements of 5G, an evolution of the IMF-A was presented in the EU project METIS I [BA15]. This, henceforth called 5G IMF-A, augments massive MIMO and small cells to the existing IMF-A. The main challenges for its implementation in FDD is the large number of antenna elements resulting in a severe reference signal overhead leading to pilot contamination. We have devised to use the Grid of Beams (GoB) concept to limit the number of antenna elements to a few directive beams spread out to cover the whole sector. This provides several potential benefits and addresses open issues 2 and 3 [AZH15b, ZAS16]. In favorable propagation conditions like Line-Of-Sight (LOS) with high Ricean factors, GoB may even achieve a similar performance to the reciprocity based TDD massive MIMO beamforming, in terms of the downlink sum-rate. In other scenarios, the performance of GoB degrades as compared to TDD beamforming [FRT+17]. Despite that, for FDD operation, GoB is being considered as the most commercially viable solution for massive MIMO[1] [KN15]. Another challenge is the availability of reliable CSIT which can be reliably resolved through channel prediction. However, reliable channel prediction over large distances is a difficult problem. Thus, in the next chapters we investigate this channel prediction problem in detail.

[1] An exemplary scenario would be to downscale 128 antenna elements to 16 GoBs.

Chapter 3 provides the underpinnings of the wireless channel. We discuss the fading phenomenon and its types. The variations in the channel stem from multipath propagation between the base station and the UE. The geometry of the environment and the transmission bandwidth influence the number of multipath components, and their corresponding delays and Angles of Arrival (AoAs). If the UE is moving slowly, the scattering environment might remain the same and hence the channel varies very slowly in time. Such a channel is easier to predict as the history of the channel provides a good fit to the future of the channel. If the UE is moving at a high speed, the scattering environment changes very quickly, especially if there is birth and/or death of multipath components due to a change in the active scatterers. Such a channel is harder to predict as the history of the channel might not provide a sufficient fit to the future. In addition, the scatterers might move as well, but this is usually a far slower process than the fading due to the movement of the UE. Moreover, in this chapter, we also take a look at the dimensionality of a real-world massive MIMO channel in a CoMP and 5G scenario. The methods to accurately and effectively model the wireless channel are also elaborated with a focus on ray-tracing methods and models used in this thesis for the simulation of systems and to access the performance of channel prediction algorithms.

Chapter 4 is dedicated to channel prediction schemes. We start off by providing a survey of the state-of-the-art prediction schemes and by distinguishing the factors which affect the performance and influence the design of the various prediction schemes. We identify that for 5G, pilot-based prediction schemes make more sense, as compared to blind schemes, as they are easier to implement and provide a reliable performance. We consider two state-of-the-art schemes in detail [MDH08, Aro11], as both of them provide a good performance in their own regard [DAZ+14]. In pilot-based prediction schemes, to predict the future CIR or channel frequency response, the channel is estimated through pilots or reference signals at several consecutive time snapshots which provides a history of the channel variation. Statistical methods are then used to utilize these estimates to form a prediction of the future channel response. The majority of channel prediction schemes proposed for MIMO systems assume that the MIMO channel is composed of independent parallel links and then employ methods devised for Single-Input and Single-Output (SISO) systems to predict these links [SLC08, WE06]. The downside of such an approach is that it does not consider the correlation between the antenna elements, which is inevitably present in a practical MIMO system [SFGK00]. In dense scattering environments, these schemes provide a prediction only over a tiny distance, on the order of a few tenths of the wavelength, and prediction beyond a wavelength is realistically not possible. Furthermore, such schemes require training over several wavelengths to accurately predict the channel [TV01]. A better utiliza-

tion of the spatial DoFs of the MIMO channel can lead to a reliable prediction up to several wavelengths. To better incorporate the spatial correlations and improve the performance of state-of-the-art channel prediction algorithms, we have proposed a novel hybrid method, derived from the two schemes, which allows catering for the noise enhancement effect present in the AutoRegressive (AR)-based models, addressing open issue 4, and also exploits the structure of the MIMO channel which is usually neglected by numerous state-of-the-art prediction schemes [AZH17].

The spatial DoFs provided by the massive MIMO systems can also be utilized to reduce the complexity of the channel prediction problem itself through curtailment into smaller spatial sub-divisions. To this end, **Chapter 5** introduces the novel concept of virtual beamforming, which enables the UE to reap the benefits of massive MIMO, irrespective of its number of antenna elements addressing open issues 5, 6 and 7 [AZH15a]. At the UE, the spatial selectivity provided by virtual beamforming is far more useful due to the large AoA spread as compared to beamforming operations at the base station side [AZH15b]. By forming a narrow beam and pointing it in an appropriate direction, we can capture the most significant sub-division of the overall channel. This results in a reduction in the number of channel components and multipath components, and in the feedback overhead as well. The energy loss incurred by forgoing a major portion of the channel is adequately compensated by the beamforming gain. Virtual beamforming also subdues the time-varying nature of the channel leading to a much smoother evolution in time, which makes it easier to predict with sophisticated prediction schemes.

Chapter 6 details the challenges and their potential solutions regarding data transmission through virtual beams. The creation of virtual beams at the UE leads to the re-transmission of data symbols from the base station, to enable the various consecutive time samples to serve as the virtual antenna elements. This re-transmission is obviously undesirable as it causes a reduction in the throughput of the system. We propose, for single antenna UEs, the utilization of circular orthogonal codes to form multiple virtual beams to partially recover the loss of throughput [AZH16]. We have also extended this concept to cover UEs with multiple antennas, which might be the more relevant case for 5G, enabling the complete recovery of re-transmission throughput loss.

In summary, we have scrutinized the feasibility of utilizing Grid of Beams at the base station side and found them to be a promising option for massive MIMO in FDD operation. Moreover, we have developed novel beamforming schemes for the UE side, which enable us to procure the benefits of massive MIMO for UEs with a limited number of antennas leading to a reduction of multipath components and channel components, the feedback overhead, and enhances the channel prediction performance. The above mentioned benefits of GoB, virtual beamforming and a

combination of both schemes were thoroughly analyzed at the link-level, utilizing measured and ray-traced channels for realistic urban macro scenarios. The results of this analysis were contributed to the Work Package 3 of METIS-I project and to the Work Package 4 of FANTASTIC5G project. Overall, the work in this thesis, in conjunction with the general work carried out in the METIS-I and FANTASTIC5G projects, can provide a first practical solution to the challenge of implementing a FDD-based massive MIMO system in a cooperative scenario. The novel virtual beamforming scheme and the specifics of its implementation for data transmission that are presented here, have been implemented in Nokia's 5G demonstrator called AMoRE (Advanced Mobile Radio Realtime Experience), and are being presented by Nokia to standardization bodies.

List of Publications

Here is the list of publications during my tenure as a doctoral researcher:

- **Patents:**
  - W. Zirwas, M.B. Amin, "*Non linear user processing*", filed US Patent No PCT/IB2016/057361, 2016.
  - W. Zirwas, M.B. Amin, "*Artificially mutually coupled antenna arrays*", filed EU Patent No 16836698.7 - 1220 PCT/FI2016050466.119, 2016.
  - W. Zirwas, M.B. Amin, "*Virtual beamforming for static users*", DE Patent App. 2015E01703DE, 2015.
- **Papers: (Peer Reviewed)**
  - M.B. Amin, W. Zirwas, M. Haardt,"*HOSVD-based denoising for improved channel prediction of weak massive MIMO channels*", IEEE Vehicular Technology Conf. (VTC 2017 Spring), Sydney, Australia. 2017.
  - M.B. Amin, W. Zirwas, M. Haardt,"*Virtual massive MIMO beamforming gains for 5G user terminals*", IEEE Vehicular Technology Conf. (VTC 2016 Fall), Montreal, Canada. 2016.
  - W. Zirwas, M.B. Amin, M. Sternad, "*Coded CSI reference signals for 5G - exploiting sparsity of FDD massive MIMO radio channels*", Workshop on Smart Antennas (WSA 2016), Munich, Germany. 2016.
  - M.B. Amin, W. Zirwas, M. Haardt, "*Design options in the context of 5G for a powerful channel prediction*", Meeting of the Wireless World Research Forum (WWRF 35), Copenhagen, Denmark. 2015.
  - M.B. Amin, W. Zirwas, M. Haardt, "*Advanced channel prediction concepts for 5G radio systems*", 2015 International Symposium on Wireless Communication Systems (ISWCS 2015), Brussels, Belgium. 2015.
  - S. Dierks, M.B. Amin, W. Zirwas, M. Haardt, B. Panzner, "*The benefit of cooperation in the context of massive MIMO*", 18th International OFDM Workshop 2014 (InOWo'14), Essen, Germany. 2014.

Chapter 2

Cooperative Communications in 5G: An Overview

Chapter Summary

In this chapter, the current status of the cooperative communications framework in the 5G cellular networks is elaborated. We also discuss the challenges involved in the realization of such a cooperative framework. The target of this thesis is to provide potential solutions to these challenges as discussed in Section 2.3.4.

The need for interference mitigation is explained in Section 2.1.1 and the optimization options available in a cellular network are discussed in Section 2.1.2. Section 2.2 describes the IMF-A, developed in the ARTIST4G project, in detail and elaborates on its building blocks while Section 2.3 details the evolution of IMF-A, by further enhancements developed in the METIS-I and FANTASTIC4G projects, to a 5G technology.

2.1 Introduction

The, so called, Interference Mitigation Framework - Advanced (IMF-A) was first developed in the European Union funded project ARTIST4G[1] [GS12] and then further enhanced in METIS-I[2] [BA15]. Currently, it is being finalized for its first deployment as a Fifth Generation (5G) technology in the ongoing European project FANTASTIC5G[3] [PM16].

2.1.1 The need for Interference Mitigation

A future 5G radio system should provide a significantly higher performance compared to the current systems like Long Term Evolution (LTE) or Long Term Evolution - Advanced (LTE-A). One of the core services defined for 5G is Mobile Broad-Band (MBB) and the primary Key Performance Indicators (KPIs) for MBB are *data throughput, latency, coverage* and *mobility*. Out of these KPIs *data throughput* is riddled with the most challenges due to the expected demand for very high data volumes and end-user data rates, as described in Table 2.1, for downlink transmission for a future 5G system.

The upcoming cellular networks will mainly be interference-limited because of densification and aggressive frequency-reuse, hence potent interference avoidance or mitigation techniques, especially ones which require minimum effort for hardware

[1]ARTIST4G - Advanced Radio Interface Technologies for Fourth Generation (4G) Systems. Duration: January 2010 - June 2012.

[2]METIS - Mobile and wireless communications Enablers for Twenty-twenty (2020) Information Society. Duration: November 2012 - April 2015. https://www.metis2020.com/.

[3]FANTASTIC5G - Flexible Air iNTerfAce for Scalable service delivery wiThin wIreless Communication networks of the 5th Generation. Duration: July 2015 - June 2017. http://fantastic5g.eu/.

Specifications	LTE	LTE-A	5G
Downlink Throughput	100 Mbps	1 Gbps	10 Gbps and higher, 1 Gbps anywhere
Uplink Throughput	50 Mbps	500 Mbps	1 Gbps and higher
Latency (round trip time)	10 ms	less than 5 ms	less than 1 ms

Table 2.1: A comparison of the primary KPIs for MBB [WD14].

and infrastructural upgrades are of great interest. To this end, a powerful IMF-A has been developed which will enable us to achieve the required coverage and data throughput [GS12]. It is the main distinguishing factor for 5G compared to what is considered in LTE or LTE-A [BA15]. Currently deployed networks, for example, LTE, avoid inter-cell interference by restricting reuse of resources in the adjacent cells but this hampers overall system performance. Hence, intuitively the best way to achieve a higher spectral efficiency is to actively use the whole spectrum at all times in all cells, leading to a reuse factor of one. Sophisticated schemes like Inter-Cell Interference Coordination (ICIC) or enhanced ICIC, have been introduced in LTE but these schemes do not completely eliminate the interference. Therefore, the gains are limited to small and medium load conditions. Hence, for a powerful inter-cell interference mitigation, strong cooperative schemes like Joint Transmission Coordinated MultiPoint (JT-CoMP) between adjacent sites, are required as shown in the ARTIST4G and METIS-I projects along with adequate optimization of other system parameters described in Table 2.2 [GS12] [BA15]. This framework together with a set of supporting enablers might brand 5G as the first *interference free* or *interference exploiting* cellular radio system.

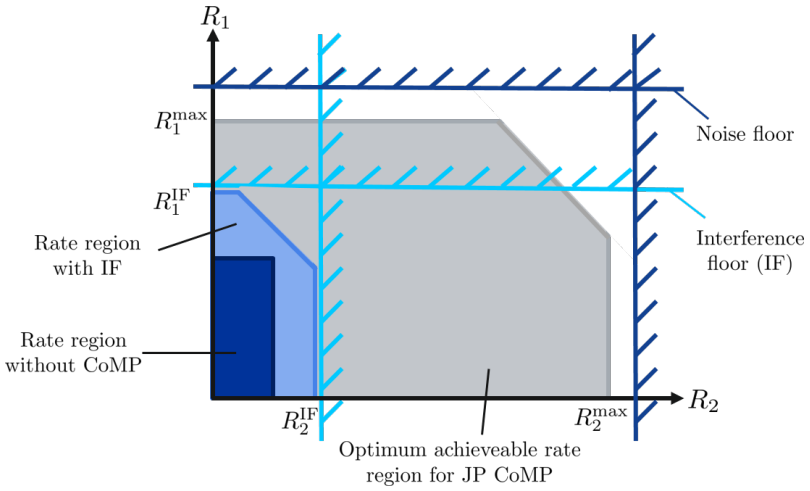


Figure 2.1: Rate regions with and without interference floor [ZMK13].

To understand the pernicious consequence of a strong inter-cell interference floor in an interference-limited scenario that is in the case of a very low noise floor, we

consider a simple example carried over from [ZMK13]. Figure 2.1 illustrates the exemplary rate region of two user User Equipments (UEs). The rate region for a conventional cellular system is represented by the small blue rectangle. If we assume that the noise floor is very low and there is no interference floor, then the achievable rate region with $R_1^{\max}$ and $R_2^{\max}$ is indicated by the large gray area. This rate is ideally only limited by the maximum Modulation and Coding Scheme (MCS). Under the assumption of a strong interference floor which is indicated by the light blue lines, the Coordinated MultiPoint (CoMP) gains are mostly hidden leaving only minor rate gains as indicated by the light blue area. Consequently, the development of a simple and robust inter-cell interference floor shaping technique was deemed very significant. Such a method was developed in ARTIST4G and will be explained in Section 2.2.2.

2.1.2 Optimization Options for a Cellular System

There are several Degrees of Freedom (DoFs) available for the optimization of a cellular system. To distinguish between these DoFs, we divide them into three separate sets of parameters based on the *time scale* at which they fluctuate and the quantities which influence these sets of parameters [GS12]. The, so called, *influence quantities* are related to the environment, data traffic and user-specific properties and are summarized in Table 2.2.

Influence Quantity	**Time Scale**	**Degrees of Freedom**
Environment	Years/month	Deployment Sectorization Antenna properties Static frequency reuse
Data Traffic	Days/hours	Electrical down tilt Hand-over parameters Switching on/off base stations Semi static (fractional) frequency reuse Semi static clustering Semi static resource allocation (among overlapping clusters) Inter-cluster interference
User-specific	Seconds/ milliseconds	Scheduling/user grouping Dynamic resource allocation (within a cluster) Spatial signal processing algorithms

Table 2.2: Structure of a three-level optimization framework considering cooperative cellular networks to avoid interference [GS12].

Out of these three parameter sets, we can consider the *environmental* parameters to be mainly static. Based on network requirements, the hardware and infrastructure for the base stations are installed at appropriate locations and then properly divided and orientated into multiple sectors. After that, the type and number of antennas, their antenna patterns, and mechanical tilting are also optimized. Once all these parameters are set, they typically remain fixed for months or even years unless some very big change, like the construction of new buildings, makes it inevitable to re-optimize them.

The second set of parameters related to *data traffic* within the network typically vary over days or hours. These traffic dependent parameters like, for example, electrical down-tilt or handover parameters, can be adapted automatically and fall under the umbrella of Self-Organizing Networks (SON). However, semi-static user related parameters like frequency-reuse, clustering and resource allocation among cooperation clusters can also be similarly categorized, as they are also managed in a time scale of hours when complexity is considered. Traffic conditions also influence the size of the cooperation clusters and the performance requirement of the CoMP techniques. For example, a higher frequency-reuse can be utilized during low traffic situations needing simple base station cooperation while a more aggressive frequency-reuse can be used during high traffic situations enabled by a strong cooperation scheme. In this aspect, cooperation clusters are formed by the coordinating base stations which can overlap during high load, to ensure a frequency-reuse factor one, where inter-cluster interference present in such systems can be reduced by, for example, electrical down-tilt. This resulted in the development of the cover shift concept which is explained in Section 2.2.1.

The third and most important set of optimization parameters is *user-specific* optimization which helps to avoid interference within a cluster by considering the positions of the users and the channel conditions. It is essential to have high performing algorithms and techniques to optimize these parameters as they vary on the scale of seconds and milliseconds and have a direct effect on the performance of a CoMP system.

2.2 Interference Mitigation Framework - Advanced

The IMF-A, developed in the ARTIST4G project, is designed for systems operating below the 6 GHz frequency range and LTE 4×2 Multiple-Input and Multiple-Output (MIMO), supporting both Time Division Duplex (TDD) and Frequency Division Duplex (FDD) modes in typical urban macro or micro scenarios with inter-site distances of approximately 500 m [GS12]. The main building blocks of the framework are sketched in Figure 2.2. At the heart of IMF-A lies a transmission

technique called JT-CoMP. The aspiration of JT-CoMP is to convert a cellular network riddled with inter-cell interference into one which exploits this interference for a better data transmission [JMZ+14].

CoMP can be defined in simple terms as a scheme which sends and receives data to and from a UE by means of several points (base stations) to ensure, that even at the cell edges, the optimum performance is achieved. The CoMP techniques, can generally be divided into two categories: *coordinated* and *cooperative* transmission. For the coordinated transmission case, the exchange of only control-plane information takes place between the base stations to avoid inter-cell interference by, for example, making beneficial user scheduling decisions. This control information could be, for example, the Signal-to-Interference-plus-Noise Ratio (SINR) or the Channel State Information (CSI). On the other hand, cooperative transmission schemes incorporate the additional exchange of the user-plane data. This data can be used in various ways depending on the required system performance as there is usually a trade-off between performance and complexity, and different backhaul requirements. Typically, for all cases, *Joint Precoding (JP)* is performed over the whole CoMP set (set of cooperating base stations). Then, simultaneous data transmission from multiple points can be executed, which is called *Joint Transmission (JT)* [IDM+11]. Alternatively, the data transmission can take place from one cell at a time depending on the quality of the link, referred to as *Dynamic Cell Selection (DCS)*. The coordinative approaches are more robust against practical impairments like outdated CSI but theoretically, cooperative techniques provide a better system performance. Between the cooperative schemes, JT-CoMP was chosen as the main scheme for IMF-A in ARTIST4G because even though it is practically challenging, it has the capability to provide the best performance through:

Macro-diversity gains: As the transmission is carried out over multiple independent fading channels, the effect of fading can be reduced through averaging.

Beamforming gains: The coherent combination of transmission from several antennas can potentially increase the average received Signal-to-Noise Ratio (SNR) at the UE.

Interference cancellation: As data transmission is simultaneously carried out from all the cooperating base stations, JT-CoMP constructively and destructively combines the superposition of several to potentially many signal components, with the aim to maximize the desired received signal at the UE while concurrently minimizing the mutual interference.

In this overview, we will elaborate on the practical challenges involved with the implementation of JT-CoMP and their potential solutions, namely the *tor-*

toise concept and the *cover shift concept*, developed in the ARTIST4G and METIS projects [GS12, BA15].

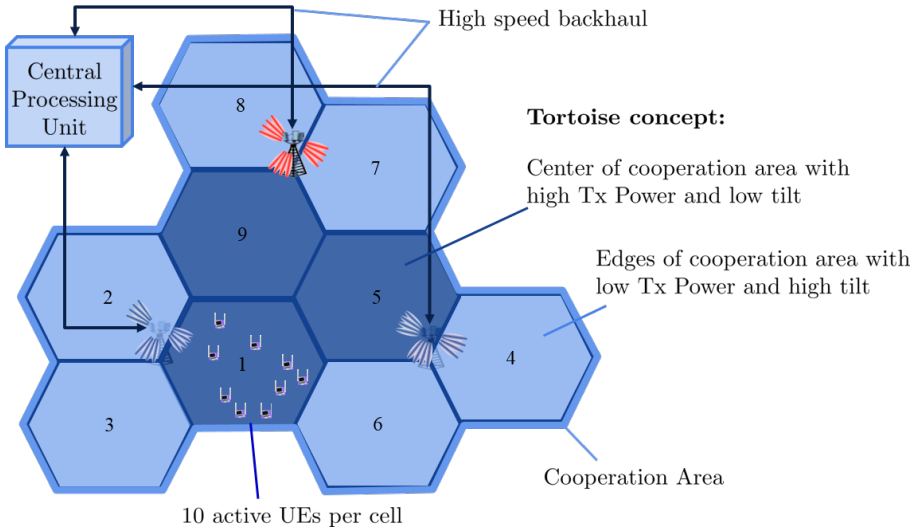


Figure 2.2: Interference mitigation framework based on cooperation areas formed by cooperating over nine cells or three sites. The tortoise concept is also illustrated with the center (dark blue) having higher transmit power and lower downtilt as compared the the edges of the cooperation area (light blue). [GS12].

2.2.1 Cover Shift Concept: Partial CoMP

Under the assumption of a network-wide base station cooperation to eliminate inter-cell interference, theoretical gains of 300 percent or more, as compared to isolated MIMO base stations, have been promised in [FKV06]. However, taking practical constraints into account, the performance gains drop to the order of 50 to 100 percent [GS12]. The practical constraints include but are not limited to feedback delay, synchronization errors, sub-optimal user selection and overhead (feedback and reference signal). In addition, even though a network-wide cooperation would eliminate the interference, it is intuitive to see that cooperation has to be limited to a few cells of the full network, forming *clusters* or *cooperation areas*. The reasoning behind this is that in a network-wide cooperation, channel information and user data would have to be exchanged between all the base stations. This is not practically feasible as it will cause the number of pilots, the backhaul traffic and the feedback overhead for the CSI reporting to explode. However, the formation of such cooperation areas intrinsically leads to residual inter-cooperation area interference [TKBH12]. The tortoise concept which is detailed in Section 2.2.2 provides a solution to reduce this inter-cooperation area interference.

The optimum method of creating a cooperation area would be to perform user-centric clustering which means that each UE in the cooperation area is served

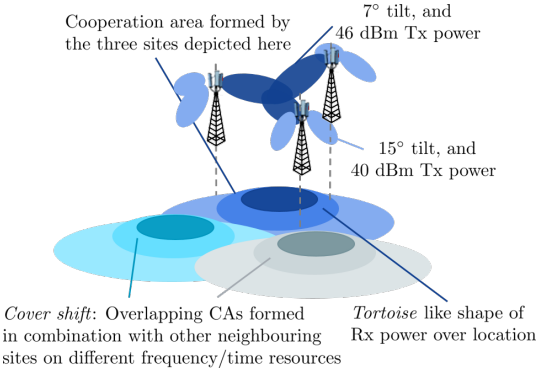
from its three strongest cells. It is evident from system level simulations that considerable potential performance gains can be achieved by doing so [MF11]. In practice, user-centric cooperation is quite challenging because it requires all the UEs in a cooperation area to be served by the same set of base stations, i.e., all UEs have the same set of the strongest cells. The probability of finding UEs associated to a single set of base stations is exceedingly low for realistic radio channel conditions with several Non-Line-Of-Sight (NLOS) components and strong shadowing.

Instead of performing optimum user-centric clustering, *partial CoMP* employs a static and structured approach by expanding the cooperation area from the commonly used three cells to over three adjacent sites with three sectors per site, resulting in a total of nine cells [ZMK13]. Such enlarged cooperation areas increase the number of UEs which contain their three strongest cells inside the cooperation area. The expansion of cooperation areas is beneficial from a practical standpoint, as only two inter-site backhaul connections between adjacent sites will be required for such an expansion. In order to limit the complexity for channel prediction and feedback schemes, the channels from the three strongest of the overall nine cells are measured by the UEs and reported. This partial reporting of channel information was the inspiration to name this scheme partial CoMP. It is also important to note that even for very large cooperation areas, there will be inter-cooperation area (out-of-cluster) interference. So, for a cooperation area comprised of nine cells, inter-cooperation area interference is detrimental and typically spoils most of the joint transmission performance gains. There are a lot of users at the edge of the cooperation area, which constitutes a large fraction of the total area because of geometrical reasons, and these users are severely affected by this inter-cooperation area interference.

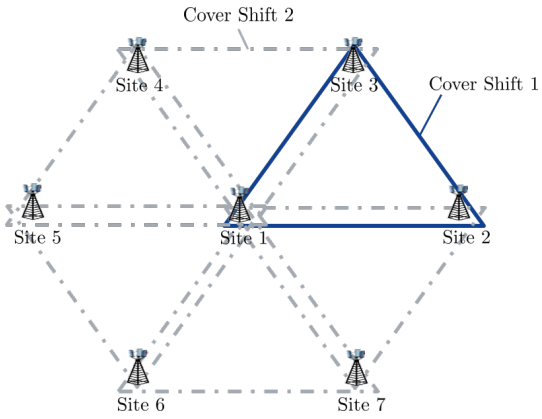
A new dimension for optimization, that is overlapping cooperation areas which have been named *cover shifts* or *super cells*, was introduced to deal with the inter-cooperation area interference [GS12]. A UE at the edge of the cooperation area can be scheduled into another cover-shift, where it is in the center of another cooperation area which utilizes different radio resources like sub-bands and/or sub-frames as shown in Figure 2.3a. The main criterion for the assignment of cooperation area-centric UEs to a certain cover shift is that it should include the three strongest cells for all the jointly served UEs. In the case of homogeneous networks, each site is involved in a total of six cover shifts as shown in Figure 2.3b [ZMK13]. Essentially this means that the overall system has a frequency reuse equal to one as each site serves all cover shifts simultaneously. Additionally, a certain data rate can always be guaranteed in the network because it can adapt the used frequency and/or time resources per cover shift [TKJ+15].

2.2.2 Tortoise Concept: Interference Floor Shaping

The use of enlarged and overlapping cooperation areas, as introduced by the cover shift concept, significantly reduces the inter-cooperation area interference but this only holds for the three strongest cells. In urban environments with a homogeneous network layout, strong path-loss variations lead to a high number of cells with a similar received power. This results in an inter-cell interference floor which does not affect the UEs at the center of the cooperation area because of their strong signal power from their serving cells but the UEs at the edge of the cooperation area suffer significantly. The challenge is the high number of interfering cells which are usually distributed over large geographical areas. To reduce interference, we can go ahead and integrate these far-off cells into the cooperation areas but this would escalate quickly into network-wide cooperation which is undesirable.



(a) Shape of received power and tilt allocation is depicted for the tortoise concept. The blue colored beams belong to the three depicted sites forming the blue cooperation area and the size of the beams indicates the transmit power. The teal and gray colored discs represent other cooperation areas utilizing different resource elements called as cover shifts.



(b) Six cover shifts for site 1. Each triangle defines the three sites for cooperation with the solid triangle defining the main cooperation area. The cover shifts utilize different frequency and/or time resources.

Figure 2.3: A visualization of the cover shift and tortoise concepts [ZMK13].

To reduce the commonly received interference power, i.e., the interference floor, originating from other cooperation areas, the cover shift concept might be extended by cell-specific vertical antenna tilting per cooperation area. This is motivated by several measurement campaigns in real networks which show that using antenna down-tilt supported by active antenna systems, interference can be limited to a few cells [JMZ+14]. The general goal is to minimize outbound interference and to maximize signal strength at the center of the cooperation area. For that purpose, wideband precoders are utilized, meaning that the same precoding weights are used

for all the sub-carriers, to semi-statically or statically form one or more wideband beams at each cell [GS12]. The concept is similar to Grid of Beams (GoB), as explained in Section 2.3.3. On the outbound beams, a strong down-tilt is applied, for example 15°, and one might reduce the transmit power of the outbound beams by, for example 3 to 6 dB, as well compared to the inbound beams, concentrating the signal power into the center of the cooperation area. In comparison, inbound beams utilize a small vertical tilt of, for example 7° [GS12]. As a result, a *tortoise* like shape of the interference power is achieved, thus significantly curtailing the leakage of interference outside the cooperation area as illustrated in Figure 2.3a.

The partial CoMP scheme and the tortoise concept are complementary schemes as each cover shift serves only the cooperation area-centric UEs benefiting from stronger coverage, whereas the UEs at the edge of the cooperation area which suffer due to heavy tilting are scheduled into another cover shift.

2.3 Evolution of the IMF-A to 5G: The OPportunistic CoMP Concept

An evolution of the interference mitigation framework to fulfill the requirements of 5G has been developed in the latest European projects called METIS-I and Fantastic5G, which we will refer to as 5G IMF-A [BA15, PM16]. It again targets the below 6 GHz frequency range for both TDD and FDD modes in typical urban macro or micro scenarios with inter-site distances of approximately 500 m. The goal is to have an even higher spectral efficiency to improve the main relevant KPIs, i.e., data throughput as well as coverage for the Mobile BroadBand (MBB) core service. Two main issues have been identified in [GS12] that upper bound the JT-CoMP gains for a typical homogeneous macro cellular 4×2 MIMO scenario. The first issue is that a high percentage of the *indoor UEs are noise limited*, as opposed to interference limited outdoor UEs, due to a high outdoor-to-indoor penetration loss of 20 dB or more. The second issue stems from the relatively broad beams formed by the $\lambda/2$-spaced four element arrays, often used in macro cells. As the beams overlap each other and experience similar path loss and small-scale fading, there is a severe correlation between the adjacent beams. As a result of this strong inter-beam interference, only a few users can be served simultaneously. To solve these issues and achieve the spectral efficiency goals, the two main ingredients of the evolution of IMF-A are *massive MIMO* and a tight cooperation with *small cells* [Zir15, BA15]. An overall schematic is provided in Figure 2.4 which shows the addition of massive MIMO arrays at the base stations and the transformation of IMF-A into a HetNet by the inclusion of small cells and Local Area Networks

(LANs).

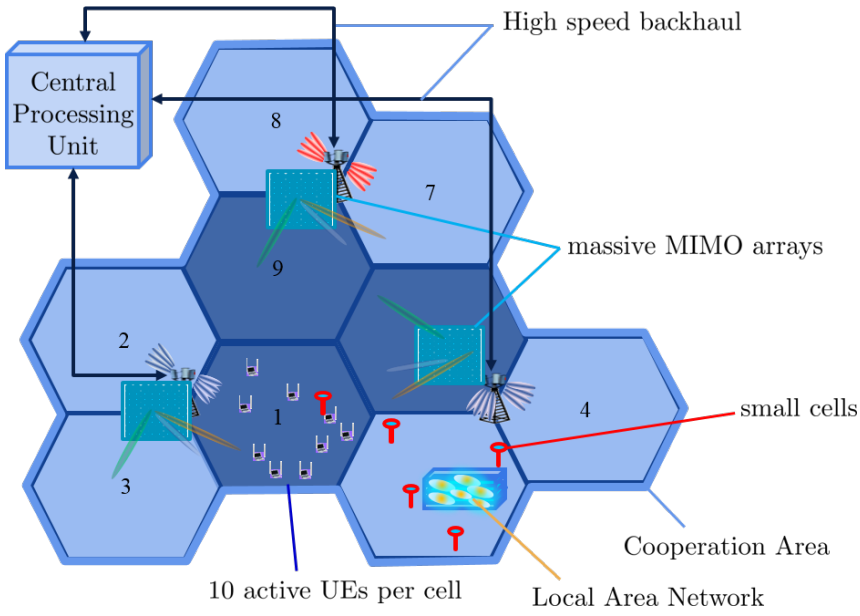


Figure 2.4: Evolution of IMF-A to 5G [JMZ+14].

2.3.1 Small Cells

The inclusion of small cells into the urban macro scenario leads to a network comprised of a combination of different cell types which is called a Heterogeneous Network (HetNet) in 3rd Generation Partnership Project (3GPP). Such a deployment poses several new challenges because a high number of small cells with differing and potentially limited backhaul connections, which results in dubious coordination capabilities, are added into the network. Additional challenges arise due to the different radio conditions for the macro and small cell layers, for example, macro cells are typically placed quite high ranging between 20 to 50 m with an approximate transmit power of 49 dBm for a 20 MHz bandwidth and therefore deliver good coverage. In comparison, small cells are normally placed below the rooftop level, at approximately 10 m, with a significantly lower transmit power of 23 to 30 dBm only. This disparity in radio conditions means that despite the shorter distance between a small cell and the UE, the macro base station is often received with a higher power than the small cell.

When a potentially large number of small cells is added to the network, it becomes tightly packed with cells, eventually becoming extremely dense [GSA15]. The complexity for a tight cooperation in such an Ultra Dense Network (UDN) increases significantly, for example, an extremely high backhaul rate of several tens to hundreds of Gbit/s might be required as shown in [JMZ+14]. In addition, the

complexity for matrix inversions required for JT-CoMP precoding or the effort for multi-cell scheduling also explodes. Hence, the simplest solution to the HetNet scenario is to allocate different carrier Radio Frequency (RF) bands (non co-channel) to the macro and the small cell layer but this leads to an inefficient use of resources. Hence, efficient co-channel deployments are preferable and can be implemented by using interference coordination techniques in either time or frequency like Inter-Cell Interference Coordination (ICIC) or enhanced ICIC [PWSF12]. Such schemes provide a fine granular coordination between the macro and the small cell layer, depending on the load conditions. The term coordination here means that a disturbance to the other layer is prevented by switching off one of the layers in time or frequency. An even better approach is to opportunistically activate only parts of the small cells when needed and is termed as OPportunistic CoMP (OP-CoMP) [Zir15].

2.3.2 Massive MIMO

Massive MIMO schemes, following the pioneering work of Larsson and Marzetta [LTEM13], have received significant attention from both the academia and industry in the last few years. Intuitively, massive MIMO can be thought of as a natural extension of conventional MIMO by upgrading from base stations equipped with a few antennas to ones rigged with a colossal number of antennas. This large provision of antennas provides additional DoFs in the spatial domain which allows the network to simultaneously serve more UEs by multiplexing the signals on the same time-frequency resource for various UEs and by directing radiated energy in an efficient manner towards intended UEs through narrower beams. In addition, it leads to a significant reduction of transmit energy and, as a result, a strong interference suppression [RPL+13]. There is no fixed upper bound on the total number of antennas and one description, for example, states that the *provisioning of antennas is ten times larger than the total number of streams served to all UEs in a cell* [JMZ+14]. Massive MIMO could be used as an *in-band backhaul scheme* as an enabler of ultra dense networks or it could be used as an *access scheme* to facilitate the evolution of IMF-A to 5G. We are interested in its role as an access scheme. The implementation of massive MIMO, and MIMO in general, as an access scheme can be carried out in two ways: Single-User MIMO (SU-MIMO) and Multi-User MIMO (MU-MIMO).

2.3.2.1 Massive SU-MIMO

In SU-MIMO, also called point-to-point MIMO, the communication takes place between a base station equipped with multiple antennas and a UE that has multiple antennas as well. Other UEs are served in separate time and/or frequency resources.

The general idea is to transmit distinct information on each of the transmitter antennas resulting in a transmit signal vector which is then multiplied by the channel matrix to form the received signal vector. Then if the condition number, defined as the ratio of the highest singular value to the lowest, of the channel matrix is low, i.e., it is well-conditioned, the distinct transmitted signals can be recovered reliably.

Although SU-MIMO is employed in current wireless systems like, for example, the IEEE 802.11ac standard, it does not effortlessly scale beyond 8×8 antennas for various reasons. The first and biggest issue is that in Line-Of-Sight (LOS) conditions, the rank of the channel matrix drops to one which means that only one data stream can be reliably supported. Hence, apart from an increase in the SNR enabled by the array gain, this dependency on propagation conditions means that data streams equivalent to the number of antennas cannot be maintained all the time resulting in a loss of throughput. The second problem is that the promised multiplexing gains of $\min(M_T, M_R)$ are not achievable at the edge of the cell due to the typically low SINRs. The third problem is that the deployment of more antennas is sometimes not possible at the UE because of physical constraints, and the added complication and cost of electronic chains required for each antenna. Similarly, rather sophisticated signal processing is required at both the transmission ends to achieve the maximum system performance. The last issue is that increasing the number of antennas results in a proportional increase in the amount of time required for training to learn the channel matrix, irrespective of whether TDD or FDD is used.

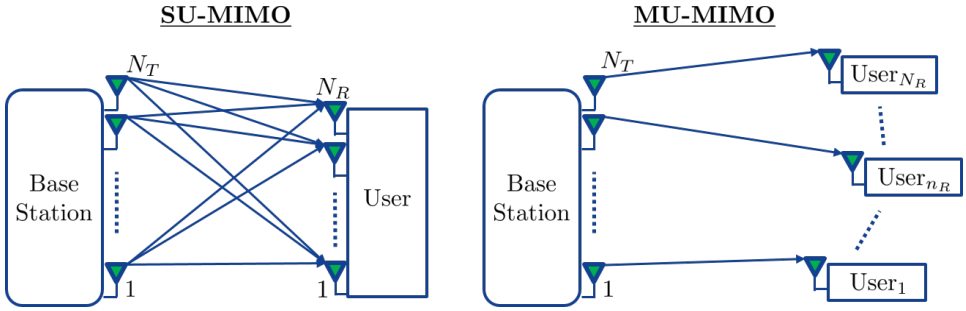


Figure 2.5: A very simple representation of the difference between SU-MIMO and MU-MIMO.

2.3.2.2 Massive MU-MIMO

A MU-MIMO system is a MIMO system which has been split up at the UE side; the base station still has M_T antennas whereas, in the simplest case, there are M_R independent and geographically distributed single-antenna UEs. A simple comparison between SU-MIMO and MU-MIMO is shown in Figure 2.5. The surprising

benefit of MU-MIMO is that even though some UEs will suffer due to poor quality channels, the sum throughput is not affected in general [GKH+07, LCG+12].

The benefit of MU-MIMO over SU-MIMO is that it is less affected by the propagation environment. If the UEs are generally more separated in the angular domain as compared to the angular resolution of the base station antenna array, MU-MIMO can provide relatively good performance even under LOS conditions. Another benefit is that single-antenna UEs can be used. The main issue with the implementation of massive MU-MIMO is still the huge time and resources spent on acquiring CSI which grows with both the addition of base station antennas and users, especially for FDD systems.

The contradictory approach adopted by [Mar15] to make massive MIMO scalable is to increase the size of the overall system and

1. the number of users is far fewer than the provision of base station antennas,
2. TDD is employed and the downlink CSI needs to be available only at the base station, and
3. simple linear schemes are employed for both precoding multiplexing on the downlink and decoding demultiplexing on the uplink.

2.3.3 Grid of Beams

The important thing to consider now is how to implement massive MIMO for FDD systems. Even though TDD is superior and the scheme of choice in literature, most of the licensed spectrum below 6 GHz is available for FDD. The basic approach uses downlink pilots to inform the UEs of the channel and then CSI is transmitted to the base station on the uplink. The problem here is that the required duration of the training interval grows proportionally to the number of base station antennas. This significantly caps the maximum number of antennas at the base station and to maintain a large ratio of antennas over served UEs, this would cause an extreme reduction in the number of UEs served. In addition, the transmission of CSI on the uplink would result in an additional overhead.

