Performance of 5G Small Cells using Flexible TDD

Catania, Davide

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

Performance of 5G Small Cells using Flexible TDD

Ph.D. Dissertation

Davide Catania

Aalborg University
Department of Electronic Systems
Fredrik Bajers Vej 7
DK – 9220 Aalborg

Thesis submitted:	July, 2015
PhD Supervisor:	Prof. Preben Mogensen Aalborg University
Assistant PhD Supervisor:	Assoc. Prof. Andrea Fabio Cattoni Aalborg University
PhD Committee:	Prof. Tuna Tuğcu Bogazici University Assoc. Prof. Daniel E. Lucani Aalborg University Juho Pirskanen Principal Researcher, Nokia, Finland
PhD Series:	Faculty of Electronic Systems, Aalborg University

ISBN: 978-87-7152-087-3

Published by:
Aalborg University Press
Skjernvej 4A, 2nd floor
DK – 9220 Aalborg Ø
Phone: +45 99407140
aauf@forlag.aau.dk
forlag.aau.dk

Curriculum Vitae

Davide Catania

Davide Catania obtained his Bachelors Degree (with Honours) in Electrical Engineering, from the University of Malta in 2008. Following this, he decided to pursue his Masters Degree in Telecommunications with focus on Protocols and Network Technologies, at the Technical University of Denmark, DTU, subsequently completing his studies in 2010. In early 2011, he spent 9 months as a Research Assistant at Aalborg University, where he was primarily involved with large scale system-level simulator development, and studies related to Cognitive Radio and WiFi. From 2012, Davide started his PhD at Aalborg University, working on the design and investigation of specific features, for a new 5G small cell concept.

Abstract

The continuous demand for better wireless data services along with the ever-increasing growth of data traffic present a significant challenge to future 5th generation (5G) cellular systems. One proven way of dealing with these challenges is via further cell densification, leading to the deployment of dense small cell systems. Given this increased number of cells and scarcity of spectral resources in prime spectrum bands, inter-cell interference becomes a significant issue to tackle.

As cell sizes shrink considerably, a smaller number of users are expected to be served by each cell. This means that the amount of resources needed for the corresponding link directions, uplink or downlink, might vary significantly over time, making fully flexible time division duplex (TDD) a desirable feature for future 5G small cells. While attractive, such a feature also presents numerous problems, since it introduces new types of interference, mainly referred to as same-entity interference. On top of that, the freedom brought along by flexible TDD induces increased inter-cell interference variation, presenting significant difficulties in estimating and adapting to future channel conditions.

The work conducted in this thesis, deals with the study of flexible TDD in small cell scenarios, investigated in the context of an envisioned 5G small cell concept. The studies are done using a discrete event system-level simulator.

The main outcome of this work shows that, if built on a suitable framework, fully flexible TDD can be a relevant technology for small cells, and allows the system to obtain significant gains over a fixed TDD system. Moreover, we see that the proper support and usage of interference suppression receivers can mitigate the severe problem of inter-cell interference variation. The study also shows that flexible TDD can be beneficial when used with higher layer reliable transport protocols such as the Transmission Control Protocol (TCP).

Resumé

Den fortsatte efterspørgsel efter bedre trådløse datatjenester og den stadigt stigende vækst i datatrafik, udgør en betydelig udfordring for fremtidige 5. generation (5G) mobilkommunikationssystemer. Den klassiske løsning på sådanne udfordringer er yderligere celle fortætning, hvilket fører til systemer bestående af mange små og tæt pakkede celler. I betragtning af det øgede antal celler og generel knaphed på frekvensspektrum i de mest benyttede frekvensbånd, bliver håndtering af inter-celle interferens et essentielt emne.

Med betydeligt mindre celler forventes også at antallet af brugere der skal serviceres per celle er mindre. Det indebærer tilsvarende at mængden af ressourcer der er nødvendige i de to linkretninger, uplink og downlink, kan variere betydeligt over tid. Anvendelsen af fuldt fleksibel time division duplex (TDD) i fremtidige 5G småcelle systemer kan derfor være en attraktiv teknik. Brugen af en sådan teknik medfører dog også en række problemer da der vil opstå en ny type krydsinterferens indenfor samme celle, baser og terminaler imellem. Tilmed bevirker den fleksible ressourcehåndtering en betydelig forøget variation i inter-celle interferens, og dermed store vanskeligheder i estimering af og tilpasning til kommunikationskanalen.

Denne afhandling omhandler studiet af fleksibel TDD i småcelle scenarier, baseret på et udtænkt 5G småcelle koncept. Undersøgelserne er udført ved hjælp af diskret event-baseret systemniveau simulering.

Det vigtigste resultat af arbejdet viser, at med det rette 5G koncept, kan fuldt fleksibel TDD være en relevant teknologi i små celler, med mulighed for at opnå betydelige gevinster i forhold til et system med fast TDD. Specifikt ser vi, at korrekt understøttelse og brug af interferensundertrykkende modtagere kan modvirke vanskelighederne i forbindelse med variationen i inter-celle interferens. Undersøgelsen viser også, at brugen af fleksibel TDD kan være gavnligt når teknikken anvendes sammen med pålidelige transportprotokoller såsom Transmission Control Protocol (TCP).

Contents

List of Abbreviations

1G 1st Generation

2G 2nd Generation

3G 3rd Generation

4G 4th Generation

5G 5th Generation

AM Acknowledged Mode

AMC Adaptive Modulation and Coding

AMPS Advanced Mobile Phone Services

AP Access Point

ASA Authorised Shared Access

CA Carrier Aggregation

CDF Cumulative Distribution Function

CDMA Code Division Multiple Access

cmWave centimetre wave

COMP Coordinated Multipoint

DF Delay Fairness Based

DL Downlink

DMRS Demodulation Reference Symbol

EDGE Enhanced Data rates for GSM Evolution

F-TDD Flexible TDD

FDD Frequency Division Duplex

FDMA Frequency Division Multiple Access

FIFO First in First Out

GP Guard Period

GPRS General Radio Packet Service

GPS Global Position System

GSM Global System for Mobile Communications

HARQ Hybrid Automatic Repeat Request

HOL Head of Line

HSPA High Speed Packet Access

ICM Interference Covariance Matrix

IRC Interference Rejection Combining

ISM Industrial, scientific and medical

LF Load Fairness Based

LOS Line Of Sight

LSA Licensed Shared Access

LTE Long Term Evolution

LTE-A Long Term Evolution Advanced

LTE-B LTE-B

LTE-eIMTA LTE Enhanced Interference Mitigation & Traffic Adaptation

MAC Medium Access Control

MCS Modulation and Coding Scheme

MIMO Multiple Input Multiple Output

MMSE Minimum Mean Square Error

mmWave millimetre wave

MRC Maximum Ratio Combining

MU-MIMO Multi User MIMO

NLOS Non-Line of Sight

NMT Nordic Mobile Telephone

OFDM Orthogonal Frequency Division Multiplexing

PDU Protocol Data Unit

PHY Physical

PMI Precoding Matrix Information

PRB Physical Resource Block

R-TDD Random TDD

RAT Radio Access Technology

RLC Radio Link Control

RTT Round Trip Time

S-TDD Static TDD

SC-FMDA Single Carrier Frequency Division Multiple Access

SDU Service Data Unit

SMS Short Messaging Service

SR Scheduling Request

SU-MIMO Single User MIMO

TCP Transmission Control Protocol

TDD Time Division Duplex

TDMA Time Division Multiple Access

UDP User Datagram Protocol

UE User Equipment

UL Uplink

UM Unacknowledged Mode

UMTS Universal Mobile Telephone Service

VoIP Voice over IP

WiMAX Worldwide Interoperability for Microwave Access

Thesis Details

Thesis Title:	Performance of 5G Small Cells using Flexible TDD
Ph.D. Student:	Davide Catania
Supervisors:	Prof. Preben Mogensen, Aalborg University
	Assoc. Prof. Andrea Fabio Cattoni, Aalborg University

The main body of this thesis consist of the following papers.

[A] Davide Catania, Marta Gatnau Sarret, Andrea Fabio Cattoni, Frank Frederiksen, Gilberto Berardinelli, Preben Mogensen, "The Potential of Flexible UL/DL Slot Assignment in 5G Systems", *Vehicular Technology Conference (VTC Fall), 2014 IEEE 80th . IEEE, 2014. s. 1-6.*

[B] Davide Catania, Marta Gatnau Sarret, Andrea Fabio Catoni, Frank Frederiksen, Gilberto Berardinelli, Preben Mogensen, "Flexible UL/DL in Small Cell TDD Systems: A Performance Study with TCP Traffic.", *2015 IEEE 81st Vehicular Technology Conference: VTC2015-Spring. 2015.*

[C] Davide Catania, Andrea Fabio Cattoni, Nurul Huda Mahmood, Gilberto Berardinelli, Frank Frederiksen, Preben Mogensen. "A Distributed Taxation Based Rank Adaptation Scheme for 5G Small Cells" *2015 IEEE 81st Vehicular Technology Conference: VTC2015-Spring. 2015.*

This thesis has been submitted for assessment in partial fulfillment of the PhD degree. The thesis is based on the submitted or published scientific papers which are listed above. Parts of the papers are used directly or indirectly in the extended summary of the thesis. As part of the assessment, co-author statements have been made available to the assessment committee and are also available at the Faculty. The thesis is not in its present form acceptable for open publication but only in limited and closed circulation as copyright may not be ensured.

Preface

The work conducted in this thesis is the result of a project carried out at the Wireless Communication Networks (WCN) section, Institute of Electronic Systems, Aalborg University, Denmark. In parallel with the work presented in this thesis, mandatory courses and teaching/working obligations were fulfilled as part of the requirements needed to obtain the Ph.D. degree.

The project was financed by Nokia Networks and the "Danish Council for Independent Research" (Det Frie Forskningsgråd, DFF), "Technology and Production Sciences" (Forskningsrådet for teknologi of Produktion, FTP), in connection with the research project "Cognitive Radio Concepts for Beyond-Femtocells". The research project has been completed under the supervision of Professor Preben Mogensen (Aalborg University), and Associate Professor Andrea Fabio Cattoni (Aalborg University).

The thesis investigates the suitability of fully flexible time division duplex for future 5G small cells, by considering both the merits and demerits brought along by such a technology.

First of all, I would like to thank my supervisor Prof. Preben Mogensen. His experience and knowledge about the subject has been instrumental in giving value to the results achieved. I would also like to express my gratitude to my co-supervisor Assoc. Prof. Andrea Fabio Cattoni, for his persistent strive for quality. His guidance has been invaluable for both the results achieved, and the personal qualities I was able to improve upon.

I would also like to thank all my colleagues at Aalborg University and Nokia Networks for creating a great working environment, where I was able to learn uncountable lessons from all aspects of life. Sincere thanks go to Gustavo Wagner and Marta Gatnau Sarret for helping me in related system-level simulator development. I would also like to thank Mads Brix for his continuous support and sharing of knowledge in both simulator development and maintenance. A special thanks also goes to Frank Frederiksen, who with his

expertise has provided me with invaluable lessons on how to look, address and tackle a specific problem at hand. Sincere thanks goes to Gilberto Berardinelli and Nurul Huda Mahmood for their patience and continuous support related to physical layer aspects. I would also like to take this opportunity to thank my closest friends, Simone Barbera, Oscar Tonelli, Michele Polignano, and Panagiotis Fotiadis for being part of this fantastic 3-year journey. Their joyful spirits have helped me enjoy and survive both the greatest and toughest times.

Special thanks goes to my girlfriend Evelina Mihailova for giving me her unconditional support during the writing of this thesis. Her affection and support were indispensable for my well being and happiness.

Last and not least, I would like to thank my family for their encouragement and support, and for always believing in my capabilities. Without their help, I would never be able to accomplish such a milestone.

This thesis is dedicated to my grandfather Joseph Zammit, who passed away during the time of writing. I will always cherish the moments we spent together in my childhood. Thanks for the countless life lessons you have delivered to me.

Davide Catania

Aalborg University, July 6, 2015

Part I

Extended Summary

Chapter 1

Introduction

1.1 Introduction

Wireless connectivity has become an integral part of our society. The undeniable success of wireless data services has had a tremendous impact on the way we conduct our daily activities. A proof of this is the widespread adoption of devices such as smart phones and tablets which are radically shaping the way we work and behave. Technological advancements at both the end user terminal and the network side continue to deliver higher unprecedented speeds and lower latencies. In turn, these improvements, not only continue to fuel a demand for better wireless data services but also open a plethora of opportunities for new innovative applications.

1.1.1 Traffic Growth

In 2014, mobile networks carried nearly 30 exabytes of traffic [1]. Studies in [2], show a 60% data traffic growth between 2013 and 2014, with video being the largest and fastest contributing traffic source of this growth. This traffic growth is not expected to stop and forecasts in [1], illustrated in Figure 1.1, predict an exponential growth in mobile data traffic in the coming years.

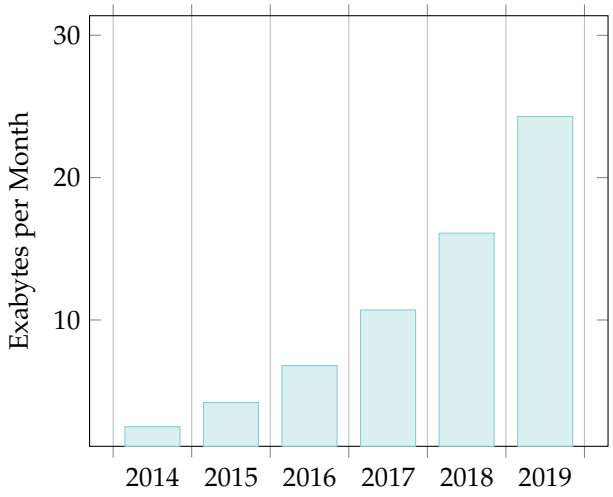


Fig. 1.1: Traffic Growth [1]

1.1.2 The Evolution of Cellular Networks

The 1st Generation (1G) commercial cellular systems emerged in the late 1970's and early 1980's, with systems such as the Nordic Mobile Telephone (NMT), Advanced Mobile Phone Services (AMPS), and Total Access Communication Systems, characterized by the usage of frequency modulated analog voice transmission and the deployment of few base stations covering large areas [3].

A digital evolution of 1G systems, arrived in the early 1990's with the 2nd Generation (2G) Global System for Mobile Communications (GSM), harmonizing various fragmented European markets and delivering improved voice quality, and Short Messaging Service (SMS). 1G and 2G were primarily designed for voice traffic, with 2G offering limited support for data traffic services via General Radio Packet Service (GPRS), providing a maximum data rate of 160 kbps. An enhancement of this technology came along with Enhanced Data rates for GSM Evolution (EDGE) via the introduction of incremental redundancy, reaching theoretical maximum speeds at around 470 kbps. GSM used Time Division Multiple Access (TDMA) and Frequency Division Multiple Access (FDMA) as its main multiplexing schemes.

In the mean time, during the mid 1990's, the Internet started witnessing increased popularity and adoption by residential end-users leading to further interest in data traffic, and richer digital media formats. It was inevitable that 3rd Generation (3G) cellular systems had to offer enhanced support for

data traffic, apart from their traditional support for voice services. The first 3G Universal Mobile Telephone Service (UMTS) standard, based on Code Division Multiple Access (CDMA), was published in 1999, offering peak data rates ranging from 384-2048 kbps [3]. An evolution of 3rd Generation (3G) systems arrived with the introduction of High Speed Packet Access (HSPA) with techniques such as Adaptive Modulation and Coding (AMC), fast dynamic scheduling and Hybrid Automatic Repeat Request (HARQ), delivering peak theoretical throughputs of 14.4 Mbps in Downlink (DL) and 5 Mbps in UL. Further revisions of HSPA, termed as HSPA+, provide further throughput and latency performance enhancements via higher order modulations, and Multiple Input Multiple Output (MIMO) support [3].

More demanding applications, and further traffic growth, led to the development of a 4th Generation (4G) system, with the specification of the Long Term Evolution (LTE) radio interface in 2009. LTE uses Orthogonal Frequency Division Multiplexing (OFDM) in DL and Single Carrier Frequency Division Multiple Access (SC-FMDA) in UL providing increased peak throughputs in the order of 300 Mbps in DL and 75 Mbps in UL, along with reduced latencies. Further enhancements to LTE were brought under Long Term Evolution Advanced (LTE-A) with the introduction of features such as Carrier Aggregation (CA) and Coordinated Multipoint (COMP), promising speeds of up to 1 Gbps in DL and 500 Mbps in UL. Nowadays, Long Term Evolution (LTE) is still in an evolutionary phase and it is expected to improve its performance via self-organizing networks, advanced receivers and further MIMO improvements. An evolved version of LTE, known as LTE-B is currently focusing its efforts on improving the performance of small cells.

A time line of these technologies is summarized and illustrated in Figure 1.2. The historical evolution of cellular systems presented above indicates that a new disruptive technology emerges every decade. This trend along with the expected data growth indicates that a 5th Generation (5G) system is expected around 2020. Projects related to 5G, such as the METIS [4] and 5GNOW [5] academic collaborative projects are already under way, and at the current time of writing there is abundant literature with regards to the requirements and visions for 5G.