For the 5G IMF-A framework, the baseline concept assumes large cooperation areas with a total of nine macro cells and any or each of these cells are equipped with an massive MIMO array. The feasible configurations for the implementation of massive MIMO are presented in Figure 1 of [LTEM13]. The most practical configuration would be to have a Uniform Rectangular Array (URA) with 16×16 or 32×16 antenna elements leading to a total of 256 or 512 antenna elements per massive MIMO array. Such a huge provision of antennas means that a straightforward implementation of FDD massive MIMO is not possible due to reasons mentioned

previously. Hence, we need a method to downscale the number of effective antennas while still retaining the benefits of massive MIMO. Such a method is provided by the GoB concept which falls under the category of fixed beamforming [SHRS08]. A pre-defined set of precoding vectors $\mathbf{V}$, is applied to the available antenna elements to generate the GoB with a certain number of equally spaced and, most of the time, narrow beams. This introduces a Spatial Division Multiple Access (SDMA) aspect to the system, as multiple UEs can be served on the same time and/or frequency resource. When applied to each sector of a macro cell, this can improve coverage and at the same time reduce inter-beam interference within each sector.

Each beam of the GoB defines one antenna port, transmitting one specific orthogonal Channel State Information Reference Signal (CSI-RS). Therefore, UEs measure the beamformed GoB as channel components instead of measuring the channels to the physical antenna elements. As an example, consider that 8 fixed beams are formed per cell resulting in a reduction in the number of channel components per cell from 256 to 8. This means that the 40 orthogonal CSI-RSs, as defined for LTE release 10, might be sufficient for massive MIMO channel estimation, and permit 5 orthogonal muting patterns for the adjacent cells, thus ensuring low downlink pilot contamination which is a major issue for massive MIMO [JAMV11].

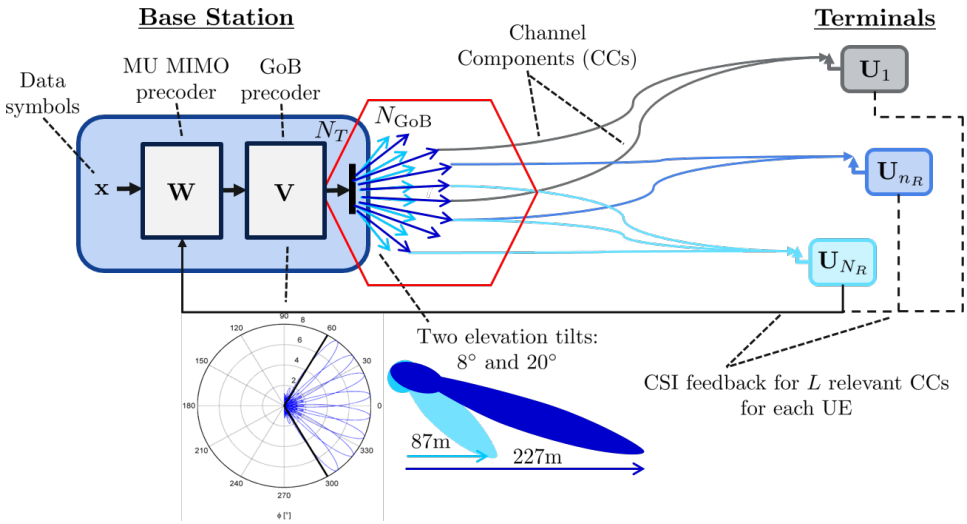


Figure 2.6: A schematic of GoB concept to convert a massive MIMO antenna array into a limited set of beams.

For 5G IMF-A, the enlarged cooperation area comprises 9 macro cells which means that if each of them supports a $16 \times 16 = 256$ antenna array, there will be a total of 2304 channel components without any beamforming. In addition, each antenna usually supports two polarizations and 5G IMF-A is targeting two elevation angles for better coverage and throughput. This results in $256 \times 2 \times 2 = 1024$ channel components per cell or 9216 channel components over the entire cooperation

area rendering the system unmanageable. Therefore, considering 8 azimuth and 2 elevation beams per polarization, reduces the number of channel components per cell from 1024 to just 32 and for the overall cooperation area from 9216 to 288. To make this even more tractable, the UEs can select the beams relevant to them based on some criteria, like a received power threshold, and estimate and report the channels only for the relevant beams resulting in further reduction in the overall complexity and feedback overhead. A simple schematic of GoB is provided in Figure 2.6. Another benefit of GoB is that it does not entail any specific system requirements that would otherwise be the case for adaptive schemes, like for example, the Joint Spatial Division and Multiplexing (JSDM) framework which requires user grouping and works only in the LOS scenarios [KTC15].

Considering a single cell, after the base station has acquired full CSI knowledge from the explicit feedback of relevant channel components from each UE, it can calculate the optimum MU-MIMO precoder matrix $\mathbf{W}$. The overall system can be described by the following received signal in the downlink for all the UEs and for one particular sub-carrier,

$$\mathbf{y} = \mathbf{HVWx} + \mathbf{z} \tag{2.1}$$

where,

- $\mathbf{H} \in \mathbb{C}^{M_R \times M_T}$ is effective channel matrix of the M_R single-antenna UEs served by the cell with $M_T = 1024$ being the total number of transmit antennas
- $\mathbf{V} \in \mathbb{C}^{M_T \times M_{\text{GoB}}}$ contains the cell-specific unitary precoders for the static wide-band beams with $M_{\text{GoB}} = 32$ being the number of GoB and $M_{\text{GoB}} > M_R$.
- $\mathbf{W} \in \mathbb{C}^{M_{\text{GoB}} \times M_R}$ is the MU-MIMO precoder matrix
- $\mathbf{x} \in \mathbb{C}^{M_R \times 1}$ is the vector of user data, and
- $\mathbf{z} \in \mathbb{C}^{M_R \times 1}$ is the vector of Additive White Gaussian Noise (AWGN).

2.3.4 Implementation Challenges for 5G IMF-A

As we have discussed throughout this chapter, inter-cell interference is a major limiting factor for cellular networks and coordination between the cells of the network is one of the most promising approaches to deal with this interference. The main enablers of the 5G IMF-A framework like JT-CoMP and massive MIMO, are riddled with challenges which must be solved for an effective implementation. A flexible scheme for the formation of cooperation clusters is required which IMF-A handles by forming enlarged cooperation areas. The interference from outside the cluster is managed by interference floor shaping and the cover-shift concept. The other unresolved challenges have been the focus of this thesis, and we have investigated and proposed potential solutions for the following challenges:

Reduction in reference signal overhead can be handled by the novel pilot design concept called 'coded CSI' reference signals [ZAS16].

Efficient feedback compression and reduction in delay is achieved through a combination of spatial processing and channel prediction farther into the future as discussed in Chapter 5.

Availability of accurate CSI at the base station is accomplished through efficient channel prediction schemes with reduced complexity as discussed in Section 4.5.

2.4 Summary

In this chapter, we have discussed the advanced interference mitigation framework which has been developed to combat interference in future cellular networks. The key components of this framework are JT-CoMP implemented in combination with proper interference floor shaping. The evolution of this concept to 5G, incorporates small cells and massive MIMO arrays at the base stations. It has been identified that the main challenge for such a framework, especially in the below 6 GHz range and FDD, is the acquisition of accurate CSI at the base station. Concepts like the Grid of Beams can help to reduce the overall dimension of massive MIMO, hence reducing the CSI feedback overhead, and channel prediction can efficiently combat feedback delay and provide reliable CSI to the base station. In addition, we have developed an enhanced channel prediction scheme to predict weak channels efficiently as discussed in Chapter 4. Moreover, we have developed novel UE-sided beamforming schemes which enable us to simplify the massive MIMO channels leading to a better channel prediction performance. These schemes are discussed in Chapters 5 and 6.

Chapter 3

Wireless MIMO-OFDM Channel Models

Chapter Summary

In the previous chapter, we have provided an overview of the status of the cooperative communications framework in 5G. As discussed in Section 2.3.4, one of the main challenges to realize such a framework is the availability of accurate and up-to-date channel information at the base stations. The best potential solution to this is channel prediction but before we delve into the details of channel prediction, we need to take a look at the channels themselves. Hence, this chapter is dedicated to wireless communication channels. In Section 3.2, we discuss the channel effects that shape the transmitted signal like, for example, attenuation, multipath fading, Doppler shift, and scattering. In Section 3.3, we detail how a downlink channel can be represented at the UE. In Section 3.4, we discuss channel modeling schemes and particularly focus on ray-tracing channel modeling tools like IlmProp and WinProp. The benefit of using ray-traced channels is that characteristics of measured channels can be emulated to test the real world performance of algorithms.

3.1 Introduction

In the previous chapter, we have discussed both the baseline Interference Mitigation Framework - Advanced (IMF-A) and its evolution to the Fifth Generation (5G). We have also discussed the building blocks of such a framework and the dominant requirement for it to work; the availability of accurate Channel State Information (CSI). Before we delve into methods to acquire reliable CSI, like for example, channel prediction schemes, we first look at the composition and general description of wireless Multiple-Input and Multiple-Output (MIMO) channels, and analyze the various methods to model such channels.

3.2 Characterization of the Channel

In a wireless communications system, the channel is the logical connection between the transmitter and the receiver which conveys the information signal through electromagnetic waves. In contrast to a physical connection, like optical fiber or cable, the wireless channel is defined by the environment in which the transmitter and the receiver reside and it is very important to characterize it because decoding the transmitted signals correctly becomes an intractable problem without correct channel information. Hence, the wireless channel is usually characterized as a mathematical model which describes the various effects experienced by the signal as it travels between the transmitter and the receiver. All these effects shape the transmitted signal in a particular manner and are usually named accordingly, for example,

attenuation, multipath fading, Doppler shift, and scattering [SA05].

3.2.1 Multipath Propagation

As the wireless propagation environment is not fixed, the signals usually travel through two or more paths on their way to the receiver which are characterized by different attributes like phase shift, attenuation, and delay. At the receiver, all the signals from these paths superpose either in a constructive or a destructive manner resulting in peaks or deep fades in the channel envelop as illustrated in Figure 3.1; a phenomenon generally known as *fading*. It can be classified into three categories based on the time scale of variation:

Path-loss is the attenuation in signal power with distance that an electromagnetic wave experiences as it propagates through space. It is typically characterized by empirical methods, collectively known as *radio wave propagation models*, for example, the Okumura-Hata model [Sey05].

Slow fading, also called large-scale fading, is generally caused by the *shadowing* effects of a large obstruction on the main signal path, resulting in the signal power to vary slowly around the mean which is often modeled using a log-normal distribution [TV05].

Fast fading, also known as small-scale fading, usually stems from the movement of the transmitter, receiver and/or the scattering sources within the *multipath* environment and imposes fast changes in the amplitude and phase of the transmitted signal. It is the most challenging to deal with and proper system design is required to cater for it or preferably take advantage of it, for example, using time diversity.

3.2.1.1 Doppler Spread and Coherence Time

The relative motion between the base station and the UE and/or the movement of the scattering objects in the channel, induce a time-varying characteristic to the channel which can be described by two parameters called *Doppler spread* and *coherence time*.

In a multipath environment, if a pure sinusoidal tone of frequency f_C is transmitted, then there will be frequency components in the received signal spectrum, known as the *Doppler spectrum* $\Phi_h(f_D)$, ranging from $f_C - f_{D,\max}$ to $f_C + f_{D,\max}$, where $f_{D,\max}$ is the maximum *Doppler shift*. The term f_D is a function of the relative velocity of the UE, and the angle between the direction of its movement and the direction of arrival of the scattered waves. Based on the value of f_D, a broadening of the received spectrum occurs which is quantified by the *Root Mean Square*

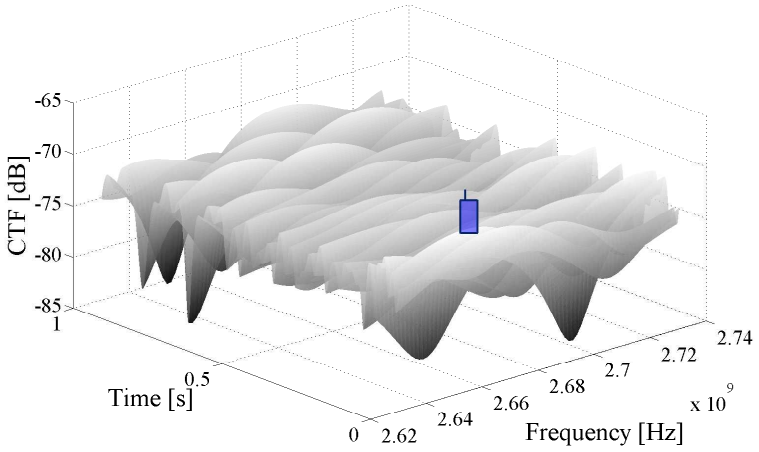


Figure 3.1: A general illustration of the effect of multipath fading. The blue rectangle represents a User Equipment (UE) and the example is generated using the IlmProp channel modeling tool.

(RMS) Doppler spread $(f_D)_{\text{RMS}}$. It is defined as the range of frequencies over which the received Doppler spectrum is essentially non-zero [Skl01],

$$(f_D)_{\text{RMS}} = \sqrt{\frac{\int_{-\infty}^{\infty} (f_D - \bar{f_D})^2 \Phi_h(f_D) \mathrm{d}f_D}{\int_{-\infty}^{\infty} \Phi_h(f_D) \mathrm{d}f_D}} \tag{3.1}$$

where $\bar{f_D}$ is the mean Doppler shift defined as,

$$\bar{f_D} = \frac{\int_{-\infty}^{\infty} f_D \Phi_h(f_D) \mathrm{d}f_D}{\int_{-\infty}^{\infty} \Phi_h(f_D) \mathrm{d}f_D} \tag{3.2}$$

The *Coherence time* $(\Delta t)_C$, which is a statistical measure of the time duration over which the Channel Impulse Response (CIR) is essentially invariant, enables the time domain characterization of the time-varying nature of the frequency dispersiveness of the wireless channel [Rap96]. The implication is that two signals arriving with a greater time separation than $(\Delta t)_C$ would be affected differently by the channel. Coherence time is the time domain dual of the RMS Doppler spread and is inversely proportional to it,

$$(\Delta t)_C \propto \frac{1}{(f_D)_{\text{RMS}}} \tag{3.3}$$

If the bandwidth of the baseband signal exceeds the maximum Doppler spectrum (B_D), the effects of Doppler spread at the receiver are negligible. Hence, the channel is characterized as a *slow fading* channel. On the contrary, if the coherence time of the channel is shorter than the reciprocal bandwidth of the baseband signal, then during the transmission of the signal, the channel will change causing distortion at the receiver. Hence, the channel is characterized as a *fast fading* channel.

3.2.1.2 Delay spread and Coherence Bandwidth

The time dispersive nature of the wireless channel is described by the *Delay spread* and the *coherence bandwidth*. The delay spread is a measure of the multipath richness of a communications channel and can be understood as the difference in the arrival time of the significant earliest and the latest multipath components. It has a profound effect on the Inter-Symbol Interference (ISI) and is usually quantified in terms of the RMS delay spread. To calculate RMS delay spread, we first need to determine the mean excess delay $\bar{\tau}$, which is given as the first moment of the power delay profile $\varphi_h(\tau)$ [TV05],

$$\bar{\tau} = \frac{\int_0^\infty \tau \varphi_h(\tau) \mathrm{d}\tau}{\int_0^\infty \varphi_h(\tau) \mathrm{d}\tau} \tag{3.4}$$

The second moment of the mean excess delay is given as,

$$\overline{\tau^2} = \frac{\int_0^\infty \tau^2 \varphi_h(\tau) \mathrm{d}\tau}{\int_0^\infty \varphi_h(\tau) \mathrm{d}\tau} \tag{3.5}$$

The RMS delay spread is defined as the second central moment of the power delay profile,

$$\tau_{\mathrm{RMS}} = \sqrt{\frac{\int_0^\infty (\tau - \bar{\tau})^2 \varphi_h(\tau) \mathrm{d}\tau}{\int_0^\infty \varphi_h(\tau) \mathrm{d}\tau}} \tag{3.6}$$

$$= \sqrt{\overline{\tau^2} - (\bar{\tau})^2} \tag{3.7}$$

The *coherence bandwidth* $(\Delta f)_C$, is a statistical measure of the range of frequencies over which two frequency components have a strong potential for amplitude correlation. Hence, if the frequency separation between two sinusoids is greater than $(\Delta f)_C$, they are distinctly affected by the channel. The coherence bandwidth is

related to the RMS delay spread as,

$$(\Delta f)_C \propto \frac{1}{\tau_{\text{RMS}}} \tag{3.8}$$

Based on the coherence bandwidth, we can classify a wireless channel into the following two categories:

Flat fading (narrowband): A channel can be qualified as *flat*, if it passes all the spectral components with roughly an equal gain and a linear phase. This means that the coherence bandwidth of the channel exceeds the transmission bandwidth of the signal.

Frequency-selective fading (wideband): In a frequency-selective channel, the different spectral components encounter uncorrelated fading. In this case, the coherence bandwidth of the channel is smaller than the transmission bandwidth of the signal.

3.3 Representation of the Channel

In the previous section, we have discussed the various types of fading inherent in wireless channels. In this section, we will present an introduction to the modeling of a wireless channel in a future 5G system. As the system incorporates multiple antennas at both ends, employs multi-carrier transmission schemes, multiple polarizations and elevation tilts, the dimensionality of the channel increases accordingly. We have elaborated the proper representation and dimension of this channel.

3.3.1 Multiple-Input and Multiple-Output

MIMO has become an essential part of current and future cellular networks. In a traditional sense, MIMO translates to a provisioning of multiple antennas at both the transmitter and the receiver, for example, point-to-point or Single-User MIMO (SU-MIMO). However, in a modern usage the term MIMO can be used for all the techniques that simultaneously support multiple data streams over the same radio channel, for example, like Multi-User MIMO (MU-MIMO) and COoperative MIMO (CO-MIMO).

Consider that we have, in the downlink of a Frequency Division Duplex (FDD) system, M_T antennas at the base station and M_R antennas at the UE, in the case of SU-MIMO or M_R single-antenna UEs in the case of MU-MIMO. The time-varying MIMO channel can then be represented by an $M_R \times M_T$ matrix $\mathbf{H}_d(t, \tau)$, which is

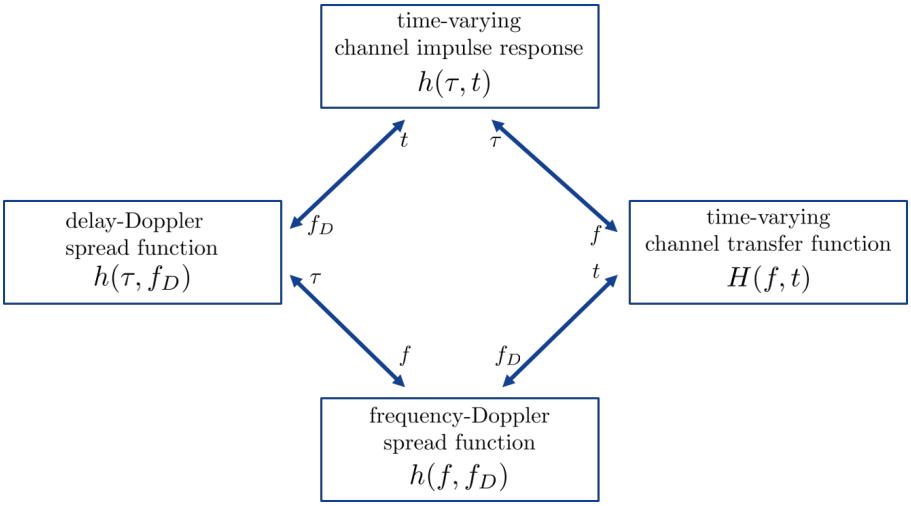


Figure 3.2: An illustration of the first set of Bello functions.

defined as,

$$\mathbf{H}_d(\tau, t) = \begin{bmatrix} h_{11}(\tau,t) & h_{12}(\tau,t) & \dots & h_{1M_T}(\tau,t) \\ h_{21}(\tau,t) & h_{22}(\tau,t) & \dots & h_{2M_T}(\tau,t) \\ \vdots & \vdots & \ddots & \vdots \\ h_{M_R1}(\tau,t) & h_{21}(\tau,t) & \dots & h_{M_RM_T}(\tau,t) \end{bmatrix} \tag{3.9}$$

where each entry $h_{m_Rm_T}(\tau, t)$ denotes the time-variant CIR between the $m_R{}^{\text{th}}$ receive and $m_T{}^{\text{th}}$ transmit antennas. It quantifies the attenuation and phase shift induced by the wireless channel in the signal that arrives at the receiver with a delay τ,

$$h_{m_Rm_T}(\tau, t) = \sum_{n_L=1}^{N_L} \alpha_{n_L}(t)\delta(\tau - \tau_{n_L}) \tag{3.10}$$

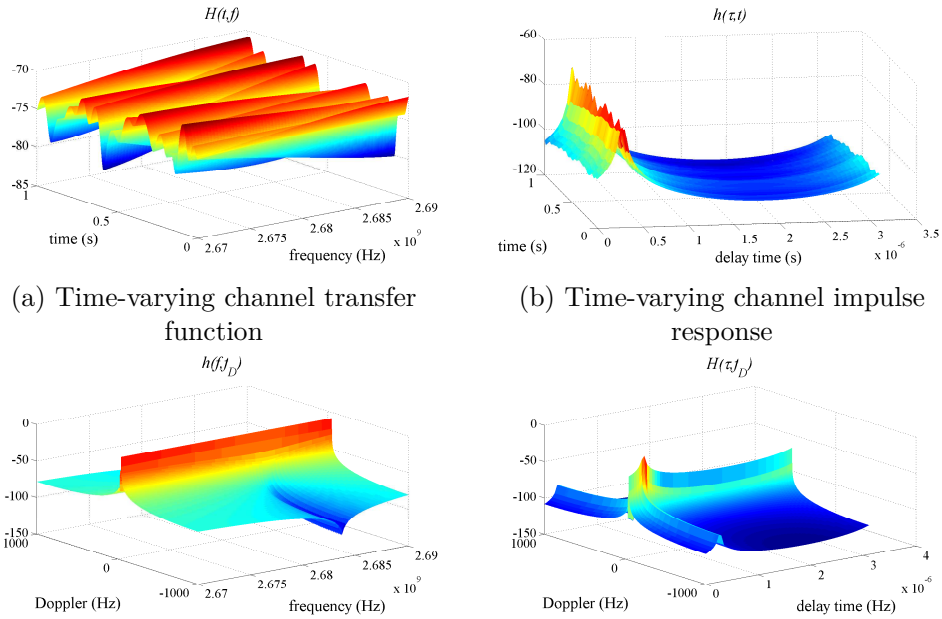
where α_{n_L} and τ_{n_L} denote the complex amplitude and delay associated with the $n_L{}^{\text{th}}$ path respectively. Note that we will mostly consider downlink transmission unless stated otherwise but we will drop the subscript d from now on for ease of notation. Utilizing the channel matrix, the overall MIMO system can now be modeled by the received signal as,

$$\mathbf{y}(t) = \int_\tau \mathbf{H}(\tau, t)\mathbf{x}(t - \tau)\mathrm{d}\tau + \mathbf{n}(t) \tag{3.11}$$

where $\mathbf{x} = [x_1(t) \ \dots \ x_{M_T}(t)]^{\mathrm{T}}$ is the transmit signal vector and $\mathbf{n}(t)$ is a vector modeling the effect of Additive White Gaussian Noise (AWGN) and interference. If we assume that the bandwidth of the transmitted signal is narrower than the coherence bandwidth, then the channel frequency response can be considered flat across frequency. The channel frequency response can be computed through a Fourier transform of the CIR as they are Fourier transform pairs with respect to the

delay variable τ according to the first set of Bello functions[1] as shown in Figure 3.2 [Bel64]. This leads to a discrete-time description of the MIMO system corresponding to Equation (3.11), given as,

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{3.12}$$



(a) Time-varying channel transfer function

(b) Time-varying channel impulse response

(c) Frequency-Doppler spread function

(d) Delay-Doppler spread function

Figure 3.3: An example of the various equivalent representations of the time-varying channel through Bello functions. The channel corresponds to the IlmProp scenario depicted in Figure 3.5 with a central frequency of 2.6 GHz and 20 MHz bandwidth.

3.3.2 Polarization

Current cellular networks, like Fourth Generation (4G) Long Term Evolution (LTE) systems, use antennas at the base station which are usually dual polarized. Polarization specifies the geometrical orientation of the electric field emitted by the antenna in the far field. The benefit of dual polarization, in addition to the spatial multiplexing gain provided by MIMO, is polarized multiplexing which can provide further improvements in spectral efficiency or transmit diversity.

For a single polarization, the equations in the previous section would hold, but for dual polarization, each antenna housing contains two antennas ports which comprise an antenna element, one using vertical polarization (v) and the other horizontal polarization (h). In this case, α_{n_L} in equation (3.10) is replaced by a 2×2

[1]Bello system functions allow the characterization of a linear time-variant function through four equivalent descriptions. An example of these descriptions for an IlmProp channel is provided in Figure 3.3.

polarimetric weight matrix,

$$\boldsymbol{\Upsilon}_{n_L} = \begin{bmatrix} \alpha_{n_L}^{\mathrm{vv}} & \alpha_{n_L}^{\mathrm{vh}} \\ \alpha_{n_L}^{\mathrm{hv}} & \alpha_{n_L}^{\mathrm{hh}} \end{bmatrix} \tag{3.13}$$

This signifies that each entry in the matrix $\mathbf{H}_d(\tau, t)$ in equation (3.9) is replaced by a polarimetric sub-matrix resulting in an increase in the total number of antennas and consequently the size of the matrix.

3.3.3 Orthogonal Frequency Division Multiplexing

It has been established that 5G systems will use some form of multi-carrier transmission and new schemes are under consideration like, for example, Filter Bank Multi-Carrier (FBMC) [PM16, PNCZ+16]. However, the potential benefits and practical impairments of these schemes are still under investigation. Hence, in this thesis, we will consider one of the most widely used multi-carrier schemes called Orthogonal Frequency Division Multiplexing (OFDM) [NP00, LHKK15].

OFDM is a subset of Frequency Division Multiplexing (FDM) in which a single wideband channel comprises numerous densely-packed and overlapping narrowband *sub-channels*. These sub-channels, also called *sub-carriers*, are orthogonal to each other and are used to carry parallel data streams without interfering with each other. The individual sub-carriers are usually modulated through a conventional modulation scheme like Quadrature Amplitude Modulation (QAM). OFDM provides several advantages over single-carrier schemes:

- It can cope with severe channel conditions like frequency-selective fading and narrowband interference without complex time-domain equalization filters. It simplifies channel equalization by dealing with numerous slow modulated narrowband sub-channels where each can be expressed as a scalar complex gain, instead of a fast modulated wideband channel.
- Due to the orthogonal nature of the sub-carriers and insertion of guard interval, it is robust against ISI.
- It exhibits a low sensitivity to time synchronization errors and can be implemented efficiently utilizing Fast Fourier Transform (FFT).
- OFDM facilitates link adaptation in multi-user systems by permitting the division of the radio resource into small time-frequency resource blocks. These blocks can use separate link adaptation parameters and can be allocated to different users.

The drawbacks of OFDM are its high sensitivity to frequency offsets and Doppler shifts, high Peak-to-Average-Power Ratio (PAPR), efficiency loss caused by the

insertion of cyclic prefix, and the requirement of a relatively complex receiver architecture.

To describe the OFDM channel, we start with the description of the single antenna case, i.e., Single-Input and Single-Output (SISO). The data stream to be transmitted is mapped onto symbols from a known constellation, and generally additional measures like encoding and interleaving are also taken for robustness against errors. These symbols are then partitioned into frequency domain vectors $\mathbf{x}_B$ of length N_{SC} so that each element in the vector corresponds to one OFDM sub-channel. The subscript B indicates that the full bandwidth of the N_{SC} sub-carriers is being considered here. The Inverse Fast Fourier Transform (IFFT) is used to convert each vector $\mathbf{x}_B$ into a vector $\mathbf{F}^{\mathrm{H}}\mathbf{x}_B$ of length N, where $\mathbf{F}$ is the $N_{\mathrm{SC}} \times N_{\mathrm{SC}}$ Fourier matrix, given as,

$$\mathbf{F}[a,b] = \frac{1}{\sqrt{N_{\mathrm{SC}}}}\, e^{-2\pi \jmath ab/N_{\mathrm{SC}}}; \quad a,b = 0,1,\ldots,N_{\mathrm{SC}}-1 \tag{3.14}$$

which implies that $\mathbf{F}\mathbf{F}^{\mathrm{H}} = \mathbf{I}_{N_{\mathrm{SC}}}$. For transmission, a *cyclic prefix* is appended to the signal, the length of which is set to surpass the maximum expected delay spread of the channel. It is then transformed by a digital-to-analog converter into a time domain analog signal which modulates the carrier. For reception, an analog-to-digital converter brings the noisy and distorted signal back into the digital domain, and the cyclic prefix is removed. If the baseband CIR of length N_L is represented by $\mathbf{g}$, then the received signal vector in time domain can be represented (by omitting the time index for simplicity) as [Aro11],

$$\mathbf{y}_t = \mathrm{circ}([\mathbf{g}^{\mathrm{T}} \quad \mathbf{0}_{1\times(N_{\mathrm{SC}}-N_L)}])\mathbf{F}^{\mathrm{H}}\mathbf{x}_B + \mathbf{n} \tag{3.15}$$

where $\mathbf{n}$ is the AWGN vector with covariance matrix $\sigma_n^2\mathbf{I}_{N_{\mathrm{SC}}}$ and considering a vector $\mathbf{a} = [a_0\, a_1\, \ldots a_{n-1}]$, the circ$(\cdot)$ operator is defined as,

$$\mathrm{circ}(\mathbf{a}) = \begin{bmatrix} a_0 & a_1 & \ldots & a_{n-1} \\ a_{n-1} & a_0 & \ldots & a_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & \ldots & a_0 \end{bmatrix} \tag{3.16}$$

Hence, circ$(\mathbf{a})$ is a square circulant matrix with $\mathbf{a}$ as its first row and its eigenvalue decomposition is given as,

$$\mathrm{circ}(\mathbf{a}) = \mathbf{F}^{\mathrm{H}}\mathrm{diag}(\sqrt{N_{\mathrm{SC}}}\mathbf{F}\mathbf{a})\mathbf{F} \tag{3.17}$$

Therefore, we can apply the FFT to get the frequency domain received signal,

$$\begin{aligned}
\mathbf{y}_B &= \mathbf{F}\mathbf{y}_t \\
&= \mathbf{F}(\mathrm{circ}([\mathbf{g}^{\mathrm{T}} \quad \mathbf{0}_{1\times(N_{\mathrm{SC}}-N_L)}])\mathbf{F}^{\mathrm{H}}\mathbf{x}_B + \mathbf{n}) \\
&= \mathbf{F}\mathbf{F}^{\mathrm{H}}\,\mathrm{diag}(\mathbf{h}_B)\,\mathbf{F}\mathbf{F}^{\mathrm{H}}\mathbf{x}_B + \mathbf{F}\mathbf{n} \qquad (3.18)\\
&= \mathrm{diag}(\mathbf{h}_B)\mathbf{x}_B + \mathbf{z}_B \qquad (3.19)\\
&= \mathrm{diag}(\mathbf{x}_B)\mathbf{h}_B + \mathbf{z}_B \qquad (3.20)
\end{aligned}$$

where $\mathbf{h}_B = \sqrt{N_{\mathrm{SC}}}\mathbf{F}(\mathrm{circ}([\mathbf{g}^{\mathrm{T}} \quad \mathbf{0}_{1\times(N_{\mathrm{SC}}-N_L)}])$, is the channel frequency response or Channel Transfer Function (CTF) for the N_{SC} sub-channels, and $\mathbf{z}_B = \mathbf{F}\mathbf{n}$ is the noise vector in the frequency domain with a covariance matrix $\sigma_z^2\mathbf{I}_{N_{\mathrm{SC}}}$.

Now we extend the OFDM system description to MIMO systems, commonly known as MIMO-OFDM, it is a popular transmission scheme for multi-user systems. Consider we have M_R single-antenna users in a multi-user scenario communicating with a base station with M_T antenna elements. The transmission and reception chains contain the same constituting blocks and steps as described for the SISO case. As the UEs are separated appropriately from each other, their channels fade independently. Hence, in the context of channel prediction or estimation, we can observe only a single user without any loss of generality leading to a Multiple-Input and Single-Output (MISO) system description. Thus, the received MISO signal for M_T transmit antennas is given as,

$$\mathbf{y}_{B,\,\mathrm{MISO}} = \sum_{m_T=1}^{M_T} \mathrm{diag}(\mathbf{x}_B^{(m_T)})\mathbf{h}_B^{(m_T)} + \mathbf{z}_B \qquad (3.21)$$

If we horizontally stack the transmitted symbol matrices for all the individual transmit antenna elements into the following matrix,

$$\mathbf{X}_B = \begin{bmatrix} \ddots & & \ddots & & \ddots & \\ & \mathbf{x}_B^{(1)} & & \mathbf{x}_B^{(2)} & \cdots & \mathbf{x}_B^{(M_T)} \\ & \ddots & & \ddots & & \ddots \end{bmatrix} \qquad (3.22)$$

and the channels into a column vector $\mathbf{h}_{B,\,\mathrm{MISO}} = [(\mathbf{h}_B^{(1)})^{\mathrm{T}} \ \ldots \ (\mathbf{h}_B^{(M_T)})^{\mathrm{T}}]^{\mathrm{T}}$, we can represent the received signal in matrix form as,

$$\mathbf{y}_{B,\,\mathrm{MISO}} = \mathbf{X}_B\mathbf{h}_{B,\,\mathrm{MISO}} + \mathbf{z}_B \qquad (3.23)$$

Our focus is on pilot-based channel prediction schemes, so we assume Pilot Symbol Assisted Modulation (PSAM) [NESB97, Cav91]. The known *reference signals* or *pilots* are regularly placed into the time-frequency grid along with the data symbols

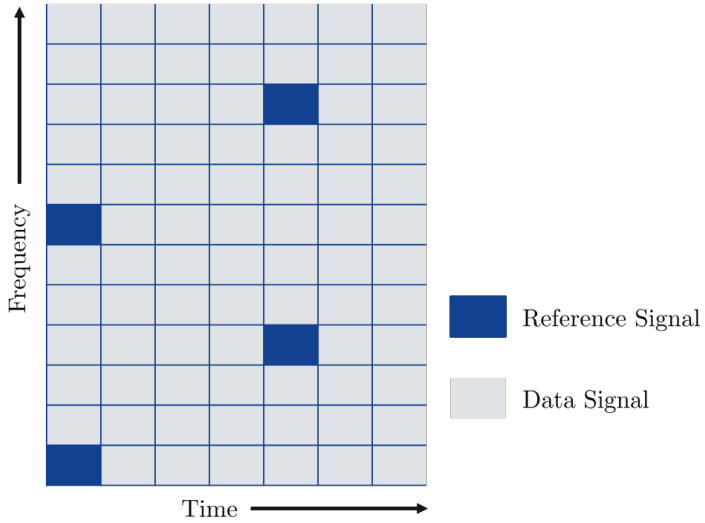


Figure 3.4: A generic representation, inspired from the reference signal placement in LTE, of the time-frequency grid in OFDM systems.

as shown in Figure 3.4. The channel prediction scheme then observes and operates on this pilot grid only and the channel estimates for all the other locations are obtained through interpolation or extrapolation. To extract the pilot symbols from the received signal $\mathbf{y}_{B,\,\mathrm{MISO}}$ spanning the complete bandwidth, we introduce a $N_{\mathrm{RS}} \times N_{\mathrm{SC}}$ *extraction matrix* $\mathbf{E}$ which upon multiplication with a vector of size $N_{\mathrm{SC}} \times 1$, extracts the appropriate N_{RS} reference signal containing Resource Elements (REs) at a time. If we denote the set of indices for the reference signal subset as $R = \{x_0, \ldots, x_{N_{\mathrm{RS}}-1}\}$, the extraction matrix contains N_{RS} ones,

$$\mathbf{E}[a, x_a] = \begin{cases} 1 & 0 \leq a < N_{\mathrm{RS}} \\ 0 & \text{Otherwise} \end{cases} \tag{3.24}$$

An example of the extraction matrix for a small number of sub-carriers (for the purpose of illustration), $N_{\mathrm{SC}} = 10$, the reference signals, $N_{\mathrm{RS}} = 4$, and $R = \{4, 6, 7, 9\}$ is as follows,

$$\mathbf{E} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{3.25}$$

It is important to note that the extraction matrix $\mathbf{E}$ corresponds to a single time slot containing reference signals. For such an upcoming time slot, we would need another matrix $\mathbf{E}$ with ones in the appropriate positions for that time slot. The extraction matrix has the following properties [Aro11]:

1. Considering an $N_{\mathrm{SC}} \times 1$ vector $\mathbf{a}$, the operation $\mathbf{E}\,\mathrm{diag}(\mathbf{a})\,\mathbf{E}^{\mathrm{H}}$ produces a $N_{\mathrm{RS}} \times N_{\mathrm{RS}}$ diagonal matrix with elements taken from $\mathbf{a}$ which correspond to the indices in R.
2. $\mathbf{E}^{\mathrm{H}}\mathbf{E}$ is an $N_{\mathrm{SC}} \times N_{\mathrm{SC}}$ diagonal matrix with ones at the x_a^{th} diagonal entries and zeros otherwise.
3. Right multiplication with $\mathbf{E}^{\mathrm{H}}\mathbf{E}$ nulls out the columns of a matrix $\mathbf{A}$ in a way that $\mathbf{A}\mathbf{E}^{\mathrm{H}}\mathbf{E}$ is all zeros except for the columns which contain a 1 in $\mathbf{E}^{\mathrm{H}}\mathbf{E}$.
4. It holds that $\mathbf{E}\mathrm{diag}(\mathbf{a}) = \mathbf{E}\mathrm{diag}(\mathbf{a})\mathbf{E}^{\mathrm{H}}\mathbf{E}$ as the non-zero columns of the matrix $\mathbf{E}\,\mathrm{diag}(\mathbf{a})$ already correspond to the indices in R.