1.1.3 5G Requirements

Increased traffic growth and heterogeneity, along with higher user expectations for faster services are the main drivers of 5G. High level requirements such as improved area capacity, better guaranteed edge rates, higher peak rates, lower latencies, faster set up times, lower energy consumption and re-

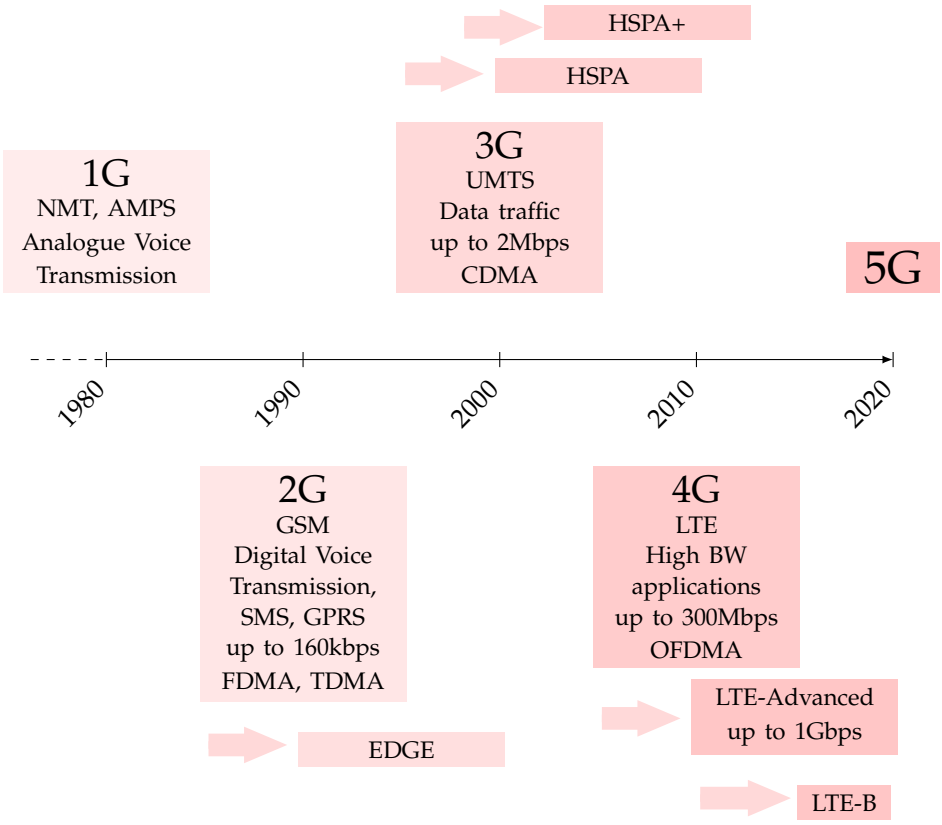


Fig. 1.2: Evolution of Cellular Systems

duced device and service costs, are some of the requirements which 5G is expected to satisfy [6–8]. Historical evolution of cellular networks dictates peak data rates in the order of 10 Gbps, latencies of 1 ms, and minimum guaranteed rates of 100 Mbps, for a forthcoming 5G system [9]. Further key requirements tied to future expected applications are mentioned in [10] indicating the need for servicing higher data volumes in dense urban areas, improving the battery life for a massive deployment of sensors and actuators, and providing better reliability for services such as traffic efficiency and safety. The huge amount of applications 5G is expected to handle, is also outlined in [11], where a set of application scenarios are classified in a hypercube according to their density, throughput and delay requirements.

At this point, it becomes seemingly challenging to satisfy all these requirements with a new single disruptive 5G Radio Access Technology (RAT), especially when one considers the wide variety of device types that are expected to be serviced by the 5G umbrella. To put matters into perspective, one can consider the diverse requirements of a sensor network where battery life is paramount, a user streaming high resolution video on his mobile device requiring high sustained data rates and a network of self driving cars requiring guaranteed latencies and network reliability for safety concerns. It is therefore envisioned by [12–14], that 5G will form a multi-tier architecture consisting of evolved existing technologies, complemented with new RAT's handling specific use cases.

1.1.4 5G Key Technology Candidates

After presenting the 5G high level requirements, the question of how to satisfy such requirements naturally arises. In principle, the system capacity can be boosted via three main approaches as depicted in Figure 1.3; by extending the system bandwidth, improving the spectral efficiency and via further cell densification. The precise contribution of each approach is undeniably a matter of debate, but the general consensus is that cell densification will be the key contributor in addressing the high level requirements of a future 5G system. The main justifications for this statement will be outlined in the following sub sections.

Spectrum Extension

As pointed out above, one approach to increase the system capacity is to simply increase the system bandwidth, a resource which is however scarce in the usable and well established cmWave frequencies. cmWave frequencies,

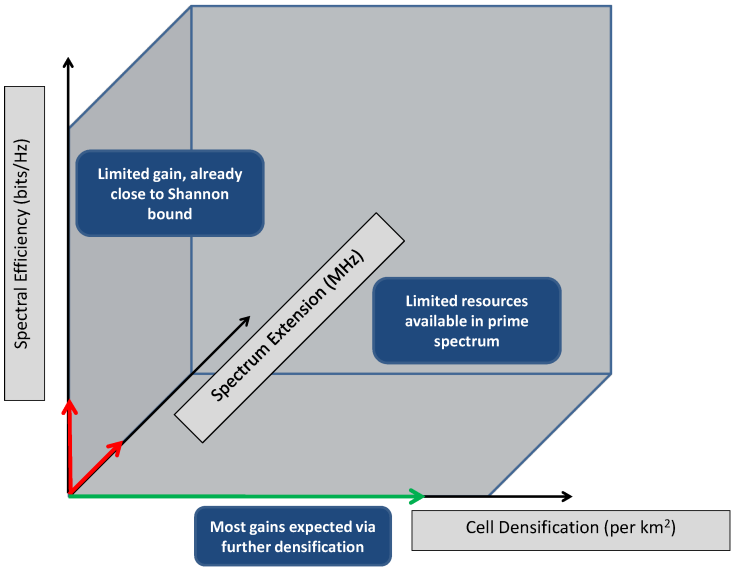


Fig. 1.3: Options for improving the system capacity

especially in the sub 3 GHz region, are typically desirable in cellular networks due to their favourable propagation characteristics. Given the severe shortage of available spectrum in this frequency region, different licensing paradigms have in fact been explored. Spectrum licensing can be either exclusive, where one operator takes exclusive control of the spectrum, unlicensed, such as the Industrial, scientific and medical (ISM) band used by technologies like Wi-Fi and Bluetooth, or shared, either via opportunistic use or via lightly licensed schemes as in the Authorised Shared Access (ASA) [15]/Licensed Shared Access (LSA) [16] concept where under a regulatory framework a number of licensed parties can share the spectrum. Alternatively, spectrum can be managed by a broker and dynamically leased based on a temporal agreement [17].

With regards to future spectrum availability at these cmWave frequencies, the FCC is currently considering the introduction of a shared spectrum band in the order of 100 MHz for small cells in the region of 3.5 GHz [18]. While additional spectrum is of course always welcome, the amount of limited spectrum which is expected to become available will not be able to fulfil the 5G's capacity growth promise.

Alternatively, ample amounts of untapped idle spectrum can be harvested from the mmWave 60 GHz bands. Initial studies on the potential of mmWave frequencies have already been made in [19, 20] showing promising gains

with the appropriate cell densification and beam directionality which can be obtained via large antenna arrays. While attractive from a spectrum availability point of view, mmWave frequencies do provide several challenges due to high propagation losses, and the necessity of Line Of Sight (LOS) between a transmitter and receiver pair.

Improving the Spectral Efficiency

An additional way to meet the capacity requirements of 5G is to improve the spectral efficiency. There is however limited gain in adopting such an approach since current systems such as LTE are already operating close to the Shannon capacity bound.

A promising technology which can further improve the spectral efficiency of a system is MIMO technology. A new complementary disruptive RAT should in fact exploit the abilities of such a key technological component since its inception. The availability of multiple antennas can be exploited to either increase the system throughput or to suppress interference at low frequencies [21], given an appropriate system design. At cmWave frequencies, a default 4x4 MIMO configuration is deemed to be feasible by 2020 [7], with higher order MIMO schemes possible at higher mmWave frequencies due to the reduced antenna sizes.

While attractive, the limited number of antennas which can be placed on a device operating at low frequencies will impose a limit on how much one can gain from this technology, meaning that a 5G system cannot fully rely on the gains from such a technology to meet its capacity demands.

Cell Densification

The biggest promise in satisfying 5G's thirst for capacity lies in cell densification. Cell densification is a well known and proven method for increasing the capacity of a system [22, 23]. Over the historical evolution of wireless networks, a 1600x gain has been obtained by reducing the cell size [23, 24]. Smaller cells enable reuse of the scarce spectrum and also entail a lower number of serviceable nodes per base station. The first generation of cells spanned over hundreds of square kilometres with newer generations incrementally reducing the cell size to a few hundred meters [6]. This trend will continue to prevail, emphasizing the importance of small cells in a future 5G system. Moreover, [25] shows that most of the traffic is generated indoors reinforcing the suitability of indoor small cells having even smaller ranges. These small

cells, sometimes referred to as femtocells [26], have in fact received considerable interest by both the industrial and academic world, dealing with problems related to cross-tier interference, interference management and coordination, access policies and load balancing. Taking it to the extreme, studies in [27, 28], have also attempted to find algorithms which identify and cache popular digital content at the femto cell base stations, improving the user experience and reducing the load in the core network.

However, introducing such a massive number of small cells brings along a number of challenges which need to be addressed. The diversity of device and traffic requirements imply the need of an-ever growing heterogeneous network, and multi-RAT association and mobility will be challenging tasks that need to be addressed in the presence of a large number of small cells [6]. Device cost, an aspect which has been mentioned earlier on, also deserves appropriate attention. In order to reach a mass market, the base station cost will need to be reduced significantly. The deployment aspects of such a large number of small cells also needs to be addressed, since it is unmanageable and infeasible to strategically deploy such systems. There are numerous reasons supporting this statement. Apart from the huge effort required in manually deploying a large number of small cells strategically, one must not overlook the availability of power, backhaul and installation permissions. Such aspects will of course play a decisive and fundamental role in the final geographical placement of the cell. It is therefore expected that these deployments will have a more random nature, potentially leading to persistently unfavourable channel conditions. The random and massive deployment of such cells also entails added inter-cell interference requiring the adoption of semi or fully distributed interference coordination or suppression techniques. A lower number of users per cell will also require appropriate flexible allocation schemes between UL and DL links, due to an increased temporal probability of traffic unbalance between the link directions.

1.2 Scope of the Thesis

As evident by the preceding section, there are numerous challenges and research opportunities along the road to 5G. The scope of this thesis is limited to study the suitability and potential of fully flexible Time Division Duplex (TDD) allowing an arbitrary UL or DL link allocation on a per frame/slot basis, as envisioned in the cmWave concept described in [8]. The cmWave concept is a disruptive 5G small cell solution operating at lower centimetre wave frequencies, aiming at providing 10 Gbps peak throughput data rates and reduced latencies in the order of 1 ms in local area scenarios. Such targets

are achieved via a suitable and shorter frame structure, inherent and proper support for interference suppression receivers, a minimum MIMO order configuration of 4x4 and fully flexible TDD support.

Fully flexible TDD allows us to quickly capture any traffic unbalance between the UL and DL links, and allocate the time slot resources accordingly. Such a feature is attractive in small cells, where an Access Point (AP)/base station typically serves a smaller number of User Equipment (UE)'s as opposed to larger macro cells. This added flexibility does however come at the cost of added inter-cell interference variation, and although typically neglected, such an aspect should not be overlooked when studying the gains originating from Flexible TDD (F-TDD). This added inter-cell interference variation can in fact have negative effects on link adaptation schemes which need to estimate the future channel conditions based on previous measurements for correct operation. Failure to do so may cause excessive unnecessary retransmissions if such a side-effect is not controlled accordingly, hindering the full potential of F-TDD. Hence, in our study, such intermediary components are modelled explicitly to assess the realistic gains originating from fully flexible TDD.

It is expected that F-TDD should offer significant gains over a fixed Static TDD (S-TDD) scheme at low load, mainly due to the high probability of temporally asymmetric links and the statistically low presence of interference. The gains at higher load are somewhat unclear due to the potential presence of inter-cell interference variation introduced by Flexible TDD (F-TDD). It is however believed that the presence of interference suppression receivers operating under a proper system design, can address this issue.

Under asymmetric traffic conditions, it is foreseen that F-TDD can offer significant gain in the lightly loaded link, and limited gain in the highly loaded link, when compared to a fixed S-TDD scheme which can predict the long term traffic asymmetry. The operation of F-TDD with higher layer protocols like TCP is also presupposed to be beneficial, since the F-TDD operation can be adjusted to accommodate the requirements of such protocols.

The scope of this thesis is to validate these hypotheses, and can be summarized as follows,

1. Assess the gains of F-TDD in the envisioned 5G small cell concept
2. Assess the role of interference suppression receivers to deal with the added inter-cell interference variation
3. Develop UL/DL link scheduling algorithms

4. Quantify the gains of F-TDD at different loads and different UL/DL traffic asymmetries
5. Investigate the potential advantages of F-TDD when operating over TCP traffic requiring reverse link feedback acknowledgements
6. Develop a performing distributed rank adaptation algorithm for the 5G small concept, which potentially induces added inter cell interference variation and reduced efficacy in dealing with such variation, and analyse the gains of F-TDD under these more realistic conditions
7. Give recommendations on whether fully flexible TDD is a suitable technology for future 5G small cells

1.3 Research Methodology

The goal of this thesis is to study the suitability of F-TDD in a future 5G small cell system as envisioned in [29]. The investigations are carried out using a bottom-up approach. Initially, tools from queueing theory are used, providing a rough expectation of the gains available from F-TDD. The core of the studies are then conducted via the usage of a discrete-event system level simulator in order to capture the interaction between various building blocks of the system and hence obtain a more realistic result with regards to the gains expected from F-TDD.

The main gains of F-TDD arise from the statistical temporal presence of traffic in a single link direction, i.e. UL or DL. Given a particular bursty traffic model and a simplified scenario, an appropriate queueing system model can be extracted, allowing us to assess the analytical gains arising from F-TDD, in terms of the expected transfer time of a file. While this approach insulates a lot of aspects, such as varying channel conditions due to interference, it provides a rough expectation of the gains arising from F-TDD. Moreover, it also serves as a validation tool ensuring the correctness of the implementation in the discrete-event system level simulator.

While valid, simply considering such a theoretical approach will entail a number of limitations and hide any potential demerits of F-TDD. In more realistic scenarios with multiple cells, the full freedom of assigning each time slot as UL or DL can in fact introduce added inter-cell interference variation, making it problematic for link and rank adaptation schemes to adapt. This situation will increase the triggering of recovery mechanisms, inducing increased retransmissions at the Hybrid Automatic Repeat Request (HARQ)

and Radio Link Control (RLC) layers. Modelling these blocks and understanding their interaction with F-TDD is therefore essential in outlining the true benefits of this feature.

Unfortunately, the development of a mathematical optimization taking into consideration the interaction of all these building blocks becomes increasingly difficult such that assumptions which mask the impact of some of the aforementioned features would need to be taken. For this reason, the work presented in this thesis is mainly conducted using a discrete-event system level simulator, on which several relevant features were implemented.

A framework revolving around the 5G small cell concept [29] was firstly developed, taking into account the envisioned frame structure allowing fully F-TDD. An appropriate indoor Multiple Input Multiple Output (MIMO) channel was then introduced, allowing us to study the role of interference suppression receivers and rank adaptation in the envisioned 5G concept. Link adaptation, delayed channel estimates and retransmission mechanisms such as Hybrid Automatic Repeat Request (HARQ) and Radio Link Control (RLC) were also introduced, in order to understand the performance impact and interaction of packet loss due to increased inter-cell interference variation introduced by F-TDD, and these retransmission mechanisms. The application layer was modelled using the bursty 3GPP FTP traffic models [30]. The different parametrizations of these models (i.e. session size, inter-arrival time) were used to study the regions over which F-TDD shows promising gains. The TCP New Reno [31] protocol was also explicitly modelled, in order to understand the benefit of F-TDD over such protocols.

The simulator is built in a modular fashion, allowing a smooth incremental introduction of features. Test-driven development is an essential aspect of the simulator development, and testing and validation of the individual blocks was carried out via unit-testing, leading to a development cycle which promoted flexibility and code re-use. Larger scale system tests were also conducted periodically ensuring that the introduction of new interacting features did not break any of the logic from the previous code.

1.4 Contributions

The main contributions of this thesis consist of work derived from the attached papers in Part III, along with additional material specifically introduced for this thesis.

In terms of additional generated material which is absent from the attached

papers, this thesis provides,

1. **A Simplified Analytical Model based on Queueing theory**

   An analytical model derived from queueing theory was extracted in order to assess the gains of flexible TDD over fixed TDD. By taking into account the effects of buffering, it was analytically shown that flexible TDD can reduce the mean expected file transfer time by half, under ideal conditions. The results from the model were also compared to simulation results and a good match was obtained.

2. **A Load-Fairness Based UL/DL Scheduling Algorithm**

   A load-fairness based UL/DL scheduling algorithm which considers the current buffer size statuses in UL and in DL, and the previously assigned TDD slot allocations was developed. Its performance was analysed and compared to the delay-fairness based UL/DL scheduling algorithm provided in Paper A. It was shown that the load-fairness based scheme can adapt better to traffic asymmetry.

3. **An in-depth analysis of Flexible TDD vs Semi-Static TDD vs Fixed TDD**

   A semi-static TDD scheme based on the LTE eIMTA TDD configurations was introduced and benchmarked against the flexible and fixed TDD schemes in order to provide further depth to the results. It was observed that in most cases, the semi-static TDD scheme operates better than the fixed TDD scheme. However, it was also shown that the provided flexible TDD scheme outperforms the semi-static TDD scheme due to its inherent limited flexibility.

The publications relevant to this thesis are attached in Part III and have been authored during the course of the PhD. The main description of the studies and their corresponding findings, is listed below,

1. **The Potential of Flexible UL/DL Slot Assignment in 5G Systems. / Catania, Davide; Gatnau, Marta; Cattoni, Andrea Fabio; Frederiksen, Frank; Berardinelli, Gilberto; Mogensen, Preben. Vehicular Technology Conference (VTC Fall), 2014 IEEE 80th . IEEE, 2014. s. 1-6.**

   This paper gives an initial insight into the potential benefits of flexible TDD within the envisioned 5G small cell concept framework. An analysis on the role of interference suppression receivers, more specifically Minimum Mean Square Error (MMSE) Interference Rejection Combining (IRC) receivers, is presented, showing their efficacy in dealing with

inter-cell interference variation. A simple flexible UL/DL algorithm is then provided, showing the gains of flexible TDD at 30% and 70% cell load. In this paper it is shown that fully flexible TDD can almost double the experienced session throughput and halve the corresponding session delay.

2. **Flexible UL/DL in Small Cell TDD Systems: A Performance Study with TCP Traffic. / Catania, Davide; Gatnau, Marta; Cattoni, Andrea Fabio; Frederiksen, Frank; Berardinelli, Gilberto; Mogensen, Preben. 2015 IEEE 81st Vehicular Technology Conference: VTC2015-Spring. 2015.**

   In this paper we study the interaction and potential benefits of flexible TDD with the TCP protocol. We first show that in a unidirectional transfer, for small file sizes, TCP requires fast acknowledgement feedback, since this allows the congestion window size to grow quickly, giving the sender the possibility to transmit more data at once. Conversely, for large file sizes, a highly asymmetric TDD configuration can be beneficial. Subsequently we show how a properly parametrized flexible UL/DL algorithm, delaying a switch in the opposite direction, can in fact follow such protocol dynamics and provide satisfactory performance for both small and large file sizes. Finally this behaviour is confirmed via multi-cell system level simulations, indeed proving that the flexibility introduced in TDD can be beneficial for TCP.

3. **A Distributed Taxation Based Rank Adaptation Scheme for 5G Small Cells. / Catania, Davide; Cattoni, Andrea Fabio; Mahmood, Nurul Huda; Berardinelli, Gilberto; Frederiksen, Frank; Mogensen, Preben. 2015 IEEE 81st Vehicular Technology Conference: VTC2015-Spring. 2015.**

   This paper proposes a distributed interference aware rank adaptation algorithm, investigated and specifically designed for the envisioned cmWave 5G small concept. The rank adaptation algorithm controls the transmission rank i.e. number of spatially multiplexed streams, in accordance to a utility function which takes into account the channel conditions and an imposed tax based on the transmission rank and the currently perceived interference levels. Extra attention in considering realistic simulation assumptions, capturing both the strong and weaker compatible aspects of the 5G small cell concept, are taken. System-level simulations show that if adequately parametrized, the algorithm can almost match the best possible outage performance and at the same time enjoy higher average peak throughputs, at both low and high load.

Additionally the follow publications have been co-authored.

1. Centimeter-wave concept for 5G ultra-dense small cells. / Mogensen, Preben; Pajukoski, Kari; Tiirola, Esa; Vihriälä, Jaakko; Lähetkangas, Eeva; Berardinelli, Gilberto; Tavares, Fernando Menezes Leitão; Mahmood, Nurul Huda; Lauridsen, Mads; Catania, Davide; Cattoni, Andrea Fabio. IEEE 79th Vehicular Technology Conference (VTC) 2014-Spring. IEEE, 2014.

2. Improving Link Robustness in 5G Ultra-Dense Small Cells by Hybrid ARQ. / Gatnau, Marta; Catania, Davide; Frederiksen, Frank; Cattoni, Andrea Fabio; Berardinelli, Gilberto; Mogensen, Preben. (ISWCS), 2014 11th International Symposium on Wireless Communications Systems. IEEE, 2014. s. 491-495

3. Dynamic Outer Loop Link Adaptation for the 5G Centimeter-Wave Concept. / Gatnau, Marta; Catania, Davide; Frederiksen, Frank; Cattoni, Andrea Fabio; Berardinelli, Gilberto; Mogensen, Preben. European Wireless (EW) 2015. IEEE, 2015. (IEEE 18th European Wireless Conference).

4. Full Duplex Communication Under Traffic Constraints for 5G Small Cells. / Gatnau, Marta; Catania, Davide; Berardinelli, Gilberto; Mahmood, Nurul Huda; Mogensen, Preben. IEEE 82nd Vehicular Technology Conference (VTC) 2015-Fall. IEEE, 2015.

5. A Multi-QoS Aggregation Mechanism for Improved Fairness in WLAN. / Gatnau, Marta; Catania, Davide; Cattoni, Andrea Fabio; Ashta, Jagjit Singh; Mogensen, Preben. Vehicular Technology Conference (VTC Fall), 2013 IEEE 78th . IEEE, 2013. s. 1-5 (IEEE Vehicular Technology Conference. Proceedings).

6. IEEE 802.11 Networks: A Simple Model Geared Towards Offloading Studies and Considerations on Future Small Cells. / Garcia, Luis Guilherme Uzeda; Rodriguez, Ignacio; Catania, Davide; Mogensen, Preben. 2013 IEEE Vehicular Technology Conference (VTC Fall). IEEE, 2013.

7. Managing inter-cell interference with advanced receivers and rank adaptation in 5G small cells. / Tavares, Fernando Menezes Leitão; Berardinelli, Gilberto; Catania, Davide; Sørensen, Troels B.; Mogensen, Preben. European Wireless (EW) 2015. IEEE, 2015. (IEEE 18th European Wireless Conference).

8. Experimental Evaluation of Interference Suppression Receivers and Rank Adaptation in 5G Small Cells. / Assefa Wassie, Dereje; Berardinelli, Gilberto; Catania, Davide; Tavares, Fernando Menezes Leitão; Sørensen, Troels B.; Mogensen, Preben. IEEE 82nd Vehicular Technology Conference (VTC) 2015-Fall. IEEE, 2015.

1.5 Thesis Outline

The thesis is organized as follows. In Chapter 2, we will give a brief overview of the envisioned 5G small concept [8], highlighting important details relevant to our study. This will include details about the frame structure and its role in supporting key technology components such as advanced receivers to suppress interference, and fully flexible TDD on a per slot/frame basis.

In Chapter 3, we will introduce in further detail the concept of flexible TDD by first presenting the available literature with regards to topic, and thereafter provide an analytical model which shows the expected gains of flexible TDD in an isolated single cell scenario. Thereafter, we will dedicate a section to assess the potential demerits which might be introduced by this feature, and show how interference suppression receivers can deal with such problems.

After having presented the flexible TDD framework, in Chapter 4, we will present two algorithms for choosing the UL and DL direction based on the instantaneous traffic conditions. We will first present a simple algorithm which triggers a switch in the link direction based on the head-of-line delay and buffer size of each link, and thereafter we will present another link scheduling algorithm which takes into account the traffic share present in each link direction, and the previous time slot allocations made to each link, resulting in a more suitable algorithm when high traffic asymmetries are present.

Chapter 5 will be dedicated to present a distributed rank adaptation algorithm that can operate on top of our 5G concept. While the gains and efficacy of flexible TDD prove to be valid in a robust air interface which dedicates all its spatial resources for interference suppression, it is important to benchmark the system in a more realistic scenario where multiple nodes could be using their spatial resources to transmit at higher rates in lightly interfered scenarios.

Chapter 6 will be focused on the performance evaluation of the system. We shall show the behaviour and gain of flexible TDD for different traffic asymmetries, different traffic loads, and different transport protocols such as User Datagram Protocol (UDP) and Transmission Control Protocol (TCP). We shall also discuss the benefits of TDD flexibility for higher level protocols requiring acknowledgements. Here we will also show the performance of fully flexible TDD versus semi-static TDD as present in LTE-TDD eIMTA.

Finally, Chapter 7 will be dedicated to outline the main concluding points of the presented work, and give an outlook with regards to potential future work.

Chapter 2

5G System Overview

In this chapter we will give a more detailed overview of our envisioned 5G concept [29] highlighting the most relevant features related to the specific studies carried out in this thesis. We shall first outline how the features introduced in this system satisfy some of the aforementioned 5G key requirements outlined in Chapter 1. In the next subsections, we will then delve deeper into the most relevant aspects to our work, by delivering further details related to the system assumptions and considered key technology components. Thereafter we will show how these key technology components influence our envisioned 5G frame design, by first showing the characteristics of the frame, the associated numerology, and finally, the related UL and DL procedures used to access the resources offered by the system.

2.1 Key Features

In this section we will specifically focus on how the main features introduced in the 5G small cell concept address some of the 5G key requirements mentioned earlier on. Some fundamental requirements which are addressed by this new RAT include the delivery of higher peak throughputs, lower latencies, improved outage performance, flexible operation, lower cost and reduced power consumption.

The need to fulfil these requirements is targeted via the introduction of several design solutions such as interference mitigation, distributed synchronization, the use of Orthogonal Frequency Division Multiplexing (OFDM)

in both UL and DL, flexible TDD, and MIMO. Appropriate support for these solutions is achieved via the introduction of a new optimized frame structure.

The new frame structure features a duration time of 0.25 ms enabling the possibility of achieving lower latencies and shortened Round Trip Time (RTT)'s. It also includes the necessary support for MIMO and advanced receivers via a special Demodulation Reference Symbol (DMRS). Flexible resource allocation, a desired feature whose applicability and performance is investigated in this work, is supported by allowing fully flexible TDD on a per frame basis, such that each frame can be allocated to UL and DL. The logical design of the frame introduces a well defined separation between the control and the data parts of the frame, providing additional benefits in terms of power consumption.

Higher peak throughputs and increased robustness against inter-cell interference, is achieved with MIMO systems used in tandem with advanced receivers utilizing Minimum Mean Square Error (MMSE) Interference Rejection Combining (IRC). In low interference scenarios, the number of transmitted spatial streams can be increased allowing for higher peak throughputs. Conversely, when the interference levels are strong, the system can decide to suppress this interference using its spatial degrees of freedom, guaranteeing a minimum throughput target. This mechanism allows for a more robust air interface [32] and aims at improving the outage performance of the system.

These aspects will be discussed in further detail in the corresponding subsections. A table summarizing the requirements and the potential key technological components used to address them is presented in Table 2.1.

2.2 System Considerations

In this section, some general design choices will be justified and motivated accordingly. In the subsequent sections, we will then expand on the most relevant key technology components and design solutions.

The first design choice relates to the usage of TDD as the principle duplexing method. The choice is motivated because of the several advantages it offers over FDD. Firstly, in TDD systems, there is no need for paired spectrum as in FDD. TDD also adapts well to unbalanced traffic scenarios via simple reconfiguration of the UL and DL resources, making it a useful feature for small cell scenarios where the number of active users is typically much lower than in macro scenarios. Some further advantages which TDD systems offer over FDD, include lower component cost, and channel reciprocity between

Requirement	Key Features
Lower Latencies	Shortened Frame Structure
Higher Peak Throughputs	MIMO, Larger bandwidths, flexible TDD
Guaranteed Minimum Throughput	Robustness to inter-cell interference via additional spatial streams and MMSE-IRC, supported by appropriately designed frame structure
Lower Power Consumption	Separation between the control and the data parts of the frame

Table 2.1: Summary of 5G small cell requirements and key features addressing them

UL and DL [33, 34].

The timing alignment of neighbour cells is also assumed, enabling the possibility of estimating the complex channels of nearby interferers. The alignment can in principle be achieved via the high precision reference clock provided by Global Position System (GPS) satellites but the typical indoor penetration losses of small cells may significantly disrupt the reliability of this approach. However, by introducing the exchange of periodical beacon messages between networks, studies conducted in [35, 36] have proved the feasibility of achieving and maintaining the required synchronization in a distributed manner.

The envisioned 5G system also considers the usage of Orthogonal Frequency Division Multiplexing (OFDM) as its preferred multiplexed scheme. While there are several candidates for future potential 5G waveforms, OFDM was selected due to some inherent advantages related to its low implementation complexity when compared to competing emerging technologies such as filter bank multi-carrier (FBMC). Moreover, OFDM lends itself easily to MIMO extension, an important key technology component in our 5G system. It is also robust to hardware impairments typical of low end devices [37, 38].

Another design choice relates to the usage of the same OFDM frame format for both UL and DL. The possibility of flexible TDD introduces additional interference types. While traditionally, only AP-UE and UE-AP interference is present, flexible TDD may potentially introduce new types of interference, known as same-entity or cross-link interference. Therefore, the usage of the

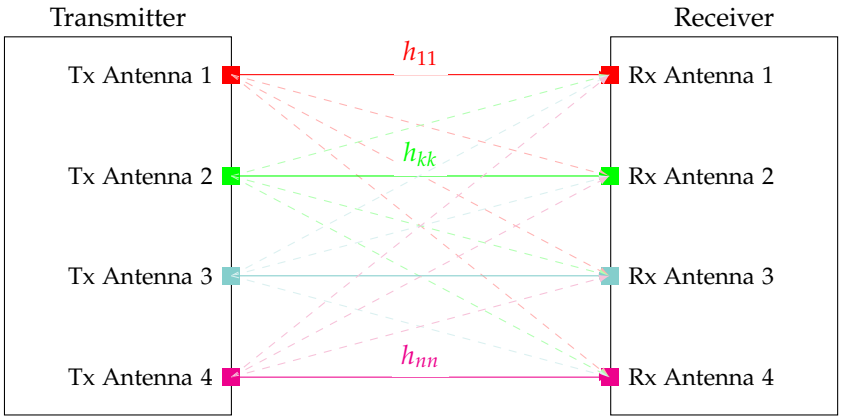


Fig. 2.1: MIMO 4x4 Scheme

same frame format in both link directions, aids in appropriate interference estimation, independently of whether the interference is originating from UL or DL.

2.3 MIMO and Advanced Receiver Support

MIMO and advanced receiver support is a central aspect of the envisioned 5G concept [29]. The availability of multiple antennas and MIMO allows us to increase our peak throughput performance in favourable conditions. This enables us to come closer to our high peak data rate requirements in future 5G systems. On the other hand our antenna resources can instead be used to suppress interference in challenging interference conditions, whenever advanced receivers capable of suppressing interference are applied. Therefore the additional antennas at our disposal give us the ability to configure the system to adapt to certain conditions, such that we can enjoy the best of both worlds, whether it is in guaranteeing a minimum throughput performance or satisfying peak throughput requirements.

In the envisioned 5G concept described in [8] it is assumed that by default, in 2020, most terminals will be at least equipped with a 4x4 antenna configuration system. Such a configuration gives a considerable amount of degrees of freedom. Figure 2.1 shows a small depiction of a 4x4 MIMO system operating over a transmitter-receiver pair. While systems operating at mmWave frequencies, can potentially increase the scale of such a MIMO configuration, due to the reduced antenna size at higher frequencies, at lower cmWave fre-

quencies, such as in the envisioned 5G small cell concept, a high order MIMO configuration is typically infeasible due to terminal size constraints. For this reason a 4x4 configuration offers a good compromise and it is expected that higher order configurations exceeding 8x8 will not be suitable for such systems operating at cmWave frequencies.

The usage of advanced interference suppression MMSE IRC receivers are expected to be operational at the AP, as well as in the UE's, within the envisioned 5G concept. A key characteristic for effectively exploiting advanced receiver techniques employing interference suppression is linked to an appropriately designed frame structure. LTE was not built from the ground up to support such functionality and the envisioned 5G frame structure described in section 2.5 attempts to overcome this limitation. The design of the frame structure and the proper support for advanced receiver techniques such as MMSE IRC becomes even more important in the light of fully flexible TDD and hence increased probability of cross-link interference, since in such a situation, the interference channel conditions can easily change from one frame to the next.

2.4 Flexible TDD Support

Flexible TDD is not a new concept and limited support for this feature has already been introduced in other systems such as LTE-TDD (in LTE Enhanced Interference Mitigation & Traffic Adaptation (LTE-eIMTA)), and WiMAX [39]. Its importance becomes even more crucial when one considers its application in small cell systems where the number of users is typically low, and the traffic requirements between UL and DL can shift more rapidly over time.

WiMAX is one system amongst others, allowing partial flexible TDD operation. WiMAX defines a superframe of 20 ms, which is then further subdivided in four frames of 5 ms each. This frame is subdivided into 8 sub frames. The frame structure of WiMAX is shown in figure 2.2. Each 5 ms frame's switching point can be varied to represent either a DL or UL heavy configuration, but the DL and UL slots are always separated by a flexible switching point.