To obtain the measurement signal $\mathbf{y}$ corresponding to the N_{RS} pilot sub-carriers, we multiply equation (3.21) by $\mathbf{E}$:

$$
\begin{aligned}
\mathbf{y} &= \mathbf{E}\mathbf{y}_{B,\,\mathrm{MISO}} \\
&= \sum_{m_T=1}^{M_T} \{(\mathbf{E}\,\mathrm{diag}(\mathbf{x}_B^{(m_T)}))\,\mathbf{h}_B^{(m_T)} + \mathbf{E}\mathbf{z}_B^{(m_T)}\} && \text{(from Equation 3.21)} \\
&= \sum_{m_T=1}^{M_T} \{(\mathbf{E}\,\mathrm{diag}(\mathbf{x}_B^{(m_T)})\,\mathbf{E}^{\mathrm{H}})(\mathbf{E}\mathbf{h}_B^{(m_T)}) + \mathbf{z}^{(m_T)}\} && \text{(from the last property of } \mathbf{E}) \\
&= \sum_{m_T=1}^{M_T} \{\mathrm{diag}(\mathbf{x}^{(m_T)})\mathbf{h}^{(m_T)} + \mathbf{z}^{(m_T)}\} \\
&= \begin{bmatrix} \ddots & & \ddots & & \ddots & \\ & \mathbf{x}^{(1)} & & \mathbf{x}^{(2)} & \cdots & \mathbf{x}^{(M_T)} \\ & & \ddots & & \ddots & & \ddots \end{bmatrix} \begin{bmatrix} (\mathbf{h}^{(1)}) \\ \vdots \\ (\mathbf{h}^{(M_T)}) \end{bmatrix} + \begin{bmatrix} (\mathbf{z}^{(1)}) \\ \vdots \\ (\mathbf{z}^{(M_T)}) \end{bmatrix} \\
&= \mathbf{X}\mathbf{h} + \mathbf{z} && (3.26)
\end{aligned}
$$

where $\mathbf{X}$ is the block diagonal *pilot matrix* which holds known orthogonal reference signals, $\mathbf{h}$ is a column vector containing the corresponding channel coefficients which we need to estimate and $\mathbf{z}$ is the corresponding noise vector.

3.3.4 Tensor Channel of 5G IMF-A

To specify the dimensions of the downlink channel in the IMF-A employing Joint Transmission Coordinated MultiPoint (JT-CoMP) and MU-MIMO, let us consider a cooperation area comprising N_C cells where each cell has M_T transmit antenna elements, N_P polarizations per antenna element[1], N_E elevation tilts[2], N_{SC} OFDM sub-carriers, and N_t OFDM symbols. Then, the overall channel as seen by M_R

[1]Typically cross-polarized antennas are used with a $\pm 45°$ orientation leading to $N_P = 2$ as discussed in Section 3.3.2

[2]The first implementation of 5G IMF-A is targeting two elevation angles as shown in Figure 2.6

single-antenna UEs in the IMF-A framework can be represented by a tensor of order seven,

$$\boldsymbol{\mathcal{H}} \in \mathbb{C}^{N_C \times N_P \times N_E \times M_T \times M_R \times N_{\mathrm{SC}} \times N_t} \tag{3.27}$$

It is evident that performing prediction on this complete tensor channel is a monumental task. To carry out efficient prediction, we need to reduce the order and dimensions of the channel tensor. The first step in this direction is provided by the 5G IMF-A framework, as it condenses the large number of transmit antennas M_T into a manageable number of grid of beams M_{GoB}. In addition, the grid of beams in the azimuth and elevation dimension can be regarded together with a corresponding assignment of antenna ports to each beam, hence eliminating the need for the elevation tilt dimension. This results in a 6^{th} order channel tensor,

$$\boldsymbol{\mathcal{H}} \in \mathbb{C}^{N_C \times N_P \times M_{\mathrm{GoB}} \times M_R \times N_{\mathrm{SC}} \times N_t}. \tag{3.28}$$

The channel prediction algorithms exploit correlations present in the various dimensions to predict the channel in the future because we cannot infer anything about the channel in a dimension where there is no correlation as the past is independent of the future. Hence, further dimensionality reduction can be achieved by individually predicting the channels for dimensions with little to no correlation. The cells belonging to separate sites exhibit almost no correlation because they are segregated by large distances in an urban scenario. Even the cells that belong to a single site might exhibit little correlation because they are usually radiating towards different directions with a distinct set of scatterers. Moreover, there is usually little correlation between the polarizations and in a MU-MIMO scenario, the users that are typically randomly distributed within the cooperation area, experience uncorrelated channels. Hence, we can consider only a single user at a time. This also holds in the case of SU-MIMO with a single user having M_R antennas elements, as UE-sided beamforming can be employed which reduces the M_R dimension to 1 as the individual channels for the UE antennas are replaced by the effective beamformed channel. In addition, we only need to consider the N_{RS} sub-carriers corresponding to reference signals as we perform pilot-aided channel prediction. Hence, for a particular cell n_C, a certain polarization n_P, and a single effective receive antenna m_R, the channel tensor has now only order 3,

$$\boldsymbol{\mathcal{H}} \in \mathbb{C}^{M_{\mathrm{GoB}} \times N_{\mathrm{RS}} \times N_t} \tag{3.29}$$

3.4 Channel Modeling Approaches and Tools

Channel modeling, in a strict sense, is a term which stands for the modeling of the CIR [Del07]. However, in a wider sense, it denotes the modeling of all the physical processes that effect a signal on its way, from the transmitter to the receiver. The motivation to develop accurate channel models is to better understand the effect of parameters such as carrier frequency, bandwidth, delay and Doppler spread, and speed of the UE on the wireless channels. It also has the following objectives from a communications research and development standpoint:

- enables better comprehension of the physics of wireless systems which is crucial to efficiently design upcoming communications schemes,
- allows us to evaluate the performance of a potential scheme in a realistic environment, and
- it provides a standard arena to compare two or more schemes.

Modeling of SISO systems is fairly easy as it is sufficient to consider just the fading statistics of the channel. This can be achieved by considering Doppler shifts and power-level distributions and the resulting channels are quite realistic with respect to these characteristics. In contrast, modeling MIMO systems is not that trivial as they incorporate the spatial dimension which leads to the consideration of additional parameters like the angles of arrival and departure, and arbitrary antenna geometries. As we are interested in MIMO or MIMO-OFDM prediction schemes is this thesis, we will discuss modeling schemes and tools for single or multi-user time-variant frequency-selective wideband MIMO channels.

The general categorization of channel models is a complicated undertaking as there are a huge number of channel models which typically employ the same modeling principle and differ only in the finer details. One way to classify channel models is to distinguish between them based on the potential applications of the synthetic channels like, for example, *link-level* or *system-level* channel modeling. An alternative categorization can be based on the modeling principle like, for example, *stochastic channel models*, *geometry-based channel models* and *directional channel models*.

3.4.1 Geometry-based Channel Models

Geometry-based channel models, also known as *deterministic* channel models, use the underlying physical phenomenon governing wireless propagation to model the channel. This modeling approach, is site dependent and suitable for link-level simulations, as it takes into consideration all the geomorphological features of a propagation environment. This requires momentous effort as the electromagnetic properties

and position of everything affecting propagation have to be known. The interaction between the electromagnetic waves and the propagation environment can be modeled in two ways:

The full-wave approach: This approach directly applies Maxwell's equations, in either differential or integral form, as they can, in combination with electromagnetic boundary conditions, provide a measurement of the field strength at all points and times. A *brute force* method can solve either formulation of Maxwell's equations, or the well-known *Method of Moments (MoM)* can be used for the integral formulation. Alternatively, methods like the *Finite Element Method (FEM)* and *Finite Difference Time Domain (FDTD)* method can be used to solve the differential formulation. The accuracy of all these methods is remarkable but comes at the expense of exorbitant computational requirements in most environments. Hence, a more practical approach is to employ approximations to Maxwell's equations as a basis for solution.

The Geometrical Theory of Diffraction (GTD): This approach approximates the planar electromagnetic waves as *rays* that follow Snell's laws for *reflection* and *transmission*, and, in addition, might also allow the appropriate inclusion of *diffraction* and *diffuse scattering* [Kel62, BK94]. Thus, GTD-based models are commonly referred to as *ray-tracing* models. These models become more accurate for higher frequencies and for environmental objects much larger than the wavelength. This is in accordance with the Rayleigh criterion which allows us to distinguish between smooth and rough reflecting surfaces and is given as,

$$\Delta q < \frac{\lambda}{8 \cdot \cos(\psi)} \tag{3.30}$$

where Δq is the standard deviation of the roughness height to a reference height and ψ is the angle of incidence. A smooth surface will lead to specular reflection whereas a rough surface will lead to diffuse scattering which is the most challenging phenomenon to describe appropriately.

There are several ray-tracing based channel models available in literature like [vDH94, FPK+06, dJH99, KGW+14], but in this work, we will consider two modeling tools that we will use later in our experiments and simulations, namely *IlmProp* and *WinProp*.

3.4.1.1 IlmProp Channel Model

IlmProp is a geometry-based channel modeling tool for wireless communications developed at the Communications Research Laboratory of TU Ilmenau [Del07,

DHS03]. IlmProp exhibits the beneficial properties of both ray-tracers and hybrid or directional channel models. It approximates the propagation as a sum of rays, like hybrid channel models but extensively models the physical environment by specifically defining the complete trajectories of the multipath components and by an accurate modeling of the 3-Dimensional (3D) polarimetric antenna radiation patterns. However, unlike traditional ray-tracers, it does not require the electromagnetic properties of all the physical objects currently in the environment. IlmProp can be used both in a measurements-based approach and in a non-measurements-based approach but we are interested in the latter approach as for measurements, we use WinProp as explained in Section 3.4.1.2. This allows us to arbitrarily define parameters to generate channels which optimally help us to investigate the performance of particular schemes in various scenarios.

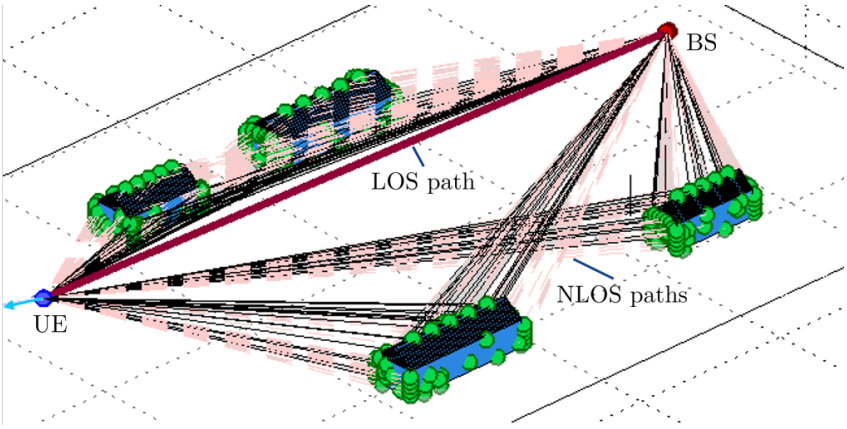


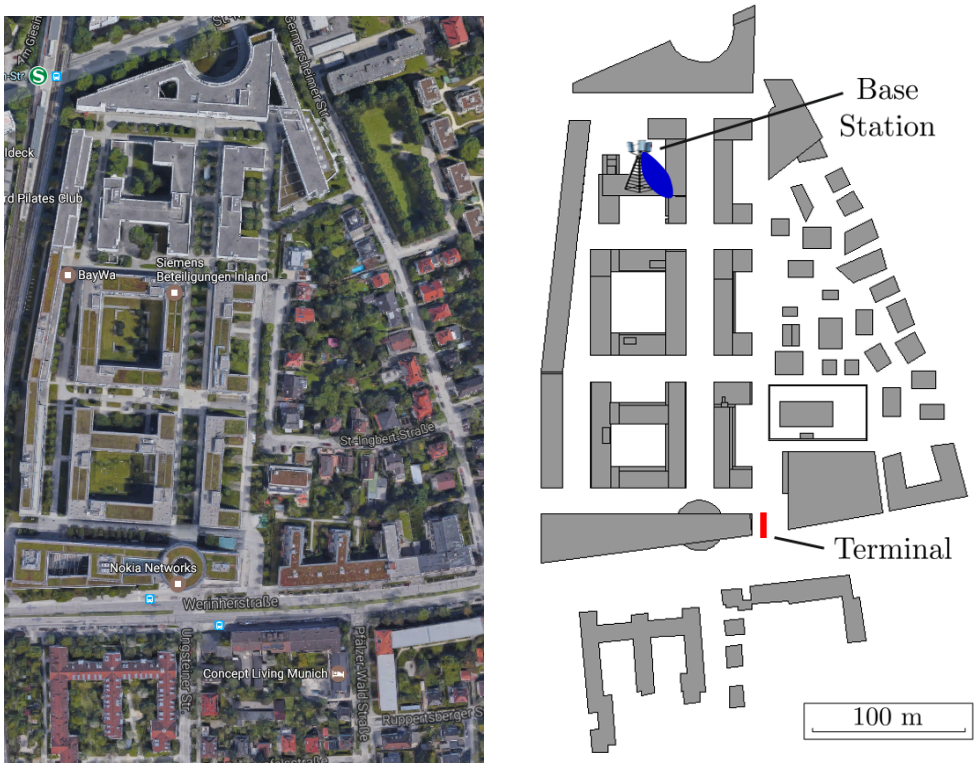
Figure 3.5: An urban macro propagation scenario generated with the IlmProp

IlmProp assumes an uplink channel with the base station as the receiver, and the UE or UEs as the transmitter(s). However, we can use the computed CIR for the downlink as well since wireless links are, in general, reciprocal [Kin63]. As an example, Figure 3.5 depicts a simple urban macro scenario where the UE is moving away from the base station. There are four buildings in the transmission path with a total of 280 scatters. The UE also sees a direct path from the transmitter. The parameters (delay spread, Doppler spread, etc.) of the channels created with the IlmProp are fairly close to real channel parameters in the corresponding scenario.

The benefit of IlmProp is that we can define all the simulation parameters like the number of antenna elements at both the base station and the UE, OFDM simulation parameters, bandwidth, time sampling frequency, and center frequency of the system differently. Thereby, we can modify one or two parameters and analyze their effect on the performance of our prospective schemes. This gives us a lot of flexibility to rigorously test our schemes in practical situations.

3.4.1.2 WinProp Channel Model and Measured Channel

WinProp - Propagation Modeling, is a simulation tool which is now a part of the FEKO software suite by Altair HyperWorks [Win13]. WinProp is a ray-tracer that provides accurate empirical propagation models for both indoor and outdoor macro scenarios. It includes rigorous 3D ray-tracing models capable of predicting the path loss, delay and angular spread. Moreover, the directional CIR, angular profile and propagation paths can be computed for both Line-Of-Sight (LOS) and Non-Line-Of-Sight (NLOS) cases.



(a) Google Maps picture of Nokia Campus in Munich

(b) WinProp model of Nokia Campus in Munich

Figure 3.6: A comparison of the map of Nokia campus and its simplified model used in WinProp

In this thesis, we use WinProp to produce ray-traced channels for a single site macro scenario at the Nokia campus in Munich as shown in Figure 3.6b. We assume a base station at a height of 30 meters, equipped with an Uniform Linear Array (ULA) of 16 standard Kathrein antennas capable of forming narrow beams with a roughly 2° Half-Power BeamWidth (HPBW). The central frequency of the system f_C is 2.6 GHz and the system bandwidth is 20 MHz. The UE moves in a straight line away from the base station, at an approximate distance of 300 m and at a height of 2.1 m, with a 1 cm resolution, resulting in a total of 50 locations (represented

by the red rectangle in Figure 3.6b). For each of these receiver locations within the propagation environment, WinProp computes trajectories and path-strengths for all multipath components as shown in Figure 3.7.

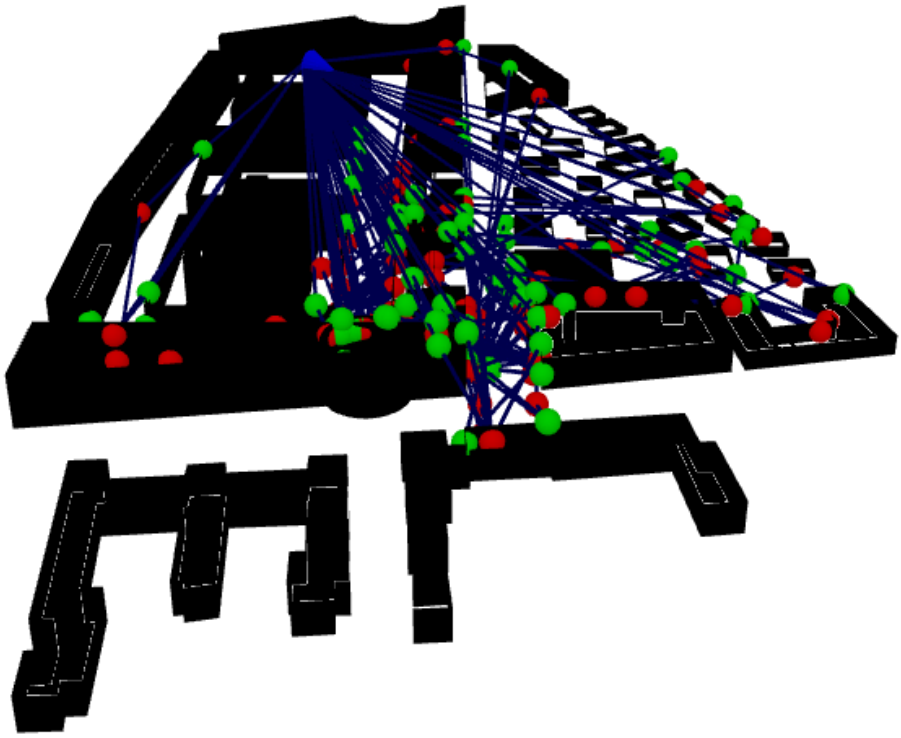
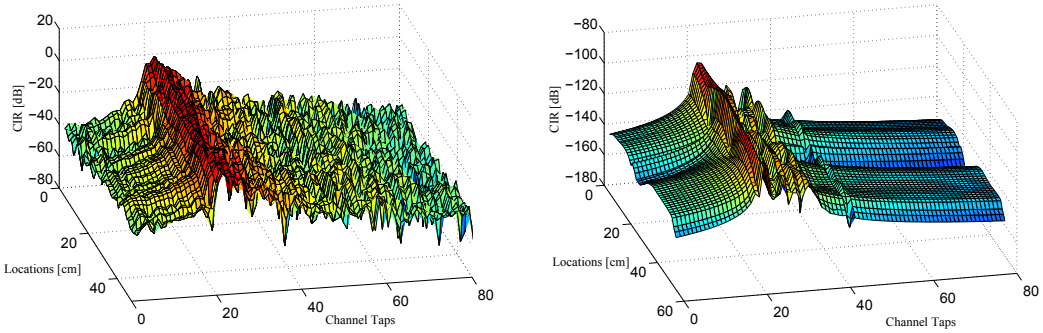


Figure 3.7: A depiction of the trajectories and last interaction points of rays arriving at one UE location.

A benefit of using WinProp with a model of the Nokia campus is that we have real-world measurements available, for the same scenario, but for a single antenna BS and UE. This allows us to validate the channels produced by WinProp and to test the real-world performance of our schemes and algorithms. In WinProp, we can add more BS antennas to simulate massive MIMO or we can define different movement resolutions and locations for the UEs to evaluate the NLOS conditions and efficiency of beamforming algorithms. This is not possible with measurements alone, as acquiring channel measurements is a time consuming and expensive process.

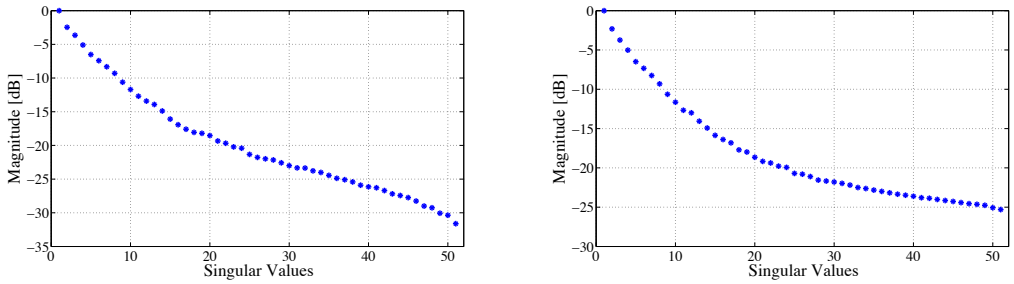
A comparison of the measured and ray-traced CIR produced by WinProp is given in Figure 3.8. The channels generated by the ray-tracer exhibit similar characteristics as compared to the real measurements which shows the capability of WinProp to model real-world scenarios. However, differences are often present in modeled channels as compared to their corresponding measured channels, for example, in this case the measurements exhibit far more variations. The main contributing factors leading to this difference are important to understand. They are typically



(a) Measured CIR (b) Ray-traced CIR

Figure 3.8: A comparison of the CIR of the measured and ray-traced channels.

the death and birth of multipath components, diffuse scatterers in the environment which are usually not modeled by the ray-tracers, or the extremely high number of multipath components comprising a typical outdoor channel component. Another important fact to consider is that ray-tracing is computationally complex, hence some trade-offs have to be made to lower the complexity like limiting the number of reflections and diffractions for all multipath components. Also, it is hard to develop accurate models which perfectly represent the environment. Some slight ray-tracing model simplifications are inevitable as shown in the comparison in Figure 3.6.



(a) Singular values of measured CIR (b) Singular values of measured CTF

Figure 3.9: A comparison of the singular values of CIR and CTF of the measured channel revealing the underlying low-rank structure in both domains.

It is interesting to check whether the same measured channel exhibits a low-rank structure. This can be accomplished through a matrix Singular Value Decomposition (SVD). The normalized singular values of the channel in both time domain and frequency domain are plotted in Figure 3.9. We can observe that it does indeed have an underlying low-rank structure in both domains. For an exemplary threshold of -20 dB, we can see that in both cases the number of significant singular values is reduced by a factor of two. This can help us to remove noise from the channel measurements and leads to a much smoother CIR or CTF. For this reason, we have

developed a Higher Order Singular Value Decomposition (HOSVD)-based denoising scheme, detailed in Section 4.5, which additionally takes the multi-dimensional structure of the massive MIMO channel into account to mitigate the effect of noise on the channel, prior to prediction with a state-of-the-art scheme.

3.5 Summary

In this chapter, we have discussed the underlying phenomena that govern the time-varying and frequency-selective nature of the wireless channel. We have also explored the overall dimension of a massive MIMO channel in a 5G setting and presented ways to properly represent it mathematically. In the latter part of the chapter, we discuss the various ways in which a wireless channel can be modeled and explicitly investigate ray-tracing based methods as they can provide a channel model which is almost identical to a measured channel in the same scenario, provided that an accurate model of the environment is available. The two ray-tracing tools considered and used in this thesis are the IlmProp and WinProp because both of them have their respective merits and benefits for different types of analysis.

Chapter 4

Channel Prediction Schemes

Chapter Summary

This chapter is dedicated to channel prediction, one of the main enablers of cooperative communications in 5G. In this chapter, we have provided a survey of channel prediction schemes in Section 4.2. Afterwards, we explore two state-of-the-art channel prediction schemes: Kalman-based channel prediction in Section 4.3 and a tensor-based prediction scheme in Section 4.4. Finally, we have detailed our novel channel prediction scheme in Section 4.5 which combines the benefits of both previously discussed schemes to a provide better prediction performance especially for channels received with poor Signal-to-Noise Ratio (SNR). This HOSVD-based denoising scheme has been published in [AZH17].

4.1 Introduction

Channel estimation and prediction schemes have been investigated and improved for a long time as they are an essential component of most practical communications systems and particularly of Multi-User MIMO (MU-MIMO) systems [LCS98, LCG+12]. For downlink transmission, in order to successfully decode the transmit signal at the User Equipment (UE), an accurate channel estimate is required. In addition, capacity or spectral efficiency gains are achieved through strong spatial multiplexing which requires an accurate knowledge of Channel State Information at the Transmitter (CSIT) for all UEs especially in Frequency Division Duplex (FDD) systems where channel reciprocity is not present. In such systems, the Channel State Information (CSI) is estimated at the receiver and then fed back to the transmitter. In practical systems, the CSIT is usually imperfect due to mobility and the corresponding out-dating or aging of the channel, which implies a mismatch between the actual and fed-back radio channels due to feedback delay [TOKM11]. *This is where channel prediction comes into play as it can efficiently combat feedback delays for uplink reporting and therefore has been identified as one of the main enablers for Coordinated MultiPoint (CoMP) and massive MIMO* [BA15]. Furthermore, large prediction horizons are a means for less frequent reports and a correspondingly low reporting overhead [TH13], or for improved power control, Multiple-Input and Multiple-Output (MIMO) precoding, multi-user scheduling, and adaptive modulation [AS14].

4.2 A Survey of Channel Prediction Schemes

Several criteria can be used to classify channel prediction schemes as discussed in the following.

The dimensionality of the channel being predicted. The channel could be a Single-Input and Single-Output (SISO) channel or a MIMO channel. Channel prediction of SISO channels, especially ones experiencing flat fading, has been studied extensively in the past. It is usually implemented by modeling the time variation in either the channel impulse response $h(\tau, t)$ or the channel transfer function $H(f, t)$ (which are interchangeable representations as governed by the Bello functions shown in Figure 3.2), as an AutoRegressive (AR) process of a certain order. This approach has been followed in [CV09,SAE03,DH07,DHHH00b,3GP17,Ekm02, HKM10, HKM10, HK07, LS08, HW98, OHH04]. The future state of the channel is then predicted using past estimates by employing a linear predictor that minimizes the Mean Square Error (MSE).

A question now arises on the motivation behind the extensive use of predictors based on AR models. To answer this, we assume the plane wave approximation which typically holds for macro scenarios. Under such an approximation, a distant cluster of reflectors or scatterers can be modeled as a damped complex sinusoid which, in turn for exponential damping as is the case for a distant Cauchy cluster, can be modeled as a stochastic AR1 process [Ekm02]. Thus, for channels with a single cluster of reflectors or scatterers, an AR1 process is a good approximation of the channel dynamics. However, in order to describe the contributions from many clusters, the AR1 processes can be summed up which leads to an Auto Regressive Moving Average (ARMA) process. But, when the poles of such an ARMA model lie near the unit circle, an AR model can render a good approximation to the dynamics of, for example, a fading channel tap. This is advantageous as it is typically harder to estimate zeros as compared to the poles [Ekm02].

Most MIMO channel prediction schemes adapt these SISO channel prediction schemes as well by individually applying these algorithms to each of the MIMO links. This is not an effective solution as correlation is typically present between the antenna elements at the transmitter or the receiver, which in this case hampers performance but can be exploited for a performance gain in a properly designed MIMO channel prediction scheme.

The availability of reference signals at the receiver. In the case of channel estimation, we can distinguish between *blind* or *pilot-based* schemes based on the availability of reference signals. Blind channel estimation schemes typically fall under the umbrella of non-coherent detection, meaning that the phase information is not available at the receiver. An equalization filter is inferred from the received signal which is the inverse of the Channel Impulse Response (CIR), rather than the estimation of the CIR itself [SHP07, MDCD02]. When it comes to channel prediction, we cannot do it in a similar fashion. Hence, current systems like Long Term

Evolution (LTE) and Long Term Evolution - Advanced (LTE-A) employ coherent detection and it is strongly believed that the Fifth Generation (5G) systems would also employ some form of coherent detection. In pilot-based estimation and prediction, the channel is observed through pilots or reference signals which are located on fixed positions in the time-frequency grid and are known to the receiver.

The type of correlation that is exploited to infer the past information to predict the future channel. In pilot-based prediction schemes, the information contained in the reference signals is available to the UE until the current time instant t. This information about the channel is usually in the form of correlations between the past channel samples [BG06]. This correlation could be in the time domain, frequency domain and/or in the spatial domain. Thus, we can further classify the pilot-based channel prediction schemes by distinguishing them based on the type of correlation that is exploited to predict the channel.

Most SISO channel prediction schemes assume the channel to be a stochastic Wide Sense Stationary (WSS) process in time and just consider correlation in the time domain by using the temporal AutoCorrelation Function (ACF). In practical systems, the ACF might be time-varying, calling for adaptive filtering schemes to track the temporal evolution. Examples of such methods are Recursive Least Squares (RLS) [DH07, MVK95], Least Mean Squares (LMS) [WC97, KAK87] and the Kalman filter [KAK87].

This concept is typically extended to MIMO and even MIMO-Orthogonal Frequency Division Multiplexing (OFDM) channel prediction schemes by treating each antenna pair and also each sub-carrier as independent SISO channels and exploiting the temporal statistics for prediction [ZJCZ08, KPI08, MCCK07, Aro11]. Such a direct application of SISO schemes does not take advantage of the spatial correlations present in the MIMO channels, which when exploited correctly, provides potential gains as proposed in [SS06]. The authors used an AR model to predict a beamspace transformed CSI and then performed an inverse transformation on the predicted CSI which reduces the effective number of rays comprising the channel, leading to improved prediction performance.

Another approach is to perform a 2-step prediction as proposed by [WE06] for MIMO and extended by [LFYH14] to MIMO-OFDM, which utilizes Minimum Mean Square Error (MMSE) filtering, both in the time-domain and in the spatial domain, to exploit the spatio-temporal characteristics of the channel for a better prediction performance.

The type of fading experienced by the channel. In practical systems, many of the UEs are mobile and thus experience fast fading, which can be further categorized into flat or frequency-selective fading. Flat fading channels are rela-

tively easy to predict as the channel does not vary a lot across the sub-carriers and even the same AR model can be used to model the neighboring sub-carriers [DHHH00b, HKM10, HK07]. Things become more challenging when considering frequency-selective channels as the channel can vary significantly from one sub-carrier to the next and needs to be modeled accordingly. OFDM transmission plays a beneficial role in this respect as it divides the frequency-selective wideband channel into several flat narrowband channels which can be predicted with ease [Aro11].

The way in which CSI is assumed to be available at the receiver. The current and past knowledge about the CSI is utilized by prediction schemes, in the form of a model which captures the dynamics of the fading channel, to predict the future CSI. The mainly used ways to model the channel lead to two broad categories of prediction schemes: parametric-model based prediction and channel frequency response based prediction.

In *parametric-model* based schemes, the channel is modeled utilizing a set of parameters which characterize the channel response, typically like the number of paths, path gains and path delays [YLCC01]. Other parameters like the Angle of Arrival (AoA) and Angle of Departure (AoD), which can be expressed in terms of azimuth and/or elevation, can also be included to make the model more comprehensive. The current or future channel response can be reconstructed in the receiver by estimating or forecasting these parameters. Intuitively, parametric-model based schemes appear to be the ideal channel prediction approaches as they model the underlying parameters causing fading in the channel. If these parameters are known and vary at a slower rate than the actual channel itself, the channel can be reliably predicted into the future.

To estimate the parameters, the first thing to determine is the number of paths, which can be achieved through several available criteria for model order selection [Hay96]. Secondly, under the assumption of frequency domain reference signals, estimating the path delays is equivalent to the estimation of AoA using an antenna array [Swi98]. Hence, super-resolution spectral estimation methods, like Estimation of Signal Parameters via Rotational Invariance Techniques (ESPRIT) [RK89,HN98, HRD08] and MUltiple SIgnal Classifier (MUSIC) algorithm [Sch86], can perform this function. Lastly, typical linear estimators can be employed to obtain the path gains from the estimates of path delays. These methods have been applied in the prediction of SISO channels in [VTR00, HW98, DXL01, AJJF99]. The SISO parametric-model schemes have also been applied to predict MIMO channels in [VS07] and to MU-MIMO channel in [SLC08], although in these cases the spatial structure of the channel is not utilized. Advanced methods utilizing the structure of the MIMO channel have been proposed in [Che07] and [Ade15].

The parametric-model based prediction schemes are typically suited to sparse channels comprised of a few dominant paths as in the case of millimeter Wave (mmWave). In this case, the number of unknown parameters is less than the dimension of the channel correlation matrix which is practically only possible in rural scenarios. This is a serious drawback as most of the users are typically located in urban areas. In addition, averaging over a long symbol sequence is required to obtain the correlation matrix for the subspace-based parameter estimation schemes. This means that the parametric-model based schemes might not be suitable for bursty communication systems like LTE-A or 5G, especially for our target case of below 6 GHz. Hence, in this thesis we will not pursue parametric-model based channel prediction schemes.

In *channel frequency response* based schemes, the channel is observed through reference signals and the schemes operate directly on the channel frequency response. Two distinctions can be made based on the availability of statistical knowledge about the channel frequency response. When no statistical information is available, the channel frequency responses are treated as deterministic but unknown and can be estimated through the Least Squares (LS) or Maximum-Likelihood (ML) methods. The drawback is that the performance is typically below average because of the unavailability of channel knowledge. On the other hand, when the channel statistics are known, Bayesian estimation and prediction schemes can be used, which treat the channel frequency responses as random variables with known statistics. Such schemes offer significant performance gains by properly exploiting this statistical information. In this case, a linear MMSE predictor can be utilized to predict, for example, SISO channels. The prediction of MIMO channels requires a special design to capitalize on the spatial correlations.

Most of the common channel frequency response based schemes employ AutoRegressive (AR) modeling to describe the statistical state of the channel frequency response [SM97, KM81]. AR based schemes work under the assumption that a weighted linear combination of the past estimates contains enough information to provide a forecast of the channel frequency response in the future [DHHH00b]. The AR coefficients are typically computed to minimize the MSE, which requires the receiver to be cognizant of the AutoCorrelation Functions (ACFs) of the channel [DH07], which are usually unknown and need to be estimated from the noisy channel estimates. Several methods are available in the literature to estimate the ACFs, for example, via a solution of the Yule-Walker equations [BB05], the Burg method which typically provides the best performance [SM05], and the covariance and modified-covariance methods [SK03].

One of the major drawbacks of channel frequency response based, especially AR based, schemes is that they suffer from the problem of error propagation which

makes them unappealing for long-range channel prediction. The error propagation problem stems from the fact that the AR model of a discrete channel is typically formulated for a single-step prediction via linear extrapolation. When prediction is performed more than one step ahead, the extrapolated channel samples at the previous time instants are used, compounding the error terms.

Another shortcoming of these schemes is their high sensitivity to noise which even makes short term prediction rather challenging [FC02]. The noise sensitivity is typically minimized by oversampling the channel, basically boosting the Signal-to-Noise Ratio (SNR). However, we have devised a scheme which utilizes subspace processing to exploit the structure of the MIMO channel to suppress the noise prior to prediction, as discussed in Section 4.5.

As one of the benchmarks to test the performance of the algorithms developed in this thesis, we will mostly use a channel frequency response based scheme. This scheme forms a state-space model of the channel knowledge through AR modeling, which is then utilized by a Kalman filter to perform channel prediction, as proposed by [Aro11]. This scheme has been essentially designed for Multiple-Input and Single-Output (MISO)-OFDM systems and has been later extended further to MIMO-OFDM and CoMP by [AS14].

4.3 Kalman Filter-based Prediction

Channel prediction is a complex process, especially for MIMO and massive MIMO systems, as it needs to incorporate the maximum amount of information possible for each of the antenna links between the transmitter and the receiver, and the frequency sub-carriers, in the case of OFDM transmission. In addition, various UEs experience different channel conditions based on their position and speed in the environment. Hence, we need an algorithm which can take all the information available about the channel ranging from fading statistics to pilot arrangements in the time-frequency grid and provide a suitable filter to predict the MIMO channels. All this can be achieved dead on by the Kalman filter, which is known to be the best linear estimator given the input characteristics. It is also capable of providing these estimates during the *transient* phase when these characteristics change [K$^+$60]. Moreover, it provides a significant complexity reduction as compared to a Wiener filter. Another benefit of using a Kalman filter for prediction is that it *provides a measure of uncertainty* about the prediction estimate, which is data dependent and directly corresponds to system performance.

State Space Model for a Single Sub-carrier

To construct the state space model for the Kalman filter, we follow [Aro11] and begin by modeling a single channel coefficient as an AR process. This channel coefficient can either be a tap in an impulse response $h(\tau, t)$ or the complex-valued channel $H(f_k, t)$ for a single pilot sub-carrier f_k, but we will only consider the later case. The local scattering environment and the velocity of the UE, determine the behavior of each coefficient, which is commonly oscillatory and justifies the use of AR models to capture the temporal variation. The channel coefficient $h_{t_i} = H(f_k, t_i)$ can be modeled as

$$h_{t_i} + a_1 h_{t_{i-1}} + \ldots + a_q h_{t_{i-q}} = u_{t_i}, \tag{4.1}$$

where q is the model order, $\{a_i\}_{i=1}^{q}$ are the AR parameters and u_{t_i} is the process noise that excites the process. When no *prior information* is available about the channel, the AR parameters can be assigned based on the approximate Doppler statistics of the environment in which the UE is residing. For example, under the assumption that the moving UE is surrounded by equally distant and evenly distributed scatterers and there is no Line-Of-Sight (LOS) between the base station and the UE, the Jakes' model can be used [JC94]. The Jakes' Doppler spectrum is defined as,

$$\Phi_H(f_D) = \begin{cases} \frac{P}{\pi f_{D,\max}} \frac{1}{\sqrt{1-(f_D/f_{D,\max})^2}} & ; \quad |f_D| \leq f_{D,\max} \\ 0 & ; \quad |f_D| > f_{D,\max}, \end{cases} \tag{4.2}$$

where P is the average received power with respect to an isotropic antenna. Jakes model leads to the following autocorrelation function,

$$\varphi_H(\Delta t) = P \cdot J_0(2\pi f_{D,\max} \Delta t_i) \tag{4.3}$$

where J_0 is the zero-order Bessel function of the first kind. An even more simplified way could be to construct AR models with flat Doppler spectra based on just the knowledge of the maximum Doppler frequency $f_{D,\max}$.

We will consider that *prior information* is available, in the form of a noise-free training segment of length N_{train}. In this case, the AR parameters can be estimated based on blocks of the available measurements through the Yule-Walker method, also known as the ACF method, or covariance and modified-covariance methods. In the analysis conducted by [Aro11] on measured channels, the Yule-Walker method outperformed the other methods in terms of prediction performance and on the agreement between theoretical and experimental values. Hence, we will use this as the method of choice for AR parameter estimation.

If we consider that the set of all the past measured channel coefficients $h_{t_i} = \{h_{t_{i-1}}, \ldots, h_{t_0}\}$ is available over a time segment $N_{\text{train}} = 100$ and a model order $q = 4$ of the system, we can define the following matrices,

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ h_{t_0} & 0 & 0 & 0 \\ h_{t_1} & h_{t_0} & 0 & 0 \\ h_{t_2} & h_{t_1} & h_{t_0} & 0 \\ h_{t_3} & h_{t_2} & h_{t_1} & h_{t_0} \\ \vdots & \vdots & \vdots & \vdots \\ h_{t_{98}} & h_{t_{97}} & h_{t_{96}} & h_{t_{95}} \\ h_{t_{99}} & h_{t_{98}} & h_{t_{97}} & h_{t_{96}} \\ 0 & h_{t_{99}} & h_{t_{98}} & h_{t_{97}} \\ 0 & 0 & h_{t_{99}} & h_{t_{98}} \\ 0 & 0 & 0 & h_{t_{99}} \end{pmatrix}, \quad \mathbf{b} = - \begin{pmatrix} h_{t_0} \\ h_{t_1} \\ h_{t_2} \\ h_{t_3} \\ h_{t_4} \\ \vdots \\ h_{t_{99}} \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{4.4}$$

where $\mathbf{A} \in \mathbb{C}^{(N_{\text{train}}+q)\times q}$ and $\mathbf{b} \in \mathbb{C}^{(N_{\text{train}}+q)\times 1}$. Under the additional assumption that these channel measurements are noise-free, the vector of AR parameters can be defined as,

$$\mathbf{a} = [a_1, \ldots, a_q]^{\mathrm{T}}. \tag{4.5}$$

Hence, by utilizing the matrices defined in (4.4), we can set up an over-determined system of equations as,

$$\mathbf{A}\mathbf{a} = \mathbf{b}, \tag{4.6}$$

A point estimate $\hat{\mathbf{a}}$ is then found by solving for $\hat{\mathbf{a}}$ in

$$(\mathbf{A}^{\mathrm{H}}\mathbf{A})\hat{\mathbf{a}} = \mathbf{A}^{\mathrm{H}}\mathbf{b} \tag{4.7}$$

$$\hat{\mathbf{a}} = (\mathbf{A}^{\mathrm{H}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{H}}\mathbf{b} \tag{4.8}$$

It is called the autocorrelation method because $\mathbf{A}^{\mathrm{H}}\mathbf{A}$ is an estimate of the autocorrelation matrix of the channel. The Toeplitz structure of $\mathbf{A}^{\mathrm{H}}\mathbf{A}$ allows us to solve the system (4.7) with the Levinson recursions.

In the case of noisy measurements, we can form a matrix $\tilde{\mathbf{A}}$ similarly as $\mathbf{A}$ in (4.4) but with noisy channel measurements instead of noise-free channel measurements. Let us assume that the complex-valued white noise present in the measured channel coefficients is of unit variance. Then we may calculate,

$$\mathbf{A}^{\mathrm{H}}\mathbf{A} \approx \tilde{\mathbf{A}}^{\mathrm{H}}\tilde{\mathbf{A}} - \eta\mathbf{I}_q \tag{4.9}$$

where η is the number of non-zero elements in each column of $\mathbf{A}$ and is equal to

N_{train} in this case, and $\mathbf{I}_q$ is an identity matrix of size $q \times q$. Since it holds that $\mathbf{A}^{\text{H}}\mathbf{b} \approx \tilde{\mathbf{A}}^{\text{H}}\mathbf{b}$ because of the whiteness of the noise, Equation (4.7) can now be approximated as,

$$(\tilde{\mathbf{A}}^{\text{H}}\tilde{\mathbf{A}} - \eta\mathbf{I}_q)\hat{\mathbf{a}} = \tilde{\mathbf{A}}^{\text{H}}\mathbf{b} \tag{4.10}$$

We can write the Equation (4.1) on the operator form as,

$$h_{t_i} = \frac{1}{1 + a_1 c^{-1} + \ldots + a_q c^{-q}} u_{t_i}, \tag{4.11}$$

or by means of the z-transform as,

$$h(z) = \frac{z^q}{z^q + a_1 z^{q-1} + \ldots + a_q} u(z) = \frac{z^q}{(z-\rho_1)(z-\rho_2)\ldots(z-\rho_q)} u(z) \tag{4.12}$$

where c^{-1} is the backward unit delay operator. After estimating the AR parameters $\hat{\mathbf{a}}$, we can calculate the corresponding poles $\{\rho_i\}_{i=1}^{q}$ of the AR process from (4.12). We use these poles to construct a state space model for the fading channel coefficient h_{t_i},

$$\mathbf{x}_{t_{i+1}}^{(1)} = \mathbf{Q}\mathbf{x}_{t_i}^{(1)} + \mathbf{m}u_{t_i}, \tag{4.13}$$

$$h_{t_i}^{(1)} = \mathbf{p}^{\text{H}}\mathbf{x}_{t_i}^{(1)}, \tag{4.14}$$

where $\mathbf{x}_{t_i}^{(1)}$ is the state vector of length q. The state space is chosen on the diagonal canonical form, so that $\mathbf{Q}$ is diagonal. The elements of $\mathbf{Q} \in \mathbb{C}^{q\times q}$, $\mathbf{m} \in \mathbb{C}^{q\times 1}$ and $\mathbf{p} \in \mathbb{C}^{q\times 1}$ are set as follows,

$$\mathbf{Q}[i,i] = \rho_i, \tag{4.15}$$

$$\mathbf{m}[i] = \prod_{\substack{j=1,\ldots,q \\ j\neq i}} (\rho_i - \rho_j)^{-1}, \tag{4.16}$$

$$\mathbf{p}[i] = \rho_i^{q-1}, \qquad i = 1, \ldots, q \tag{4.17}$$

where $\mathbf{Q}[i,i]$ represents the i-th diagonal element of the matrix $\mathbf{Q}$, $\mathbf{m}[i]$ and $\mathbf{p}[i]$ represent the element at index i of the vectors $\mathbf{m}$ and $\mathbf{p}$ respectively.

State Space Model for Multiple Sub-carriers

By following the steps in the previous section, we have formed the state space model for one particular sub-carrier. To model the complete SISO channel, we need to model each sub-carrier in the channel. We can track w parallel pilot sub-carriers if they are correlated. This correlation depends on the spacing between them and the coherence bandwidth of the channel. To track w sub-carriers, we set up a block

diagonal state space,

$$\begin{aligned}
\mathbf{x}_{t_{i+1}}^{(2)} &= \{\mathrm{diag}(\mathbf{Q},\ldots,\mathbf{Q})\}\mathbf{x}_{t_i}^{(2)} + \{\mathrm{diag}(\mathbf{m},\ldots,\mathbf{m})\}\mathbf{u}_{t_i}^{(2)}, && (4.18)\\
&= \{\underbrace{\mathbf{I}_w \otimes \mathbf{Q}}_{\mathbf{A}}\}\mathbf{x}_{t_i}^{(2)} + \{\underbrace{\mathbf{I}_w \otimes \mathbf{m}}_{\mathbf{B}}\}\mathbf{u}_{t_i}^{(2)}, && (4.19)\\
&= \mathbf{A}\mathbf{x}_{t_i}^{(2)} + \mathbf{B}\mathbf{u}_{t_i}^{(2)}, && (4.20)\\
\mathbf{h}_{t_i}^{(2)} &= \{\mathrm{diag}(\mathbf{p},\ldots,\mathbf{p})\}\mathbf{x}_{t_i}^{(2)} && (4.21)\\
&= \{\underbrace{\mathbf{I}_w \otimes \mathbf{p}}_{\mathbf{C}}\}\mathbf{x}_{t_i}^{(2)} = \mathbf{C}\mathbf{x}_{t_i}^{(2)}, && (4.22)
\end{aligned}$$

where $\otimes$ represents the Kronecker product and $\mathbf{x}_{t_i}^{(2)}$ is the state vector which now has a length qw. $\mathbf{A} \in \mathbb{C}^{qw\times qw}$ is diagonal and $\mathbf{B} \in \mathbb{C}^{qw\times w}$ and $\mathbf{C} \in \mathbb{C}^{w\times qw}$ are block diagonal matrices with w blocks each. Ideally, we should employ different state space models for the individual sub-carriers, but as long as the narrowband assumption holds, which is practically the case for OFDM systems, the same model can be used for all w sub-carriers which means that the same AR parameters and thus the same poles. The covariance matrix of the sub-carriers, $\mathbf{R}_{\mathbf{h}^{(2)}} = \mathrm{E}\{\mathbf{x}_{t_i}^{(2)}\mathbf{x}_{t_i}^{(2)^{\mathrm{H}}}\}$, can be directly estimated from the noisy measurements of the time-frequency channel. On the other hand, the calculation of the process noise covariance matrix, $\mathbf{L} = \mathrm{E}\{\mathbf{u}_{t_i}^{(2)}\mathbf{u}_{t_i}^{(2)^{\mathrm{H}}}\}$ and the prior $\mathbf{\Gamma}_0 = \mathrm{E}\{\mathbf{x}_{t_0}^{(2)}\mathbf{x}_{t_0}^{(2)^{\mathrm{H}}}\}$, from the covariance matrix of the sub-carriers $\mathbf{R}_{\mathbf{h}^{(2)}}$ is not trivial but can be calculated explicitly by using Theorem 4.2.1 in [Aro11].

State Space Model for Multiple Transmit Antennas

We can continue this modeling strategy to model a MISO system by augmenting a third hierarchical level to the state space model. According to the authors of [Aro11], the model of a MISO system corresponds to a MIMO system, where each receive antenna is modeled separately, which does not result in a loss of generality or achievable performance in the context of channel prediction. The state space models can be set up simultaneously for each of the receiver antennas and the overall complexity of the receiver scales linearly with the number of antennas. The state space model for one receive antenna is thus given as,

$$\begin{aligned}
\mathbf{x}_{t_{i+1}}^{(3)} &= \mathrm{diag}(\mathbf{A}_1,\ldots,\mathbf{A}_{M_T})\mathbf{x}_{t_i}^{(3)} + \mathrm{diag}(\mathbf{B}_1,\ldots,\mathbf{B}_{M_T})\mathbf{u}_{t_i}^{(3)}, && (4.23)\\
&= \mathbf{F}\mathbf{x}_{t_i}^{(3)} + \mathbf{G}\mathbf{u}_{t_i}^{(3)}, && (4.24)\\
\mathbf{h}_{t_i}^{(3)} &= \mathrm{diag}(\mathbf{C}_1,\ldots,\mathbf{C}_{M_T})\mathbf{x}_{t_i}^{(3)} = \mathbf{H}\mathbf{x}_{t_i}^{(3)}, && (4.25)
\end{aligned}$$

where M_T is the number of transmit antennas. The matrices corresponding to the different transmit antenna elements $\mathbf{A}_1, \ldots, \mathbf{A}_{M_T}$, $\mathbf{B}_1, \ldots, \mathbf{B}_{M_T}$ and $\mathbf{C}_1, \ldots, \mathbf{C}_{M_T}$ are individually calculated as,

$$\begin{aligned} \mathbf{A}_i &= \text{diag}(\mathbf{Q}_i, \ldots, \mathbf{Q}_i) \\ &= \mathbf{I}_w \otimes \mathbf{Q}_i \qquad (4.26) \\ \mathbf{B}_i &= \text{diag}(\mathbf{m}_i, \ldots, \mathbf{m}_i) \\ &= \mathbf{I}_w \otimes \mathbf{m}_i \qquad (4.27) \\ \mathbf{C}_i &= \text{diag}(\mathbf{p}_i, \ldots, \mathbf{p}_i) \\ &= \mathbf{I}_w \otimes \mathbf{p}_i \qquad (4.28) \end{aligned}$$

utilizing corresponding matrix $\mathbf{Q}_i$, and vectors $\mathbf{m}_i$ and $\mathbf{p}_i$, where $i = 1, \ldots, M_T$. The model matrices are now of size $\mathbf{F} \in \mathbb{C}^{qwM_T \times qwM_T}$, $\mathbf{G} \in \mathbb{C}^{qwM_T \times wM_T}$ and $\mathbf{H} \in \mathbb{C}^{wM_T \times qwM_T}$, which is rather large, in general, but the matrices are sparse and thus require little memory for storage. Under the assumption of independence between the M_T transmitting antennas, the process noise covariance matrix $\mathbf{\Xi} = \text{E}\{\mathbf{u}_{t_i}^{(3)} \mathbf{u}_{t_i}^{(3)^{\text{H}}}\}$ and the prior $\mathbf{\Pi}_0 = \text{E}\{\mathbf{x}_{t_0}^{(3)} \mathbf{x}_{t_0}^{(3)^{\text{H}}}\}$ can be accordingly constructed as,

$$\mathbf{\Xi} = \text{diag}(\mathbf{L}_1, \ldots, \mathbf{L}_{M_T}) \qquad (4.29)$$

$$\mathbf{\Pi} = \text{diag}(\mathbf{\Gamma}_1, \ldots, \mathbf{\Gamma}_{M_T}) \qquad (4.30)$$

The final step is to model the measurements, in which the fading channel is observed in Additive White Gaussian Noise (AWGN) through the reference signals,

$$\begin{aligned} \mathbf{y}_{t_i} &= \mathbf{X}_{t_i} \mathbf{h}_{t_i}^{(3)} + \mathbf{z}_{t_i} \qquad \text{(inserting value of } \mathbf{h}_{t_i}^{(3)} \text{ from Equation 4.25)} \quad (4.31) \\ &= \underbrace{\mathbf{X}_{t_i} \mathbf{H}}_{\mathbf{J}_{t_i}} \mathbf{x}_{t_i}^{(3)} + \mathbf{z}_{t_i} = \mathbf{J}_{t_i} \mathbf{x}_{t_i}^{(3)} + \mathbf{z}_{t_i} \qquad (4.32) \end{aligned}$$

where $\mathbf{X}_t \in \mathbb{C}^{w \times wM_T}$ is the matrix that holds the time-varying reference symbols as introduced in equation (3.26) and $\mathbf{z}_{t_i}$ is AWGN with the covariance matrix $\mathbf{R} = \text{E}\{\mathbf{z}_{t_i} \mathbf{z}_{t_i}^{\text{H}}\}$. The overall model that can accurately model the qwM_T channel coefficients in the system is now given as,

$$\begin{aligned} \mathbf{x}_{t_{i+1}}^{(3)} &= \mathbf{F} \mathbf{x}_{t_i}^{(3)} + \mathbf{G} \mathbf{u}_{t_i}^{(3)}, \qquad (4.33) \\ \mathbf{y}_{t_i} &= \mathbf{J}_{t_i} \mathbf{x}_{t_i}^{(3)} + \mathbf{z}_{t_i} \qquad (4.34) \end{aligned}$$

where the vectors $\mathbf{x}_0^{(3)}$, $\mathbf{u}_{t_i}^{(3)}$ and $\mathbf{z}_{t_i}$ are white, zero-mean, and Gaussian with covariance matrices $\mathbf{\Pi}_0 > 0, \mathbf{\Xi} > 0$ and $\mathbf{R} > 0$ (meaning that the matrices are positive definite).

Once we have modeled the state space as shown above, we apply the Kalman filter algorithm presented in Table 4.1 of [Aro11] to estimate or predict the channel. From the state estimate vector we can compute the channel estimate for one receive antenna by,

$$\hat{\mathbf{h}}_{t_i} = \mathbf{H}\mathbf{x}_{t_i}^{(3)}, \tag{4.35}$$

and the P-step prediction estimate by,

$$\hat{\mathbf{h}}_{t_{i+P}} = \mathbf{H}\mathbf{F}^P\mathbf{x}_{t_i}^{(3)}. \tag{4.36}$$

4.3.1 Performance Assessment of Kalman Prediction on Measurements

We have applied this predictor in our publications [DAZ+14], [AZH15b] and [AZH15a] and have shown that when the channel has a reasonable number of multipath components, this predictor performs quite well.

Here, as an example consider the local-area scenario in [DAZ+14]. Specific local-area measurements have been performed by Umer Zeeshan in a cooperation project between NSN and TU Dresden (TUD) under the supervision of Michael Grieger. The measurements were carried out with the TUD LTE testbed within a single room of a typical office building. The base station and the UE are both static and placed at a height of 1.62 m and at a distance of 5.46 m apart from each other. The total system bandwidth is 20 MHz and the total number of pilots N_{RS} is 200. The central RF-frequency f_c is 2.68 GHz and the measurement is carried out for a duration of 2.1 seconds with a sampling time of 1 ms. The measurements were made for different scenarios, like the basic static case, and for cases where different numbers of people are moving between the base station and the UE. A basic depiction of the measurement setup and the scenario where a person is moving around while waving a big board in his hand is illustrated in Figure 4.1

The moving people act as scatters and affect the fading behavior of the channel as can be observed in Figure 4.2. The channel is almost static when there is no movement between the base station and the UE (case (a)). However, when a person is moving around (case (b)) or additionally waving a big board (case (c): move-board scenario), we see that it induces some severe fading at select pilot sub-carriers and time instances.

The time variance of the radio channel is one issue, but it is more interesting to see how well state-of-the-art prediction, like the Kalman filter scheme being discussed in Section 4.3, can exploit the relatively slow channel fluctuations. For evaluation, we track $w = 8$ parallel pilot sub-channels at a time with a spacing of

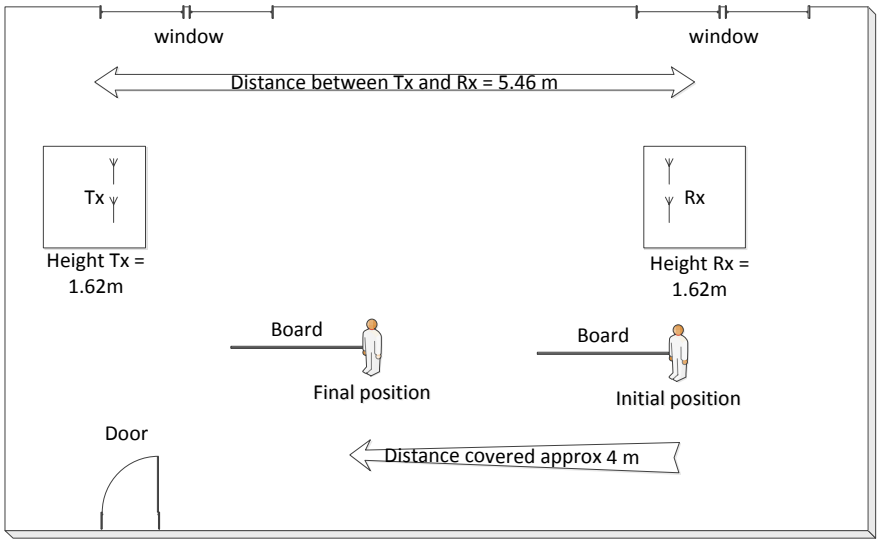


Figure 4.1: A general illustration of the move-board measurement scenario

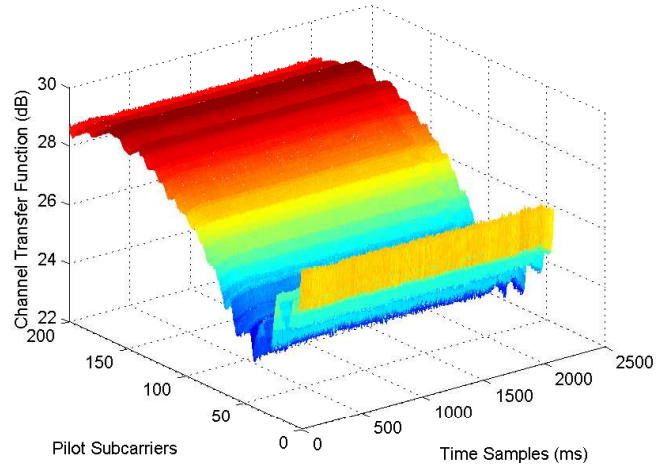
90 kHz between them which is less than the coherence bandwidth. As there are 200 pilot sub-carriers in total, we need $200/8 = 25$ parallel Kalman filters to track the complete channel. Similarly, to ensure that the channel remains stationary during the prediction and hence the state space model remains valid, we train over a 100 ms time segment and then estimate/predict the channel for up to 100 ms. We model our channel coefficients on an AR process of order 1, as we found out that such a model represents these indoor channels appropriately.

We use the normalized mean square error (NMSE) to characterize the performance of the predictor as it provides a direct relative measure for the degradation due to channel prediction. The NMSE is computed as an expectation over all time slots for every pilot sub-channel,

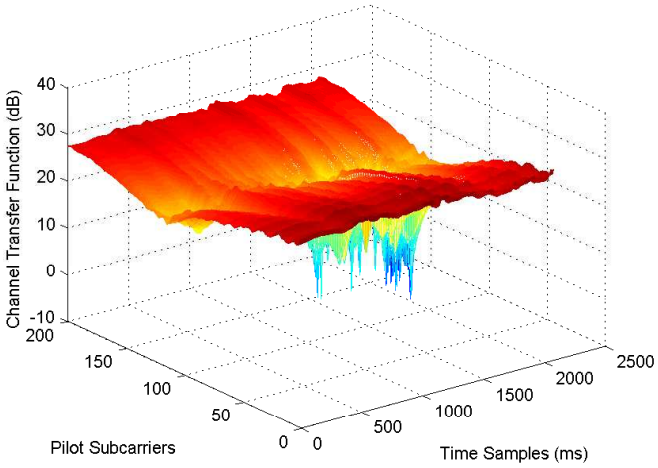
$$\text{NMSE} = \sigma_E^2 = \frac{\mathrm{E}\{\|\mathbf{e}\|^2\}}{\mathrm{E}\{\|\mathbf{h}\|^2\}} = \frac{\mathrm{E}\{\|\mathbf{h} - \hat{\mathbf{h}}\|^2\}}{\mathrm{E}\{\|\mathbf{h}\|^2\}} \tag{4.37}$$

where $\mathbf{h} \in \mathbb{C}^{100\times 1}$ is the true channel and $\hat{\mathbf{h}} \in \mathbb{C}^{100\times 1}$ is the estimated channel, for a particular sub-carrier under consideration.

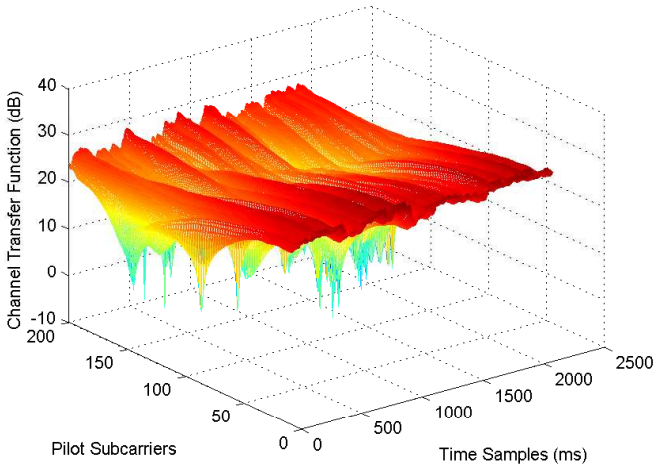
Figure 4.3 shows the NMSE performance versus the prediction time for the different investigated scenarios. It was shown in [GS12] that an NMSE below -10 dB is required to have successful link adaptation and/or scheduling. We can see that the NMSE is below -10 dB until a prediction horizon of 18 ms for the static case (case (a)) and 8 ms with one person moving (case (b)). This means that in a local-area, a relatively large prediction horizon with a correspondingly low feedback overhead seems to be possible.



(a) Case (a): Static Scenario



(b) Case (b): One person is moving



(c) Case (c): A person is moving with a board

Figure 4.2: The channel transfer functions (CTFs) $H(f,t)$ for different indoor measurement scenarios. The pilot spacing is 15 kHz.

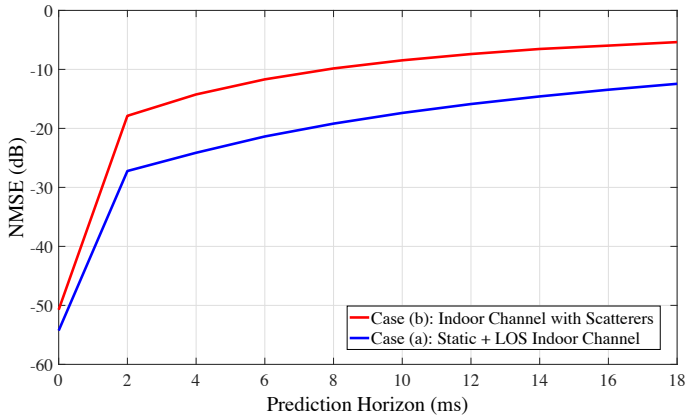


Figure 4.3: NMSE performance of various scenarios at different prediction horizons

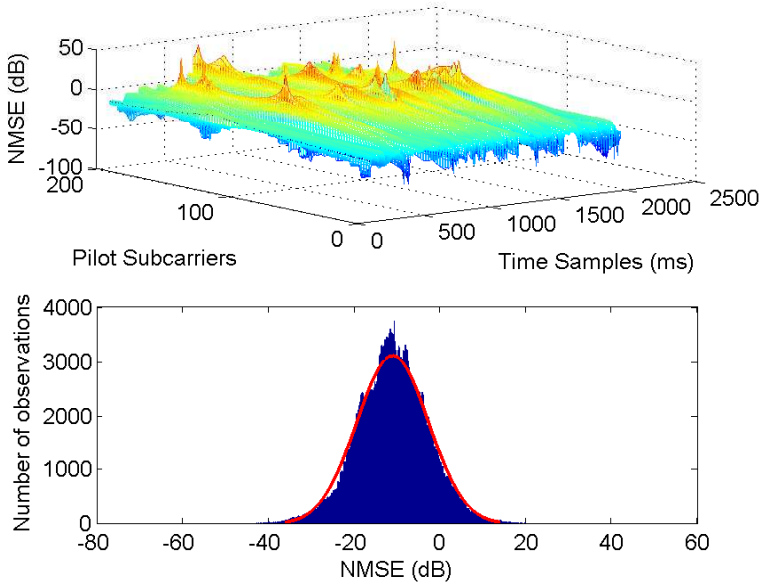


Figure 4.4: NMSE at a prediction horizon of 8 ms for the move-board scenario (case (c)).

Figure 4.4 shows the NMSE for various time samples and pilot sub-channels at a prediction horizon of 8 ms for the move-board scenario (case (c)). We can observe that at most of the sub-carriers, the NMSE is quite similar with a mean value of -11 dB. This scenario has the worst performance out of all the three scenarios, but a much better prediction performance seems to be possible by proper scheduling. We can avoid to schedule users in sub-channels containing frequency notches with corresponding fast channel fluctuations, hence a higher NMSE. We see from the histogram that the NMSE in the logarithm domain is close to Gaussian distributed.

4.4 HOSVD Based Channel Prediction

We will now discuss a subspace based channel prediction scheme, proposed by [MDH08], which takes the channel frequency response in tensor form and operates on it in a higher order subspace to transform the tensor elements. The prediction is then performed in this transformed domain providing substantial performance gains as compared to prediction conducted directly on the channel coefficients.

Let us represent the sampled channel frequency response of a time-variant frequency-selective massive MIMO channel as a four-dimensional tensor $\boldsymbol{\mathcal{H}} \in \mathbb{C}^{M_T \times M_R \times N_{\mathrm{RS}} \times N_t}$, where M_T and M_R are the number of antenna elements at the base station and at the UE, while N_{RS} and N_t are the number of snapshots, sampled in the frequency and the time domain corresponding to the reference (pilot) signals, respectively. Interpolation and extrapolation can be used to get the complete channel for all time samples and all sub-carriers N_{SC}. In practical systems, different polarizations, and/or elevation angles are also present, which can add other dimension(s) resulting in a five- or six-dimensional channel tensor. However, in this chapter, we only consider one polarization and no angular variation in the elevation domain, and assume that the past noisy channel estimates are available. Thus, grouping together the last N_t estimates for all antennas and frequencies gives us the estimated channel tensor $\tilde{\boldsymbol{\mathcal{H}}} \in \mathbb{C}^{M_T \times M_R \times N_{\mathrm{RS}} \times N_t}$. The noisy channel estimates $\tilde{\boldsymbol{\mathcal{H}}}$ can be modeled as,

$$\tilde{\boldsymbol{\mathcal{H}}} = \boldsymbol{\mathcal{H}} + \boldsymbol{\mathcal{E}} \tag{4.38}$$

where $\boldsymbol{\mathcal{E}}$ is the noise term of appropriate size. It is introduced by the channel estimator, and contains independent and identically distributed complex Gaussian entries with average power σ_e^2. Let $\tilde{h}_{m_T, m_R, f_j, t_k}$ denote the channel tensor element corresponding to the channel between the m_T-th transmit antenna, and the m_R-th receive antenna at baseband frequency $f_j = j\Delta f$ and time $t_k = k\Delta t$, where Δf is the sub-carrier spacing and Δt is the time sampling interval of the system under consideration.

The Higher Order Singular Value Decomposition (HOSVD) of such a four-dimensional channel tensor is given as [LMV00, Tuc66],

$$\tilde{\boldsymbol{\mathcal{H}}} = \boldsymbol{\mathcal{S}} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \times_3 \mathbf{U}_3 \times_4 \mathbf{U}_4 \tag{4.39}$$

where every $\mathbf{U}_n$; $n = 1, \ldots, 4$ is a unitary matrix of size $I_n \times I_n$ (where $I_n = M_T, M_R, N_{\mathrm{RS}}, N_t$ for $n = 1, 2, 3, 4$ respectively) consisting of the n-mode singular vectors, $\boldsymbol{\mathcal{S}} \in \mathbb{C}^{I_1 \times I_2 \times \ldots \times I_4}$ is the core tensor, and $\times_n$ denotes the n-mode product of a tensor with a matrix (please refer to Appendix A for further details). Note that the HOSVD decomposition of a two-dimensional tensor, which is a matrix, reduces

to the Singular Value Decomposition (SVD) decomposition.

In order to perform tensor-based prediction, we sub-divide the channel tensor $\tilde{\boldsymbol{\mathcal{H}}}$ into several sub-tensors. Let $\tilde{\boldsymbol{\mathcal{H}}}(t_i)$ denote such a sub-tensor containing the elements $\tilde{h}_{m_T,m_R,f_j,t_k}$ where $m_T \in \{1,2,...,M_T\}$, $m_R \in \{1,2,...,M_R\}$, $f_j \in \{-\frac{N_{\mathrm{RS}}}{2}+1, -\frac{N_{\mathrm{RS}}}{2}+2, \ldots, -\frac{N_{\mathrm{RS}}}{2}\}$ and $t_k \in \{i-N_{t_0}+1, i-N_{t_0}+2, \ldots, i\}$. Here, t_i denotes the current time and N_{t_0} the number of sub-tensor time snapshots. The values N_{t_0} and N_{RS} are chosen in such a way that the terms $N_{t_0}\Delta t$ and $N_{\mathrm{RS}}\Delta f$ stay within the coherence time and coherence bandwidth of the channel respectively, while both the antenna dimensions are kept fixed. Hence, we can consider that the bases $\mathbf{U}_n, n \neq c$ where c is the prediction dimension (in this case time snapshots), to remain constant in the vicinity of the observed and current channel coefficients. The n-th basis matrix can be calculated as the matrix of the left singular vectors of the n-mode unfolding. From these bases, we can generally compute a transformed tensor $\boldsymbol{\mathcal{A}}(t_i)$ estimated from the known channel sub-tensor $\tilde{\boldsymbol{\mathcal{H}}}(t_i)$, being considered here as [MDH08],

$$\begin{aligned} \boldsymbol{\mathcal{A}}(t_i) &= \tilde{\boldsymbol{\mathcal{H}}}(t_i) \times_1 \mathbf{U}_1^{\mathrm{H}} \times_2 \mathbf{U}_2^{\mathrm{H}} \times_3 \mathbf{U}_3^{\mathrm{H}} \\ &= \boldsymbol{\mathcal{S}} \times_4 \mathbf{U}_4 \quad \in \mathbb{C}^{M_T \times M_R \times N_{\mathrm{RS}} \times N_{t_0}} \end{aligned} \tag{4.40}$$

We can track the evolution of the channel in the time domain through the tensors $\boldsymbol{\mathcal{A}}(t_{i-N_0+1}), \boldsymbol{\mathcal{A}}(t_{i-N_0+2}), \ldots, \boldsymbol{\mathcal{A}}(t_i)$, where N_0 is the number of sub-tensors. In addition, employing a linear prediction filter such as a Wiener filter, denoted as $F(\cdot)$, these sub-tensors can be used to predict the channel at a future time instant $t_p = t_{i+P}$ where P is the prediction horizon, as [MDH08],

$$\hat{a}_{m_T,m_R,f_j,t_p} = F(a_{m_T,m_R,f_j,t_k}(t_{i_0})) \tag{4.41}$$

for each value of m_T, m_R, f_j and t_k, where $t_{i_0} \in \{i-N_0+1, i-N_0+2, \ldots, i\}$, and $\hat{(\cdot)}$ denotes predicted values. Various filtering methods $F(\cdot)$ can be used to predict the transformed tensor coefficients in the future ranging from linear filters, non-linear filters to even adaptive filters [MDH08].

As the time dimension of each sub-tensor corresponds to the coherence time of the channel and it is usually composed of multipath components whose directions of arrival and departure change slowly in time, the channel will span a common signal subspace so that a reduced number r_n, where $n = 1, 2, 3$, of basis vectors per tensor dimension n is sufficient to represent it. Here, $0 < r_n \leq I_n$, are the ranks of the different matrix unfoldings, called n-ranks. Therefore, we can reliably predict $\boldsymbol{\mathcal{H}}$ utilizing a low rank approximation of $\hat{\boldsymbol{\mathcal{A}}}$, computed generally as,

$$\hat{\boldsymbol{\mathcal{H}}}(t_p) = \boldsymbol{\mathcal{A}}^{[s]}(t_p) \times_1 \mathbf{U}_1^{[r_1]} \ldots \times_2 \mathbf{U}_2^{[r_2]} \times_3 \mathbf{U}_3^{[r_3]} \tag{4.42}$$

where $\mathbf{U}_n^{[r_n]} \in \mathbb{C}^{I_n \times r_n}$, $n = 1, 2, 3$, is the matrix comprised of r_n n-mode singular vectors, which define the signal subspace corresponding to the r_n strongest n-mode singular values, and $\hat{\boldsymbol{\mathcal{A}}}^{[s]}(t_p) \in \mathbb{C}^{r_1 \times r_2 \times r_3 \times N_{t_0}}$ is the predicted transformed tensor, truncated to correspond to the signal subspace singular vectors.

Overall, the prediction of the time-variant frequency-selective 4D channel tensor, is partitioned into $M_T \cdot M_R \cdot N_{\mathrm{RS}}$ SISO channel predictions.

4.4.1 Performance Analysis of Tensor-based Prediction

In order to assess the performance benefit, we use synthetic channel frequency responses generated with the IlmProp [Del07, DHS03], a geometry based channel modeling tool for wireless communications discussed in Section 3.4.1.1. We use the same simple urban macro scenario depicted in Figure 3.5, as the propagation environment for the discussion of the simulation results. In this scenario, the 4-element Uniform Linear Array (ULA) mounted on the UE is moving away from the base station with a speed of 10 km/h , while a second 16-ULA is acting as the base station. There are four buildings in the transmission path with a total of 280 scatters. The UE also sees a direct path from the transmitter. The parameters (delay spread, Doppler spread, etc.) of the channels created with the IlmProp are fairly close to real channel parameters in the corresponding scenario. The OFDM simulation parameters consist of $N_{\mathrm{RS}} = 64$ reference signals over a 20 MHz bandwidth and a time sampling frequency of 0.5 ms. The center frequency of the system f_c is 5 GHz.

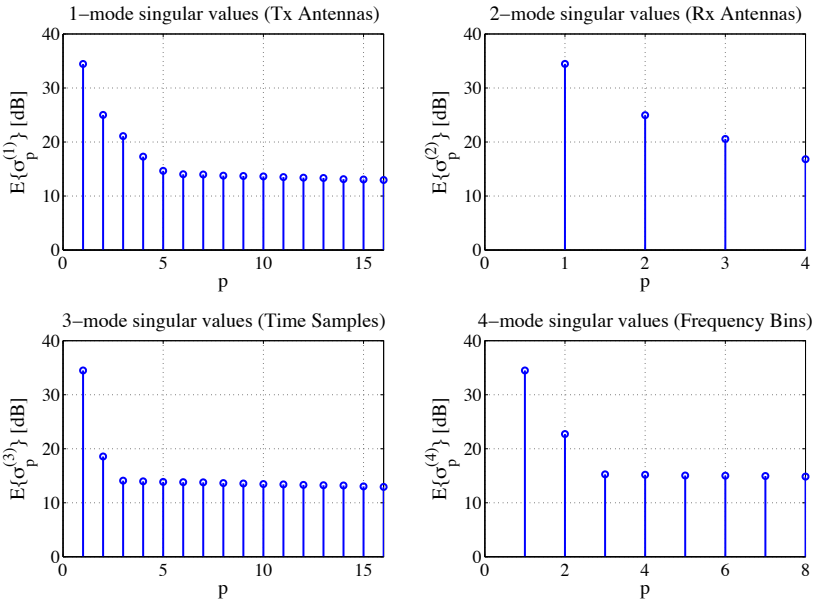


Figure 4.5: Higher-order singular values of the tensor corresponding to the n-th dimension of the channel sub-tensor.

The values of the massive MIMO channel tensor n-mode singular values in all dimensions for the IlmProp macro scenario are displayed in Figure 4.5. The first n-mode singular value corresponds to the strongest singular vector in each dimension and is significantly higher than the other singular values. Singular vectors corresponding to weak singular values can be neglected, for example, in the transmit antenna dimension (1-mode) we can neglect the singular vectors starting from the 5-th. The reason for this is that the first r_n singular vectors corresponding to the r_n largest singular values lie in the signal subspace while the remaining $(I_n - r_n)$ can be expected to lie in the noise subspace. Since the noise is white Gaussian, it is expected that the noise contributes the same value to all singular values defining the cut-off threshold (as can be observed in Figure 4.5 that all weak singular values have the same power). Since there are I_n n-mode singular values in the n-th tensor dimension, the cut-off threshold $\sigma_{\text{th}}^{(n)}$ for that dimension can be estimated as [MDH08],

$$\sigma_{\text{th}}^{(n)} = \sqrt{\frac{\sigma_e^2}{I_n}} \tag{4.43}$$

where σ_e^2 is the average power of the noise term in equation (4.38). Neglecting the singular values corresponding to the noise subspace and using, for this particular example, $r_1 = 4, r_2 = 3, r_3 = 2$ and $r_4 = 2$ singular vectors in each dimension corresponding to the signal subspace in equation (4.42), enables us to predict the channel utilizing its low-rank approximation.

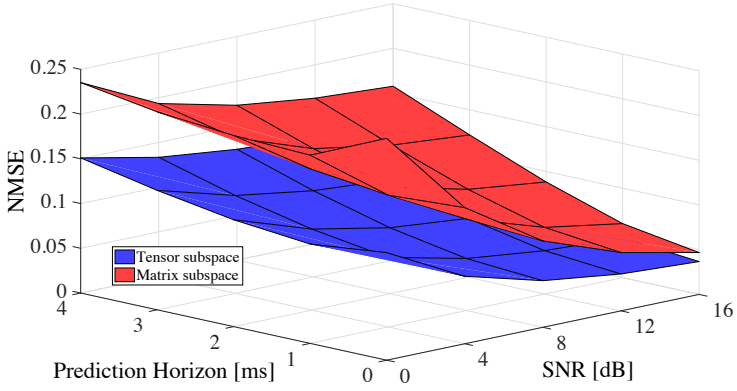


Figure 4.6: NMSE as a function of SNR and prediction horizon for tensor and matrix subspace based prediction schemes.

In Figure 4.6, we have plotted the NMSE as a function of the SNR and the prediction horizon, for the HOSVD based prediction scheme under consideration and a matrix subspace based prediction scheme from [MDSH06]. We can clearly see that prediction performed on the transformed tensor elements, rather than on a matricised version of the channel, provides intrinsic spatial noise reduction leading to a

better prediction performance. Moreover, the results are worse for the case when prediction is performed directly on the channel coefficients (please see [MDH08]). On the downside, this scheme invokes additional complexity due to the tensor transformations and HOSVD computation.

4.5 Combined Scheme: HOSVD followed by Kalman filter-based Prediction

In practical systems, the SNR or Signal-to-Interference-plus-Noise Ratio (SINR) is usually not that high and noise plays a significant role in prediction performance. It has been observed that in a MIMO or massive MIMO system, there are several weak channels with very low SNR compared to the strongest channels, but they still need to be estimated or predicted for good coverage and spectral efficiency. Correspondingly, in Grid of Beams (GoB) systems, it has been noted that a UE observes more than one serving beam and some of these beams could be very weak, in terms of the SNR, compared to the strongest one. For a better prediction of such weak channels, we propose an HOSVD based denoising step to clean the noise in the channels before using any of the state-of-the-art prediction schemes. This scheme and the corresponding results have been published in [AZH17].