LTE-eIMTA also introduces a limited set of flexibility in allocating its resources. This is done by defining a set of 7 TDD configurations with different DL-UL traffic asymmetries. These configurations are shown in Table 2.2. The configurations can be reconfigured every x ms where studies conducted in [40] have shown that there is additional benefit in lowering this reconfiguration time. The available frame configurations allow the DL to UL asymmetry

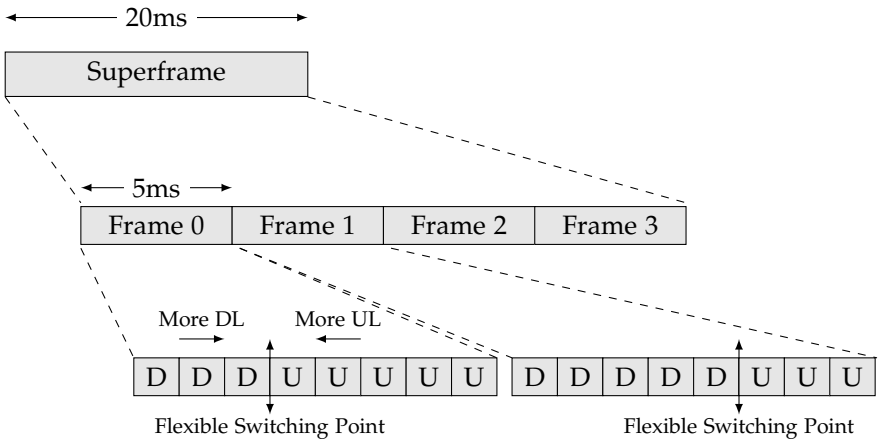


Fig. 2.2: WiMAX Frame Structure

to vary from 40% to 90%. Some of the subframes in Table 2.2 are aligned, while others are not, introducing the problem of cross-link interference, an aspect which introduces numerous challenges with regards to interference estimation, given LTE's different multiple access schemes in UL and DL.

Configuration	Subframe Number										DL-UL Asymmetry
	0	1	2	3	4	5	6	7	8	9	
0	D	S	U	U	U	D	S	U	U	U	$0.25 - 0.40^1$
1	D	S	U	U	D	D	S	U	U	D	$0.50 - 0.60^1$
2	D	S	U	D	D	D	S	U	D	D	$0.75 - 0.80^1$
3	D	S	U	U	U	D	D	D	D	D	$0.67 - 0.70^1$
4	D	S	U	U	D	D	D	D	D	D	$0.78 - 0.80^1$
5	D	S	U	D	D	D	D	D	D	D	$0.89 - 0.90^1$
6	D	S	U	U	U	D	S	U	U	D	$0.38 - 0.50^1$

Table 2.2: LTE-TDD Configurations, D = Downlink Frame, U = Uplink Frame, S = Special Frame

Understanding the merits and demerits of a fully flexible TDD system within this envisioned 5G concept lies at the heart of this thesis. Unlike other systems, allowing limited flexibility, the envisioned 5G system allows full flexibility on a per frame basis, such that each frame can be allocated as UL or DL. This flexibility provides benefits in terms of traffic adaptation but will also undoubtedly create additional inter-cell interference variation, since a cell's neighbours could be shifting its allocation from UL to DL at its convenience

[1]Valid if special subframe is considered as DL frame

to deal with its own traffic unbalance scenario. This issue has an impact and presents numerous challenges to link adaptation, or any algorithm which relies on the short-term stability of the channel conditions, and therefore the applicability of flexible TDD needs to be carefully investigated. This drawback, as we will see in the coming chapters, can somehow be dealt with by the usage of MMSE-IRC receivers, for which our frame structure, as it will be seen in Section 2.5, provides appropriate support to operate sufficiently well.

2.5 Envisioned 5G Frame Structure

A new frame structure, used as a base for our investigations, was proposed in [33]. This frame structure serves as a foundation for the envisioned 5G small cell concept and was specifically designed for local area systems. The structure was influenced by the previously mentioned design solutions and physical layer technology components, which are expected to be an important part of a new 5G RAT.

A pictorial representation of the frame structure, is shown in Figure 2.3. The frame has a short duration of 0.25 ms and is logically divided as follows; out of the 14 symbols available in the frame, 2 symbols are reserved for the DL and UL control channel, and 1 symbol is reserved for the DMRS. The rest of the symbols are dedicated to data part of the frame.

The frame explicitly introduces time separation between the control and the data part of the frame, allowing a device to turn-off its receiver chain for the rest of the frame, if no information is received in the corresponding control part [41]. This reduces the power consumption of the devices, a key desired feature for future 5G systems.

Fully flexible TDD allocation, another desired feature of future 5G systems, is also allowed in the given frame structure. Each frame can in fact be configured to operate in either UL or DL, and the link direction can be switched on a per-frame basis, giving the system the possibility to swiftly capture any traffic fluctuations present in the system. Please note that the link direction is assumed to be stable within each frame, meaning that no switch in the link direction is allowed within the frame itself.

Support for interference mitigation, an important design solution for small cells typically experiencing high inter-cell interference levels, is introduced via two mechanisms; interference coordination, and interference suppression via advanced receivers.

Interference coordination is enabled by strategically selecting and avoiding frequency channels over which to transmit data, in an attempt to lower the perceived interference of tightly coupled cells. The envisioned 5G system enables the support of this feature by splitting the available spectral resources into different frequency resource blocks, and by allowing the selection of the frequency resource blocks used for transmission.

Good use of interference suppression receivers is enabled via several system assumptions and features of the frame. For effective operation of advanced interference suppression receivers such as MMSE IRC, an updated estimation of the Interference Covariance Matrix (ICM) needs to be available at the receiver, such that the receiver can adjust its weights accordingly to suppress interference. However, the introduction of flexible TDD operation, can significantly alter the perceived interference conditions from one frame to the next. In order to account for this, and provide updated estimates of the ICM to the receiver, the proposed frame structure introduces a DMRS symbol. During this symbol, all the nodes scheduled in the data part will transmit simultaneously, allowing the MMSE IRC capable receivers, to distinguish and identify the desired and interfering channel matrix, such that the appropriate weights can be applied to suppress this interference. This mechanism allows for accurate ICM estimation independently of the link direction applied at each cell. Please note that this approach is only possible due to the usage of the same frame format in UL and in DL, and the introduced restriction of not allowing the system to switch the link direction within the frame itself. The combined use of flexible TDD and advanced interference suppression receivers is in fact much more challenging in LTE-TDD eIMTA due to the absence of these assumptions and lack of underlying system support.

2.5.1 5G Numerology

In this section we will describe the assumed numerology [41] for the envisioned 5G system.

The frame shown in Figure 2.3 has a duration of 0.25 ms, and is sub-divided in time and in frequency. The frame consists of 14 symbols, and each Physical Resource Block (PRB) is assumed to span a whole frame in time, and 165 sub-carriers in frequency. The sub-carrier spacing and corresponding symbol time, is assumed to be 60 kHz and 16.67 us respectively. With this configuration, each PRB approximately represents 10 MHz. The system contains 20 PRB's for an assumed system bandwidth of 200 MHz.

A short comparative overview of the 5G and LTE-TDD numerology is sum-

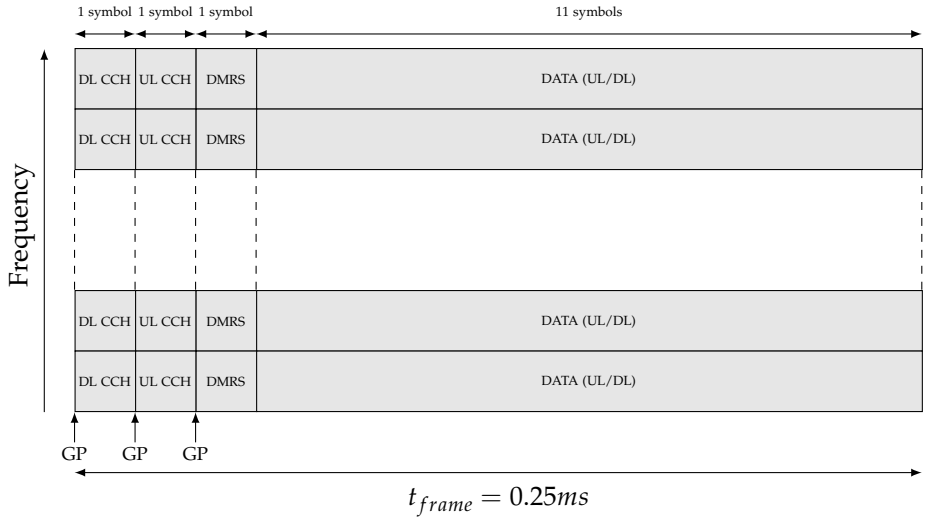


Fig. 2.3: 5G Frame Structure

marized in table 2.3.

A quick inspection at Table 2.3 shows numerous differences between the two systems. The 5G frame duration is four times shorter than LTE-TDD's frame duration, making the 5G system more attractive from a latency and RTT minimization point of view. The shrinking of the frame duration was done by increasing the sub carrier spacing from 15 kHz to 60 kHz, hence obtaining a shorter symbol time. Shortening the symbol time increases the relative overhead of the cyclic prefix to the symbol duration but since this frame structure was designed for small cell systems, lower delay spreads are expected, and therefore compared to LTE-TDD, the cyclic prefix can be shortened significantly without worrying too much about inter-symbol interference. Within the UL-DL switching points of the frame itself, a Guard Period (GP) is typically required to avoid power leakage between the transmitter and the receiver of a node. Given the lower transmit powers of small cell systems and hence the shorter associated rise and fall times, along with the evolution of technological components, it is assumed that the GP's duration can also be shrunk considerably when compared to LTE-TDD.

Additional differences between the two systems lie in the granularity of the systems' PRB's. The LTE-TDD system is much more granular than the envisioned 5G system. Since the 5G system targets smaller cells with fewer number of UE's, the granularity of the system's resources can be relaxed in

	LTE-TDD	5G
Number of Symbols per Frame	14	14
Sub carrier Spacing	15 kHz	60 kHz
Symbol Time	66.67 us	16.67 us
CP duration	4.7 us (short)	1 us
GP duration	66.67 us (min)	0.89 us
Frame Length	1ms	0.25 ms
Subcarriers per PRB	12	165
PRB Allocation (BW)	180 kHz	10 MHz
System Bandwidth	1.4 up to 20 MHz	100 or 200 MHz
TDD Flexibility	Set of TDD Configurations (DL:UL, 2:3 to 9:1)	Full Flexibility per Frame
HARQ Processes	up to 15	4

Table 2.3: Numerology Differences between 5G Frame Structure and LTE-TDD [33, 41]

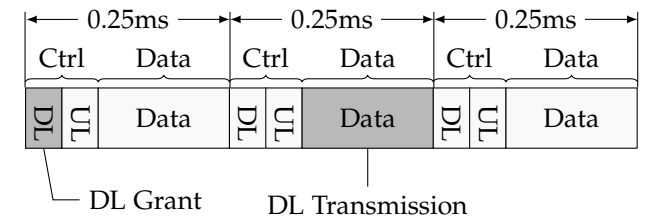
favour of simplified resource management. The shorter duration of the frame also compensates for this reduced granularity, as there are additional opportunities for allocating resources in the same given amount of time.

Other characteristics which differ between the two system are related to the increased TDD flexibility offered by the envisioned 5G system and the reduced number of HARQ processes employed. The reduced number of HARQ processes decreases the amount of buffer circuitry required, hence lowering the cost of the end user device [7], another key requirement of the 5G system.

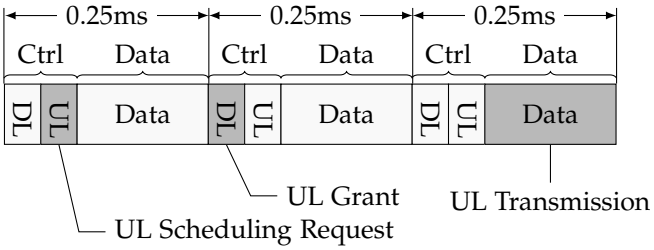
2.6 Downlink and Uplink Access Procedures

After having described the desired key technology components and their impact on the frame structure, in this section, we shall briefly describe the operation and signalling involved for a DL and UL transmission to take place, highlighting how the available system resources are accessed. The DL and UL procedures' timing diagrams are respectively depicted in Figures 2.4a and 2.4b. As in previous systems, it is assumed that the AP will be the final decision maker for allocating resources and scheduling users.

For a downlink transmission to occur the AP simply schedules a Grant in the Downlink control channel, and the corresponding transmission will take place on the subsequent frame, allowing suitable time for processing of the DL grant by the UE. The scheduling grant will include transmission param-



(a) DL Procedure



(b) UL Procedure

Fig. 2.4: DL and UL Access Procedures (DMRS symbol ommitted for clarity)

eter information such as the Modulation and Coding Scheme (MCS) to use, the PRB's on which the transmission is scheduled, the transmission rank and Precoding Matrix Information (PMI). It is immediately noticeable that given the short frame structure discussed previously, the minimum RTT can also be shortened significantly.

The uplink transmission procedure takes an additional frame to be completed. The UE wishing to transmit will send a Scheduling Request (SR) in the uplink control channel. Subsequently the AP will receive this information, and in the best case scenario, opts for scheduling the UE via an UL grant sent in the DL control channel. The UE can then conduct its uplink transmission on the subsequent frame.

In total, for a DL transmission spanning only 1 frame, there is a 0.5 ms delay from when it is granted permission to transmit, until the corresponding transmission completes. For a similar UL transmission, there is a 0.75 ms delay from when the user requests a resource via a SR and the completion of the UL transmission.

2.7 Wrap-up

In this chapter we have described the most relevant key technology components envisioned by the 5G small cell concept [29], by showing how they address some of the key requirements of a future 5G system. Thereafter we have outlined the main system assumptions, and provided some detail in relation with the possibilities of MIMO and advanced receiver support, and flexible TDD. Once the technologies have been outlined, we presented the 5G frame structure, whose design and characteristics enable the usage of these same technologies. The design choices for the frame structure were highlighted and justified, showing how the proposed 5G frame structure satisfies the required design criteria for appropriate operation in 5G small cells. The differences between the envisioned 5G small concept and current RAT's such as LTE-TDD, were also highlighted in terms of characteristics, structure, and frame numerology. Once the system design and 5G frame structure was established, a section dedicated to the access procedures was presented, showing how the resources available to the system can be accessed. In the next chapter we will shift our focus on flexible TDD, the core topic of this work, keeping in mind the 5G framework described in this chapter.

Chapter 3

Flexible TDD

Fully flexible (dynamic) TDD [1] introduces a very simple paradigm allowing the system to freely assign each time slot as UL or DL. Given this flexibility, the system can then capture the instantaneous traffic conditions, and react accordingly. There are several advantages of utilizing a fully Flexible TDD approach. The availability of F-TDD can avoid manual TDD configuration, requiring long term statistical information related to the DL and UL traffic profiles experienced by a certain deployment. Even if such information is known and applied accurately, F-TDD can still offer the advantage of capturing rapidly varying traffic conditions, a situation which might not be uncommon in small cells where the number of served users is typically low. These advantages have already been acknowledged in some of the technologies available today, offering various degrees of flexibility.

While intuitively attractive, F-TDD also poses some challenges which need to be overcome. In particular, F-TDD introduces new types of interference, such as AP-AP, and UE-UE interference. These are typically termed as cross-link interference in the literature, and introduce significant problems which hinder the full potential of F-TDD. Apart from the issue of cross-link interference, F-TDD introduces additional inter-cell interference variation. This behaviour creates a significant complications when estimating the conditions of the channel, since future channel conditions can be temporally varying on a frame basis, due to the freedom available in scheduling the link direction.

The threat of increased inter-cell interference variation can in principle be

[1] Please note that the terms flexible and dynamic are used interchangeably when referring to flexible TDD

mitigated by the usage of advanced receivers such as MMSE IRC, which with the appropriate frame structure, have the ability to suppress interference independently from the source of the interferer. As a further step, we therefore assess the suitability and impact of such receivers in relation to the problem mentioned above, considering a realistic multi-cell system to investigate whether the flexibility brought along by F-TDD outweigh the presented demerits of such a system.

In this chapter, a short and general survey on the previous work related to F-TDD, will first be presented. Subsequently, we shall from a traffic perspective, analytically assess the expected gains of F-TDD over a fixed Static TDD (S-TDD) scheme which can predict the long term average traffic share between UL and DL. Thereafter, the demerits of F-TDD will be outlined. Further insight on the problem of inter-cell interference variation will be provided, and the negative effects of this behaviour will be quantified. Finally, we will show how the introduction of MMSE IRC interference suppression receivers can minimize this demerit introduced by F-TDD.

3.1 Literature Review

Cellular systems were historically designed for voice services such that symmetric links were the preferred resource allocation configuration in the system. FDD was therefore assumed to be a very viable duplexing scheme. Over time, data traffic started dominating the main share of mobile traffic, leading to an increased interest towards TDD systems, allowing system designers to configure their resources flexibly and in accordance to the respective DL and UL traffic requirements. Literature related to F-TDD started becoming increasingly popular with the arrival of UMTS-TDD and recently after with WiMAX systems. Subsequently, renewed interest in F-TDD, emerged with the approval of the dynamic TDD study item for LTE-TDD [42], a particularly attractive feature for smaller cells, allowing the temporal reconfiguration of the TDD pattern used. For these reasons, one can find ample literature related to this topic.

Figure 3.1 attempts to categorize the reviewed literature into the two main study topics of dynamic TDD, and the relevant proposed solutions. In most cases the studies tackle the traffic adaptation and interference aspect of dynamic TDD. The traffic adaptation aspect deals with how we should allocate our time slot resources to follow the traffic dynamics accurately and efficiently. The interference aspect relates to the new types of AP-AP, and UE-UE interference arising from the flexibility offered by TDD. Most litera-

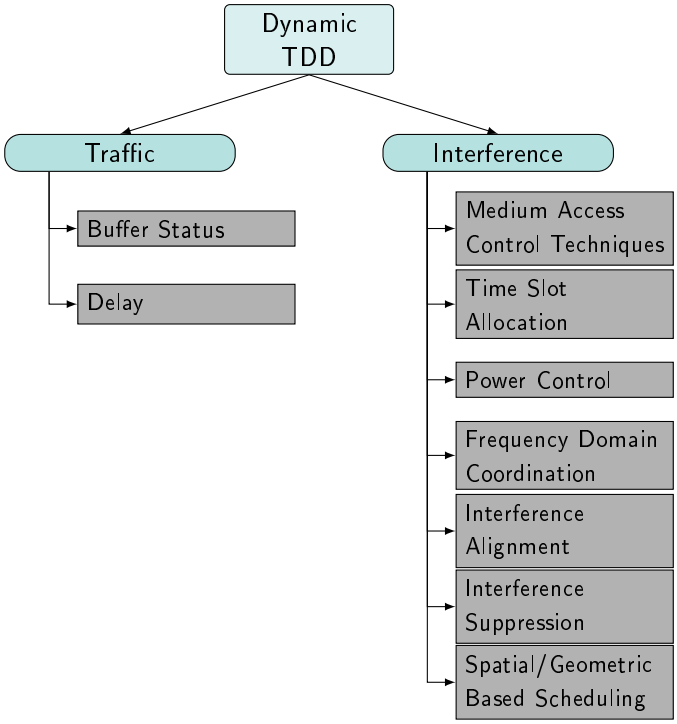


Fig. 3.1: Dynamic TDD Taxonomy

ture focuses on AP-AP interference in outdoor small cells, due to the typical presence of LoS conditions, and hence strong coupling, between two outdoor base stations.

3.1.1 Traffic Based Adaptation

Studies in [43–45], consider the gains of dynamic TDD from a traffic perspective. In all these cases, a single isolated cell with bursty UL and DL traffic is considered, and the time slot resources are allocated using the flexibility provided by the slot configurations in the LTE Enhanced Interference Mitigation & Traffic Adaptation (LTE-eIMTA) feature [46].

In [43], the traffic share between UL and DL is periodically assessed. The configuration which represents most closely this traffic share is then chosen. It is shown that the faster the slot reconfiguration time, the better the session throughput performance, since the TDD allocations can capture the UL and DL traffic dynamics of the cell more accurately. In this study it is also noted

that the defined LTE eIMTA patterns have limited flexibility and gains in the UL direction. A similar approach is taken in [44] but the configuration is chosen based on the weighted sum of the instantaneous UL/DL buffer status and the previous UL/DL traffic statuses. Again it is shown that the system benefits from faster TDD slot reconfiguration, especially at low loads. It is noted that at higher loads the gains of dynamic TDD start diminishing and that the larger the packet size, the smaller the sensitivity to faster reconfiguration times, due to the longer time required to transmit the packet. Similar findings are reported in [45] with the added observation that as the traffic load increases, the fluctuation of the traffic ratio decreases.

In [47], the performance of dynamic TDD is assessed in the presence of both real-time and non-real time traffic by using a unified utility function previously discussed in [48]. The utility function is parametrized in such a way that it considers not only the buffer status, but also the delay and packet deadline requirements of Voice over IP (VoIP) traffic. The authors finally benchmark their algorithm against the proposed scheduling algorithms provided in [44, 45] showing that they can retain the same performance in terms of average packet throughput while considerably lowering the packet drop rate.

The general outcome of these studies indicate that from a traffic perspective, there is a benefit in exploiting the characteristics of F-TDD, even if in the considered systems, the flexibility is limited and constrained by the TDD configurations patterns available in LTE-eIMTA. Most of the gains are reported for low to medium load, and a faster TDD reconfiguration, resulting in improved agility and flexibility, offers additional benefits in some specific traffic scenarios.

3.1.2 Interference Based Adaptation

Other works, mainly deal with the same-entity interference problem introduced by dynamic TDD. In [40], the flexibility introduced by dynamic TDD is exploited and the problem of same-entity interference is mitigated via appropriate cell clustering. Neighbouring cells are grouped inside a cluster, such that all cells within that cluster use the same TDD slot configuration. This TDD configuration can then be collectively reconfigured by all the cells inside that cluster based on longer term traffic statistics.

Conversely, the new interference modes introduced by dynamic TDD are exploited constructively in [49] via a technique termed as random time slot opposing. This technique resembles frequency hopping, but is conducted

in the time domain and can be used to avoid jamming. A Medium Access Control (MAC) scheme aiding in the mitigation of inter-cellular interference is discussed in [50, 51] via the concept of a busy burst. At each time slot, after a successful reception, a receiver will transmit a busy burst in a reserved mini slot. This allows the receiver to indicate its presence to neighbouring, potentially harmful, transmitter entities, such that inter-cell interference can be controlled in a distributed and collaborative fashion.

In [52], the authors investigate the link throughput performance of dynamic TDD at variable loads, claiming large gains at low load, and diminishing gains at high load. Gains are in fact present until network saturation is reached. Flexible TDD is then tested over an interference aware scheduling algorithm proposed in [53]. The end result shows that in the presence of an interference aware scheduler, one can experience improved dynamic TDD performance at even higher loads, with the net effect of prolonging the network operating load at which dynamic TDD offers superior performance over fixed static TDD.

Various time slot allocation algorithms taking into account the interference and traffic conditions are also proposed. A flexible frame structure with length N_f slots is considered in [54]. The frame can be configured such that it allocates at most $N_f - 1$ slots in one direction and one slot in the other direction. The problem is formalized as an optimization problem targeted at reducing the packet delay experienced by a traffic burst, taking into account the perceived interference levels and the traffic load. Based on this optimization, a game theoretic framework based distributed algorithm is proposed. The algorithm is benchmarked against a fixed (allocating equal UL and DL time slot resources) and random TDD scheme showing F-TDD gains whenever different DL/UL symmetries are present.

A heuristic dynamic TDD scheme coordinating the transmission modes of neighbouring cells is proposed in [55]. The possible transmission modes of multiple cells are analysed and the combination providing the maximum utility, the maximum log average throughput, is chosen. The authors state that the dynamic TDD gains depend on whether the network is congested or not, and if instantaneously there is a high asymmetry between UL and DL, large gains are expected. A cooperative decentralized switching point algorithm is presented in [56]. The nodes adjust their switching point based on exchanged interference pricing information, observing spectral efficiency gains in a simplified 2-cell scenario. Another time slot allocation algorithm is proposed in [57], showing how an algorithm which is aware of the channel and traffic conditions can adapt to an appropriate UL/DL time slot allocation, allocating few slots to the direction having bad channel conditions and low

traffic shares.

Focus on the introduced AP-AP interference is given in [58–63]. Cross-subframe interference is mitigated via coordination and power control in [58, 62]. In [58], uplink traffic bitmaps are exchanged between a set of strongly coupled cells, and in cross-subframe slots, uplink transmissions are protected via quantized downlink power controlled transmissions in neighbouring base stations, with the result of reducing the impact of harmful AP-AP interference and also mitigating inter-cell interference. A direct study on the benefit of DL power control is conducted in [62], showing considerable gains in the UL SINR, at the expense of slightly degrading the DL SINR distribution. The feasibility of DL power control is justified via the usage of a phantom cell architecture [64, 65] which separates the control and the data plane. In this architecture, control signalling related to dynamic DL/UL slot reconfiguration and dynamic power control is delivered by an overlay macro cell. Frequency domain interference coordination techniques for dynamic TDD phantom cells are studied in [61], yielding significant performance improvement in SINR at very high traffic load conditions. In this study, it is assumed that the base station is full duplex capable, and the separation between UL and DL is done via a FDD-like radio resource assignment scheme, with the frequency division being flexible as it is configurable by the phantom cell. In [59], the AP-AP interference issue is tackled by exploiting the extra degrees of freedom available in MIMO using interference alignment techniques.

Strategic time scheduling of the UEs, based on their geometric position, is discussed in [63]. A general TDD framework dividing the UL and DL portions of a frame via an adaptive switching point is defined. The UL allocations are then divided into UL allocations for UEs located on the outer edge of the cell and those located on the inner part of the cell. The inner and outer UEs are then scheduled separately in time, resulting in a gain in SIR when compared to a static and dynamic interference-unaware TDD system. In [60], cross-link interference is dealt with via sectored antennas along with time slot allocation algorithms. The considered TDD frame structure consists of a fixed part for UL and DL, and a flexible UL/DL allocation region. Focus is again placed on the outage performance in UL during the flexible UL period, since neighbouring cells could be in DL, and a receiving AP can suffer from high AP-AP interference. The degrees of freedom allowed by time scheduling along with the increased number of sectored antennas shows SIR gains when compared to an omni directional antenna based system which makes no attempt at actively improving the minimum SIR outage performance.

Apart from the classical problems of traffic scheduling and same-entity interference, some of the practical problematic aspects of dynamic TDD imple-

mentation have also been discussed. In [66], the authors discuss the interaction between dynamic TDD and the HARQ feedback framework present in LTE. Per-frame TDD reconfiguration, can in fact introduce ambiguity in HARQ signalling in LTE. By using partial HARQ bundling schemes, logically separating the fixed and flexible parts of the LTE-TDD subfame and introducing a new decision table based on the actual size of the association set, i.e. number of transmissions to be acknowledged/negatively acknowledged, the authors claim a 10% spectrum efficiency gain arising from reduced overhead and retransmission probability.

In general, the literature reviewed in this sub-section, stresses upon finding adequate solutions, to address the detrimental impact of cross-link interference introduced by dynamic TDD. In order to deal with this problem, a number of interference coordination techniques are proposed. Amongst others, time-slot allocation algorithms introducing scheduling constraints, were proposed numerous times, limiting both the harmful impact of interference along with the same flexibility introduced by dynamic TDD. Other solutions attempt to deal with this problem in the spatial domain. Techniques considering power control, sectorized antennas and interference alignment have in fact also been proposed to combat the problem of cross-link interference. The wide variety of solutions indicates that the problems brought along by dynamic TDD can be effectively dealt with, justifying the consideration of such a feature in a future 5G system.

3.1.3 Investigation of Flexible TDD in the 5G Concept

In this thesis, F-TDD is studied in the context of the envisioned 5G small cell concept [29]. When compared to the presented literature, there are some inherent differences in terms of opportunities and challenges, mainly attributed to the considered indoor small cell scenario, newly proposed frame structure, and support for new features.

Throughout the work which will be described hereinafter, full flexibility on a per slot basis is assumed, as opposed to some of the aforementioned works which consider a frame structure which supports a fixed and flexible portion. This allows additional freedom and possibilities in terms of traffic adaptation.

The problem of AP-AP interference is also less critical in our scenario, since we assume identical frame formats and similar link budgets in UL and DL. In an indoor small cell scenario, with randomly deployed nodes having similar transmission powers in UL and in DL, all kinds of interference are equally harmful i.e. (AP-AP, UE-UE, AP-UE or UE-AP). Instead a more challenging

problem relates to the introduced inter-cell interference variation which can change from one frame to the next, leading to increased complications for adaptation schemes which require stable interference and channel conditions. In our studies, the interference aspect of dynamic TDD is dealt with via the usage of interference suppression capable receivers.

3.2 Expected Gains from Flexible TDD

In this section, an analytical approach to find the maximum achievable gain of fully F-TDD over fixed S-TDD from a traffic perspective, will be presented. The potential presence of data buffering will also be taken into account when considering this gain.

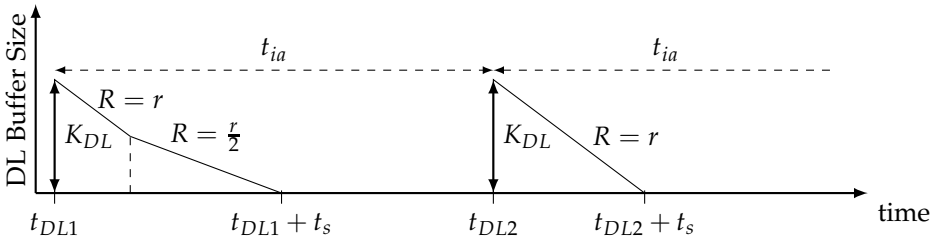
Let us consider the simplified scenario of a single cell, consisting of an access point (AP) and a user equipment (UE), each having an independent traffic profile. We further assume that the traffic profile of both link directions, follows a Poisson process, representing a bursty traffic model as specified in the 3GPP FTP Traffic Model 1 [30], with session file size K bits, and an exponentially distributed session inter-arrival time, t_{ia}.

For the sake of simplicity, a F-TDD scheme with the following characteristics is introduced. If the data to be serviced, is present only in UL or DL, a channel rate of r is assumed. Alternatively, if both UL and DL data is present, the scheme allocates an equal division of resources. For a session transfer encompassing multiple time slots, the long term subdivision of the time slot resources will effectively provide a channel rate of $\frac{r}{2}$ to each link direction.

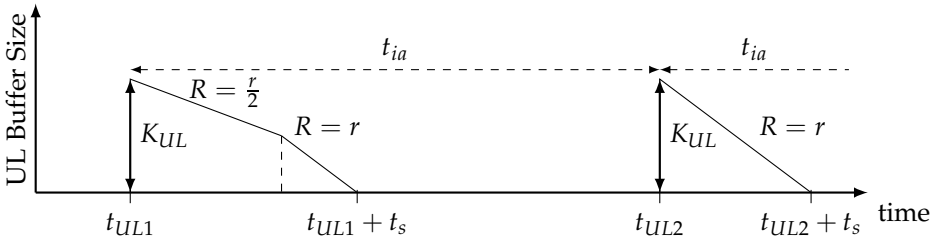
This means that the time t_s, required to service a session, can be represented as $\frac{K}{R}$, with R depending on the buffer conditions in DL and UL, K_{DL} and K_{UL} respectively. The rate R varies based on whether both DL and UL directions have data to be transmitted or not. Under the assumption that $t_s > t_{TTI}$, the rate of each link, R, is given as,

$$R = \begin{cases} 0 & \text{if } K_{DL} = 0 \text{ and } K_{UL} = 0 \\ \frac{r}{2} & \text{if } K_{DL} > 0 \text{ and } K_{UL} > 0 \\ r & \text{if } K_{DL} > 0 \text{ and } K_{UL} = 0 \\ r & \text{if } K_{DL} = 0 \text{ and } K_{UL} > 0 \end{cases} \tag{3.1}$$

An example of the expected behaviour is depicted in Figure 3.2.



(a) DL Buffer Status



(b) UL Buffer Status

Fig. 3.2: Buffer Status with Traffic Arrivals in Uplink and Downlink

For comparison, a fixed S-TDD scheme, whose time slots are configured to match the long term traffic share between UL and DL is considered. Under the assumption of equal traffic in UL and DL, a 1:1 (DL:UL) TDD allocation is assumed for the fixed S-TDD scheme. This means that in the presence of data in the buffers, the rate R for the S-TDD scheme is always $\frac{r}{2}$ for each link direction, independently of whether DL and UL data are concurrently present or not.

Intuitively, at low load, the gains of F-TDD are mainly attributed to the statistical presence of either UL or DL data. At higher load, the probability of concurrent UL and DL data increases. However, it is expected that statistically, lower buffer occupancy levels are present in the F-TDD system. The actual session throughput performance of each session is affected by buffering, and the sojourn time of each file is composed of the waiting time in the buffer and the service time in the channel. It is therefore believed, that F-TDD should give gains even at higher loads.

In order to assess the gains of F-TDD over fixed S-TDD in the presence of buffering, we apply models from the well established field of queuing theory [67], to express the mean sojourn time of each file for both TDD schemes.

Let us start by assuming the session arrival rates, λ_{DL} and λ_{UL}, for the individual DL and UL traffic profiles, and equal session sizes in DL and in UL such that, $K_{DL} = K_{UL}$. Assuming that the traffic follows a Poisson process and is identical in DL and in UL, we can establish the following relationships,

$$K = K_{DL} = K_{UL} \tag{3.2}$$

$$\lambda_{DL} = \lambda_{UL} \tag{3.3}$$

$$\lambda_{tot} = \lambda_{DL} + \lambda_{UL} \tag{3.4}$$

As mentioned earlier on, the S-TDD scheme is configured to operate in a 1:1 (UL:DL) fashion to match the assumed long term traffic share. Therefore, the session service rate for the S-TDD scheme can be assumed to be deterministic with a rate $\frac{r}{2K}$, for each of the DL and UL link directions. Under these assumptions, an M/D/1 queueing model, with session arrival rate λ_{DL}, and session service rate $\frac{r}{2K}$ can be established for the DL link direction. Similarly, an M/D/1 queueing model, with session arrival rate λ_{UL}, and session service rate $\frac{r}{2K}$, can be established for the UL link direction. Under the assumption of $\lambda_{DL} = \lambda_{UL}$, the two queueing models are identical.

For the F-TDD scheme, a deterministic session service rate $\frac{r}{K}$ can be assumed if the aggregated UL and DL traffic, is serviced opportunistically. Considering that the aggregated traffic represents a Poisson process with a total arrival rate $\lambda_{tot} = \lambda_{DL} + \lambda_{UL}$, the F-TDD scheme, can be modelled as an M/D/1 queue with session arrival rate λ_{tot}, and session service rate $\frac{r}{K}$.

The assumed M/D/1 [2] queueing models for the S-TDD and F-TDD schemes are shown in Figure 3.3.