As discussed in Section 4.4, massive MIMO channels exhibit a multi-dimensional structure, but most of the channel prediction algorithms are derived based on two-dimensional models. Consequently, the multi-dimensional data have to be unfolded as a vector or a matrix to be processed. These operations result in a loss of structure, which could lead to several issues like a decrease in performance and a lack of robustness. To preserve this structural information, it is beneficial to represent the massive MIMO channel as a multi-dimensional array, called a tensor. In this case, we use the truncated HOSVD to exploit the structural information and denoise the channels prior to prediction with state-of-the-art two-dimensional methods.

In [MDH08] as discussed in Section 4.4, the authors proposed to perform prediction on the individual elements of the truncated transformed tensor using Wiener filters which can be computationally expensive. Our approach is to do the channel denoising and prediction separately. The benefit is that the denoising step can be considered as an additional module and depending on the SNR of the received channel, for example, in a cooperative scenario, can be switched on or off. This would considerably reduce the complexity and provide an overall adaptive solution. In addition, any state-of-the-art channel predictor can be utilized without any modifications.

Assume that we have access to the 4D noisy estimated channel tensor $\tilde{\boldsymbol{\mathcal{H}}} \in$

$\mathbb{C}^{M_T \times M_R \times N_{\mathrm{RS}} \times N_t}$ at the receiver, where M_T and M_R are the number of antenna elements at the base station and at the UE, while N_{RS} and N_t are the number of snapshots, sampled in the frequency and the time domain corresponding to the pilot signals, respectively. We can divide this tensor into several sub-tensors, in the time and frequency dimensions, so that each sub-tensor stays within the coherence time and coherence bandwidth of the channel. We can then calculate the HOSVD of each channel sub-tensor using equation (4.39), which gives us the core tensor $\boldsymbol{\mathcal{S}}$ and the factor matrices $\mathbf{U}_n$ comprised of basis vectors in each dimension. As the time dimension of each sub-tensor corresponds to the coherence time of the channel and it is usually composed of multipath components whose directions of arrival and departure change slowly in time, the channel will span a common signal subspace so that a reduced number r_n, where $n = 1, 2, \ldots, 4$, of basis vectors per tensor dimension n is sufficient to represent it. Here, $0 < r_n \leq I_n$, denotes the number of dominant components of the different matrix unfoldings. Therefore, using a low rank approximation of $\boldsymbol{\mathcal{S}}$, it is possible to denoise the channel as [AZH17],

$$\hat{\boldsymbol{\mathcal{H}}} = \boldsymbol{\mathcal{S}}^{[r_1, r_2, r_3, r_4]} \times_1 \mathbf{U}_1^{[r_1]} \times_2 \mathbf{U}_2^{[r_2]} \times_3 \mathbf{U}_3^{[r_3]} \times_4 \mathbf{U}_4^{[r_4]} \tag{4.44}$$

where $\mathbf{U}_n^{[r_n]} \in \mathbb{C}^{I_n \times r_n}$, $n = 1, 2, \ldots, 4$, is the matrix consisting of r_n n-mode singular vectors (defining the signal subspace) corresponding to the r_n strongest n-mode singular values, and $\boldsymbol{\mathcal{S}}^{[r_1, r_2, r_3, r_4]} \in \mathbb{C}^{r_1 \times r_2 \times r_3 \times r_4}$ is the core tensor consisting of elements corresponding to the signal subspace singular vectors.

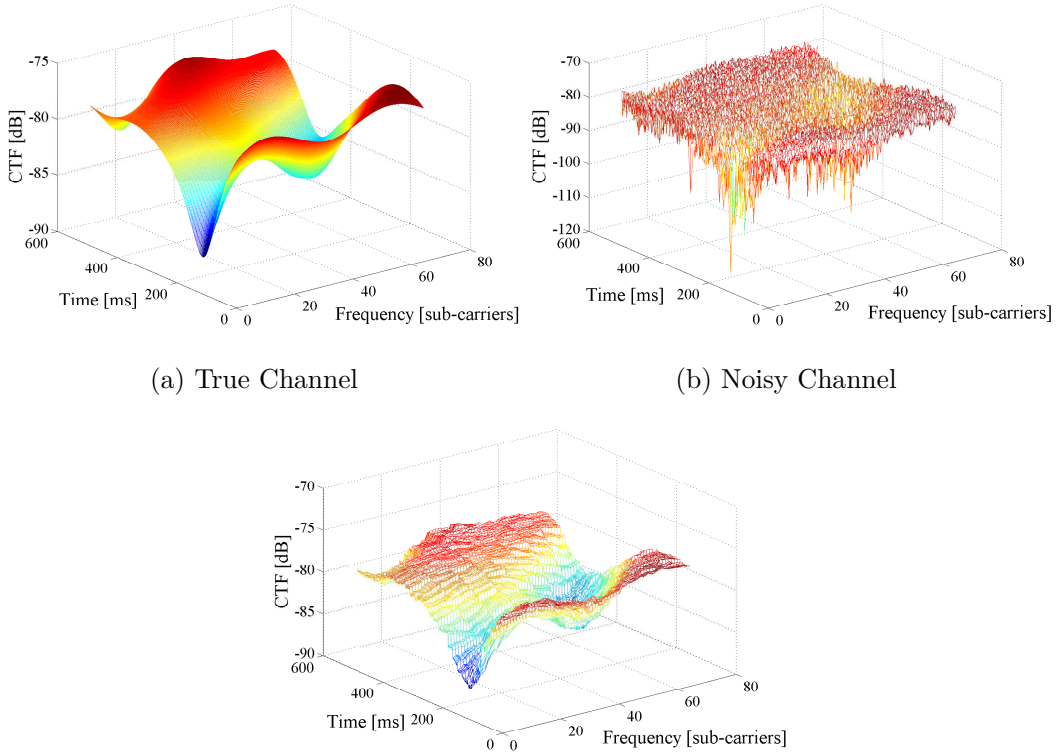
Since we have defined the tensor model and the HOSVD of the massive MIMO channel, let $\tilde{h}_{m_T, m_R, f_j, t_k}$ denote the truncated tensor element corresponding to the denoised (or cleaned) channel between the m_T-th transmit and the m_R-th receive antenna at baseband frequency $f_j = j\Delta f$ and time $t_{t_k} = k\Delta t$.

Here, the state-of-the-art prediction scheme of choice is the Kalman filter-based channel prediction as discussed in Section 4.3. We construct a state space model for the Kalman filter by modeling the truncated tensor element as an AR process. The denoised tensor channel coefficient $\tilde{h}_{m_T, m_R, f_j, t_k}$ at time t_k can be modeled as

$$\hat{h}_{m_T, m_R, f_j, t_k} + a_1 \hat{h}_{m_T, m_R, f_j, t_k - 1} + \ldots + a_q \hat{h}_{m_T, m_R, f_j, t_k - q} = u_t \tag{4.45}$$

where q is the model order and is equal to 2 in our simulation example, $\{a_1, \ldots, a_q\}$ are the AR parameters, and u_t is the process noise that excites the process. The estimation of AR parameters based on blocks of measurements, the formulation of the three-hierarchical state space models to model the overall MIMO tensor, and the prediction scheme are detailed already in Section 4.3. Overall, the prediction of the time-variant frequency-selective denoised massive MIMO channel tensor is

partitioned into $M_T \cdot M_R \cdot N_{\mathrm{RS}}/w$ SISO channel predictions, where w is the number of sub-carriers modeled together in the Kalman based prediction scheme. The value of w is set to 8 in the simulations.



(a) True Channel

(b) Noisy Channel

(c) Cleaned Channel

Figure 4.7: A representation of the denoising effect of the HOSVD processing. The channel transfer function $H(f,t)$ for one particular transmit and receive antenna pair is shown here.

4.5.1 Performance Benefits of HOSVD-based Denoising

To analyze the benefit of such an HOSVD based denoising, we utilize the same simulation setup as described in Section 4.4.1. Discarding the singular values that most likely correspond to the noise subspace, leads to clean channels which are very close to the respective true channels. This can be very helpful for channel prediction. An example of this cleaning effect is shown in Figure 4.7 for an SNR of 10 dB.

The NMSE is used to characterize the performance as it provides a direct relative measure for the degradation due to channel prediction. The NMSE of the massive

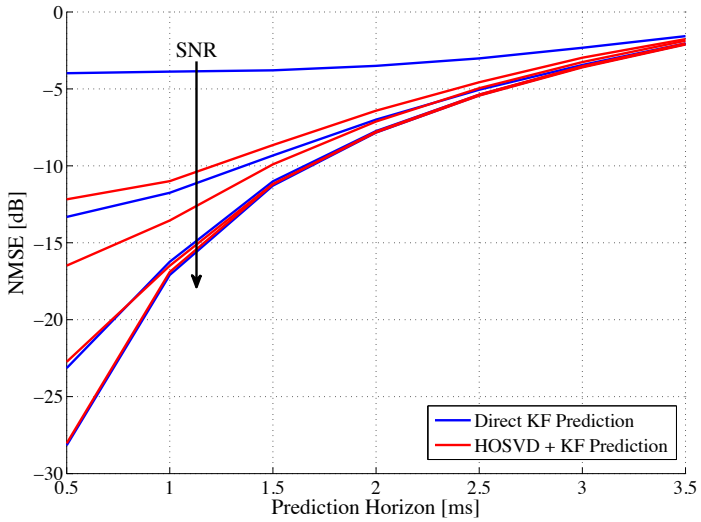


Figure 4.8: Normalized Mean Square Error (NMSE) comparison of prediction performance of noisy channels vs denoised channels using HOSVD-based subspace processing. SNR for the different curves is 0 dB, 10 dB, 20 dB and 40 dB.

MIMO channel tensor is given as,

$$\text{NMSE} = \frac{\text{E}\{\|\boldsymbol{\mathcal{H}} - \hat{\boldsymbol{\mathcal{H}}}\|^2\}}{\text{E}\{\|\boldsymbol{\mathcal{H}}\|^2\}}. \tag{4.46}$$

Figure 4.8 shows a comparison of the prediction performance of the Kalman filter for the noisy channel and for the denoised channel, for the scenario under consideration. The four curves correspond to an SNR of 0 dB, 10 dB, 20 dB and 40 dB from top to bottom, respectively. The performance difference is bigger for low SNR values, where the proposed method reduces the NMSE significantly. At 0 dB SNR, the HOSVD based prediction outperforms the direct Kalman filter-based prediction by 7 dB at 1 ms prediction horizon. This is due to the fact that the denoised channel contains the singular vectors corresponding to the strongest singular values (signal subspace) while neglecting most of the noise subspace. Although the Kalman filter can also cater for noisy measurements as observed in Figure 4.8, the pre-denoising step definitely improves the performance at low SNR. It should be noted that if the noise level value is not determined appropriately, we might neglect a part of the signal subspace or include the noise and therefore increase prediction errors resulting in less gain. At higher SNRs, we do not see any gain because the noise does not play a significant role and the inherent noise filtering of the Kalman filter is enough to provide a good prediction.

We can see in Figure 4.8 that at a 1 ms prediction horizon and 10 dB SNR, the NMSE is approximately -13 dB. For the same value of prediction horizon and

SNR, we can observe in Figure 4.6 that the NMSE is 0.05, which corresponds to approximately -13 dB as well. Hence, both solutions provide a similar performance with regards to the NMSE. However, the benefit of the proposed scheme is that the HOSVD based denoising step can be turned off at high SNRs, as it provides no further gain. Moreover, in practical scenarios, the Wiener filter suffers from various problems like, for example, the need for large memory filters and the recalculation of the filter coefficients at each prediction step. With a state-space formulation, such as the one incorporated in a Kalman filter, such problems can be mitigated [K$^+$60]. This leads to an overall lower complexity solution from a practical standpoint.

4.6 Summary

In this chapter, we have surveyed the distinguishing design features, performance and shortcomings of various state-of-the-art channel prediction schemes. We have identified that channel frequency response based channel prediction schemes, despite being relatively conventional and old school as compared to parametric-model based schemes, are more suitable for bursty transmission which is usually the case for cellular traffic. Thus, they still might be the scheme of choice for future 5G networks. We look at two state-of-the-art channel frequency response based schemes in detail, namely Kalman filter-based prediction and a tensor based prediction scheme. We have then proposed a hybrid scheme which combines the benefits of each scheme, the noise reduction capability of the tensor framework and the low complexity of the Kalman filter-based predictor, while overcoming the drawbacks to potentially provide a more well rounded, better performing and useful scheme.

Chapter 5

Virtual Beamforming

Chapter Summary

This chapter motivates the benefits of beamforming at the UE side through a novel scheme called virtual beamforming. We observe a reduction in the number of channel components and also in the number of multipath components within each channel component. As a consequence, we see an improvement in the overall channel prediction performance irrespective of the prediction algorithm used. The results are further improved with the application of Grid of Beams (GoB) at the transmitter side. We have already introduced the GoB concept in Section 2.3.3 and in Section 5.3, a comprehensive description of virtual beamforming is presented. The contents of this chapter have been published in [AZH15a, AZH15b].

5.1 Introduction

It has been identified that User Equipments (UEs) with interference suppression capabilities allow for less complex coordination schemes at the base stations, and that fundamental enablers like channel prediction should be considered in order to overcome the limitation due to feedback delays [GS12]. Interference cancellation capabilities can even be enhanced if the network assists the UEs in this task. Moreover, a joint optimization of base station clustering, beamforming design and received signal processing at the UE can be exploited to achieve higher spectral and energy efficiency.

The beam pattern of an antenna array depends on various parameters like the inter-element distance, number of antenna elements, orientation of the array, and the beamforming scheme to name a few. The beamwidth of the mainbeam-lobe of the beam pattern is physically limited for a given set of parameters and exhibits an inverse relationship with the number of antenna elements, and consequently the aperture of the array. At the UE, we have limited space and power, which will limit the maximum number of antenna elements at the UE side (especially below 6 GHz). This translates to a limited beamforming gain, poor interference rejection capabilities, a relatively large Half-Power BeamWidth (HPBW), inadequate channel prediction quality due to a high number of received multipath components, and a bad condition number of the channel matrix. Hence, we need to make use of or imitate the physical features of an array artificially to reduce the beamwidth and/or increase the beamforming gain.

By controlling the phase and relative amplitude of the signal at each antenna element in an array, a beamformer allows us to achieve spatial selectivity at the transmitter, the receiver or both transmission ends. This can help us in reducing the number of multipath components, and potentially channel components as well,

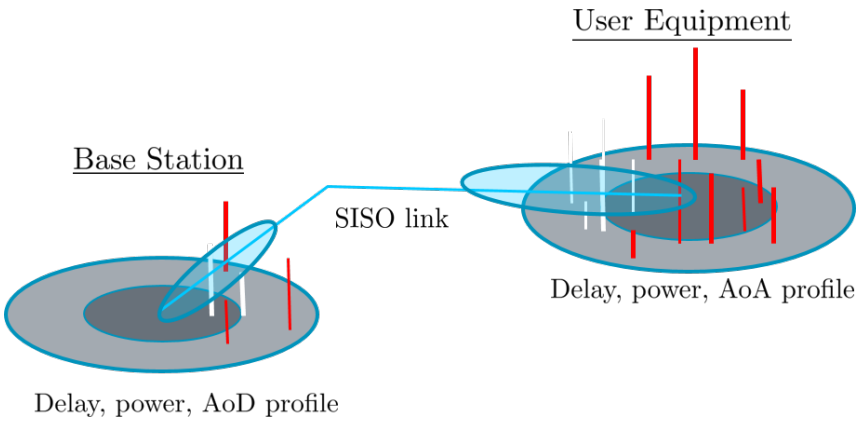
in the channel frequency response, hence making the channel prediction problem easier.

A number of beamforming techniques is available in the literature and the one being considered here for UE-sided beamforming is the novel virtual beamforming concept [ZH13, AZH15a]. We will assess virtual beamforming by itself at the UE side, and also in combination with the Grid of Beams (GoB) concept, discussed in Section 2.3.3, at the base station side. Virtual beamforming is analogous to the Synthetic-Aperture RADAR (SAR) schemes, which are quite well-known and well-developed in the RADAR literature [MPIY+13]. A virtual beamformer is equipped with a few, or even a single antenna element, but utilizes the movement of the UE to generate a virtual Uniform Linear Array (ULA) with a larger aperture. This enables us to reap the benefits of large arrays with a limited number of physical antenna elements, as in the case of cellular UEs. An inherent drawback of virtual beamforming is that it requires strict time synchronization and the data symbols need to be re-transmitted. These drawbacks are discussed in detail in Section 5.3 and in Chapter 6.

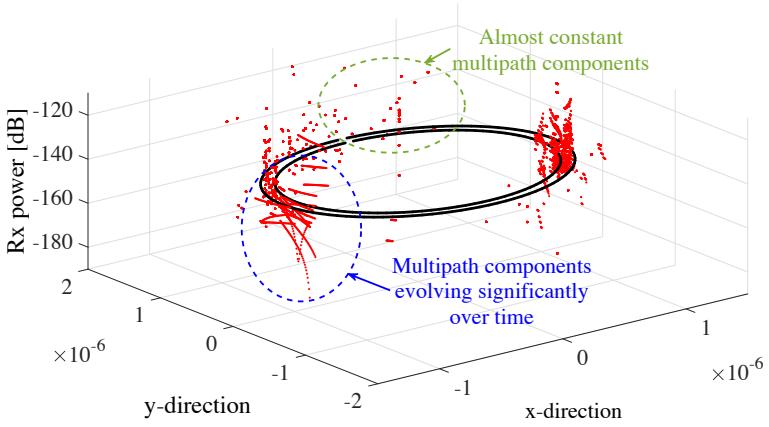
5.1.1 Benefits of UE-Sided Beamforming

The availability of steerable beams with very narrow HPBW at the UE can be beneficially exploited in various ways. For example, in the case of downlink transmission and especially for Frequency Division Duplex (FDD), one of the main benefits would be the ability of the UE to adapt autonomously to changing channel and interference conditions, without involving any control information exchange with the base station, therefore allowing fast reactions to channel variations. In addition, a narrow beam might reduce the number of multipath components from several hundreds to a few tens, thereby improving channel predictability as well as reducing overhead for feedback of Channel State Information (CSI), without the need of any central coordination [ZTKW16]. Furthermore, the number of relevant channel components might be reduced to some extent by suitably pointing the receive beam, which again reduces feedback overhead and provides further robustness, e.g., for massive MIMO precoding including Joint Transmission Coordinated MultiPoint (JT-CoMP) (discussed in Section 2.2) over several distributed sites [AZH15b]. In that sense, narrow beamforming at the UE side will quite possibly be the main enabler in 5G to achieve the intended performance gains on the system level.

Beamforming at the UE side is more effective in reducing the number of multipath components as compared to the typical implementation of beamforming only at the base station side, especially in typical urban macro Non-Line-Of-Sight (NLOS) scenarios. The reason for this is the high spread of the AoA at the UE which



(a) An illustration of the comparison of the Angle of Departure (AoD) spread at base station with the Angle of Arrival (AoA) spread at the UE. The power and delay of the various paths are also depicted for both transmission ends.



(b) Received power over AoA and delay in microseconds for all multipath components in a ray-traced channel at the Nokia campus.

Figure 5.1: A depiction of the benefits of UE-sided beamforming to reduce the number of multipath components by selecting a spatial subset of the channel.

means that a narrow beam pointed in the direction with the strongest multipath components will filter out a high number of multipath components from the other directions. In contrast, even though the AoD spread at the base station is smaller when the GoB concept is used, in NLOS scenarios the reflections of the beams are spread out in many directions and suffer from power degradation, reducing the effectiveness of the beamforming scheme. An illustration is provided in Figure 5.1a where we can see the typical comparison of the AoD spread at the base station side as compared to the AoA spread at the UE side. The center of the gray circles represents the location of the base station and the UE, at the respective sides, and the radii of the gray circles represent various delay values. The blue circles represent the beams on both sides. The location of the vertical bars represents the AoD and

the AoA profiles, at the respective sides, with the length of the bar representing the power of that particular multipath component. The observable multipath components after beamforming are represented with white bars, whereas those represented with red bars are filtered out. It is apparent that a stronger reduction in the number of multipath components is possible at the UE side.

A supporting practical example to the general illustration is shown in Figure 5.1b, which shows the received power of all multipath components as a function of the AoA and delay of the received signal for ray-traced channels generated using WinProp (please see Section 3.4.1.2). The UE is placed at the center $(0, 0)$, the red dots show the evolution over 50 time samples of the relative direction of the last interaction points of the multipath components. The radius from the center defines the delay values as shown by the two circles. We can see that some multipath components are evolving quite significantly over time, indicating a strong Doppler shift, while others are almost constant. A suitably pointed narrow beam can take advantage of these almost constant multipath components and produce a channel which is, as a consequence, easier to predict.

5.2 A Primer on Beamforming

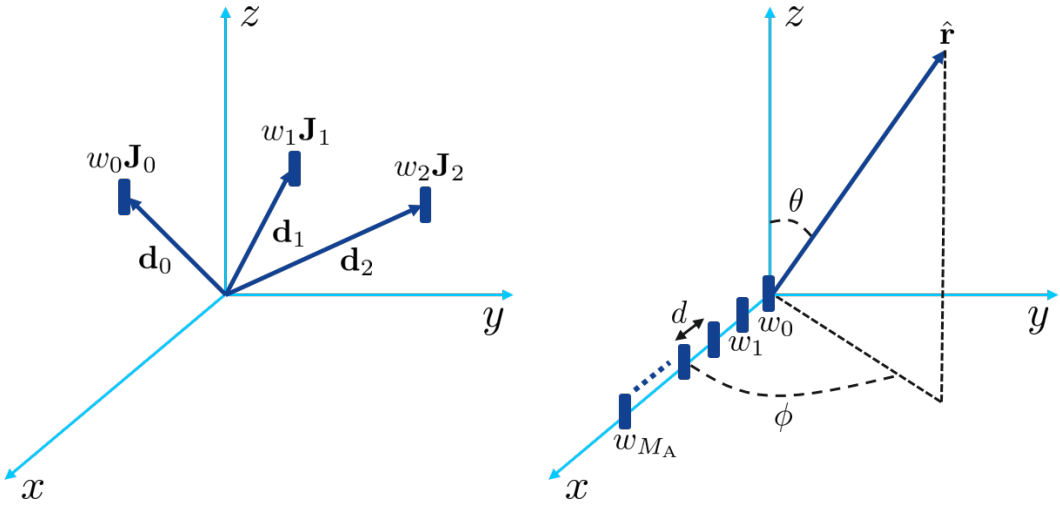
Before we delve into the specifics of the novel virtual beamforming scheme, let us go through the basics of beamforming.

A fundamental property of an antenna array is that phase shifts are introduced into the radiation vectors by the relative displacements of the antenna elements with respect to each other. These phase shifts add up constructively in some directions and destructively in other directions leading to a directional radiation pattern for the antenna array. This stems from the *translational phase-shift* property of Fourier transforms, according to which, a translation in time or space turns into a phase shift in the Fourier domain.

Let us consider a generic array comprised of identical antenna elements that are placed at the positions $\mathbf{d}_0, \mathbf{d}_1, \mathbf{d}_2, \ldots$ in 3-Dimensional (3D) space, with the amplitudes of the excitation currents given as $w_0, w_1, w_2, \ldots$ as shown in Figure 5.2a. Please note that $\mathbf{d}_0$ can be placed at the origin and w_0 can be set to 1 without loss of generality.

The current density vector $\mathbf{J}$ has a magnitude equal to the electric current density and it points in the direction of flow of the electric charge[1]. Thus, the current density vector of the n-th antenna is $\mathbf{J}_n(\mathbf{r}) = w_n \mathbf{J}(\mathbf{r} - \mathbf{d}_n)$. Based on this translated current, the corresponding radiation vector, which is the 3D Fourier

[1]Note that $\mathbf{J}$ is a vector but it is mostly represented with a capital letter in the literature. That notation has been followed here as well.



(a) Several antennas translated to different locations and fed with different excitation amplitudes.

(b) A exemplary linear array along the x-axis. Here, θ is the co-elevation angle and ϕ is the azimuth angle.

Figure 5.2: Various array configurations.

transform of the current density, is given as [Orf16],

$$\mathbf{F}_n(\mathbf{k}) = w_n e^{\jmath \mathbf{k}^{\mathrm{T}} \mathbf{d}_n} \mathbf{F}(\mathbf{k}) \tag{5.1}$$

where $\mathbf{k} = k\hat{\mathbf{r}} = \frac{2\pi}{\lambda}\hat{\mathbf{r}}$ is called the *wave vector* with $\hat{\mathbf{r}}$ being the radial unit vector, $\mathbf{F}(\mathbf{k})$ is the radiation vector of a single antenna element at the origin and $\mathbf{F}_n(\mathbf{k})$ is the radiation vector of the n-th element of the array[1]. The term $e^{\jmath \mathbf{k}^{\mathrm{T}} \mathbf{d}_n}$ stems from the translational phase shift property. The total radiation vector is now given as,

$$\mathbf{F}_{\text{total}}(\mathbf{k}) = \mathbf{F}_0 + \mathbf{F}_1 + \mathbf{F}_2 + \ldots = w_0 e^{\jmath \mathbf{k}^{\mathrm{T}} \mathbf{d}_0} \mathbf{F}(\mathbf{k}) + w_1 e^{\jmath \mathbf{k}^{\mathrm{T}} \mathbf{d}_1} \mathbf{F}(\mathbf{k}) + w_2 e^{\jmath \mathbf{k}^{\mathrm{T}} \mathbf{d}_2} \mathbf{F}(\mathbf{k}) + \ldots \tag{5.2}$$

As the factor $\mathbf{F}(\mathbf{k})$, is common to all the terms, we can rewrite the above equation as,

$$\mathbf{F}_{\text{total}}(\mathbf{k}) = \underbrace{\left(w_0 e^{\jmath \mathbf{k}^{\mathrm{T}} \mathbf{d}_0} + w_1 e^{\jmath \mathbf{k} \Delta \mathbf{d}_1} + w_2 e^{\jmath \mathbf{k}^{\mathrm{T}} \mathbf{d}_2} + \ldots \right)}_{A(\mathbf{k})} \mathbf{F}(\mathbf{k}) \tag{5.3}$$

$$\mathbf{F}_{\text{total}}(\mathbf{k}) = A(\mathbf{k}) \mathbf{F}(\mathbf{k}) \tag{5.4}$$

The term $A(\mathbf{k})$ is called the *array factor* which can also be denoted as $A(\hat{\mathbf{r}})$ or $A(\theta, \phi)$. The array factor, in itself, encapsulates all the translational phase shifts and the relative weighting coefficients of the array elements. Equation (5.4) tells us that the combined effect of an array comprised of identical antenna elements, is to

[1]Once again, the vectors $\mathbf{F}$ and $\mathbf{F}_n$ are represented with capital letters to keep in sync with the literature.

modify the radiation vector of a single antenna element by the array factor. This property is known as *array pattern multiplication* [Orf16].

5.2.1 Uniform Linear Arrays

Let us consider a one-dimensional array with equally-spaced antenna elements along the x-axis placed at locations $x_n, n = 0, 1, 2, \ldots, M_\mathrm{A} - 1$, where M_A is the total number of elements in the array, as shown in Figure 5.2b. The array factor is then given as,

$$A(\theta, \phi) = \sum_{n=0}^{M_\mathrm{A}-1} w_n e^{\jmath \mathbf{k}^\mathrm{T} \mathbf{d}_n} = \sum_n w_n e^{\jmath k_x x_n} \tag{5.5}$$

where $\mathbf{d}_n = x_n \hat{\mathbf{x}}$. If we set $k_x = k \sin(\theta) \cos(\phi)$ and as the array is comprised of equally-spaced elements leading to $x_n = nd$, where d is the distance between the elements, the array factor can now be expressed as,

$$A(\theta, \phi) = \sum_{n=0}^{M_\mathrm{A}-1} w_n e^{\jmath knd \sin(\theta) \cos(\phi)} \tag{5.6}$$

The angular dependence of the array factor stems from the term $k_x d = kd \sin(\theta) \cos(\phi)$. Hence, we can define the variable [Orf16],

$$\psi = k_x d = kd \sin(\theta) \cos(\phi) \tag{5.7}$$

The variable ψ is called the *digital wave number* and it is a normalized version of the wave number k_x. It is measured in the units of radians per (space) sample. Now, we can also define the array factor in the digital wave number space as,

$$A(\psi) = \sum_{n=0}^{M_\mathrm{A}-1} w_n e^{\jmath \psi n} \tag{5.8}$$

If we set the polar angle to $\theta = 90°$, then for an array along the x-axis, we can measure its array factor only in the xy-plane, leading to,

$$\psi = kd \cos(\phi) \tag{5.9}$$

In this case, the azimuthal angle varies over $-\pi \leq \phi \leq \pi$, but as the array response is symmetric in ψ, it can be evaluated only for $0 \leq \phi \leq \pi$. To make such an array even simpler, assume one having equal weights for all antenna elements, and therefore, the name *uniform* linear array. If the array consists of M_A isotropic

elements at locations $x_n = nd$, where $n = 0, 1, \ldots, M_\mathrm{A} - 1$, we can define,

$$\mathbf{w} = [w_0, w_1, \ldots, w_{M_\mathrm{A}-1}]^\mathrm{T} = \frac{1}{M_\mathrm{A}}[1, 1, \ldots, 1]^\mathrm{T} \tag{5.10}$$

so that the sum of the weights is unity. The corresponding array factor is given as,

$$A(\psi) = \frac{1}{M_\mathrm{A}} \sum_{n=0}^{M_\mathrm{A}-1} e^{\jmath\psi n} = \frac{1}{M_\mathrm{A}} \frac{e^{\jmath M_\mathrm{A}\psi} - 1}{e^{\jmath\psi} - 1} \tag{5.11}$$

This array factor can be viewed as a window-based narrow-beam design using a rectangular window and is the spatial analog of a low-pass Finite Impulse Response (FIR) averaging filter in discrete-time digital signal processing [Orf16]. Another way to express $A(\psi)$ is as follows,

$$A(\psi) = \frac{\sin(\frac{M_\mathrm{A}\psi}{2})}{M_\mathrm{A}\sin(\frac{\psi}{2})} e^{\jmath(M_\mathrm{A}-1)\psi/2} \tag{5.12}$$

5.2.2 Fundamentals of Linear Arrays

When designing ULAs, one needs to understand and consider the basic fundamentals for a successful design. To this end, we will briefly explain concepts like visible region, grating lobes, beamwidth, and steering angle.

Visible Region and Grating Lobes

By the definition of ψ in equation (5.9), we can observe that the array factor $A(\psi)$ is periodic in ψ with period 2π. Hence, it is sufficient to know it within $-\pi \leq \psi \leq \pi$, a single Nyquist interval. If we vary the azimuthal angle ϕ from 0° to 180°, the value of $\psi = kd\cos(\phi)$, varies in the range,

$$-kd \leq \psi \leq kd \tag{5.13}$$

This range of variation of ψ is called the *visible region* and its total width is $\psi_\mathrm{vis} = 2kd$. It is dependent on the quantity $kd = \frac{2\pi}{\lambda}d$. For example, if $d < \lambda/2$, it means that $kd < \pi$ leading to a visible region $\psi_\mathrm{vis} < 2\pi$. Hence, the visible region is less than one Nyquist interval. On the other hand, when $d = \lambda/2$, $kd = \pi$ and the visible region covers the whole Nyquist interval $\psi_\mathrm{vis} = 2\pi$.

In the case of $d > \lambda/2$, $kd > \pi$ and the width of the visible region is more than one Nyquist interval $\psi_\mathrm{vis} > 2\pi$. This means that the values of $A(\psi)$ are over-specified and repeat over the visible region leading to *grating lobes*. These are basically mainbeam lobes pointing in undesired directions. They are the spectral images produced by the sampling process and in ψ-space, these images fall in other

Nyquist intervals as opposed to the central one.

The number of complete Nyquist intervals that fit within the width of the visible region dictates the number of grating lobes in an array pattern, that is, $\psi_{\text{vis}} = kd/\pi = 2d/\lambda$. Typically, grating lobes are undesirable and can be completely avoided by setting $d \leq \lambda/2$.

Array Pattern Steering

The array along the x-axis that we have discussed so far, has been designed to have its maximum directive gain at broadside, or in other words, at $\phi = 90°$. The array factor $A(\psi)$ has its maximum at $\psi = kd\cos(\phi) = 0$, i.e., $|A|_{\max} = |A(0)|$. However, the main motivation to design antenna arrays is to steer the array pattern, without physically rotating it, towards some desired direction, say ϕ_0 leading to a corresponding wave number $\psi_0 = kd\cos(\phi_0)$.

This task of steering the beam can be accomplished by wave number translation in ψ-space. This implies that the broadside pattern $A(\psi)$ needs to be replaced by the translated pattern $A(\psi - \psi_0)$. Consequently, we can define the steered array factor as,

$$A'(\psi) = A(\psi - \psi_0) \tag{5.14}$$

and the translated wave number variable as,

$$\psi' = \psi - \psi_0 = kd(\cos(\phi) - \cos(\phi_0)) \tag{5.15}$$

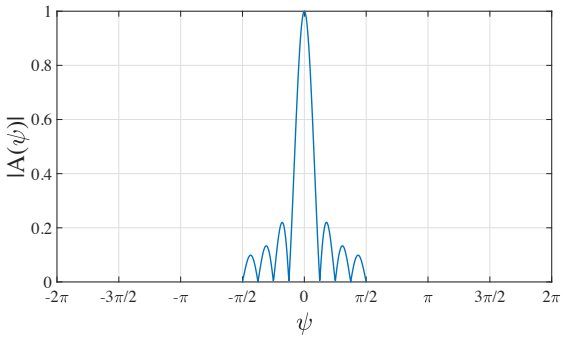
In order to calculate the steered array weight coefficients w'_n, we can make use of the translation theorem of Fourier transforms, according to which the weight coefficients of the translated pattern $A'(\psi)$ are given as [Orf16],

$$w'_n = w_n e^{-\jmath\psi_0 n} \tag{5.16}$$

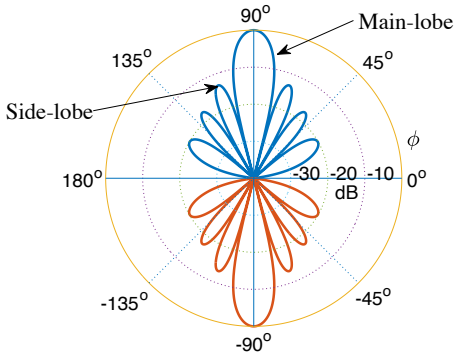
The steered array is also occasionally called a *phased* or *scanning* array due to the progressive phase factors $e^{-\jmath\psi_0 n}$ present in the weights in equation (5.16). The steered array factor is now given as,

$$\begin{aligned} A'(\psi) &= \sum_{n=0}^{M_{\text{A}}-1} w'_n e^{\jmath\psi n} = \sum_{n=0}^{M_{\text{A}}-1} w_n e^{-\jmath\psi_0 n} e^{\jmath\psi n} \\ &= \sum_{n=0}^{M_{\text{A}}-1} w_n e^{\jmath(\psi-\psi_0)n} = \sum_{n=0}^{M_{\text{A}}-1} w_n e^{\jmath\psi' n} = A(\psi') \end{aligned} \tag{5.17}$$

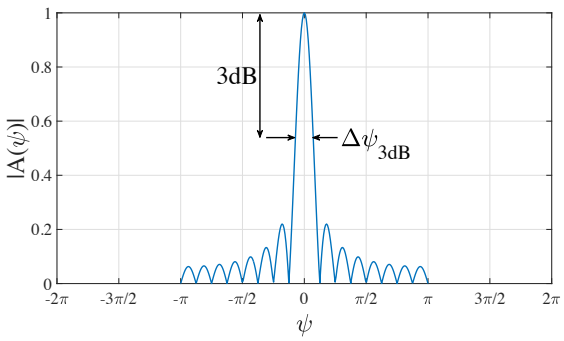
Hence, the of maximum of $A'(\psi)$ coincides with the maximum of $A(\psi')$, which occurs at $\psi = \psi_0$, or at the azimuthal angle $\phi = \phi_0$.



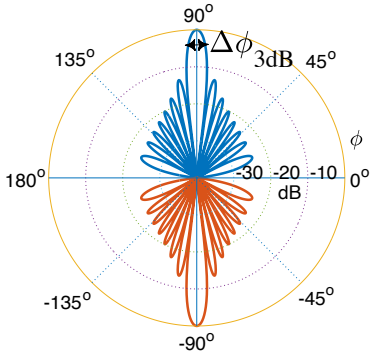
(a) Array factor and a representation of the visible region ψ_{vis} for $d = \lambda/4$.



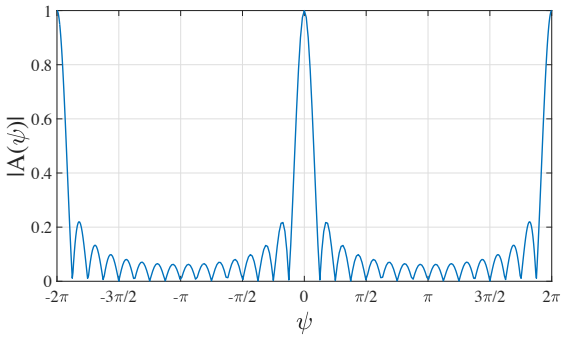
(b) Azimuthal beam pattern for $d = \lambda/4$.



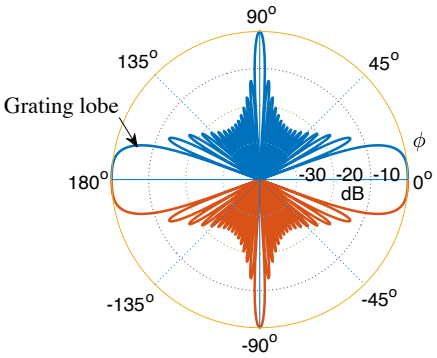
(c) Array factor and a representation of the visible region ψ_{vis} for $d = \lambda/2$. The 3dB bandwidth $\Delta\psi_{\text{3dB}}$ is also depicted.



(d) Azimuthal beam pattern for $d = \lambda/2$. The 3dB bandwidth $\Delta\phi_{\text{3dB}}$ is also depicted.



(e) Array factor and a representation of the visible region ψ_{vis} for $d = \lambda$.



(f) Azimuthal beam pattern for $d = \lambda$. The presence of grating lobes can be observed.

Figure 5.3: Array factor and azimuthal beam patterns for varying elemental spacing of a 16-element ULA. The fundamental concepts of arrays are illustrated.

In the case of a steered array, the concepts of visible region carries over with minor adjustment. Hence, as the azimuthal angle ϕ varies from 0° to 180°, the translated wave number ψ' of equation (5.15) varies over the shifted visible region,

$$-kd(1+\cos(\phi_0)) \leq \psi' \leq kd(1+\cos(\phi_0)) \tag{5.18}$$

The total width of the visible region remains $2kd$. Now, the condition for the absence of grating lobes also needs a little modification and is obtained with the help of the following inequality,

$$|\psi'| \leq kd|(\cos(\phi)+\cos(\phi_0))| \leq kd(|\cos(\phi)|+|\cos(\phi_0)|) \leq kd(1+|\cos(\phi_0)|) \tag{5.19}$$

Hence, ψ' must remain strictly less than 2π to ensure no grating lobes, which results in the sufficient condition,

$$\begin{aligned} kd(1+|\cos(\phi_0)|) &\leq 2\pi & (5.20)\\ d &\leq \frac{2\pi}{k(1+|\cos(\phi_0)|)} \\ d &\leq \frac{\lambda}{(1+|\cos(\phi_0)|)} & (5.21) \end{aligned}$$

Beamwidth

When it comes to the beam pattern of a uniform array, two major features are its beamwidth and its side-lobe level. The beamwidth refers to the width of the mainbeam lobe where the power drops by a certain value. That value is, typically set at 3 dB, leading to the definition of Half-Power BeamWidth (HPBW). In ψ-space, it is defined for an array with large M_A as,

$$\Delta\psi_{3\mathrm{dB}} = 0.886\frac{2\pi}{M_\mathrm{A}} \tag{5.22}$$

where $\frac{2\pi}{M_\mathrm{A}}$ represents half of the base of the main-lobe as the first nulls in the array factor concerning the main-lobe occur at $\pm\psi_1 = \pm\frac{2\pi}{M_\mathrm{A}}$. The HPBW in angle space $\Delta\phi_{3\mathrm{dB}}$ can be calculated by differentiating the equation (5.9) and is given as,

$$\Delta\phi_{3\mathrm{dB}} = 0.886\frac{\lambda}{M_\mathrm{A}d} \tag{5.23}$$

Here, $M_\mathrm{A}d$ is the effective aperture of the array. It is important to note that the beamwidth gets narrower with increasing M_A. However, the side-lobe level for $M_\mathrm{A} > 6$ becomes independent of M_A and is capped at a value of 13 dB for uniform arrays. In order to achieve lower side-lobe levels, proper non-uniform weights must

be designed.

For the steered array, the HPBW in angle space is given as [Orf16],

$$\Delta\phi_{\mathrm{3dB}} = \begin{cases} \dfrac{0.886}{\sin(\phi_0)} \dfrac{\lambda}{M_{\mathrm{A}} d} b, & 0^\circ \leq \phi_0 \leq 180^\circ \\ 2\sqrt{0.886 \dfrac{\lambda}{M_{\mathrm{A}} d}} b, & \phi_0 = 0, 180^\circ \end{cases} \tag{5.24}$$

Here, b is a so-called *broadening* factor, the value of which depends on the choice of the window function and the side-lobe level. For the uniform arrays considered so far, $b = 1$. Moreover, the presence of the value $\sin(\phi_0)$ in the denominator means that the beamwidth increases as the beam is steered from broadside towards end-fire.

5.3 Virtual Beamforming

To avail the benefits of narrow beams, one would require numerous antennas at the UE, in order to have a large aperture $M_{\mathrm{VB}} d$, where M_{VB} is the number of virtual antenna elements. For centimeter as well as millimeter Wave (mmWave) systems, another important ingredient of future Fifth Generation (5G) systems, a massive number of antenna elements can be placed at the base station as well as at the UE-side due to their very short wavelength λ, of only a few centimeters to even a few millimeters. On the contrary, in the case of sub-6 GHz systems, for example, at 2.6 GHz Radio Frequency (RF) frequency, the wavelength λ is about 0.12 m, i.e., of similar size as a typical smart phone. Thus, with typical $\lambda/2$-spacing, an exemplary array of $10 \times 10 = 100$ antenna elements would already end up in an array size of $0.3 \times 0.3\,\mathrm{m}^2$, which is way beyond the size of a typical user device. Moreover, finite battery power as well as limited space for placing a high number of RF-chains including power amplifiers, filters, analog-to-digital converters etc., are further reasons to expect UEs with a few to a very few antenna elements. So even for the 5G time frame, one can expect just two to maybe eight antenna elements per UE.

For that reason the question arises, *how to make the best use of a limited number of UE antennas with a limited overall aperture size of about one to a few λ?* Our answer to this question is virtual beamforming [ZH13, AZH15a], which is a scheme that allows us to reap the benefits of a much larger virtual ULA at the UE, even if it has only a single physical antenna element.

Virtual beamforming is a scheme which allows us to utilize the movement of the UE to generate a virtual ULA [ZH13, AZH15a]. The idea behind it is to combine

$M_{\mathrm{VB}} \in \mathbb{N}$ adjacent received signal measurements. Ideally, the UE is moving at a constant speed and on a straight line during the measurement phase to provide equidistant measurements. But, this is not a limiting factor as acceleration and haphazard movements of the UE can be catered for easily, through a proper adaptation of the phase changes. The UE can then consider these measurements, which have been acquired over time, as measurements acquired by the antenna elements of an ULA. A schematic is provided in Figure 5.4. This enables the creation of large virtual massive MIMO arrays at the UE side with a limited number of physical antenna elements. For example, if the UE has 4 physical antennas the combination of 8 previous received signal measurements results in a virtual massive MIMO array of size $8 \times 4 = 32$. The downsides to this approach are that the UE has to wait for the time it takes to receive M_{VB} signals before it can use the virtual ULA and the same data symbols have to be transmitted over all the M_{VB} time slots. This retransmission is not an issue in the case of channel prediction as the reference signals are repeated periodically. However, it poses a challenge for data transmission as it translates to a reduction in throughput by M_{VB}, but we have provided a potential remedy for this in Chapter 6.

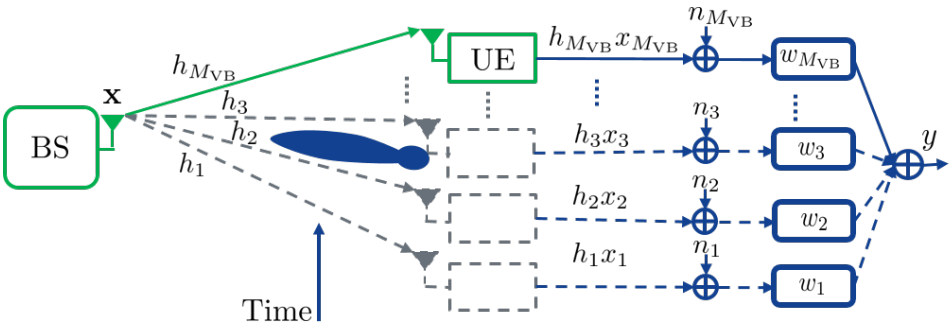


Figure 5.4: A schematic of the virtual beamforming scheme. The green link shows the current location of the UE whereas the the gray links depict the previous locations of the UE where measurements have been stored to perform virtual beamforming.

As an example, consider a downlink Orthogonal Frequency Division Multiplexing (OFDM) system with central RF frequency $f_C = 2.6$ GHz and a 15 kHz subcarrier spacing. For estimation and prediction of Multiple-Input and Multiple-Output (MIMO)-OFDM channels, Channel State Information Reference Signals (CSI-RSs) are inserted, per Long Term Evolution (LTE) specifications [3GP11, Wan13], at every 6^{th} sub-carrier in the frequency domain and every 5 or 10 ms in the time domain. For a particular link between transmit antenna m_T (or a GoB beam m_{GoB}) and receive antenna m_R (which is the only antenna in the case of single antenna UEs), the elements of the sampled time-variant channel frequency response

matrix at the UE after down-conversion and bandpass filtering are given as,

$$H(n_{\mathrm{RS}}, n_t) = H(f_C + 6 \cdot n_{\mathrm{RS}}\Delta f, t_0 + n_t \Delta t_{\mathrm{CSI}}). \tag{5.25}$$

where $n_{\mathrm{RS}} \in 1, \ldots, N_{\mathrm{RS}}$ is the index of the reference signals, $n_t \in 1, \ldots, N_t$ is the time index, Δf is the spacing between the reference signals in frequency, and Δt_{CSI} is the spacing between the reference signals in time. The UE is assumed to be moving with a constant speed along a straight line which leads to regular CSI-RSs. Note that irregular movements of the UE can be used for virtual beamforming as well, if known, for example, from the corresponding sensor data of the UE. For a mobility independent performance evaluation of the scheme, the calculations can also be carried out in the spatial domain. This results in the channel matrix having elements $H(n_{\mathrm{RS}}, n_L) = H(n_{\mathrm{RS}}, n_{L0} + n_L \Delta n_L)$ with $n_L \in 1, \ldots, N_L$ being the UE locations according to Figure 5.4.

We can calculate the virtually beamformed vectors as,

$$\mathbf{h}_{n_t} = \mathbf{H}(1 \ldots N_{\mathrm{RS}}, n_t \ldots n_t + M_{\mathrm{VB}} - 1)\boldsymbol{\omega} \in \mathbb{C}^{N_{\mathrm{RS}} \times 1}, \tag{5.26}$$

for $n_t = 1 \ldots N_t - M_{\mathrm{VB}} + 1$. To get the complete effective beamformed channel after N_t time samples, we combine all the vectors $\mathbf{h}_{n_t}$ into one matrix $\mathbf{H}_{\mathrm{VB}} \in \mathbb{C}^{N_{\mathrm{RS}} \times N_t - M_{\mathrm{VB}} + 1}$. Here,

$$\boldsymbol{\omega} = [\omega_1, \ldots, \omega_{M_{\mathrm{VB}}}]^T \in \mathbb{C}^{M_{\mathrm{VB}} \times 1} \tag{5.27}$$

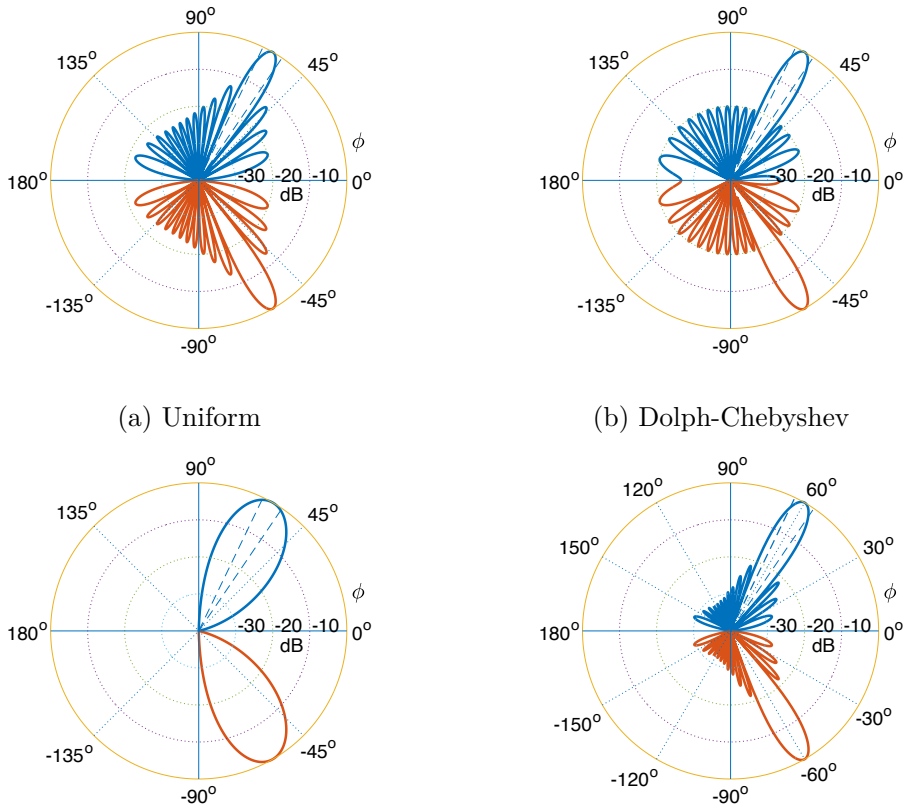
is the weight vector which is applied to generate each $\mathbf{h}_{n_t}$.

5.3.1 Selection of Weights for the Virtual Beamformer

Generally, the optimum virtual complex weight vector $\boldsymbol{\omega}$ should be chosen according to the number of virtual antenna elements, the relative distance between two measurement locations, overall Rx power, delay and power distribution of multipath components in a certain beam direction, interference power from other cells or sites, etc. However, as we focus mainly on a single link, we typically use a fixed phase shift ψ_0 between adjacent measurements. The primary issue in choosing the weights for a fixed look direction is that one would like to have a very narrow main-lobe and very small side-lobes. However, this is not possible as these are conflicting requirements and the beamforming weights need to be designed around the trade-off between main-lobe width and side-lobe level. There are several ways to choose the weights by utilizing window designs like *uniform, Dolph-Chebyshev, Taylor-Kaiser* and *binomial*, each of which offers a certain performance in regards to this conun-

drum. The uniform array weights provide the narrowest main-lobe but also with the highest side-lobes whereas the Dolph-Chebyshev weights provide the narrowest main-lobe for a given side-lobe level, e.g., of 20 dB. The binomial array weights, on the other hand, provide the widest main-lobe but with totally no side-lobes [Orf16]. A comparison is provided in Figure 5.5. As we are after the narrowest beams to simplify the channels as much as possible, we choose the uniform scheme giving us the weight vector with elements,

$$\omega_{m_{\mathrm{VB}}}(\psi_0) = \frac{1}{M_{\mathrm{VB}}} e^{-\jmath \psi_0 m_{\mathrm{VB}}}. \tag{5.28}$$



(a) Uniform

(b) Dolph-Chebyshev

(c) Binomial

(d) Taylor-Kaiser

Figure 5.5: A comparison of the beamwidth of various beamforming weights for an ULA with $d = \lambda/2$ and $M_{\mathrm{A}} = 16$.

The received power is an important criterion as a conventional omni-directional reference receiver will collect all the power, from all multipath components, and from all directions, while a narrow beam might receive only a small portion of that power. We can choose the phase shift ψ_0 in such a manner that the beam is steered towards the significant multipath components in terms of received power. This leads to a general reduction in the number of multipath components forming the virtually

beamformed channel $\mathbf{H}_{\mathrm{VB}}$, but still includes the significant multipath components comprising the channel and allows us to simultaneously exploit sufficient virtual beamforming gains.

For geometrical reasons, multipath components in the broadside direction are less effected by Doppler as compared to multipath components coming from the end-fire directions. Hence, it is beneficial to search for suitable beams in the broadside direction first, especially as it offers beam patterns with typically the lowest HPBW.

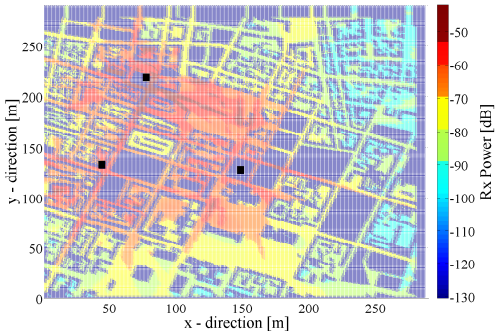
Once we have chosen an appropriate $\psi_0 = kd\cos(\phi_0)$, we can calculate the complex-valued virtual beam pattern $A(\theta, \phi_0)$ over the azimuth angle $\phi_0 \in [-\pi, \ldots, \pi]$ as the far field superposition of all virtual antenna elements transmitting a $\omega = \frac{1}{M_{\mathrm{VB}}}[1, 1, \ldots, 1]^{\mathrm{T}} \in \mathbb{C}^{M_{\mathrm{VB}} \times 1}$ vector in the direction ϕ_0. The beam pattern generated by the virtual ULA is symmetric on both sides of the UE.

5.4 Evaluation of the Benefits of GoB and Virtual beamforming

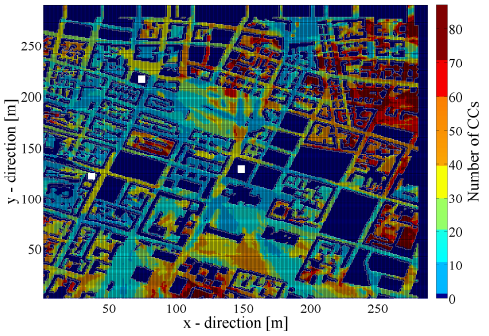
5.4.1 Number of Channel Components

The Interference Mitigation Framework - Advanced (IMF-A), introduced in Section 2.2, uses interference floor shaping to decouple cooperation areas formed by the nine cells of the three cooperating sites. The advantage of this is that one can focus primarily on a single cooperation area for the evaluation of different aspects, like the parameter of interest here, the received power over various outdoor UE locations. The results for such a typical cooperation area can be found in Figure 5.6a, based on a WinProp ray-tracing simulation in the Munich city center (Karlsplatz area). Parameters are close to the 3rd Generation Partnership Project (3GPP) case 1, i.e., an RF frequency of 2.6 GHz, a 20 MHz bandwidth, LTE parameters, a 30 m height of the base station sites and a tilt of 7°. Each cell forms a GoB of 8 equally spaced narrow beams with an HPBW of 2 degrees in the horizontal and 6 degrees in the vertical direction. The path loss is calculated via the accumulation of the power of all the beams from all 3 sites. To generate such narrow beams, massive MIMO antenna arrays with $16 \times 32 = 512$ elements are assumed.

It is interesting to see how many of the channel components or beams from all 3 sites will typically be received by the UEs at different locations, which is shown in Figure 5.6b. Defining relevant channel components as those which fall within a certain received power window, typically less than 20 dB below that of the strongest channel component, we can see that for most of the inner part of the cooperation area, the number of relevant channel components is below 20 or even below 10



(a) The received power in dB for JT-CoMP.



(b) The number of relevant channel components within a 20 dB power window.

Figure 5.6: A depiction of the benefit of GoB over a three site cooperation area with real measurements for the outdoor UEs in the Munich city center. The black and white squares mark the location of the base stations, respectively.

in some cases. In contrast, at the border area of the cooperation area, the UEs might receive 50 or more relevant channel components. The reason is that at the cooperation area border, there are no really strong serving channel components, but instead the UEs experience an interference floor. Fortunately, with the cover shift concept from [ZMK13], cooperation area border UEs will be scheduled into another cover shift, so that these UEs become the center UEs within another cooperation area.

So far we have evaluated the benefits of just GoB, to help reduce the number of relevant channel components. In order to evaluate the benefit of additional UE-sided beamforming in the form of virtual beamforming, we consider a single site scenario at the Nokia campus in Munich. We have employed WinProp (discussed in Section 3.4.1.2) to produce a ray-traced channel with a 1 cm resolution linear grid resulting in a total of 50 UE locations (see *Prediction Area 2* in Figure 5.7). The base station has an ULA of 16 Kathrein antennas, allowing again to form narrow beams of about 2° HPBW.

For a single site case, we can carefully calibrate the virtual beam direction to find the best reduction in channel components possible, in combination with the GoBs at the transmitter side. Figure 5.8 shows the number of relevant channel components for a virtual beam formed by combining 32 antenna elements, steered in various directions. We can observe that the number of relevant channel components can be halved by the application of virtual beamforming. For example, when pointed in the right direction like 2 or 3 degrees in this case, at a reasonable threshold level of less than 20 dB.

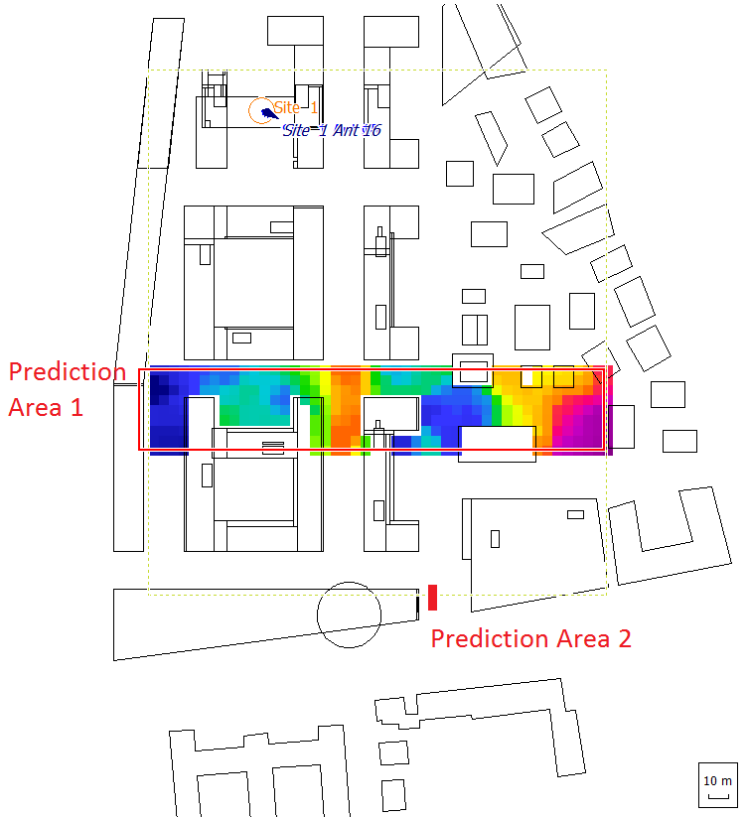


Figure 5.7: WinProp ray-tracing scenario for the Nokia Campus in Munich.

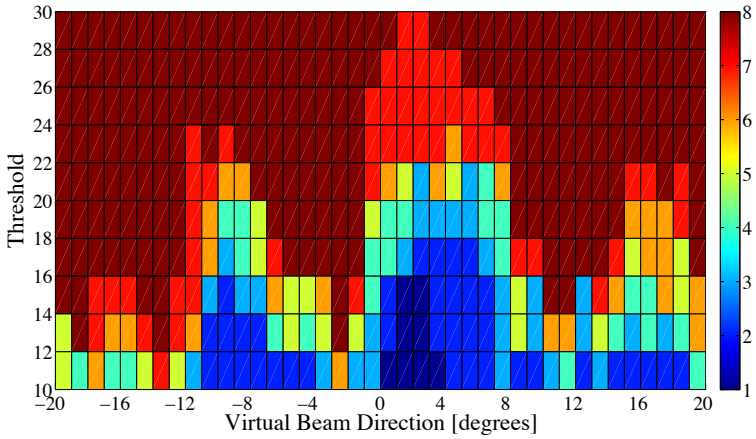


Figure 5.8: Color-coded number of relevant channel components for various virtual beamforming directions and varying threshold levels

5.4.2 Number of Multipath Components

Minimizing the number of channel components is important to reduce the CSI reporting overhead and to increase the overall robustness of the cooperation or massive MIMO precoding. However, another significant issue is the number of multipath components within each channel component, which seems to be of great

importance for the CSI predictability. To get an idea of the benefit of massive MIMO techniques like the GoB and virtual beamforming, to help reduce the number of relevant multipath components, a sub-area of the Nokia campus in Munich (see *Prediction Area 1* in Figure 5.7) has been ray-traced on a 5 m × 5 m grid in WinProp, resulting in approximately 200 UE locations. The base station again has an ULA of 16 standard Kathrein antennas.

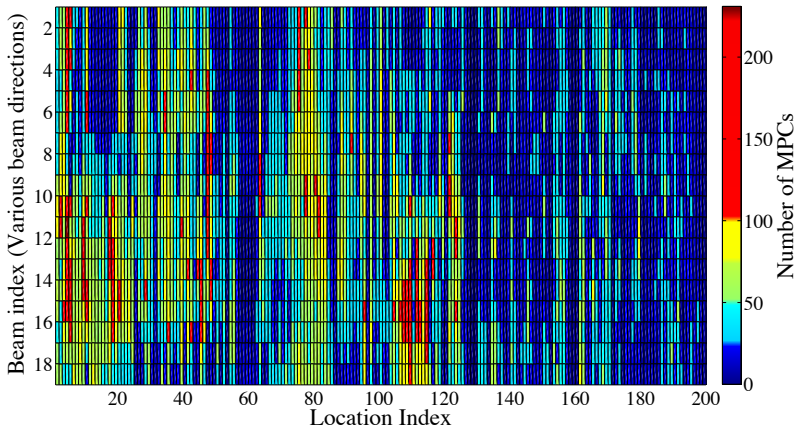


Figure 5.9: Number of multipath components over location 1 to 200 for beam index 1 to 18. Here, MPC is short for multipath component.

As a first test, we form a narrow beam at the base station and then steer it in 10° steps from −90° to +90°, resulting in 18 beams overall. Our goal is to identify the variation in the number of multipath components for all 200 UE locations, with no beamforming performed at the UE side and for each of the steered beams. The result is shown in Figure 5.9 and we can see that the number of multipath components varies quite significantly from a few multipath components to more than 100 multipath components. This means that in a NLOS scenario, the direction of the serving beam at the base station and the location of the UE within the environment, play an important role in terms of the number of multipath components. The location of the UE can be beneficially exploited by employing UE-sided beamforming.

As a further result, see Figure 5.10 where the number of multipath components has been evaluated for a varying threshold defining the relevance of a multipath component. Again, there is a great difference between the different beams being the parameter of the various curves. It can be observed that the serving beams (all the GoB beams received with enough power to serve UEs) experience a strong reduction in the multipath components, something like 10 multipath components, in the case of a threshold of −20 dB. For the reflected beams (all the GoB beams received with a lower received power than acceptable for a reliable communica-

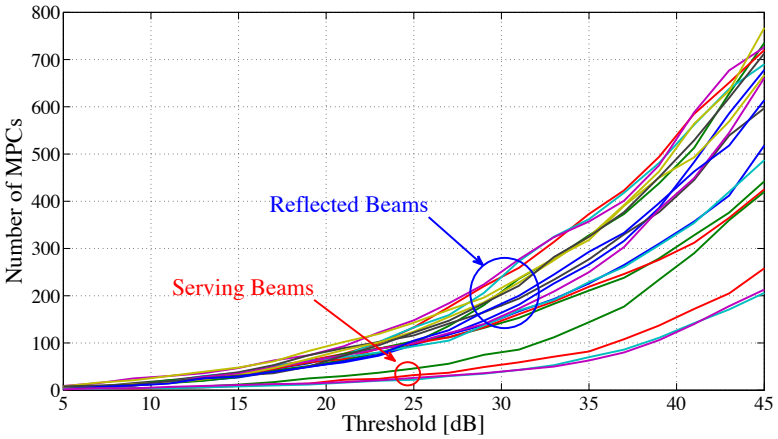


Figure 5.10: Number of multipath components for decreasing threshold from −5 to −45 dB with beam index as parameter for a single measurement location. Here, MPC is short for multipath component.

tion), massive MIMO techniques bring much smaller gains despite the very narrow transmit-beams, comprising on the order of 100 multipath components. This means that the massive MIMO beamforming, like GoB at the transmitter side, will reduce the number of multipath components mainly for the serving beams.

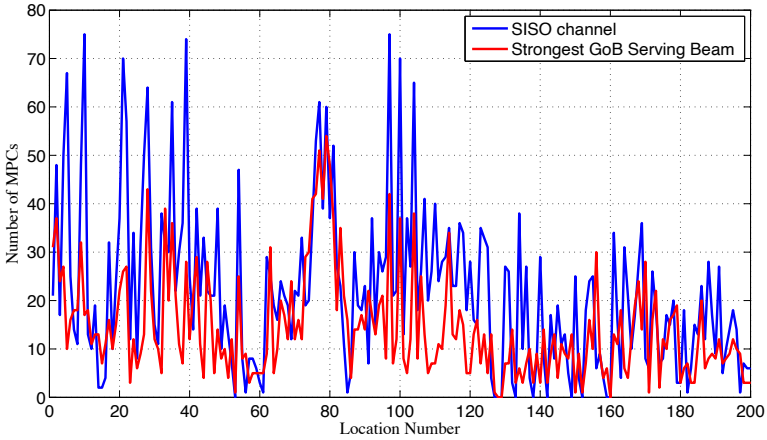


Figure 5.11: Number of multipath components for a Single-Input and Single-Output (SISO) channel versus the strongest serving beam over all 200 locations. Here, MPC is short for multipath component.

Figure 5.11 provides the number of multipath components for all 200 UE locations for the strongest serving beam (red line) versus that of a SISO channel (blue line) at the same location. Almost for all locations, with minor exceptions, the GoB beam has a lower number of multipath components, typically ending up at 10 to 30 multipath components. However, this reduction in the number of multipath components due to massive MIMO beamforming for the serving beam(s) is rela-

tively modest and it is even lower for the reflected beams. It can be explained by the relatively small AoD at base station cells, i.e., even very narrow beams excite a large portion of the multipath components and only a few of them are suppressed. In contrast, at the UE side with a large AoA, beamforming is much more effective in down selecting multipath components (as discussed in Section 5.1.1) which can be verified in Figure 5.12 for the *Prediction Area 2* of Figure 5.7, with the optimum beam direction. Unfortunately, placing many antennas at the UE side is very challenging which solidifies the importance of schemes like virtual beamforming.

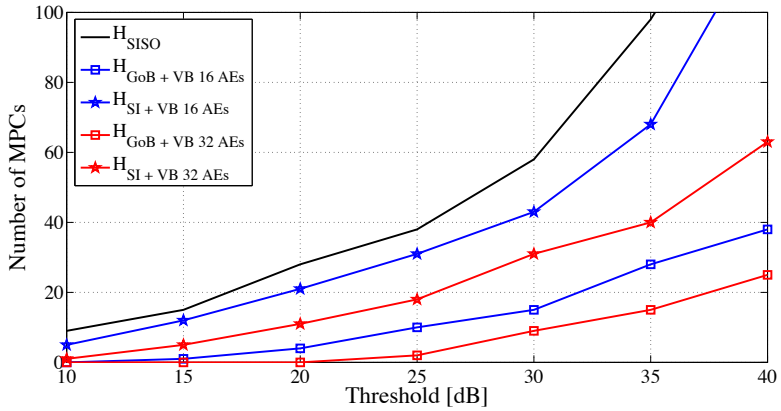


Figure 5.12: Number of multipath components for a SISO channel versus the strongest serving beam over all 200 locations. Here, MPC is short for multipath component.

5.4.3 Channel Prediction Performance

In order to appraise the benefit of reducing the number of multipath components on the performance of channel prediction, we again consider the *Prediction Area 2* of Figure 5.7. The single antenna UE is slowly moving, and by storing and combining several measurement time instances, it forms a virtual antenna array of size 16 or 32 elements [AZH15a]. Note, alternatively we can consider UEs with 16 or 32 physical antenna elements. Each of the 16 base station ULA elements is a standard Kathrein antenna comprising 16 antenna elements each, resulting in a massive MIMO antenna array at the base station of size $16 \times 16 = 256$; placed at a typical macro site height of 30 m. The base station-UE distance is about 300 m and the UE moves along a line of 50 points with a resolution of 1 cm at a height of 2.19 m.

We again use parameters close to the 3GPP case 1. The reduction in the number of multipath components directly affects the time-variance of the relevant taps resulting in a far smoother evolution of the Channel Impulse Response (CIR). For that reason, we have applied the most simple linear prediction per tap of the CIR

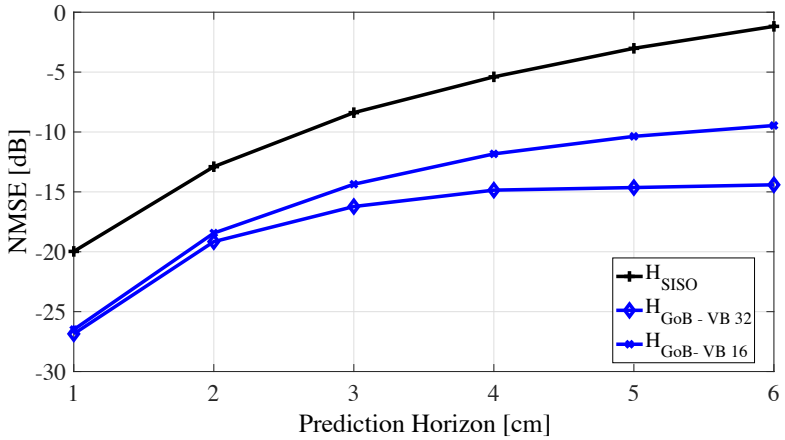


Figure 5.13: NMSE vs prediction horizon for various effective channels

by calculating the amplitude and phase difference of two consecutive time slots and extrapolating this linearly. As usual for linear prediction, the long term performance is degrading fast. But, for the interesting prediction horizon of about 10 ms and mobile speeds up to 15 km/h, corresponding to about 0.3λ to 0.4λ at an RF frequency of 2.6 GHz, this very simple scheme achieves a very good performance for the combination of beamforming schemes which results in the smallest number of multipath components. It achieves an extremely low Normalized Mean Square Error (NMSE) of approximately -20 dB at a prediction horizon of 2 cm (corresponding roughly to 0.2λ), for all the beamforming schemes under consideration. It also achieves an NMSE of -15 dB at prediction horizons up to 6 cm (0.6λ) for a combination of GoB at the transmitter side and virtual beamforming over 32 antenna elements at the receiver side.

5.5 Summary

In this chapter, we have introduced the novel UE-sided beamforming scheme, called virtual beamforming. We have shown the benefit of virtual beamforming to reduce the number of channel components and multipath components, in combination with base station-sided beamforming scheme like GoB, and have shown that a significant improvement is possible in channel prediction performance, with a proper system design. Virtual beamforming helps to subdue the abrupt variations present in a time-varying and frequency-selective channel, resulting in smoother channel frequency responses which are easier to predict by any channel prediction schemes.

Chapter 6

Data Transmission over Virtual Beams

Chapter Summary

This chapter details the challenges encountered when considering virtual beams for data transmission. A potential solution is provided in Section 6.3, in the form of multiplexing of user data by coded transmission and reception over virtual beams. This approach enables us to partially overcome the inherent re-transmission penalty of virtual beamforming. The re-transmission penalty arises from the need to re-transmit data or pilot symbols as the virtual beam is formed by combining measurements over from a single antenna UE. Moreover, we also present a model for combining virtual and physical antenna arrays in Section 6.4, which can help further reduce the re-transmission penalty. This work has been published in [AZH16].

6.1 Introduction

Virtual beamforming, as discussed in Section 5.3, enables the generation of a (virtual) Uniform Linear Array (ULA) at the User Equipment (UE) side, even for single (physical) antenna UEs [ZH13, AZH15a]. The idea behind it is to utilize the movement of the UE and store the $M_{\mathrm{VB}} \in \mathbb{N}$ adjacent received signal measurements, which can then be thought of as the signal received at the various antenna elements of an array. A weighting vector $\boldsymbol{\omega} = [\omega_1, \ldots, , \omega_{M_{\mathrm{VB}}}]^T \in \mathbb{C}^{M_{\mathrm{VB}} \times 1}$, is then applied to these measurements to generate the beamformed channel. The complex weight vector $\boldsymbol{\omega}$ should be chosen according to the number of virtual antenna elements, the relative distance between two measurement locations, and the desired application of virtual beamforming.

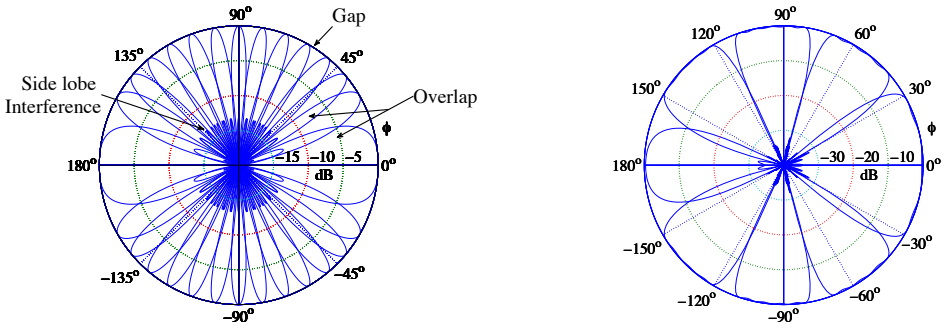
In the previous chapter, we have already established that virtual beams provide a better prediction performance with larger prediction horizons. However, this performance gain is only for a part of the channel in the beam direction. We essentially eliminate all the other parts of the channel that we cannot predict reliably. Hence, as we need information about the complete channel, the question arises, *how to incorporate the virtual beams for user data transmission?* There are two potential answers to this question:

- One way to go about this could be to predict or estimate the channel for one beam direction, then rotate the beam in another direction and predict or estimate the channel for that direction. We continue doing this until we have covered the complete 360° around the UE. We then combine these channels to get the complete predicted channel for the UE, and feed this channel information back to the base station for precoding and data transmission.
- Another way, and a potentially better solution, is to incorporate orthogonal coding to simultaneously employ multiple virtual beams for data transmission.

We will discuss both these solutions in detail in the following sections.

6.2 Combine Virtual Beams to Get Complete Channel

The simplest way to design the complex weights for virtual beamforming is to use a fixed phase shift ψ_0 between adjacent locations giving us the weighting vector with elements $\omega_{m_{\mathrm{VB}}}(\psi) = e^{-\jmath\psi_0 m_{\mathrm{VB}}}$. For a given ψ_0 (refer to equation (5.15)), we can calculate the complex-valued virtual beam pattern $A'(\psi)$ by utilizing equation (5.17).



(a) Uniformly weighted beams

(b) Sector beams of width 30° each

Figure 6.1: Multiple virtual beams rotated in various azimuthal directions are overlaid on top of each other to capture the complete channel. The idea is to showcase the inherent challenges and problems of this approach. The co-elevation angle as depicted in Figure 5.2b, is fixed at $\theta = 90°$.

A conventional omni-directional reference receiver will collect all the power, from all the multipath components and from all directions, while a narrow beam might receive only a small portion of that power. This allows us to divide the channel into parts which can be predicted with fewer errors and for a longer time because the beamformed channel is comprised of very few multipath components [AZH15a]. To estimate the complete channel, we need to rotate the beam in all directions to capture all multipath components and then combine them using a suitable combining strategy. A straightforward summation would be easy to implement, but ignores the side-lobes which will typically cause inter-beam interference. Additionally, there is overlap between the adjacent beams which would result in some multipath components contributing twice to the combined channel as shown in Figure 6.1a. There are gaps between adjacent beams as well due to the elliptical shape of typical beam patterns. Due to the difference in beamforming gain in these gaps, some multipath components will contribute with lower power as compared to others in the combined

channel. If we move the beams further apart to reduce the overlap, the size of the gaps would increase, which is highly undesired.

Ideally, we would like to have rectangular beams which could cover all 360° with minimal overlap and coverage gaps but it is not possible to design such beams. However, we can design sector beams which are based on the inverse discrete-space Fourier transforms of the array factor and are very close to rectangular beams as shown in Figure 6.1b [Orf16]. Such beams allow us to set very low stop-band attenuation values to minimize inter-beam interference from side-lobes. The sectors can also be designed to have minimal gaps and overlaps between adjacent beams. The downside is that such sector beams have wide lobes and thus the benefits of UE-sided beamforming, like a reduction in the number of channel components and multipath components, are not achieved properly.

From a hardware perspective, the UE needs to be capable of simultaneously forming the beams in all directions, predicting the channel for each individual beam and then combining them together at each time slot so that there is no additional time penalty to virtual beamforming. Another important factor to consider is the width of the individual beams to be combined to get the complete channel. Very narrow beams give us far better prediction performance [AZH15a, AZH15b], but more beams need to be predicted individually and then combined together resulting in higher complexity. Furthermore, in case of narrow beams, potentially a high number of virtual beams have to be combined so that estimation errors can be expected to add up to large error terms. Hence, it makes sense to find a solution that enables reliable and efficient use of the virtual beams directly for data transmission.