Irrespective from the scheduling mechanism, at steady state, the mean sojourn time $E[T]$, for an M/D/1 system is given by [68],

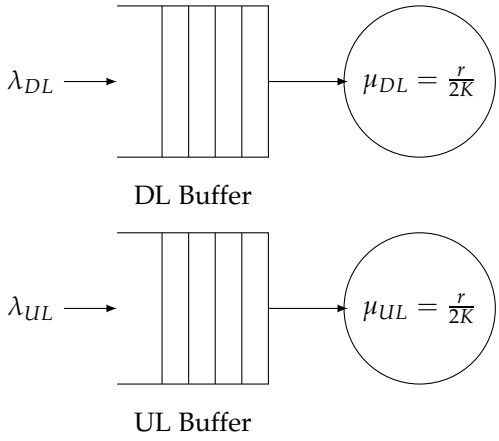
$$E[T] = \left(\frac{2-\rho}{1-\rho}\right)\frac{1}{2\mu} \tag{3.5}$$

where ρ represents the utilization and is given by,

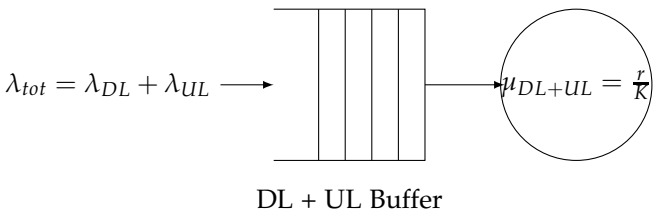
$$\rho = \frac{\lambda}{\mu} \tag{3.6}$$

From equation 3.6, one can immediately notice that as $\lambda = \mu$, and therefore $\rho = 1$, the sojourn time goes to infinity.

[2] Markovian Arrival Rates, Deterministic Service Time, Single Server Model



(a) Queueing Model for S-TDD scheme



(b) Queueing Model for F-TDD scheme

Fig. 3.3: Queueing Models

For the fixed S-TDD scheme, the values for μ and ρ in DL and UL are identical, and the queues can therefore be treated independently. For DL [3], the mean waiting time can be derived as follows,

$$\mu_{DL} = \frac{r}{2K} \tag{3.7}$$

$$\rho_{DL} = \frac{2\lambda_{DL}}{r} \tag{3.8}$$

$$E[T]_{STDD} = \left(\frac{2-\rho}{1-\rho}\right)\frac{K}{r} \tag{3.9}$$

For the F-TDD scheme,

$$\mu_{DL+UL} = \frac{r}{K} \tag{3.10}$$

$$\lambda_{tot} = \lambda_{UL+DL} = 2\lambda_{DL} \tag{3.11}$$

$$\rho_{UL+DL} = \frac{2\lambda_{DL}}{r} \tag{3.12}$$

$$E[T]_{FTDD} = \left(\frac{2-\rho}{1-\rho}\right)\frac{K}{2r} \tag{3.13}$$

From these equations, one can observe that the mean sojourn time of the F-TDD scheme is always half the one of the S-TDD scheme.

In order to compare this result with our simulation tool, the session delay experienced by each file is inspected. This delay represents the time difference from when the file was created until it was entirely received. A fixed MCS operation is assumed, and a file size of 8 MB is chosen to reduce the impact of the channel access time over the whole file transmission. In an attempt to account for overhead and other practicalities implemented in the simulator, we first inspect the transmission rate of a file in a system with traffic in only one direction, and extract the two channel rates, $R_{STDD} = 479Mbps$, and $R_{FTDD} = 956Mpbs$.

The 1:1 (UL:DL) fixed S-TDD is then compared against the simplified F-TDD scheme. The values of λ_{DL} and λ_{UL} are equally varied to capture the file delay performance over a traffic load range. Finally the analytical sojourn time, and the file delay experienced by each packet in the simulation [4] are compared. The results are shown in Figure 3.4.

[3] Under the assumption of identical traffic in UL and DL, the result is identical for UL

[4] The simulation was run for 90 seconds, in an attempt to approach the steady-state behaviour

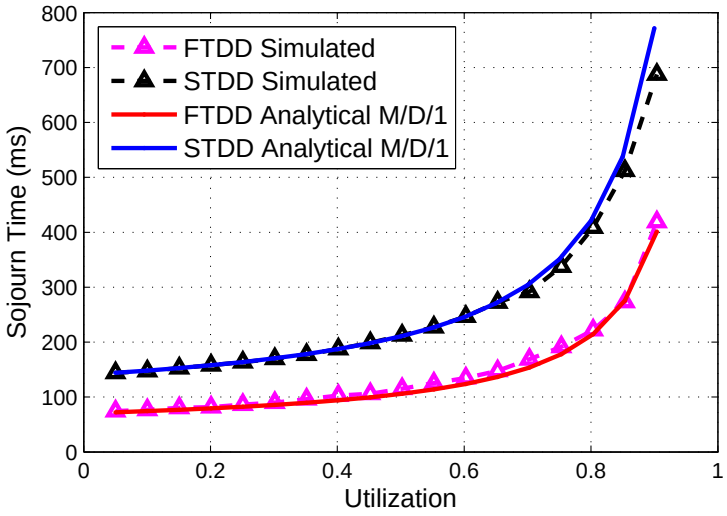


Fig. 3.4: Sojourn Time Results, Analytical versus Simulated

The results shown in Figure 3.4 indicate a good match between the analytical and simulated results, with a slight deviation at high utilization. From Figure 3.4, one can see that if buffering is considered, the sojourn time in a F-TDD scheme is halved when compared to the S-TDD scheme. At high loads, the gains of F-TDD reduce slightly, since the analytical approach assumes steady-state operation which might take a very long time to reach in practice, but nonetheless considerable gains are observed. Please note that in addition to the practical limitation on the simulation time, the system level simulation models various aspects such as radio channel variations, signalling delays and other intricate details, such that a perfect match can not be expected.

For completeness sake, the simulated average session throughput is shown in Figure 3.5 confirming the gains observed above.

It is important to note the assumption of deterministic service time is only feasible in the single cell scenario assumed above. In a real system, the service time varies based on a number of factors. Most importantly one can deduce that over time, the F-TDD scheme will have a higher probability of experiencing varying interference conditions (see Section 3.3.2), making the assumption of deterministic service time invalid.

In fact, if the service time is assumed to take an exponential time instead of a deterministic one, the expected sojourn time increases. For illustrative purposes, we hypothetically assume an exponential service rate for the F-TDD

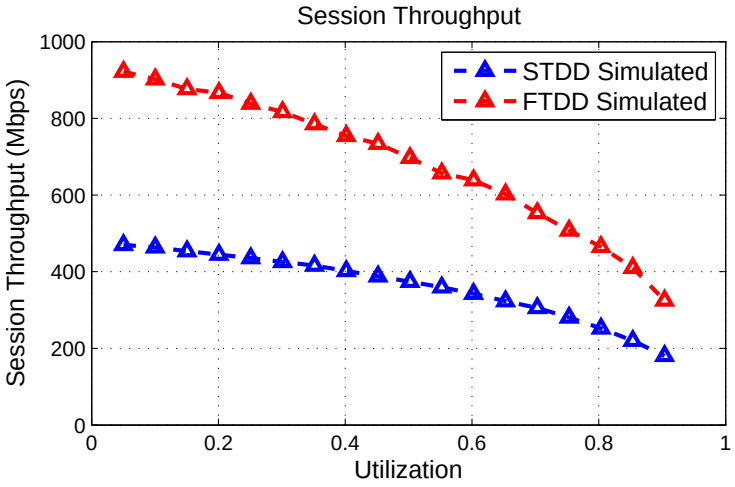


Fig. 3.5: Average Session File Throughput Simulation Results

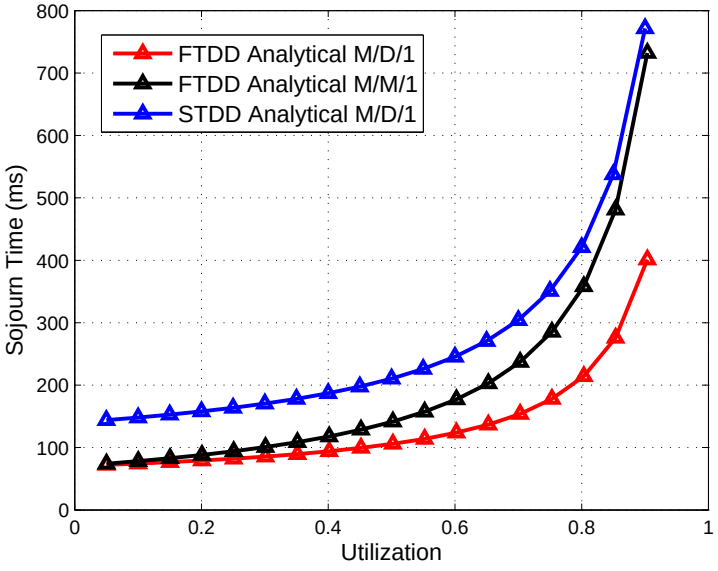


Fig. 3.6: Mean Sojourn Time for M/M/1 vs. M/D/1

scheme in order to show this behaviour. A comparison of the fixed S-TDD scheme under the M/D/1 model assumption, and the F-TDD scheme under the M/D/1 and M/M/1 model assumptions, is shown in Figure 3.6, confirming that at higher values of ρ, the mean sojourn time of the M/M/1 model increases considerably when compared to the M/D/1 model.

3.3 The Demerits of Flexible TDD

As seen from section 3.2, in a simple single cell scenario, F-TDD offers significant gains over fixed S-TDD. The reason for this is that unlike a static fixed TDD system, it has the flexibility to follow and accommodate the instantaneous traffic demands.

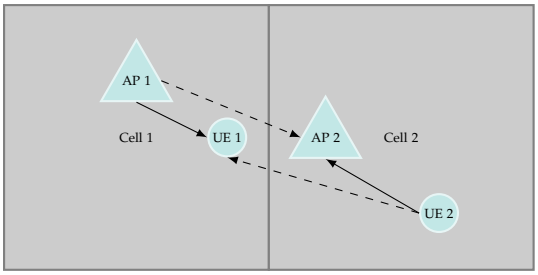
From a radio perspective and in a realistic multi-cell system where inter-cell interference plays an important role there are however some challenges to address when considering the usage of F-TDD. In the following subsections we will present some of these aspects, estimate their impact and subsequently investigate potential tools which can mitigate the effect of these challenges.

3.3.1 Main Challenges

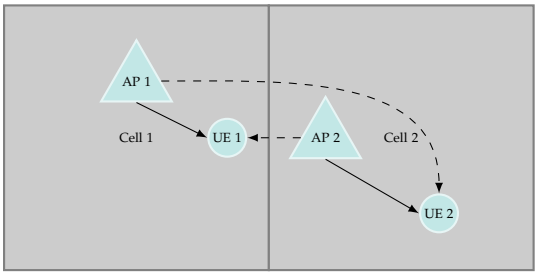
In a fully synchronized fixed TDD system, an AP transmission typically interferes with a UE's reception, and a UE transmission interferes with an AP's reception. In the absence of mobility this behaviour stabilizes and makes the perceived channel conditions predictable since all cells are either in DL or in UL. In order to capture the instantaneous traffic requirements, F-TDD changes this paradigm, and all the cells in a network can either be in UL or DL. Apart from AP-UE and UE-AP interference, such flexibility also introduces AP-AP and UE-UE interference increasing and eventually diversifying the number of interference levels that can be experienced by a victim receiving node.

Figure 3.7 illustrates a very simple example showing this problem. Consider UE1 in Figure 3.7a with cell 1 operating in DL and cell 2 operating in UL. The perceived interference at UE1 is significantly lower than in Figure 3.7b, where cell 2 switches to DL. With F-TDD the condition displayed in Figure 3.7 is perfectly legal, and can occur from one frame to the next without restrictions. If one adds even more cells in the picture, the impact of this inter-cell interference variation can be even larger, and therefore the future channel conditions of UE1 are more difficult to predict. This poses a significant problem to link

and rank adaptation schemes especially when considering that the scheduled rank and MCS transmission parameters are signalled one frame in advance due to control channel processing delays (for further details please refer to section 2.6).



(a) Cell 1 in Downlink - Cell 2 in Uplink



(b) Cell 1 in Downlink - Cell 2 in Downlink

Fig. 3.7: Inter-cell Interference Variation

3.3.2 Inter-cell Interference Variation

In this subsection we will attempt to quantify the impact of this added inter-cell interference variation. We have acknowledged the threat of increased inter-cell interference variation brought along by F-TDD. It is important to note that in a fixed S-TDD system, milder inter-cell interference variation can also be experienced. This is subject to traffic profile present in neighbouring cells, which can mute or unmute their time slots based on whether they have any traffic to transmit. On the other hand, as evidenced by the previous section, the flexible link direction switching allowed by F-TDD introduces a more severe source of inter-cell interference variation. It is however important to appreciate that for a system with a given absolute offered load, the F-TDD scheme can service its data more quickly, potentially reducing the experienced interference, at least temporally.

To understand better the behavioural differences of fixed S-TDD and F-TDD,

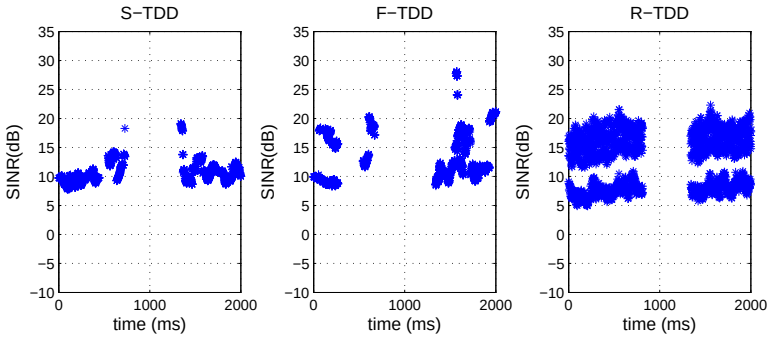


Fig. 3.8: SINR Time Trace

let us consider a small cell scenario consisting of 20 cells arranged in a 10x2 grid fashion, where each cell consists of a single AP and single UE. We compare a fixed S-TDD 1:1 scheme[5], a F-TDD scheme and a Random TDD (R-TDD) scheme which randomly chooses the link direction.

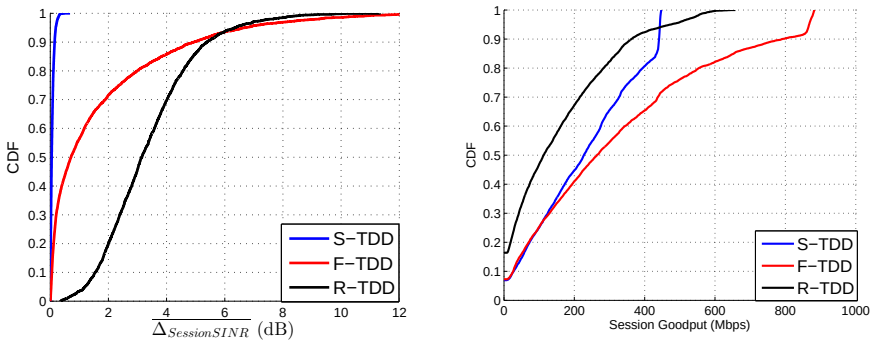
To place additional stress on the inter-cell interference variation problem, an absolute load of 150 Mbps, representing around 60% resource utilization, is offered to each node. A lightly loaded system would decrease the probability of concurrent cell operation substantially, such that very mild inter-cell interference is experienced. Conversely a very highly loaded system will tend to reduce the traffic fluctuation ratio, as also observed in [45], leading to a F-TDD scheme with a potentially convergent TDD allocation.

The actual traffic generation is conducted using the finite buffer bursty 3GPP traffic model [30], assuming equal UL and DL load, and sessions with a file size of 2 MBytes. In this investigation it is also assumed that all the nodes are equipped with a 4x4 MIMO transceiver, and operate with a Minimum Mean Square Error (MMSE) Maximum Ratio Combining (MRC) receiver whose model is attached in Appendix B of this thesis.

A simple SINR trace for a particular node is illustrated in Figure 3.8, immediately capturing the added SINR variability introduced by the F-TDD schemes. The SINR variability, as expected, is more pronounced in the R-TDD case.

In order to concisely capture the experienced SINR variability over all the received sessions (for all the deployed nodes), the information from the SINR time traces needs to be aggregated and compressed. Each SINR time trace is first segmented into the corresponding received sessions. Please note that a session is fragmented into N packets, and is transmitted/received over

[5] 1 DL slot, followed by 1 UL slot



(a) $\overline{\Delta_{SessionSINR}}$, MMSE-MRC Receiver **(b)** Session Goodput, MMSE-MRC Receiver

Fig. 3.9: $\overline{\Delta_{SessionSINR}}$ and Corresponding Impact on Session Goodput

multiple, potentially dis-contiguous time slots, t_i. For each session received, the absolute difference in SINR from one reception to the next, Δ_{SINR}, is stored. Thereafter, the raw absolute Δ_{SINR} values are averaged as shown in Eq. 3.14, extracting a single averaged $\overline{\Delta_{SessionSINR}}$ value for each session received.

$$\overline{\Delta_{SessionSINR}} = \frac{\sum_{i=0}^{N-1} |SINR(t_{i+1}) - SINR(t_i)|}{N} \tag{3.14}$$

The Cumulative Distribution Function (CDF) of the averaged $\overline{\Delta_{SessionSINR}}$ is shown for each of the individual TDD schemes in Figure 3.9a, reconfirming the added SINR variation introduced by the F-TDD scheme.