6.3 Orthogonally Coded Data Transmission over Virtual Beams

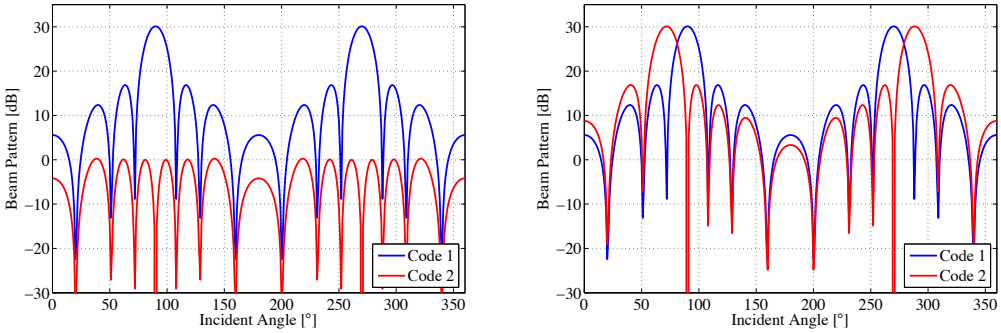
One of the biggest drawbacks of virtual beamforming is that it requires the retransmission of M_{VB} data symbols, resulting in a reduction in throughput, by a factor of M_{VB}. This means that there is a trade-off between prediction performance and data throughput. To counter this loss in throughput, we devise the use of multiple virtual beams segregated through orthogonal codes (for which a schematic is provided in Figure 6.3) [AZH16].

The main challenge for multiplexing of user data by coded transmission over virtual beams is that at the UE, the code is inevitably coupled with the beam shape. Often orthogonal codes have values which change sign from one value to the next, for example, $+1$ or -1 in Hadamard codes. Such codes can result in a complete reversal of the phase shift between adjacent antenna elements, which

results in undesired beam shapes and a loss of gain as shown in Figure 6.2a for an antenna spacing of $\lambda/10$. Here, we denote the first row of the Hadamard matrix of size 32 as *code* 1 (all ones code) and can observe that it has a desirable beam shape with the look direction at 90°. On the contrary, *code* 2, which is another row of the Hadamard matrix, has no preferred direction and cannot be used as a suitable beam for data transmission. One way to get around this problem is to use nearly orthogonal circular codes of the form,

$$\mathbf{c}_{n_{\mathrm{OC}}} = e^{-\jmath 2\pi \, n_{\mathrm{OC}} \, \mathbf{m}_{\mathrm{VB}}/M_{\mathrm{VB}}} \tag{6.1}$$

where, $n_{\mathrm{OC}} = 0, 1, \ldots, N_{\mathrm{OC}}$, with N_{OC} being the total number of codes or coded virtual beams, $\mathbf{m}_{\mathrm{VB}} = [1, 2, \ldots, M_{\mathrm{VB}}]^{\mathrm{T}}$, and M_{VB} being the total number of (virtual) antenna elements. This kind of circular code can give us usable beams, which are still fully orthogonal to each other, at least at the transmitter side, as shown in Figure 6.2b. Hence, we will use circular codes for our user data multiplexing scheme.



(a) Hadamard Code

(b) Circular Code

Figure 6.2: A comparison of beam patterns of two codes over a 32 element virtual array having a $\lambda/10$ spacing for different types of codes. Here, 0 dB represents the gain of an isotropic radiator.

6.3.1 User Data Multiplexing Using Two Orthogonal Codes

To explain the main idea, we first start with a very basic setup with a single antenna transmitter and a single antenna UE in an urban macro scenario. The UE moves on a straight line with a constant speed. The transmit symbols d_1 and d_2 are simultaneously transmitted repeatedly over M_{VB} time slots, where M_{VB} is the number of measurements that will be stored at the UE to create the virtual antenna elements. At each time slot, the transmit symbol d_1 is multiplied by the respective code bit from the code word $\mathbf{c}_1$ of length M_{VB}. Similarly, the transmit symbol d_2 is multiplied by the code word $\mathbf{c}_2$. As in an Orthogonal Frequency Division Multiplexing (OFDM) system, the sub-carriers are orthogonal, hence they can carry

separate transmit symbols and the following process can be separately applied to each of the sub-carriers. The transmission vector of length $M_{\mathrm{VB}} \times 1$ is then given as,

$$\mathbf{x} = \mathbf{c}_1\, d_1 + \mathbf{c}_2\, d_2. \tag{6.2}$$

A simple schematic is provided in Figure 6.3, where the green link is the actual location of the UE and the gray dashed lines show the past locations and stored measurements for the UE over time.

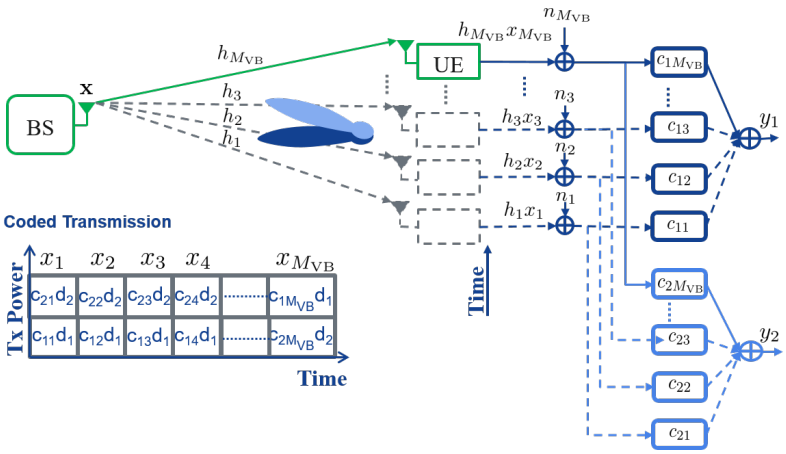


Figure 6.3: A simple schematic of the coded data transmission and reception over virtual beams using two orthogonal codes

When the UE has moved for M_{VB} time slots, we can combine the measurements from these locations to generate the virtual beams. As the codes are coupled with the beams, we use the codes themselves as the receive-beamforming weights, hence $\boldsymbol{\omega}_1^{\mathrm{H}} = \mathbf{c}_1^{\mathrm{H}}$ and $\boldsymbol{\omega}_2^{\mathrm{H}} = \mathbf{c}_2^{\mathrm{H}}$ (Figure 6.2b). The received beam signal for code 1 is given as,

$$\begin{aligned} y_1 &= \mathbf{c}_1^{\mathrm{H}}(\mathrm{diag}(\mathbf{h})\mathbf{x} + \mathbf{n}) \\ &= \mathbf{c}_1^{\mathrm{H}}(\mathrm{diag}(\mathbf{h})(\mathbf{c}_1\, d_1 + \mathbf{c}_2\, d_2) + \mathbf{n}) \\ &= \mathbf{c}_1^{\mathrm{H}}(\mathrm{diag}(\mathbf{h})\mathbf{c}_1\, d_1 + \mathrm{diag}(\mathbf{h})\mathbf{c}_2\, d_2 + \mathbf{n}) \\ &= \underbrace{\mathbf{c}_1^{\mathrm{H}}(\mathrm{diag}(\mathbf{h})\mathbf{c}_1)}_{A}\, d_1 + \underbrace{\mathbf{c}_1^{\mathrm{H}}(\mathrm{diag}(\mathbf{h})\mathbf{c}_2)}_{XT_{12}}\, d_2 + \underbrace{\mathbf{c}_1^{\mathrm{H}}\mathbf{n}}_{n_1} \end{aligned} \tag{6.3}$$

where, $\mathbf{h} \in \mathbb{C}^{M_{\mathrm{VB}} \times 1}$ is the vector containing the time-variant channel coefficients of the channel transfer function $H(f,t)$ (refer to the Bello functions shown in Figure 3.2) for one particular sub-carrier. Similarly, the beam for the second code is

given as,

$$\begin{aligned} y_2 &= \mathbf{c}_2^{\mathrm{H}}(\mathrm{diag}(\mathbf{h})\mathbf{x} + \mathbf{n}) \\ &= \mathbf{c}_2^{\mathrm{H}}(\mathrm{diag}(\mathbf{h})(\mathbf{c}_1\, d_1 + \mathbf{c}_2\, d_2) + \mathbf{n}) \\ &= \mathbf{c}_2^{\mathrm{H}}(\mathrm{diag}(\mathbf{h})\mathbf{c}_1\, d_1 + \mathrm{diag}(\mathbf{h})\mathbf{c}_2\, d_2 + \mathbf{n}) \\ &= \underbrace{\mathbf{c}_2^{\mathrm{H}}(\mathrm{diag}(\mathbf{h})\mathbf{c}_1)}_{XT_{21}}\, d_1 + \underbrace{\mathbf{c}_2^{\mathrm{H}}(\mathrm{diag}(\mathbf{h})\mathbf{c}_2)}_{B}\, d_2 + \underbrace{\mathbf{c}_2^{\mathrm{H}}\mathbf{n}}_{n_2}. \end{aligned} \tag{6.4}$$

where A and B are the desired signal parts whereas XT_{12} and XT_{21} are the sum cross-talk levels, respectively, for data symbols d_1 and d_2, and n_1 and n_2 are the corresponding noise samples.

We observe that the individual additional noise terms in each of the measurements are added together and amplified by the beamformer. For Independent Identically Distributed (IID) noise values, this leads to the same overall noise power as known for a physical antenna array. But, there is a penalty for the virtual versus the physical antenna array, as in the case of virtual beamforming, the received-power per time slot is limited to that of the physically available antenna elements. So for the user data multiplexing, virtual beamforming provides no to very little beamforming Signal-to-Noise Ratio (SNR) gains.

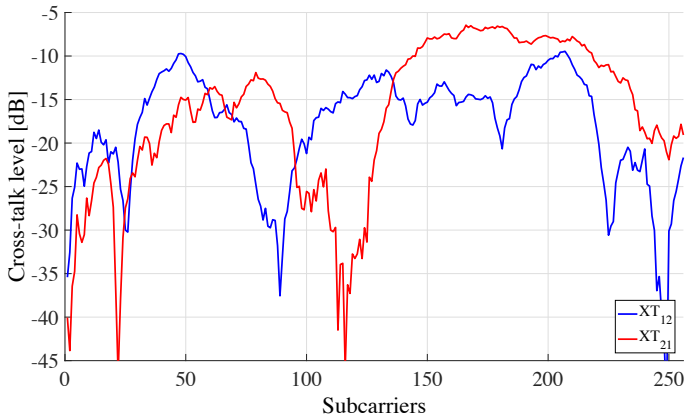


Figure 6.4: Sum cross-talk levels XT_{12} and XT_{21} over sub-carriers using codes of length 32 for a measured channel.

The codes are orthogonal at the transmitter but after passing through the time-variant channel, they are no longer perfectly orthogonal at the receiver of the UE. This leads to the cross-talk terms XT_{12} and XT_{21}. These sum cross-talk terms can be calculated as the channel $\mathbf{h}$ is known because of prior channel estimation and/or channel prediction performed at the UE side. In contrast, the base station transmitter knows only the reported Channel State Information (CSI) per virtually beamformed radio channel and therefore cannot estimate this cross-talk level.

If the cross-talk level is low, we can formulate a system of linear equations as,

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} A & XT_{12} \\ XT_{21} & B \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \end{bmatrix}. \tag{6.5}$$

It can be solved for the desired symbols d_1 and d_2 through a Least Squares (LS) estimation or if the second-order statistics of the noise are known, through a Minimum Mean Square Error (MMSE) estimation.

Figure 6.4 shows the sum cross-talk levels for an urban macro measured channel with noise. The channel measurements were acquired at the Nokia campus in Munich as described in Section 3.4.1.2. A sum cross-talk level of 0 dB means that the codes became non-orthogonal after passing through the channel and the undesired symbol in a certain receive beam has the same magnitude as the desired one. Hence, the desired data symbol cannot be resolved properly from the received signal. We can see that cross-talk levels below -15 dB or even -20 dB can be achieved, which would insure reliable communication over the coded virtual beams. The cross-talk goes above -10 dB for some of the sub-carriers. However, through proper user scheduling, the areas with high cross-talk levels can be avoided. This way every UE might be able to constantly remain below the minimum cross-talk level threshold most of the time.

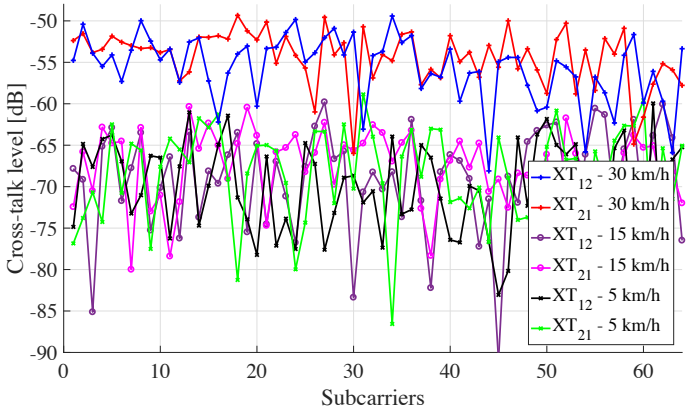


Figure 6.5: Sum cross-talk level XT_{12} and XT_{21} over sub-carriers using codes of length 32 for a ray-traced IlmProp channel. The various curves correspond to different velocities of the UE.

In order to assess the affect of channel variation on cross-talk levels, we have plotted the cross-talk levels for varying UE velocities in Figure 6.5. The SNR is 0 dB. This time we have used the ray-traced IlmProp channel from the geometrical scenario described in Section 3.4.1.1. To analyze the effect of varying velocity with measured channels is not possible as it would require a new measurement campaign. On the other hand, the IlmProp allows us to vary the velocity of the UE with ease;

a feature which is not present (at the time of writing) in WinProp ray-tracing tool. However, it should be noted that the overall cross-talk levels are far lower as compared to our measured channel because the IlmProp scenario is very simple compared to the Nokia campus scenario. We can see that for pedestrian velocities up to 15 km/h, the cross-talk levels are fairly similar. However, as the velocity increases, the cross-talk level gets worse as well. This means that there is a direct correlation between the variations in the channel and the cross-talk level.

6.3.2 User Data Multiplexing Using Four Orthogonal Codes

We can now expand the same setup of a single antenna transmitter and a single antenna UE in an urban macro scenario to four orthogonal codes. Therefore, we can transmit 4 data symbols over M_{VB} time slots. This increases the efficiency from $2/M_{\mathrm{VB}}$ to $4/M_{\mathrm{VB}}$. We again consider a single sub-carrier. The transmission vector of length $M_{\mathrm{VB}} \times 1$ is given as,

$$\mathbf{x} = \mathbf{c}_1\, d_1 + \mathbf{c}_2\, d_2 + \mathbf{c}_3\, d_3 + \mathbf{c}_4\, d_4. \tag{6.6}$$

Similar to the case with two codes, when the UE has moved for M_{VB} time slots, we combine the measurements from these locations to generate the virtual beams. The received beam signal for code 1 is given as,

$$\begin{aligned}
y_1 &= \mathbf{c}_1^{\mathrm{H}}(\mathrm{diag}(\mathbf{h})\mathbf{x} + \mathbf{n}) \\
&= \mathbf{c}_1^{\mathrm{H}}(\mathrm{diag}(\mathbf{h})(\mathbf{c}_1\, d_1 + \mathbf{c}_2\, d_2 + \mathbf{c}_3\, d_3 + \mathbf{c}_4\, d_4) + \mathbf{n}) \\
&= \mathbf{c}_1^{\mathrm{H}}(\mathrm{diag}(\mathbf{h})\mathbf{c}_1\, d_1 + \mathrm{diag}(\mathbf{h})\mathbf{c}_2\, d_2 \\
&\quad + \mathrm{diag}(\mathbf{h})\mathbf{c}_3\, d_3 + \mathrm{diag}(\mathbf{h})\mathbf{c}_4\, d_4 + \mathbf{n}) \\
&= \underbrace{\mathbf{c}_1^{\mathrm{H}}(\mathrm{diag}(\mathbf{h})\mathbf{c}_1)}_{A}\, d_1 + \underbrace{\mathbf{c}_1^{\mathrm{H}}(\mathrm{diag}(\mathbf{h})\mathbf{c}_2)}_{XT_{12}}\, d_2 \\
&\quad + \underbrace{\mathbf{c}_1^{\mathrm{H}}(\mathrm{diag}(\mathbf{h})\mathbf{c}_3)}_{XT_{13}}\, d_3 + \underbrace{\mathbf{c}_1^{\mathrm{H}}(\mathrm{diag}(\mathbf{h})\mathbf{c}_4)}_{XT_{14}}\, d_4 + \underbrace{\mathbf{c}_1^{\mathrm{H}}\mathbf{n}}_{n_1}
\end{aligned} \tag{6.7}$$

Similarly, the received beams for the other three codes are given as,

$$\begin{aligned}
y_2 &= \underbrace{\mathbf{c}_2^{\mathrm{H}}(\mathrm{diag}(\mathbf{h})\mathbf{c}_1)}_{XT_{21}}\, d_1 + \underbrace{\mathbf{c}_2^{\mathrm{H}}(\mathrm{diag}(\mathbf{h})\mathbf{c}_2)}_{B}\, d_2 \\
&\quad + \underbrace{\mathbf{c}_2^{\mathrm{H}}(\mathrm{diag}(\mathbf{h})\mathbf{c}_3)}_{XT_{23}}\, d_3 + \underbrace{\mathbf{c}_2^{\mathrm{H}}(\mathrm{diag}(\mathbf{h})\mathbf{c}_4)}_{XT_{24}}\, d_4 + \underbrace{\mathbf{c}_2^{\mathrm{H}}\mathbf{n}}_{n_2}
\end{aligned} \tag{6.8}$$

$$y_3 = \underbrace{\mathbf{c}_3^{\mathrm{H}}(\mathrm{diag}(\mathbf{h})\mathbf{c}_1)}_{XT_{31}}\, d_1 + \underbrace{\mathbf{c}_3^{\mathrm{H}}(\mathrm{diag}(\mathbf{h})\mathbf{c}_2)}_{XT_{32}}\, d_2 + \underbrace{\mathbf{c}_3^{\mathrm{H}}(\mathrm{diag}(\mathbf{h})\mathbf{c}_3)}_{C}\, d_3 + \underbrace{\mathbf{c}_3^{\mathrm{H}}(\mathrm{diag}(\mathbf{h})\mathbf{c}_4)}_{XT_{34}}\, d_4 + \underbrace{\mathbf{c}_3^{\mathrm{H}}\mathbf{n}}_{n_3} \tag{6.9}$$

$$y_4 = \underbrace{\mathbf{c}_4^{\mathrm{H}}(\mathrm{diag}(\mathbf{h})\mathbf{c}_1)}_{XT_{41}}\, d_1 + \underbrace{\mathbf{c}_4^{\mathrm{H}}(\mathrm{diag}(\mathbf{h})\mathbf{c}_2)}_{XT_{42}}\, d_2 + \underbrace{\mathbf{c}_4^{\mathrm{H}}(\mathrm{diag}(\mathbf{h})\mathbf{c}_3)}_{XT_{43}}\, d_3 + \underbrace{\mathbf{c}_4^{\mathrm{H}}(\mathrm{diag}(\mathbf{h})\mathbf{c}_4)}_{D}\, d_4 + \underbrace{\mathbf{c}_4^{\mathrm{H}}\mathbf{n}}_{n_4} \tag{6.10}$$

where A, B, C and D are the desired signal parts and now each code has cross-talk from the other three codes. Once again, the sum cross-talk levels at the UE, can be calculated as the channel $\mathbf{h}$ is known at the UE receiver. If the cross-talk level is low, we can formulate a system of linear equations,

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} A & XT_{12} & XT_{13} & XT_{14} \\ XT_{21} & B & XT_{23} & XT_{24} \\ XT_{31} & XT_{32} & C & XT_{34} \\ XT_{41} & XT_{42} & XT_{43} & D \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \\ n_3 \\ n_4 \end{bmatrix} \tag{6.11}$$

which can be solved, similarly to the two-code case, for the desired data symbols d_1, d_2, d_3 and d_4 through least-squares or MMSE estimation.

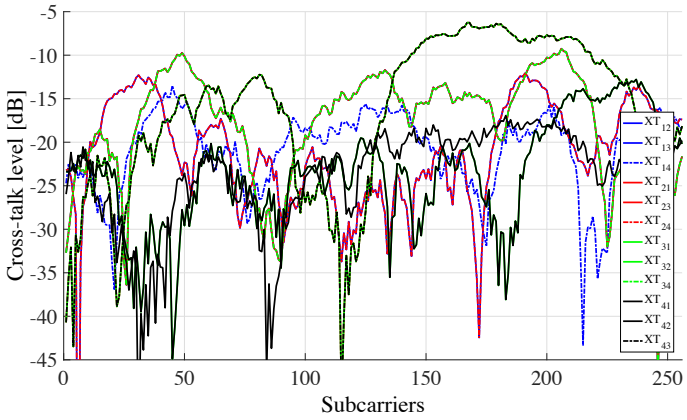


Figure 6.6: Sum cross-talk level over sub-carriers between 4 codes of length 32 for a measured channel.

Figure 6.6 shows the sum cross-talk level for four coded multiplexed streams for an urban macro measured channel with noise. We can see a similar trend as in the

two code case, where a cross-talk level below −20 dB can be achieved quite easily by properly adapting the beams to the sub-carriers.

6.4 Orthogonally Coded Data Multiplexing over Combined Beams

A virtual antenna array of size 32, as evaluated above for a single antenna UE, would need the multiplexing of 32 data streams to achieve the same efficiency as a physical antenna array of the same size. An extension of the coded data multiplexing scheme, from 4 parallel streams to 16 or 32 orthogonal streams, is obviously challenging. In addition, beamforming SNR gains are lower or even zero for virtual arrays (when compared to physical arrays), as the transmit power has to be shared between the codes. The user data transmission, therefore, mainly benefits from the improved suppression of the interferer and the larger channel prediction horizon. Note, as the Channel State Information Reference Signals (CSI-RSs) do not have to share codes, but repeat the same signal multiple times, channel estimation itself profits also from SNR gains.

The above mentioned issues suggest combining the benefits of the two worlds by enhancing a limited set of physical UE antennas with virtual beamforming. For example, an 8 antenna UE would require only 4 retransmissions to form a combined physical-virtual 32 element antenna array, which additionally can achieve a 9 dB SNR gain. In this case, full efficiency is already achieved by the multiplexing of 4 data streams over four orthogonal codes of length four, transmitted in four time slots.

Instead of transmitting four streams to one UE, one might alternatively serve four streams to four different UEs which are spatially located far apart from each other. This would relax the fixed coupling of code and beamforming weights. In that case, the streams will be separated by Multi-User MIMO (MU-MIMO) precoding instead of the orthogonal transmit-codes. In practice, the number of active UEs might often be limited and a suitable mode adaptation might be needed.

In the following, we will investigate in more detail another potential setup, which combines low complexity with large virtual beamforming gains. It is based on the observation that the 32 virtual antenna elements array discussed so far, uses a very low virtual antenna spacing of just 1 cm (equal to $\lambda/10$ for the given Radio Frequency (RF) frequency of 2.6 GHz). The channel prediction quality is mainly affected by the Half-Power BeamWidth (HPBW) of the virtual beamformer, which itself depends on the aperture of the virtual array. This is defined by the distance between the outer most antenna elements and will be unchanged by down-sampling

of antenna elements, for example, to a $\lambda/2$-spacing.

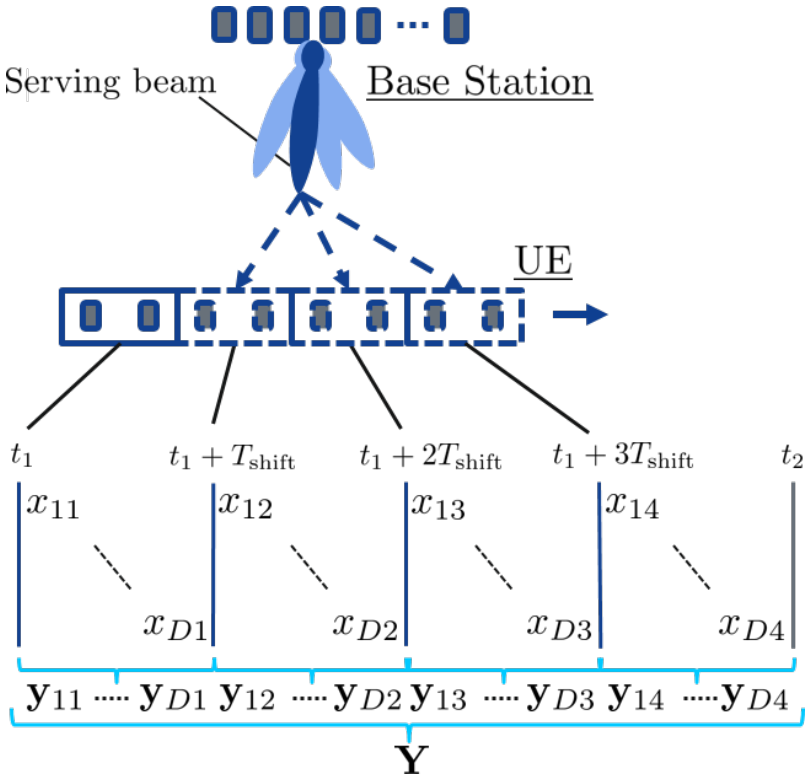


Figure 6.7: Order of data transmission over one base station beam to a combined 8 element antenna, consisting of 2 physical UE antennas and 4 virtual steps

An exemplary setup is given in Figure 6.7 for a 2-antenna element UE plus *four* virtual time steps forming a *virtual eight element array* with $\lambda/2$-spacing. The figure additionally illustrates the overall timing diagram of the proposed concept. The base station transmits, for example, over one beam of a fixed Grid of Beams (GoB) the complex scalar transmit-signal x_{ij}, comprising the elements of the four orthogonal codes c_{1j} to c_{4j} corresponding to the four virtual beams and the data symbols d^i_{1j} to d^i_{4j}. It is given as,

$$x_{ij} = [c_{1j}\ c_{2j}\ c_{3j}\ c_{4j}][d^i_{1j}\ d^i_{2j}\ d^i_{3j}\ d^i_{4j}]^{\mathrm{T}} \tag{6.12}$$

where,

- The second subscript j is related to the time instant under consideration, i.e., t_1, $t_1 + T_{\text{shift}}$, $t_1 + 2T_{\text{shift}}$ or $t_1 + 3T_{\text{shift}}$, where T_{shift} is the fixed time split between different positions of the UE which is moving at a constant velocity.
- The first subscript i corresponds to the superscript in the transmit symbols and indicates the selected data symbols from the block of data for the parallel data streams d_{1j} to d_{4j}. The length of the data blocks denoted by D has to

be adapted to the time T_{shift} and the symbol timing so that a continuous data transmission can be achieved. Correspondingly, the transmission of $x_{(i+1)j}$ compared to x_{ij} is shifted by the timing of one symbol t_{symbol}.

For the given example, the UE receives x_{ij} with the two receive antennas leading to the received vector,

$$\mathbf{y}_{ij} = [h_{1j}x_{ij} \quad h_{2j}x_{ij}]^{\mathrm{T}} + \mathbf{n}_j. \tag{6.13}$$

where x_{ij} is affected by the corresponding channel coefficients for the different antenna elements and the received noise vector $\mathbf{n}_j$. The UE stores all the received vectors $\mathbf{y}_{ij}$ into the matrix $\mathbf{Y}$ to apply, after full reception, the combined beamformers to all the data sets.

Let us assume a UE moving with a speed of 43 km/h, then one has to set T_{shift} to a value of 10 ms to achieve a λ spacing between two transmission steps, or equivalently, the $\lambda/2$ spaced combined virtual antenna array of Figure 6.7. For this scheme, the HPBW would be just 6 degrees compared to 17 degrees for the 32 element virtual antenna array.

6.5 Summary

In this chapter, we discussed the challenges posed by virtual beamforming for data transmission, that is, the re-transmission penalty which causes a loss of throughput. We have devised an orthogonal coding based scheme to reduce the effect of the re-transmission penalty by forming more than one orthogonal virtual beams. We have also proposed the potential extension of the scheme for multiple antenna UEs.

Chapter 7

Conclusions and Outlook

7.1 Conclusions

The results presented in this thesis hold significance to cellular system designers as we move towards the standardization and implementation of Coordinated Multi-Point (CoMP) and massive MIMO in Fifth Generation (5G) and beyond 5G networks. We have presented a framework that supports an effective implementation of massive MIMO systems in Frequency Division Duplex (FDD) operation.

A detailed description of the proposed Interference Mitigation Framework - Advanced (IMF-A) and its evolution for 5G, which additionally supports massive MIMO and small cells, is presented in Chapter 2. The IMF-A was designed with the goal to mitigate interference between cooperation areas. This enables the development of high-performing advanced transmission and reception schemes focusing on a single cooperation area. However, the inclusion of massive MIMO is challenging for FDD systems as the scheme is, in essence, defined for Time Division Duplex (TDD) systems. In TDD systems, the increasing number of antennas leads to *channel hardening*, which essentially means that the fading channels behave as deterministic channels because the channel fluctuations average out over the antennas. This means that there is no need to transmit reference signals and estimate the channel as the current realization of the channel is close to its average value. This is not the case for FDD systems and the channel needs to be estimated for each of the antenna elements. For pilot-assisted channel estimation, this means that a huge number of pilot or reference signals needs to be transmitted, taking away precious resources. As a solution to the exploding reference signal overhead for FDD massive MIMO systems, we have proposed to use the Grid of Beams (GoB) concept, which enables us to squeeze the large number of massive MIMO antennas into a few effective beams. These beams are spread out to cover the entire sector, similar to a traditional Chinese fan, with minimal overlap. The benefit is that these beams comprise the effective channel components and the User Equipments (UEs) only

need to estimate the few strongest channel components, resulting in a significant reduction in Channel State Information (CSI) reporting overhead.

The underlying challenge for a practical and efficient implementation of both CoMP and massive MIMO, is the large number of channel components and the huge number of multipath components present within each channel component. Channel prediction has been identified as the scheme which could effectively resolve this issue. Channel prediction algorithms available in the literature (an overview is provided in Section 4.2), whether parametric-model based or channel frequency response based, directly utilize the channel knowledge available in its respective format, to predict the channel in the future. This imposes an upper bound on the prediction performance, which is usually worse than the required performance in terms of the Normalized Mean Square Error (NMSE) for a reliable communication that is typically dictated by the physical conditions and geometry of the environment within which the base station and the UE reside. As a part of this thesis, we have analyzed and detailed two state-of-the-art channel prediction algorithms in Section 4.3 and Section 4.4. In Section 4.5, we have devised a scheme which combines the best of both state-of-the-art algorithms and works very well, specifically in the low Signal-to-Noise Ratio (SNR) region. The basic idea is to perform a Higher Order Singular Value Decomposition (HOSVD) on the channel tensor and mitigate the noise by means of a low-rank approximation prior to predicting the channel with the Kalman filter. This gives, for our particular example scenario, a 7 dB prediction NMSE gain as compared to just the Kalman filter-based prediction scheme. Another benefit is that, in the case of high SNR, the HOSVD step can simply be omitted to lower the complexity of the overall system.

The most significant contribution comes in the form of a method which enables UEs to form narrow beams, as if the UEs were equipped with massive MIMO arrays, even if the UEs only have a single physical antenna element. This method is called virtual beamforming and is introduced in Chapter 5. The general idea is that we exploit the movement of the UE and store the received signals for several time slots. These measurements are then perceived as if they were received by the individual elements of a physical antenna array. The drawbacks are obvious. For example, as we need to transmit the same signal over the time slots forming the virtual beam, we have to wait for the time it takes to gather the measurements. Therefore, we need good time synchronization and the UE is moving at ideally a constant speed which we either know or can estimate. On the other hand, if applied appropriately, the benefits can be enormous as well. For example, virtual beamforming allows us to simplify or reduce the physical limitations of the channel prediction problem itself by effectively reducing the number of multipath components in each channel component. This enables better prediction performance and a reduction in the CSI

reporting overhead in FDD systems.

Lastly, in Chapter 6 we have proposed a solution to transmit user data over the virtual beams. Virtual beamforming directly applied to user data can be very inefficient as the data symbols need to be transmitted multiple times. To overcome this challenge we have proposed parallel transmission over a set of coded virtual beams. We have shown that many parallel coded beams can be supported reliably, which allows us to downgrade the re-transmission penalty of data symbols, inherent in virtual beamforming systems. We have also proposed a method to combine an arbitrary number of physical and virtual antenna elements to form virtual massive MIMO arrays. The benefit of having more physical antennas is that fewer re-transmissions are required to form an overall large effective massive MIMO array, giving a better throughput.

7.2 Future Work

For future work, it would be interesting to evaluate other schemes to form even narrower beams as compared to virtual beamforming. The challenge is that the Half-Power BeamWidth (HPBW) of a Uniform Linear Array (ULA) is dictated mainly by its aperture, and partially by the number of antenna elements and the spacing between them. Thus, using linear methods, we can only achieve the HPBW provided by virtual beamforming or any other beamforming scheme utilizing physical antennas. For this reason, one has to look for other ways to achieve narrower beams. We have already proposed two ways to achieve an even narrower HPBW.

First, we have proposed a scheme to artificially induce similar effects as physical mutual coupling in virtual antennas [ZA16a, AZH15a]. It has been known in the literature for quite some time and was recently shown by [IN14] that mutual coupling can induce super-directivity. This results in extremely slender beams with very low HPBW. The problem is that it requires huge input currents to achieve super-directivity in physical antennas and results in large Ohmic losses. We do not have this problem with virtual antennas and artificially induced mutual coupling, as it is basically post-processing of the received signals.

Second, we have proposed a novel UE-sided nonlinear processing scheme which also achieves super-directivity like effects [ZA16b]. In this scheme, we use two separate beamformers and multiply their respective beamformed received signals together, either in the frequency domain or in the time domain. This gives us the equivalent of very narrow beams, in terms of the reduction in the number of multipath components in a given direction. The main caveat though is that nonlinear processing comes with many challenges and it is not trivial to deal with,

for example, it would require a proper modification of the transmit symbols as well. We have already analyzed the benefits it has on multipath component and feedback reduction, both on the link-level and on the system-level. However, the main mathematical aspects like the SNR of the resulting received signal, its higher-order statistics and decoding of the desired symbols still require further study. For this reason, both of these schemes have been excluded from this thesis.

In the analysis presented in this thesis, we have only considered a single active UE at a time and looked at the performance improvements for various parameters only at the link-level. For implementation purposes and practical adoption of schemes like virtual beamforming, it would be essential to have a look at the overall system-level performance, at least within a single cooperation area. To this end, our schemes have been implemented in Nokia's 5G system-level simulator called AMoRE (Advanced Mobile Radio Realtime Experience), which analyses and showcases the latest advancements in 5G.

Appendix A

Definitions

n-mode product of a Tensor with a Matrix

The general description of an n-mode product $\boldsymbol{\mathcal{H}} \times_n \mathbf{U}_n$ of a tensor $\boldsymbol{\mathcal{H}} \in \mathbb{C}^{I_1 \times I_2 \times \ldots \times I_N}$ with a matrix $\mathbf{U} \in \mathbb{C}^{J_n \times I_n}$ is an $(I_1 \times I_2 \times \ldots \times I_{n-1} \times J_n \times I_{n+1} \times \ldots \times I_N)$ tensor whose elements are calculated as,

$$(\boldsymbol{\mathcal{H}} \times_n \mathbf{U})_{i_1,\ldots,i_{n-1},j_n,i_{n+1},\ldots,i_N} = \sum_{i_n} \boldsymbol{\mathcal{H}}_{i_1,i_2,\ldots,i_N} \mathbf{u}_{j_n,i_n} \tag{A.1}$$

where the subscripts $i_1, i_2, \ldots, i_N$ denote the indices of the corresponding tensor dimensions.

Properties of the Core Tensor

The two significant properties of the core tensor $\boldsymbol{\mathcal{S}}$ are as follows:

Orthogonality: if $a \neq b$, two sub-tensors $\boldsymbol{\mathcal{S}}_{i_n=a}$ (the n-th dimension is set to the fixed value a) and $\boldsymbol{\mathcal{S}}_{i_n=b}$ are orthogonal for all possible values of n, a, and b.

Ordering: the n-mode singular values of $\boldsymbol{\mathcal{H}}$ denoted as $\sigma_p^{(n)}$ are defined as the higher-order norms of the core tensor $\|\boldsymbol{\mathcal{S}}_{i_n=p}\|_{\mathrm{H}}$ where $p = 1, 2, \ldots I_N$. They exhibit the following property for all values of n:
$\|\boldsymbol{\mathcal{S}}_{i_n=1}\|_{\mathrm{H}} \geq \|\boldsymbol{\mathcal{S}}_{i_n=2}\|_{\mathrm{H}} \geq \ldots \|\boldsymbol{\mathcal{S}}_{i_n=I_N}\|_{\mathrm{H}} \geq 0$.