The net effect of this added SINR variability causes channel estimation errors that affect the link adaptation, inducing increased HARQ retransmissions. Simulation results show that in this particular case, in the fixed TDD case, 5% of the packets need to be retransmitted at least once, while in the F-TDD case, 14% of the packets need to be retransmitted. The situation gets even worse with the R-TDD scheme requiring up to 36% of the packets to be retransmitted.

These additional retransmissions do affect the experienced end-user session goodput as shown in Figure 3.9b. While the F-TDD scheme allows an end-user to experience higher average peak throughputs, the added retransmissions due to channel estimation errors impose a small performance penalty at the lower end of the goodput CDF curve.

The envisioned 5G concept discussed in chapter 2 assumes the availability

of MIMO and MMSE-IRC receivers which have been shown to effectively suppress inter-cell interference in [21]. Unlike MMSE-MRC, MMSE-IRC attempts to project any interfering signals to a non-orthogonal subspace with respect to the desired signal, with the end goal of improving the SINR. Given the availability of an appropriately designed frame structure such as the one explained in Section 2.5, MMSE-IRC can suppress inter-cell interference independently from the source, hence also reducing the experienced inter-cell interference variation, aiding in more accurate channel estimation in the presence of a F-TDD scheme.

If we repeat the previously analysis, moving from an MMSE-MRC to an MMSE-IRC receiver[6], it is possible to see from Figure 3.10, that the $\overline{\Delta_{SessionSINR}}$, is reduced dramatically. This reduction has a direct impact on the number of induced HARQ retransmissions as shown in Table 3.1, reducing the number of retransmissions from 14% to 0.8% for the considered F-TDD scheme. In turn, this also improves the perceived session goodput as shown in Figure 3.11, allowing us to exploit the full advantages of F-TDD without being affected by excessive retransmissions caused by the increased inter-cell interference variation introduced by this same feature.

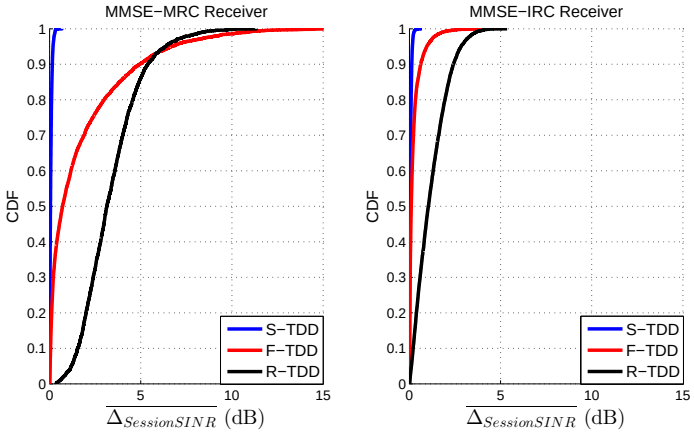


Fig. 3.10: $\overline{\Delta_{SessionSINR}}$ - MMSE-MRC vs. MMSE-IRC Receiver

[6]For this investigation, a rank 1 transmission is assumed, meaning that the antenna resources available are dedicated for interference suppression

	MMSE-MRC	MMSE-IRC
S-TDD	5%	0.5%
F-TDD	14%	0.8%
R-TDD	36%	3%

Table 3.1: Number of HARQ Retransmissions

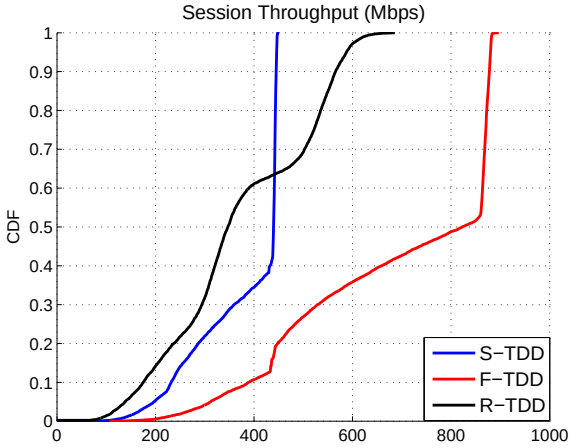


Fig. 3.11: Session Goodput - MMSE-IRC Receiver

3.4 Wrap-up

In this chapter we have described the main principles of F-TDD, by first outlining the merits and demerits of this feature. A comprehensive literature review on the topic was presented, eventually highlighting how the problems mentioned in previous work, relate to or differ from our specific investigations. The expected gains of F-TDD over different traffic loads were then assessed by using tools from queueing theory, and a cross-validation between this analysis and the F-TDD implementation in the system-level simulator was conducted. Subsequently, the complications resulting from F-TDD were presented, and an in-depth analysis quantifying the negative inter-cell interference variation effects brought along by this feature was conducted. Finally, we have seen how the usage of MMSE-IRC receivers can help in mitigating these effects.

Chapter 4

UL/DL Scheduling

4.1 Flexible UL/DL Algorithms

In the previous chapter we have seen how MMSE-IRC receivers can significantly reduce the impact of inter-cell interference variation, a potential threat introduced by F-TDD systems. The next natural question, is how and on which criteria we should assign our time slot resources, to UL or DL respectively. In the following sections we provide two possible F-TDD schemes to address this problem.

The first F-TDD scheme, the Delay Fairness Based (DF) scheme, consists of a simple algorithm operating on two input parameters; the Head of Line (HOL) delay, and the buffer size in each link direction (UL and DL). The HOL delay represents the waiting time of the oldest packet in the queue, while the buffer size simply relates to the amount of bits in the buffer. The algorithm decides the link direction based on the following criteria. If data is present in only one link direction, that link direction is scheduled. If data is present in both link directions, the input parameters mentioned above start playing a more decisive role. The HOL delay and buffer size of each link direction are inspected and compared to a corresponding threshold, for each of the input parameters. If the HOL delay or buffer size thresholds are exceeded, a switch in the link direction is conducted. As we will see when discussing the algorithm in further detail, the HOL delay threshold plays a more prominent role in effecting the switch. The main purpose of the algorithm is therefore aimed at guaranteeing a switch in the link direction, if a packet experiences a HOL delay exceeding the defined threshold, and is hence referred to as the

Delay Fairness Based (DF) based scheme.

Another F-TDD scheme, denoted as the Load Fairness Based (LF) algorithm, is also presented in this chapter. The LF based scheme operates on the instantaneous traffic share between DL and UL, and the previous slot allocations which have been made. A metric based on these values is calculated for each link direction, and the metric with the highest value determines the link direction of the next TDD slot. Unlike the DF based algorithm, the LF based scheme is focused on allocating the time slot resources based on the traffic share between UL and DL. Allocation to a very lightly loaded link, is however guaranteed by the consideration of the previous slot allocations.

In this chapter we will first give a general overview on the link scheduling framework. In order to enable the usage of the F-TDD schemes, information related to the current buffer status in each link is required. The relevant procedures used in supplying the algorithms with updated UL and DL information will therefore be discussed. Thereafter we will explain in detail, the mechanisms of the two proposed F-TDD algorithms, showing the impact of certain parametrizations on their behaviour. Finally, a comparative analysis between the two proposed F-TDD algorithms will be provided.

4.2 Link Scheduling

Before delving deeper into the corresponding F-TDD allocation schemes, we will describe how the link scheduling algorithms operate within the envisioned 5G system explained in Chapter 2. The link direction is chosen based on current conditions in UL and in DL, and any related information needs to be available to the UL/DL link scheduler, the entity responsible for taking the link direction decision. The flow of such information will therefore firstly be discussed. Following this, the general scheduling framework will be presented, and the specific role and scope of the UL/DL link scheduling entity will be outlined.

4.2.1 Information Flow

The UL/DL link scheduling entity is located in the AP. It is assumed that the AP decides the link direction on a per frame basis. In order to do so it is essential that the AP has both local DL information and remote UL information. Most importantly, this implies that UL information needs to be made available to the AP frequently and accurately, such that the link direction

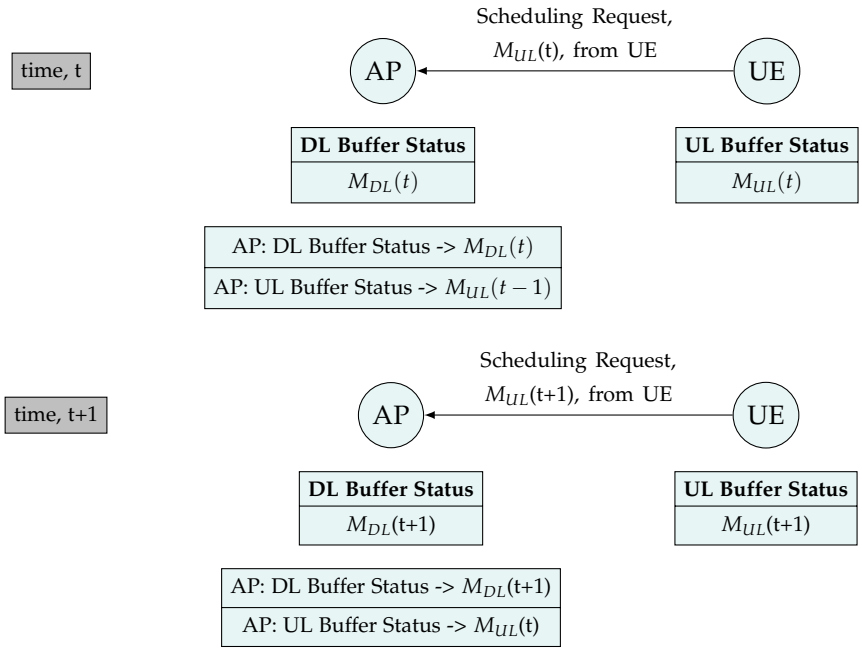


Fig. 4.1: Buffer Status Update Procedure

decision is based on actual DL and UL conditions.

Let us exemplify how such remote UL information is made available to the AP by considering one important input parameter, the buffer status information M. In this case, the DL buffer status information, M_{DL}, is available locally at the AP, while the UL buffer status information, M_{UL}, needs to be embedded within a Scheduling Request message and signalled over the UL control channel, or potentially piggybacked with data information in the data part of the frame. For the sake of simplicity, in this work we assume that such UL information is ideally embedded in the Scheduling Request message. A timing diagram showing the interaction of the link scheduling and the UL information signalling procedure is illustrated in Figure 4.1.

In this example, the AP will make a decision at time instance t, to schedule either the UL or DL direction, based on the information it has available. The DL information is always available instantly, since it is locally present in the AP. The UE buffer status information M_{UL}, is delayed by one time step t (one frame in our case), if a scheduling request with updated buffer status information is sent on every frame. Therefore when conducting a link direction decision, the AP will have at its disposal the latest updated DL

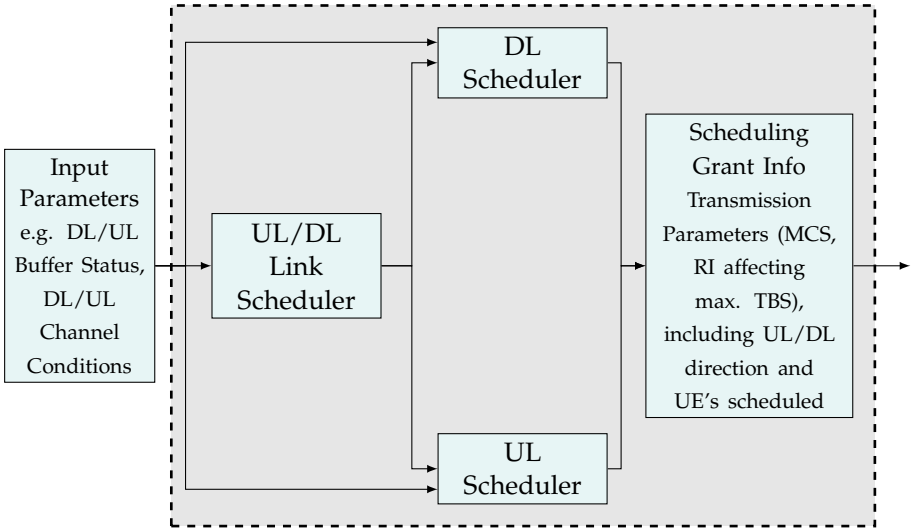


Fig. 4.2: Scheduling Framework

buffer status information M_{DL}, and at best, the UL buffer status information M_{UL}, from one time-step (frame) ago.

4.2.2 Scheduling Framework

The considered scheduling framework is depicted in Figure 4.2. The framework is composed of multiple scheduling entities each handling a specific scheduling task. The complete scheduling framework, highlighted in grey in Figure 4.2, will receive a set of input parameters. These inputs contain information related to traffic and channel conditions in both UL and DL. Examples of these could include, buffer status information, current head-of-line delay in each link direction, and the latest experienced DL and UL SINR conditions. Please note that the precise need and usage of this input information is dependent on the corresponding F-TDD algorithm and will be explained in further detail in the subsequent sections. This input information is then fed into the UL/DL link scheduler, which decides the link direction, UL or DL.

The decision taken by the link scheduler is then passed to the corresponding scheduling entity, DL or UL scheduler, which chooses a set of scheduled users along with the corresponding transmission parameters. These decisions are then formed into a scheduling grant message, which is sent over the DL control channel of the frame.

4.3 Delay-Fairness Based Algorithm

4.3.1 Algorithm Description

We first consider a simple DF based algorithm to allocate the UL and DL time slot resources. The main objectives of the algorithm lie in exploiting F-TDD on a per slot basis while attempting to guarantee a minimum time to start servicing a packet.

The algorithm operates on two inputs, the amount of bits in each of the buffers, K, and the head-of-line delay in each of the buffers, d. The buffer status information M, is therefore composed of K and d, such that for direction D, $M_D = \{K_D, d_D\}$. The algorithm operates on the following principles. If there is no data available in any of the buffers, the frame is muted. If data is available in only one direction, that direction is scheduled.

If data is available in both directions, the buffer status information M, starts playing a more important role. If a certain direction D has been scheduled in the previous frame, the opposite link direction D''s buffer size and head-of-line delay are inspected, and if any of those values exceed any of the respective predefined thresholds, $Th_{BufferSize}$ and $Th_{HOLDelay}$, a switch in the link direction is conducted, and the direction D' is scheduled on the following frame. A flowchart summarizing the behaviour of the algorithm is depicted in Figure 4.3. The corresponding pseudo-code is shown in Algorithm 1.

4.3.2 Algorithm Behaviour

Unlike a rigid fixed S-TDD scheme, in the absence of traffic in one direction, the DF based algorithm can exploit all the available empty time slots to accommodate the link direction instantaneously having traffic.

The predefined thresholds have the following effects on the algorithm: a short $Th_{HOLDelay}$ will cause fast switching in the link direction, while a larger $Th_{HOLDelay}$ will delay the link direction switch. Therefore, a large $Th_{HOLDelay}$, prioritizes the currently scheduled link direction, and delays the time to start serving a link direction which has just received a packet in its buffers. The opposite behaviour is experienced with shorter $Th_{HOLDelay}$ values.

For a large arriving payload burst, a small $Th_{BufferSize}$ will cause an immediate switch in the opposite link direction, while a larger $Th_{BufferSize}$, will only conduct a switch in the link direction, once the amount of bits in the buffer

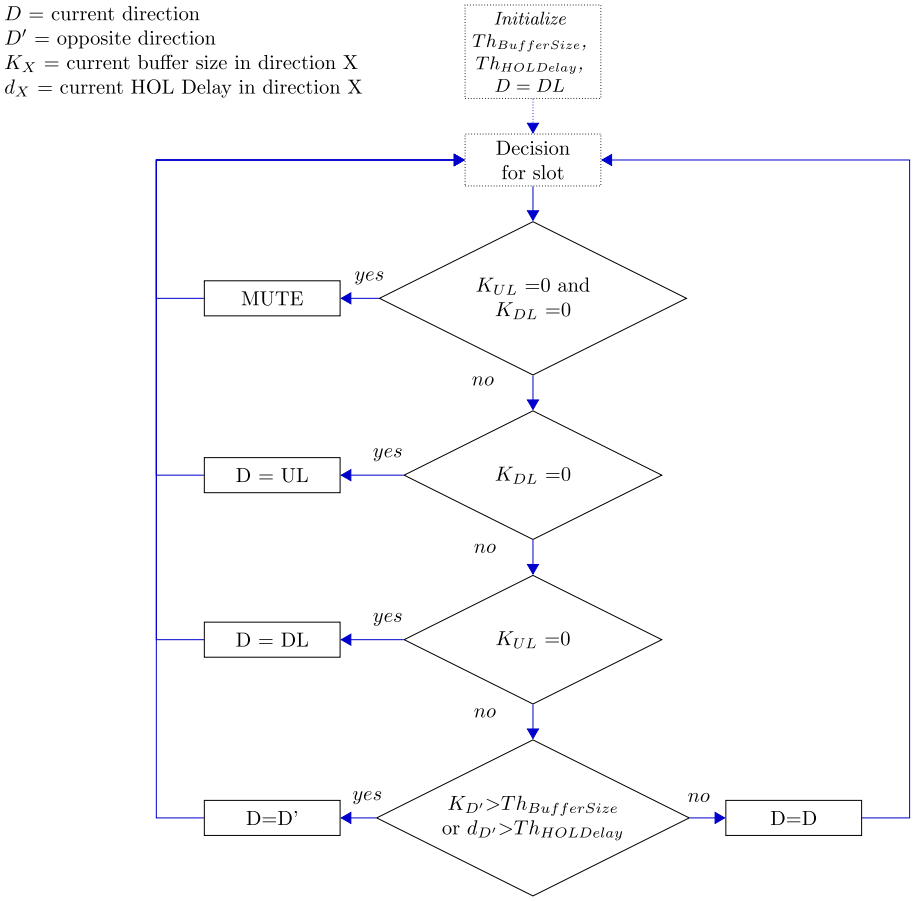


Fig. 4.3: DF-Based Algorithm Flowchart

Algorithm 1 DF-Based Algorithm

```
Initialize thresholds, Th_BufferSize, Th_HOLDelay
for each time slot do
    if K_UL = 0 and K_DL = 0 then
        D = Mute
    else if K_UL > 0 and K_DL = 0 then
        D = Uplink
    else if K_UL = 0 and K_DL > 0 then
        D = Downlink
    else if K_UL > 0 and K_DL > 0 then
        if K_D' > Th_BufferSize or d_D' > Th_HOLDelay then
            Switch Link Direction, D = D'
        else
            Keep Current Link Direction, D = D
        end if
    end if
end for
```

has grown significantly. The $Th_{BufferSize}$ threshold can therefore serve as a protection mechanism to avoid buffer overflow, if a small buffer is present.