List of Abbreviations

3D	3-Dimensional
3GPP	3rd Generation Partnership Project
4G	Fourth Generation
5G	Fifth Generation
ACF	AutoCorrelation Function
AoA	Angle of Arrival
AoD	Angle of Departure
AR	AutoRegressive
AWGN	Additive White Gaussian Noise
CDMA	Code Division Multiple Access
CIR	Channel Impulse Response
CO-MIMO	COoperative MIMO
CoMP	Coordinated MultiPoint
CRB	Cramer Rao Bound
CSI	Channel State Information
CSI-RS	Channel State Information Reference Signal
CSIT	Channel State Information at the Transmitter
CTF	Channel Transfer Function
DCS	Dynamic Cell Selection
DoFs	Degrees of Freedom
ESPRIT	Estimation of Signal Parameters via Rotational Invariance Techniques
FBMC	Filter Bank Multi-Carrier
FD	Full-Duplex
FDD	Frequency Division Duplex
FDM	Frequency Division Multiplexing
FDTD	Finite Difference Time Domain

FEM	Finite Element Method
FFT	Fast Fourier Transform
GoB	Grid of Beams
GTD	Geometrical Theory of Diffraction
HetNet	Heterogeneous Network
HOSVD	Higher Order Singular Value Decomposition
HPBW	Half-Power BeamWidth
ICIC	Inter-Cell Interference Coordination
ICN	Information Centric Networking
IFFT	Inverse Fast Fourier Transform
IID	Independent Identically Distributed
IMF-A	Interference Mitigation Framework - Advanced
IoT	Internet of Things
IP	Internet Protocol
ISI	Inter-Symbol Interference
JP	Joint Precoding
JSDM	Joint Spatial Division and Multiplexing
JT	Joint Transmission
JT-CoMP	Joint Transmission Coordinated MultiPoint
KPIs	Key Performance Indicators
LMS	Least Mean Squares
LOS	Line-Of-Sight
LS	Least Squares
LTE	Long Term Evolution
LTE-A	Long Term Evolution - Advanced
MBB	Mobile BroadBand
MCS	Modulation and Coding Scheme
MIMO	Multiple-Input and Multiple-Output
MISO	Multiple-Input and Single-Output
ML	Maximum-Likelihood
MMSE	Minimum Mean Square Error
mmWave	millimeter Wave
MoM	Method of Moments
MSE	Mean Square Error
MU-MIMO	Multi-User MIMO

MUSIC	MUltiple SIgnal Classifier
NLOS	Non-Line-Of-Sight
NMSE	Normalized Mean Square Error
OFDM	Orthogonal Frequency Division Multiplexing
OP-CoMP	OPportunistic CoMP
PAPR	Peak-to-Average-Power Ratio
PSAM	Pilot Symbol Assisted Modulation
QAM	Quadrature Amplitude Modulation
RE	Resource Element
RF	Radio Frequency
RLS	Recursive Least Squares
RMS	Root Mean Square
SDMA	Spatial Division Multiple Access
SINR	Signal-to-Interference-plus-Noise Ratio
SISO	Single-Input and Single-Output
SNR	Signal-to-Noise Ratio
SON	Self-Organizing Networks
SU-MIMO	Single-User MIMO
SVD	Singular Value Decomposition
TDD	Time Division Duplex
TDMA	Time Division Multiple Access
UDN	Ultra Dense Network
UE	User Equipment
ULA	Uniform Linear Array
URA	Uniform Rectangular Array
WSS	Wide Sense Stationary
ZF	Zero Forcing

List of Symbols

a, b, c	Scalers
$\mathbf{a}, \mathbf{b}, \mathbf{c}$	Column Vectors
$\mathbf{A}, \mathbf{B}, \mathbf{C}$	Matrices
$\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{B}}, \boldsymbol{\mathcal{C}}$	Tensors
$h, \mathbf{h}, \mathbf{H}, \boldsymbol{\mathcal{H}}$	Channel coefficient, vector, matrix and tensor
N_C	Number of Cells
M_T	Number of Transmit antennas
M_R	Number of Receive antennas
M_{GoB}	Number of Grid of Beams
N_U	Number of served Users
N_{SC}	Number of Sub-Carriers
N_t	Number of time slots
M_{VB}	Number of Virtual Beamforming antennas
N_{RS}	Number of Reference Signals
N_L	Number of Paths/Taps
$\mathbf{y}$	Received Signal Vector
$\mathbf{W}$	MU-MIMO Precoding Matrix
$\mathbf{V}$	GoB Precoding Matrix
$\mathbf{x}$	Data Symbol Vector
$\mathbf{n}$	Noise Vector in Time Domain
$\mathbf{z}$	Noise Vector in Frequency Domain
τ	Delay
α	Complex Amplitude
ρ	Signal-to-Noise Ratio of the Received Signal
$(\Delta t)_C$	Coherence Time
$(\Delta f)_C$	Coherence Bandwidth
f_D	Doppler Frequency

f_C	Center Frequency of the System
$\mathbf{F}$	Fourier Matrix
$\mathbf{U}$	Matrix Unfoldings of the Channel Tensor
$\boldsymbol{\mathcal{S}}$	Core Tensor of the HOSVD of Channel Tensor

References

[3GP09] 3GPP, "Evolved universal terrestrial radio access (E-UTRA); physical channels and modulation," 3GPP TR 36.211 - Release 8, Tech. Rep., 2009. 6

[3GP11] ——, "Evolved universal terrestrial radio access (E-UTRA); downlink multiple input multiple output (MIMO) enhancement for LTE-advanced," 3GPP TR 38.871 - Release 11, Tech. Rep., 2011. 93

[3GP17] ——, "Study on scenarios and requirements for next generation access technologies," 3GPP TR 38.913 - Release 14, Tech. Rep., 2017. 4, 57

[AASS+14] A. Adhikary, E. Al Safadi, M. K. Samimi, R. Wang, G. Caire, T. S. Rappaport, and A. F. Molisch, "Joint spatial division and multiplexing for mm-wave channels," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1239–1255, 2014. 8, 9

[ABC+14] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, June 2014. 4

[Ade15] R. O. Adeogun, "Channel prediction for mobile MIMO wireless communication systems," Ph.D. dissertation, Victoria University of Wellington, New Zealand, 2015. 59

[AJJF99] J. B. Andersen, J. Jensen, S. H. Jensen, and F. Frederiksen, "Prediction of future fading based on past measurements," in *Proceedings of IEEE VTS 50th Vehicular Technology Conference, VTC Fall 1999.*, vol. 1, 1999, pp. 151–155 vol.1. 59

[ANAC13] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing - the large-scale array regime," *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6441–6463, 2013. 9

[Aro11] D. Aronsson, "Channel estimation and prediction for MIMO OFDM systems: Key design and performance aspects of Kalman based algorithms," Ph.D. dissertation, Uppsala University, Sweden, 2011. 11, 42, 44, 58, 59, 61, 62, 65, 67

[AS14] R. Apelfrojd and M. Sternad, "Design and measurement-based evaluations of coherent JT-CoMP: a study of precoding, user grouping and resource allocation using predicted CSI," *EURASIP Journal on Wireless Communications and Networking*, 2014. 56, 61

[AZH15a] M. B. Amin, W. Zirwas, and M. Haardt, "Advanced channel prediction concepts for 5G radio systems," in *Proceedings of the Twelfth International Symposium on Wireless Communication Systems (ISWCS 15)*, 2015. 12, 67, 82, 83, 92, 101, 104, 105, 106, 119

[AZH15b] ——, "Design options in the context of 5G for a powerful channel prediction," in *Proceedings of the 35th Wireless World Research Forum (WWRF 35)*, 2015. 10, 12, 67, 82, 83, 106

[AZH16] ——, "Virtual massive MIMO beamforming gains for 5G user terminals," in *Proceedings of IEEE Vehicular Technology Conference (VTC 2016 Fall)*, 2016. 12, 104, 106

[AZH17] ——, "HOSVD-based denoising for improved channel prediction of weak massive MIMO channels," in *Proceedings of IEEE Vehicular Technology Conference (VTC 2017 Spring)*, 2017. 12, 56, 75, 76

[BA15] P. Baracca and D. Aziz, "METIS deliverable 3.3 - final performance results and consolidated view on the most promising multi-node/multi-antenna transmission technologies," The METIS 2020 Project, Tech. Rep., 2015. [Online]. Available: https://www.metis2020.com/wp-content/uploads/deliverables/METIS_D3.3_v1.pdf 10, 16, 17, 21, 24, 56

[BB05] K. E. Baddour and N. C. Beaulieu, "Autoregressive modeling for fading channel simulation," *IEEE Transactions on Wireless Communications*, vol. 4, no. 4, pp. 1650–1662, 2005. 60

[BDHM01] R. Becher, M. Dillinger, M. Haardt, and W. Mohr, "Broadband wireless access and future communication networks," *Proceedings of the IEEE*, vol. 89, no. 1, pp. 58–75, Jan 2001. 3

[Bel64] P. Bello, "Characterization of randomly time-variant linear channels," vol. 11, pp. 360 – 393, 01 1964. 40

[Ber15] D. Beres, "Nikola Tesla predicted smartphones in 1926," 2015. [Online]. Available: http://www.huffingtonpost.com/2015/07/10/tesla-quotes_n_7771358.html 2

[BG06] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 884–893, March 2006. 58

[BHL+14] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 74–80, February 2014. 4

[BJ+13] E. Björnson, E. Jorswieck *et al.*, "Optimal resource allocation in coordinated multi-cell systems," *Foundations and Trends in Communications and Information Theory*, vol. 9, no. 2–3, pp. 113–381, 2013. 6

[BK94] V. A. Borovikov and B. E. Kinber, *Geometrical theory of diffraction*. IET, 1994, no. 37. 48

[BLM+14] N. Bhushan, J. Li, D. Malladi, R. Gilmore, D. Brenner, A. Damnjanovic, R. Sukhavasi, C. Patel, and S. Geirhofer, "Network densification: the dominant theme for wireless evolution into 5G," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 82–89, 2014. 5

[BLM15] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: 10 myths and one grand question," *Computing Research Repository (CoRR)*, vol. abs/1503.06854, 2015. [Online]. Available: http://arxiv.org/abs/1503.06854 6

[BMK13] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," *ACM SIGCOMM Computer Communication Review*, vol. 43, no. 4, pp. 375–386, 2013. 7

[Cav91] J. K. Cavers, "An analysis of pilot symbol assisted modulation for rayleigh fading channels (mobile radio)," *IEEE Transactions on Vehicular Technology*, vol. 40, no. 4, pp. 686–693, 1991. 43

[CCLM13] J. Choi, Z. Chance, D. J. Love, and U. Madhow, "Noncoherent trellis coded quantization: A practical limited feedback technique for massive MIMO systems," *IEEE Transactions on Communications*, vol. 61, no. 12, pp. 5016–5029, December 2013. 8

[Che07] M. Chen, "Radio channel prediction based on parametric modeling," Ph.D. dissertation, Chalmers University of Technology, Sweden, 2007. 59

[CJS+10] J. I. Choi, M. Jain, K. Srinivasan, P. Levis, and S. Katti, "Achieving single channel, full duplex wireless communication," in *Proceedings of the sixteenth annual international conference on Mobile computing and networking*. ACM, 2010, pp. 1–12. 7

[CL14] J. Chen and V. K. N. Lau, "Two-tier precoding for FDD multi-cell massive MIMO time-varying interference networks," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1230–1238, June 2014. 9

[CLB14] J. Choi, D. J. Love, and P. Bidigare, "Downlink training techniques for FDD massive MIMO systems: Open-loop and closed-loop training with memory," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 802–814, Oct 2014. 8

[CLK15] J. Choi, D. J. Love, and T. Kim, "Trellis-extended codebooks and successive phase adjustment: A path from LTE-advanced to FDD massive MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 4, pp. 2007–2016, April 2015. 8

[CLRH+14] I. Chih-Lin, C. Rowell, S. Han, Z. Xu, G. Li, and Z. Pan, "Toward green and soft: a 5G perspective," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 66–73, February 2014. 4

[CV09] M. Chen and M. Viberg, "Long-range channel prediction based on non-stationary parametric modeling," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 622–634, Feb 2009. 57

[DAZ+14] S. Dierks, M. B. Amin, W. Zirwas, M. Haardt, and B. Panzner, "The benefit of cooperation in the context of massive MIMO," in *Proceedings of 18th International OFDM Workshop, InOWo 2014*, Aug 2014, pp. 1–8. 6, 11, 67

[Del07] G. Del Galdo, "Geometry-based channel modeling for multi-user MIMO systems and applications," Ph.D. dissertation, Ilmenau University of Technology, 2007. 47, 48, 73

[DGK+13] P. Demestichas, A. Georgakopoulos, D. Karvounas, K. Tsagkaris, V. Stavroulaki, J. Lu, C. Xiong, and J. Yao, "5G on the horizon: Key

challenges for the radio-access network," *IEEE Vehicular Technology Magazine*, vol. 8, no. 3, pp. 47–53, Sept 2013. 4

[DH07] A. Duel-Hallen, "Fading channel prediction for mobile radio adaptive transmission systems," *Proceedings of the IEEE*, vol. 95, no. 12, pp. 2299–2313, 2007. 57, 58, 60

[DHHH00a] A. Duel-Hallen, S. Hu, and H. Hallen, "Long-range prediction of fading signals," *IEEE Signal Processing Magazine*, vol. 17, no. 3, pp. 62–75, May 2000. 9

[DHHH00b] ——, "Long range prediction of fading signals: Enabling adaptive transmission for mobile radio channels," *IEEE Signal Processing Magazine*, vol. 17, pp. 62–75, 2000. 57, 59, 60

[DHS03] G. Del Galdo, M. Haardt, and C. Schneider, "Geometry-based channel modelling of MIMO channels in comparison with channel sounder measurements," *Advances in Radio Science - Kleinheubacher Berichte*, vol. 50, no. 5, p. 117–126, October 2003. 48, 73

[dJH99] Y. L. C. de Jong and M. H. A. J. Herben, "Experimental verification of ray-tracing based propagation prediction models for urban microcell environments," in *Proceedings of IEEE VTS 50th Vehicular Technology Conference, VTC Fall 1999.*, vol. 3, 1999, pp. 1434–1438 vol.3. 48

[DS10] M. Duarte and A. Sabharwal, "Full-duplex wireless communications using off-the-shelf radios: Feasibility and first results," in *Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers (ASILOMAR), 2010.* IEEE, 2010, pp. 1558–1562. 7

[DXL01] L. Dong, G. Xu, and H. Ling, "Prediction of fast fading mobile radio channels in wideband communication systems," in *Proceedings of Global Telecommunications Conference, GLOBECOM 2001.*, vol. 6, 2001, pp. 3287–3291 vol.6. 59

[Ekm02] T. Ekman, "Prediction of mobile radio channels - modeling and design," Ph.D. dissertation, Uppsala University, Sweden, 2002. 57

[FC02] C. Y. Fung and S. Chan, "Estimation of fast fading channel in impulse noise environment," in *Proceedings of IEEE International Symposium on Circuits and Systems, ISCAS 2002.*, vol. 4. IEEE, 2002, pp. IV–IV. 61

[FG98] 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, pp. 311–335, 1998. 7

[FKV06] G. J. Foschini, K. Karakayali, and R. A. Valenzuela, "Coordinating multiple antenna cellular networks to achieve enormous spectral efficiency," *IEEE Proceedings - Communications*, vol. 153, no. 4, pp. 548–555, August 2006. 6, 21

[FPK+06] T. Fugen, M. Porebska, S. Knorzer, J. Maurer, and W. Wiesbeck, "Verification of 3D ray tracing with measurements in urban macrocellular environments," in *Proceedings of First European Conference on Antennas and Propagation 2006*, Nov 2006, pp. 1–6. 48

[FRT+17] J. Flordelis, F. Rusek, F. Tufvesson, E. G. Larsson, and O. Edfors, "Massive MIMO performance - TDD versus FDD: What do measurements say?" *IEEE Transactions on Wireless Communications (submitted)*, 2017. [Online]. Available: https://arxiv.org/pdf/1704.00623.pdf 10

[GERT15] X. Gao, O. Edfors, F. Rusek, and F. Tufvesson, "Massive MIMO performance evaluation based on measured propagation data," *IEEE Transactions on Wireless Communications*, vol. 14, no. 7, pp. 3899–3911, 2015. 8

[GKH+07] D. Gesbert, M. Kountouris, R. W. Heath, C.-B. Chae, and T. Salzer, "Shifting the MIMO paradigm," *IEEE Signal Processing Magazine*, vol. 24, no. 5, pp. 36–46, 2007. 28

[Gol05] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005. 7

[GP15] 5G-PPP, "The 5G infrastructure public private partnership: The next generation of communication networks and services," Tech. Rep., 2015. [Online]. Available: https://5g-ppp.eu/wp-content/uploads/2015/02/5G-Vision-Brochure-v1.pdf 3

[GPA+15] O. Galinina, A. Pyattaev, S. Andreev, M. Dohler, and Y. Koucheryavy, "5G multi-RAT LTE-WiFi ultra-dense small cells: performance dynamics, architecture, and trends," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 6, pp. 1224–1240, 2015. 5

[GS12] D. Gesbert and T. Svensson, "ARTIST4G Deliverable D1.4 - Interference avoidance techniques and system design," ARTIST4G consortium, Tech. Rep., July. 2012. [Online]. Available: http://publications.lib.chalmers.se/records/fulltext/212923/local_212923.pdf xv, 10, 16, 17, 18, 19, 21, 22, 24, 68, 82

[GSA15] A. G. Gotsis, S. Stefanatos, and A. Alexiou, "Ultra Dense Networks: The new wireless frontier for enabling 5G access," *Computing Research Repository (CoRR)*, vol. abs/1510.05938, 2015. [Online]. Available: http://arxiv.org/abs/1510.05938 25

[Hay96] S. Haykin, *Adaptive Filter Theory (3rd Ed.)*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1996. 59

[HK07] A. Heidari and A. K. Khandani, *Adaptive modeling, prediction, and tracking of wireless fading channels*. Department of Electrical and Computer Engineering, University of Waterloo, 2007. 57, 59

[HKM10] A. Heidari, A. K. Khandani, and D. McAvoy, "Adaptive modelling and long-range prediction of mobile fading channels," *IET Communications*, vol. 4, no. 1, pp. 39–50, January 2010. 57, 59

[HM00] M. Haardt and W. Mohr, "The complete solution for third-generation wireless communications: two modes on air, one winning strategy," *IEEE Personal Communications*, vol. 7, no. 6, pp. 18–24, Dec 2000. 3

[HN98] M. Haardt and J. A. Nossek, "Simultaneous schur decomposition of several nonsymmetric matrices to achieve automatic pairing in multidimensional harmonic retrieval problems," *IEEE Transactions on Signal Processing*, vol. 46, no. 1, pp. 161–169, Jan 1998. 59

[HRD08] M. Haardt, F. Roemer, and G. Del Galdo, "Higher-order SVD-based subspace estimation to improve the parameter estimation accuracy in multidimensional harmonic retrieval problems," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, pp. 3198–3213, July 2008. 59

[HRTA14] E. Hossain, M. Rasti, H. Tabassum, and A. Abdelnasser, "Evolution toward 5G multi-tier cellular wireless networks: An interference management perspective," *IEEE Wireless Communications*, vol. 21, no. 3, pp. 118–127, June 2014. 4

[HT09] H. Holma and A. Toskala, *LTE for UMTS-OFDMA and SC-FDMA based radio access*. John Wiley & Sons, 2009. 3

[HW98] J.-K. Hwang and J. H. Winters, "Sinusoidal modeling and prediction of fast fading processes," in *Proceedings of IEEE GLOBECOM 1998*, vol. 2, 1998, pp. 892–897 vol.2. 57, 59

[IDM+11] R. Irmer, H. Droste, P. Marsch, M. Grieger, G. Fettweis, S. Brueck, H. P. Mayer, L. Thiele, and V. Jungnickel, "Coordinated multipoint: Concepts, performance, and field trial results," *IEEE Communications Magazine*, vol. 49, no. 2, pp. 102–111, February 2011. 6, 20

[IN14] M. Ivrlac and J. Nossek, "The multiport communication theory," *IEEE Circuits and Systems Magazine*, vol. 14, no. 3, pp. 27–44, 2014. 119

[IR15] ITU-R, "Final acts WRC-15," in *World Radiocommunication Conference, Geneva*, 2015. 4

[ITU15] ITU, "IMT vision – framework and overall objectives of the future development of IMT for 2020 and beyond," Recommendation ITU-R M.2083-0, Tech. Rep., 2015. 4

[JAMV11] J. Jose, A. E. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems." *IEEE Transactions on Wireless Communications*, vol. 10, no. 8, pp. 2640–2651, 2011. 29

[JC94] W. C. Jakes and D. C. Cox, Eds., *Microwave Mobile Communications*. Wiley-IEEE Press, 1994. 62

[JCK+11] M. Jain, J. I. Choi, T. Kim, D. Bharadia, S. Seth, K. Srinivasan, P. Levis, S. Katti, and P. Sinha, "Practical, real-time, full duplex wireless," in *Proceedings of the 17th annual international conference on Mobile computing and networking*. ACM, 2011, pp. 301–312. 7

[JMCN15] Z. Jiang, A. F. Molisch, G. Caire, and Z. Niu, "Achievable rates of FDD massive MIMO systems with spatial channel correlation," *IEEE Transactions on Wireless Communications*, vol. 14, no. 5, pp. 2868–2882, May 2015. 9

[JMZ+14] V. Jungnickel, K. Manolakis, W. Zirwas, B. Panzner, V. Braun, M. Lossow, M. Sternad, R. Apelfrojd, and T. Svensson, "The role of small cells, coordinated multipoint, and massive MIMO in 5G," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 44–51, May 2014. xv, 5, 6, 20, 23, 25, 26

[K+60] R. E. Kalman *et al.*, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960. 61, 79

[KAK87] H. Kanai, M. Abe, and K. Kido, "Accurate autoregressive spectrum estimation at low signal-to-noise ratio using a phase matching technique," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, no. 9, pp. 1264–1272, Sep 1987. 58

[Kel62] J. B. Keller, "Geometrical theory of diffraction," *JOSA*, vol. 52, no. 2, pp. 116–130, 1962. 48

[KGW+14] D. W. Kifle, L. C. Gimenez, B. Wegmann, I. Viering, and A. Klein, "Comparison and extension of existing 3D propagation models with real-world effects based on ray-tracing," *Wireless Personal Communications*, vol. 78, no. 3, pp. 1719–1738, Oct. 2014. [Online]. Available: http://dx.doi.org/10.1007/s11277-014-1910-0 48

[Kin63] R. King, *Fundamental electromagnetic theory*, ser. Dover books for engineers. Dover Publications, 1963. [Online]. Available: https://books.google.pl/books?id=A17xAAAAMAAJ 49

[KM81] S. M. Kay and S. L. Marple, "Spectrum analysis - a modern perspective," *Proceedings of the IEEE*, vol. 69, no. 11, pp. 1380–1419, 1981. 60

[KN15] N. Kundargi and K. Nieman, "Massive MIMO vs FD-MIMO: Defining the next generation of MIMO in 5G (panel discussion)," in *Proceedings of Global Telecommunications Conference, GLOBECOM 2015*, Dec 2015. 10

[KPI08] K. J. Kim, M. O. Pun, and R. A. Iltis, "Channel prediction for limited feedback precoded MIMO-OFDM systems over time-varying fading channels," in *Proceedings of 42nd Annual Conference on Information Sciences and Systems, 2008*, March 2008, pp. 972–977. 58

[KTC15] M. Kurras, L. Thiele, and G. Caire, "Interference mitigation and multiuser multiplexing with beam-steering antennas," in *Proceedings of 19th International ITG Workshop on Smart Antennas, WSA 2015.*, March 2015, pp. 1–5. 30

[LCG+12] L. Liu, R. Chen, S. Geirhofer, K. Sayana, Z. Shi, and Y. Zhou, "Downlink MIMO in LTE-advanced: SU-MIMO vs. MU-MIMO," *IEEE Communications Magazine*, vol. 50, no. 2, 2012. 28, 56

[LCS98] Y. G. Li, L. J. Cimini, and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," in *Proceedings of IEEE International Conference on Communications, ICC 1998.*, vol. 3, Jun 1998, pp. 1320–1324 vol.3. 56

[LFYH14] L. Liu, H. Feng, T. Yang, and B. Hu, "MIMO-OFDM wireless channel prediction by exploiting spatial-temporal correlation," *IEEE Transactions on Wireless Communications*, vol. 13, no. 1, pp. 310–319, 2014. 58

[LHKK15] A. Loch, M. Hollick, A. Kuehne, and A. Klein, "OFDMA for wireless multihop networks: From theory to practice," *Pervasive and Mobile Computing*, vol. 23, pp. 104–121, 2015. 41

[LMV00] L. D. Lathauwer, B. D. Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1253–1278, 2000. [Online]. Available: http://dx.doi.org/10.1137/S0895479896305696 71

[LS08] X. Long and B. Sikdar, "A wavelet based long range signal strength prediction in wireless networks," in *Proceedings of IEEE International Conference on Communications, ICC 2008*, May 2008, pp. 2043–2047. 57

[LTEM13] E. G. Larsson, F. Tufvesson, O. Edfors, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, 2013. 6, 26, 28

[Mar10] 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. 6, 7

[Mar15] ——, "Massive MIMO: An introduction," *Bell Labs Technical Journal*, vol. 20, pp. 11–22, 2015. 3, 5, 7, 28

[MCCK07] C. Min, N. Chang, J. Cha, and J. Kang, "MIMO-OFDM downlink channel prediction for IEEE 802.16e systems using Kalman filter," in *Proceedings of IEEE Wireless Communications and Networking Conference, 2007*, March 2007, pp. 942–946. 58

[MDCD02] B. Muquet, M. De Courville, and P. Duhamel, "Subspace-based blind and semi-blind channel estimation for OFDM systems," *IEEE Transactions on Signal Processing*, vol. 50, no. 7, pp. 1699–1712, 2002. 57

[MDH08] M. Milojevic, G. Del Galdo, and M. Haardt, "Tensor-based framework for the prediction of frequency-selective time-variant MIMO channels," in *Proceedings of International ITG Workshop on Smart Antennas, WSA 2008.*, 2008. 11, 71, 72, 74, 75

[MDSH06] M. Milojevic, G. Del Galdo, V. Stankovic, and M. Haardt, "Subspace-based framework for the prediction of MIMO OFDM channels," in *Proceedings of 11-th International OFDM Workshop (InOWo 2006)*, 2006. 74

[MF11] P. Marsch and G. P. Fettweis, *Coordinated Multi-Point in Mobile Communications: From Theory to Practice*, 1st ed. New York, NY, USA: Cambridge University Press, 2011. 22

[MPIY+13] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. P. Papathanassiou, "A tutorial on synthetic aperture radar," *IEEE Geoscience and Remote Sensing Magazine*, vol. 1, no. 1, pp. 6–43, March 2013. 83

[MPT+13] P. Mogensen, K. Pajukoski, E. Tiirola, E. Lahetkangas, J. Vihriala, S. Vesterinen, M. Laitila, G. Berardinelli, G. W. Da Costa, L. G. Garcia *et al.*, "5G small cell optimized radio design," in *IEEE Globecom Workshops, 2013.* IEEE, 2013, pp. 111–116. 5

[MVK95] M. M. Milosavljevic, M. D. Veinovic, and B. D. Kovacevic, "Estimation of nonstationary AR model using the weighted recursive least square algorithm," in *Proceedings of 1995 International Conference on Acoustics, Speech, and Signal Processing*, vol. 2, May 1995, pp. 1432–1435 vol.2. 58

[NAAC14] J. Nam, A. Adhikary, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing: Opportunistic beamforming, user grouping and simplified downlink scheduling," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 876–890, 2014. 9

[Nam14] J. Nam, "Fundamental limits in correlated fading MIMO broadcast channels: Benefits of transmit correlation diversity," in *Proceedings of IEEE International Symposium on Information Theory (ISIT), 2014.* IEEE, 2014, pp. 2889–2893. 9

[NESB97] R. Nilsson, O. Edfors, M. Sandell, and P. O. Borjesson, "An analysis of two-dimensional pilot-symbol assisted modulation for OFDM," in

Proceedings of IEEE International Conference on Personal Wireless Communications, 1997. IEEE, 1997, pp. 71–74. 43

[Nok16] Nokia, "5G use cases and requirements," Tech. Rep., 2016. [Online]. Available: http://resources.alcatel-lucent.com/asset/200010 xv, 3, 4

[NP00] R. V. Nee and R. Prasad, *OFDM for wireless multimedia communications.* Artech House, Inc., 2000. 41

[OBB+14] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, M. Schellmann, H. Schotten, H. Taoka, H. Tullberg, M. A. Uusitalo, B. Timus, and M. Fallgren, "Scenarios for 5G mobile and wireless communications: the vision of the METIS project," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 26–35, May 2014. 4

[OHH04] G. E. Oien, H. Holm, and K. J. Hole, "Impact of channel prediction on adaptive coded modulation performance in rayleigh fading," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 3, pp. 758–769, May 2004. 57

[Orf16] S. J. Orfanidis, *Electromagnetic Waves and Antennas*, 2016. [Online]. Available: http://eceweb1.rutgers.edu/~orfanidi/ewa/ 86, 87, 88, 89, 92, 95, 106

[PM16] N. Pratas and N. Mahmood, "FANTASTIC5G Deliverable D4.1 - technical results for service specific multi-node/multi-antenna solutions," FANTASTIC5G consortium, Tech. Rep., June. 2016. [Online]. Available: http://fantastic5g.eu/wp-content/uploads/2016/06/FANTASTIC-5G_D4.1_Final.pdf 16, 24, 41

[PNCZ+16] A. I. Perez-Neira, M. Caus, R. Zakaria, D. L. Ruyet, E. Kofidis, M. Haardt, X. Mestre, and Y. Cheng, "MIMO signal processing in offset-QAM based filter bank multicarrier systems," *IEEE Transactions on Signal Processing*, vol. 64, no. 21, pp. 5733–5762, Nov 2016. 41

[PWSF12] K. I. Pedersen, Y. Wang, B. Soret, and F. Frederiksen, "eICIC functionality and performance for LTE HetNet co-channel deployments," in *Proceedings of IEEE Vehicular Technology Conference, VTC Fall 2012.*, Sept 2012, pp. 1–5. 26

[Que15] O. Queseth, "METIS deliverable 8.4 - METIS final project report," The METIS 2020 Project, Tech. Rep., 2015. [Online]. Avail-

able: https://www.metis2020.com/wp-content/uploads/deliverables/METIS_D8.4_v1.pdf 4

[Rap96] T. S. Rappaport, *Wireless communications: principles and practice*, 1996, vol. 2. 36

[RHSK08] D. Rhee, H. G. Hwang, Y. J. Sang, and K. S. Kim, "Multiuser adaptive transmission technique for time-varying frequency-selective fading channels," *Elsevier Signal Processing*, vol. 88, no. 8, pp. 2095–2107, Aug. 2008. [Online]. Available: http://dx.doi.org/10.1016/j.sigpro.2008.02.014 9

[Rik15] K. Rikkinen, "DUPLO Deliverable D6.3 - full-duplex radios for local access, final report," DUPLO Project, Tech. Rep., 2015. [Online]. Available: http://www.fp7-duplo.eu/images/docs/Deliverables/D6{_}3%20v.1.0.pdf 7

[RK89] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 7, pp. 984–995, 1989. 59

[RL14] X. Rao and V. K. N. Lau, "Distributed compressive CSIT estimation and feedback for FDD Multi-User massive MIMO systems," *IEEE Transactions on Signal Processing*, vol. 62, no. 12, pp. 3261–3271, June 2014. 8

[RPL+13] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, Jan 2013. 26

[RTMR06] M. L. Roberts, M. A. Temple, R. F. Mills, and R. A. Raines, "Evolution of the air interface of cellular communications systems toward 4G realization," *IEEE Communications Surveys and Tutorials*, vol. 8, no. 1, pp. 2–23, First 2006. 2

[SA05] M. K. Simon and M.-S. Alouini, *Digital communication over fading channels*. John Wiley & Sons, 2005, vol. 95. 35

[SAE03] Z. Shen, J. G. Andrews, and B. L. Evans, "Short range wireless channel prediction using local information," in *Conference Record of the Thirty-Seventh Asilomar Conference on Signals, Systems and Computers, 2004.*, vol. 1. IEEE, 2003, pp. 1147–1151. 57

[SBT11] S. Sesia, M. Baker, and I. Toufik, *LTE- the UMTS long term evolution: from theory to practice*. John Wiley & Sons, 2011. 3

[Sch86] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, 1986. 59

[Sch16] F. Schaich, "FANTASTIC5G Factsheet," FANTASTIC5G consortium, Tech. Rep., 2016. [Online]. Available: http://fantastic5g.eu/wp-content/uploads/2015/09/FANTASTIC-5G_Factsheet_v1.1.pdf 4

[Sey05] J. S. Seybold, *Introduction to RF propagation*. John Wiley & Sons, 2005. 35

[SFGK00] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Transactions on Communications*, vol. 48, no. 3, pp. 502–513, Mar 2000. 11

[SHP07] C. Shin, R. W. Heath, and E. J. Powers, "Blind channel estimation for MIMO-OFDM systems," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 2, pp. 670–685, 2007. 57

[SHRS08] S. Saur, H. Halbauer, A. Rueegg, and F. Schaich, "Grid-of-beams (GoB) based downlink multi-user MIMO," Alcatel-Lucent, Contribution to IEEE 802.16m C802.16m-08/487, 2008. 29

[SK03] S. Semmelrodt and R. Kattenbach, "Investigation of different fading forecast schemes for flat fading radio channels," in *Proceedings of IEEE Vehicular Technology Conference, VTC Fall 2003.*, vol. 1, Oct 2003, pp. 149–153 Vol.1. 60

[Skl01] B. Sklar, *Digital communications*. Prentice Hall Upper Saddle River, 2001, vol. 2. 36

[SKM+10] M. Sawahashi, Y. Kishiyama, A. Morimoto, D. Nishikawa, and M. Tanno, "Coordinated multipoint transmission/reception techniques for LTE-advanced [coordinated and distributed MIMO]," *IEEE Wireless Communications*, vol. 17, no. 3, 2010. 6

[SLC08] H. Shirani-Mehr, D. N. Liu, and G. Caire, "Channel state prediction, feedback and scheduling for a multiuser MIMO-OFDM downlink,"

Computing Research Repository (CoRR), vol. abs/0811.4630, 2008. [Online]. Available: http://arxiv.org/abs/0811.4630 11, 59

[SM97] P. Stoica and R. L. Moses, *Introduction to spectral analysis*. Upper Saddle River, N.J. Prentice Hall, 1997. [Online]. Available: http://opac.inria.fr/record=b1092427 60

[SM05] D. Schafhuber and G. Matz, "MMSE and adaptive prediction of time-varying channels for OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 4, no. 2, pp. 593–602, March 2005. 60

[SRH13] B. Song, F. Roemer, and M. Haardt, "Flexible coordinated beamforming (FlexCoBF) for the downlink of multi-user MIMO systems in single and clustered multiple cells," *Elsevier Signal Processing*, vol. 93, no. 9, pp. 2462–2473, Sep 2013. [Online]. Available: http://dx.doi.org/10.1016/j.sigpro.2013.03.012 6

[SS06] T. Svantesson and A. L. Swindlehurst, "A performance bound for prediction of MIMO channels," *IEEE Transactions on Signal Processing*, vol. 54, no. 2, pp. 520–529, 2006. 58

[SSG+14] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1637–1652, 2014. 7

[Swi98] A. L. Swindlehurst, "Time delay and spatial signature estimation using known asynchronous signals," *IEEE Transactions on Signal Processing*, vol. 46, no. 2, pp. 449–462, Feb 1998. 59

[Taf13] R. Tafazolli, "Why 5G," 2013. [Online]. Available: https://docbox.etsi.org/Workshop/2013/201311_FUTUREMOBILESUMMIT/03_UNIofSURREY_TAFAZOLLI.pdf xv, 3

[TH13] K. T. Truong and R. W. Heath, "Effects of channel aging in massive MIMO systems," *Computing Research Repository (CoRR)*, vol. abs/1305.6151, 2013. [Online]. Available: http://arxiv.org/abs/1305.6151 9, 56

[TKBH12] L. Thiele, M. Kurras, K. Börner, and T. Haustein, "User-aided sub-clustering for CoMP transmission: Feedback overhead vs. data rate

trade-off," in *Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Nov 2012, pp. 1142–1146. 21

[TKJ+15] L. Thiele, M. Kurras, S. Jaeckel, S. Fähse, and W. Zirwas, "Interference-floor shaping for liquid coverage zones in coordinated 5G networks," in *Proceedings of 49th Asilomar Conference on Signals, Systems and Computers*, Nov 2015, pp. 1102–1106. 22

[TOKM11] L. Thiele, M. Olbrich, M. Kurras, and B. Matthiesen, "Channel aging effects in CoMP transmission: gains from linear channel prediction," in *2011 Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Nov 2011, pp. 1924–1928. 56

[Tuc66] L. R. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, no. 3, pp. 279–311, 1966. [Online]. Available: http://dx.doi.org/10.1007/BF02289464 71

[TV01] P. D. Teal and R. G. Vaughan, "Simulation and performance bounds for real-time prediction of the mobile multipath channel," in *Proceedings of the 11th IEEE Signal Processing Workshop on Statistical Signal Processing*, 2001, pp. 548–551. 11

[TV05] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005. 35, 37

[TZH16] O. Taghizadeh, J. Zhang, and M. Haardt, "Transmit beamforming aided amplify-and-forward MIMO full-duplex relaying with limited dynamic range," *Elsevier Signal Processing*, vol. 127, no. C, pp. 266–281, Oct 2016. [Online]. Available: https://doi.org/10.1016/j.sigpro.2016.02.026 7

[vDH94] G. van Dooren and M. Herben, "A deterministic approach for the modeling of wave propagation around buildings," *Journal of Electromagnetic Waves and Applications*, vol. 8, no. 2, pp. 175–194, 1994. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1163/156939394X00029 48

[VS07] J. Vanderpypen and L. Schumacher, "MIMO channel prediction using ESPRIT based techniques," in *Proceedings of IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC 2007*. IEEE, 2007, pp. 1–5. 59

[VTR00] R. Vaughan, P. Teal, and R. Raich, "Short-term mobile channel prediction using discrete scatterer propagation model and subspace signal processing algorithms," in *Proceedings Vehicular Technology Conference, VTC fall 2000.*, vol. 2. IEEE, 2000, pp. 751–758. 59

[Wan13] J. Wannstrom, "LTE-Advanced," 2013. [Online]. Available: http://www.3gpp.org/technologies/keywords-acronyms/97-lte-advanced 6, 93

[WC97] W.-R. Wu and P.-C. Chen, "Adaptive AR modeling in white gaussian noise," *IEEE Transactions on Signal Processing*, vol. 45, no. 5, pp. 1184–1192, May 1997. 58

[WD14] D. Warren and C. Dewar, "Understanding 5G: Perspectives on future technological advancements in mobile," GSMA Intelligence, Tech. Rep., 2014. [Online]. Available: https://www.gsmaintelligence.com/research/?file=141208-5g.pdf&download 17

[WE06] I. C. Wong and B. L. Evans, "Exploiting spatio-temporal correlations in MIMO wireless channel prediction," in *IEEE Globecom 2006*, Nov 2006, pp. 1–5. 11, 58

[WHG+14] C.-X. Wang, F. Haider, X. Gao, X.-H. You, Y. Yang, D. Yuan, H. Aggoune, H. Haas, S. Fletcher, and E. Hepsaydir, "Cellular architecture and key technologies for 5G wireless communication networks," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 122–130, 2014. 5

[Win13] WinProp, 2013. [Online]. Available: http://www.altairhyperworks.com/product/FEKO/WinProp-Propagation-Modeling 50

[XYM14] Y. Xu, G. Yue, and S. Mao, "User grouping for massive MIMO in FDD systems: New design methods and analysis," *IEEE Access*, vol. 2, pp. 947–959, 2014. 9

[YLCC01] B. Yang, K. B. Letaief, R. S. Cheng, and Z. Cao, "Channel estimation for OFDM transmission in multipath fading channels based on parametric channel modeling," *IEEE Transactions on Communications*, vol. 49, no. 3, pp. 467–479, 2001. 59

[ZA16a] W. Zirwas and M. B. Amin, "Artificially mutually coupled antenna arrays," 2016, filed EU Patent No 16836698.7 - 1220 PCT/FI2016050466. 119

[ZA16b] ——, "Nonlinear user processing," 2016, filed US Patent No PCT/IB2016/057361. 119

[ZAS16] W. Zirwas, M. B. Amin, and M. Sternad, "Coded CSI reference signals for 5G-exploiting sparsity of FDD massive MIMO radio channels," in *Proceedings of the 20th International ITG Workshop on Smart Antennas, WSA 2016.* VDE, 2016, pp. 1–8. 10, 31

[ZH13] W. Zirwas and M. Haardt, "Channel prediction for B4G radio systems," in *Proceedings of 77th IEEE Vehicular Technology Conference, VTC Spring 2013.*, 2013. 83, 92, 104

[Zir15] W. Zirwas, "Opportunistic CoMP for 5G massive MIMO multilayer networks," in *WSA 2015; 19th International ITG Workshop on Smart Antennas*, March 2015, pp. 1–7. 24, 26

[ZJCZ08] L. Zhang, Z. Jin, W. Chen, and X. Zhang, "An improved adaptive channel prediction for MIMO-OFDM systems," in *Proceedings of Third International Conference on Communications and Networking in China, 2008*, Aug 2008, pp. 1008–1012. 58

[ZMK13] W. Zirwas, W. Mennerich, and A. Khan, "Main enablers for advanced interference mitigation," *Transactions on Emerging Telecommunications Technologies*, vol. 24, no. 1, pp. 18–31, 2013. [Online]. Available: http://dx.doi.org/10.1002/ett.2567 xv, 17, 18, 22, 23, 97

[ZTKW16] W. Zirwas, L. Thile, M. Kurras, and G. Wunder, "Flexible 5G below 6 GHz mobile broadband radio air interface," in *5G workshop, IEEE VTC Spring 2016*, 2016. 83