Given the presence of data in both link directions, and given that the traffic offered to the system consists of large payload bursts requiring multiple frames to complete, the algorithm, independently of any parametrization, will converge to a 1:1 UL:DL TDD allocation. This occurs because the experienced buffering time will exceed any reasonably defined $Th_{HOLDelay}$ threshold, causing an alternating UL/DL switch. In general, this condition will occur whenever the payload service time T_s, is considerably greater than the defined $Th_{HOLDelay}$ threshold. This convergent behaviour will create an oscillating but potentially predictable interference generation pattern, with the disadvantage of allocating resources equally, even if on a long term average one link has much more traffic to serve. Please note that a 1:1 UL:DL TDD allocation can also be enforced by parametrizing the algorithm with a $Th_{HOLDelay}$ threshold shorter or equal to the frame length T_f.

4.3.3 Performance Evaluation

In order to validate the operation of the DF based algorithm, system-level simulations are carried out. The system is loaded with a traffic load representing a cell resource utilization of around 55%, with equal load in UL and in DL. This configuration is also replicated for different two file sizes represent-

Table 4.1: Simulation Parameters for DF-Based Scheme Evaluation

Traffic Parameters	
Cell Offered Load	400 Mbps
Downlink Offered Load	200 Mbps
Uplink Offered Load	200 Mbps
Approximate Resource Utilization	55%
File Sizes	200 kB, 2 Mb
Algorithm Parameters	
$Th_{HOLDelay}$	1, 2, 4 ms
$Th_{BufferSize}$	1, 4 MB

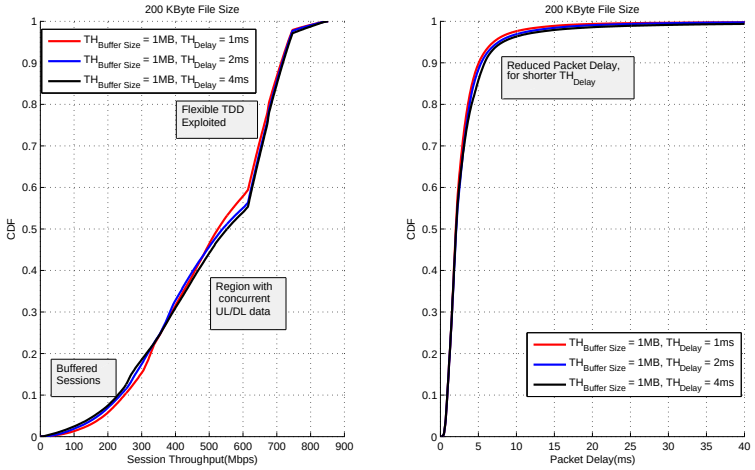
ing a session, in order to confirm the algorithm's behaviour and dependence on different file sizes. Please note that these files sizes are fragmented to smaller packets when effecting the actual transmissions.

The offered load is strategically chosen as defined above, in order to show the gains achievable by F-TDD while at the same time show the effects of the corresponding $Th_{HOLDelay}$ and $Th_{BufferSize}$ parametrizations, which only become of active importance during the concurrent availability of DL and UL data. The algorithm is tested with different $Th_{HOLDelay}$ and $Th_{BufferSize}$ thresholds, and the session throughput of the file and the individual packet delays of the fragmented files are inspected. It is important to note that these individual packet delays are also affected by the buffering time, and the packet scheduling is done on a First in First Out (FIFO) basis. The relevant parameters for the simulation are summarized in Table 4.1 and the results for the different parametrizations are shown in Figures 4.4 and Figures 4.5.

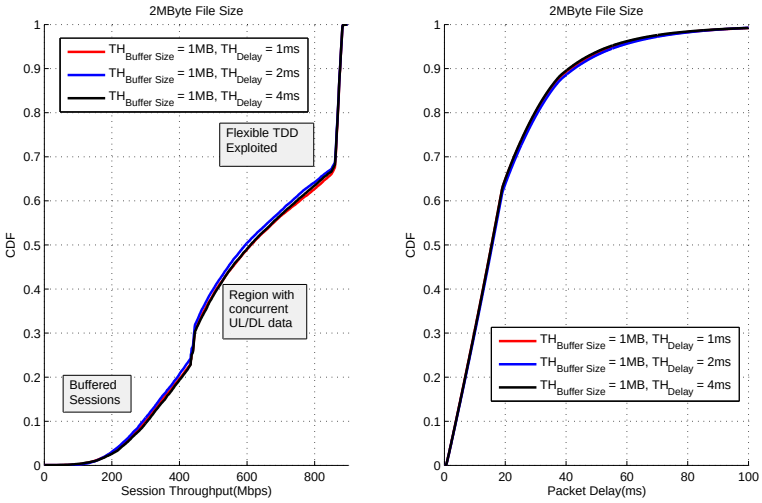
A rough inspection at Figures 4.4 and 4.5 illustrates that the algorithm is almost independent of the parametrization applied, due to the buffering effect mentioned earlier on. Moreover one can immediately notice,

- a payload buffering region, specifically where low session throughput is obtained,
- a region where UL/DL traffic are concurrently available in the buffers,
- and an additional region at the high end of the session throughput curve showing the ability of flexible TDD to exploit the empty time slots, when instantaneously only UL or DL traffic is available.

In Figure 4.4, the $Th_{BufferSize}$ is set to 1 MB, such that it is expected that for

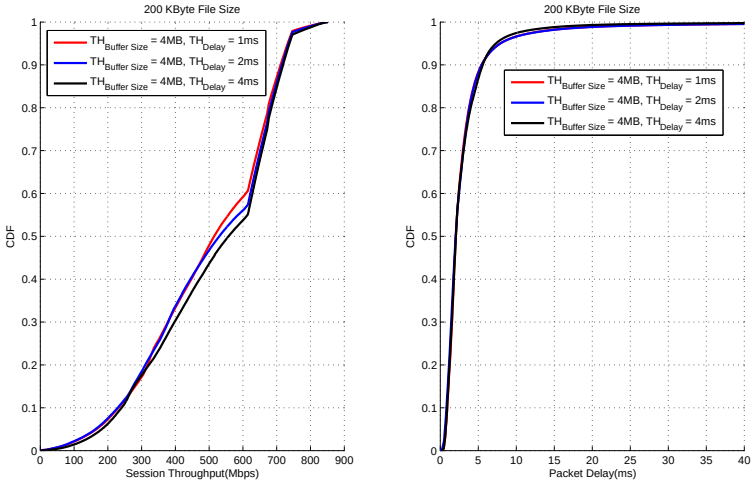


(a) File Size - 200 kB

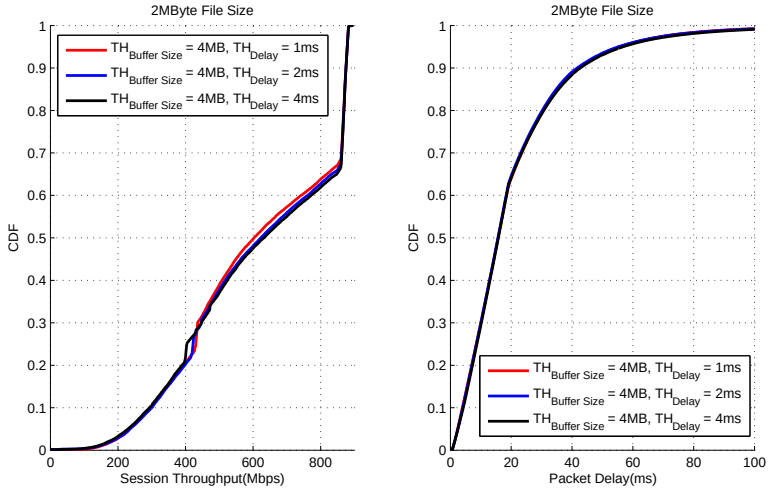


(b) File Size - 2 MB

Fig. 4.4: Individual Packet Delay for different Parametrizations of the DF-Based Algorithm, $Th_{BufferSize}$ = 1 MB



(a) File Size - 200 kB



(b) File Size - 2 MB

Fig. 4.5: Individual Packet Delay for different Parametrizations of the DF-Based Algorithm, $Th_{BufferSize}$ = 4 MB

file sizes greater than this value, the algorithm should converge to a 1:1 allocation, almost instantly. This can observed in Figure 4.4b, where a switch is effected immediately, and after some time the buffering delay exceeds any of the predefined thresholds. Figure 4.4a shows that for the smaller 200 kB file size, one can observe a slight difference in behaviour for the different investigated parametrizations. In this case, a low $Th_{HOLDelay}$ will tend to affect a switch more frequently, ensuring fairness, and reducing the time until the packet is first serviced. This however sacrifices the achievable throughput of ongoing sessions which might be close to termination, resulting in a small penalty gain in session throughput, when compared to parametrizations which delay the switch direction.

In Figure 4.5, the $Th_{BufferSize}$ threshold is raised to 4 MB, allowing us to study the effect of the $Th_{HOLDelay}$ threshold parameter. While for the 2 MB payload file size in Figure 4.5b the effect is almost negligible, Figure 4.5a shows similar trends to what is observed in Figure 4.4a. Again, one can observe that fast link direction switching, a property achieved by reducing the $Th_{HOLDelay}$ parameter, can ensure fairness between the candidate payloads to be transmitted, at the cost of slight session throughput degradation for already ongoing sessions.

The effect of the $Th_{HOLDelay}$ parameter starts playing an important role in the presence of uni-directional transfers operating over higher layer protocols requiring acknowledgement feedbacks, such as Transmission Control Protocol (TCP). This has been studied in Paper B, [69]. In this case, the DF-based algorithm presented above benefits from a larger $Th_{HOLDelay}$, especially at large payloads, since it does not over-prioritize TCP ACK's which are typically small, inhibiting useful TDD slot resources which can be used for data transfer. This loss due to TCP ACK over-prioritisation becomes especially wasteful once the TCP congestion window size has grown sufficiently, allowing data to be transferred over multiple TDD slots.

Moreover, the algorithm presented above does not incur any penalties due to long $Th_{HOLDelay}$ thresholds in the initial phase of TCP, the phase at which TCP ACK's are particularly important to allow the growth of the congestion window. This is true because once the congestion window has been exhausted, TCP will block the transfer of any additional data if a TCP ACK has not been received, and the algorithm will simply adapt and allocate resources to the link direction containing data in its buffers, in this case the link containing the TCP ACK.

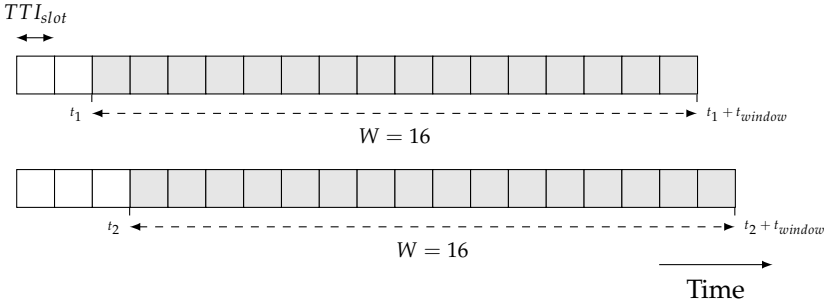


Fig. 4.6: Sliding Window

4.4 Load-Fairness Based Algorithm

4.4.1 Algorithm Description

As observed in the preceding section, for large bursts of traffic, such as file transfers, the delay experienced by the packets is directly correlated to the file size and the required buffering time. For a large file transfer, the throughput experienced whilst delivering that file is more meaningful than the delay experienced by the individual fragmented packets of that file, since the file needs to be received in its entirety for it to be useful.

The proposed LF based algorithm decides the link direction based on the current traffic share and previous slot allocation share between the two links. This achieves a subtle balance with regards to the fairness of the allocations based on the traffic load in each direction and the minimum amount of slots which need to be given to the lightly loaded link direction.

The allocation of resources to the particular link direction is done by computing a priority metric on each of the directions, UL and DL. The metric ξ is computed as follows,

$$\xi_{DL} = \left(\frac{K_{DL}}{K_{Total}}\right)\left(\frac{W - L_{DL}}{W}\right) \tag{4.1}$$

$$\xi_{UL} = \left(\frac{K_{UL}}{K_{Total}}\right)\left(\frac{W - L_{UL}}{W}\right) \tag{4.2}$$

The ratio $\frac{K_D}{K_{Total}}$ denotes the ratio of the link's traffic over the total traffic. The term on the right in Eqs. 4.1 and 4.2 is responsible for considering the

previous slots assigned to each direction over a sliding window W (shown in Figure 4.6) with an arbitrary number of slots. L_{DL} and L_{UL}, denote the number of previous slots allocated to DL and UL respectively over the sliding window W. The following conditions should be outlined to capture the essence of this algorithm concisely. If L equals the window size W, the metric for that particular direction goes down to 0, outlining that since all previous slots have been allocated to a single direction, it is now time to allocate slots to the other direction, no matter how low its traffic share is. On the other hand, if L is equals to 0, the algorithm will simply compute a metric, which is solely dependent on the traffic share of each link. Therefore the scheduling priority metric depends on the traffic share of each link, and the amount of previously assigned slots to each link. In general, the higher the traffic share, the higher the priority, and the more slots previously assigned to a particular link, the lower its priority. Figure 4.7 depicts the aforementioned characteristics graphically showing how the metric varies with varying traffic share and previously assigned slots. The pseudo code for the LF based scheme is given in Algorithm 2.

Algorithm 2 LF-Based Algorithm

```
Initialize window size W, L_UL = 0, L_DL = 0
for each time slot do
    if K_UL = 0 and K_DL = 0 then
        D = Mute
    else if K_UL > 0 and K_DL = 0 then
        Register UL direction in sliding window W, L_UL = L_UL + 1
        D = Uplink
    else if K_UL = 0 and K_DL > 0 then
        Register DL direction in sliding window W, L_DL = L_DL + 1
        D = Downlink
    else if K_UL > 0 and K_DL > 0 then
        Compute ξ_UL and ξ_DL
        if ξ_UL > ξ_DL then
            D = Uplink
            Register UL direction in sliding window W, L_UL = L_UL + 1
        else if ξ_UL < ξ_DL then
            D = Downlink
            Register DL direction in sliding window W, L_DL = L_DL + 1
        end if
    end if
    Slide Window
end for
```

A large value of W will represent a more accurate and granular representa-

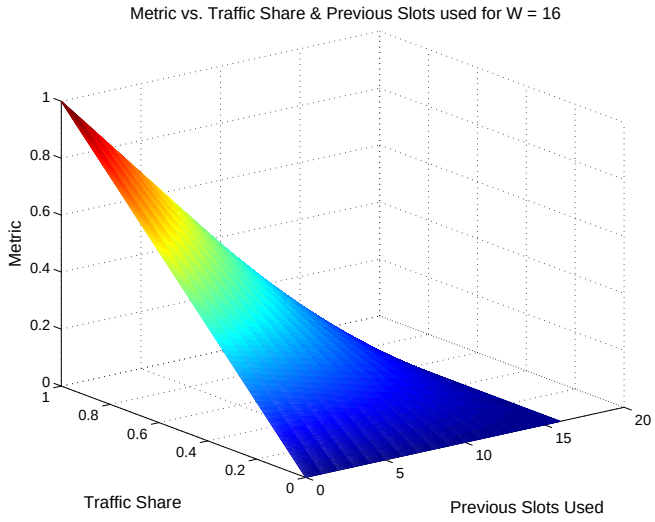


Fig. 4.7: Priority Metric vs. Traffic Share $\frac{K_D}{K_{Total}}$ and Previously Assigned Slots L for W = 16

tion of the traffic share, but setting W to be large will also mean that if the traffic share of a particular link is very low, it will wait longer before it is scheduled.

The direction D having the largest calculated ξ metric is selected for scheduling. This calculation is only carried out when traffic is present in both directions. If no traffic is present in either direction, the scheduler will mute the slot, while if traffic is present in only one direction, the scheduler will naturally schedule the direction having traffic. Summarizing,

$$D = \begin{cases} MUTE & \text{if } K_{DL} = 0 \text{ and } K_{UL} = 0 \\ DL & \text{if } K_{DL} > 0 \text{ and } K_{UL} = 0 \\ UL & \text{if } K_{DL} = 0 \text{ and } K_{UL} > 0 \\ \underset{D}{\operatorname{argmax}}\, \xi_D & \text{if } K_{DL} > 0 \text{ and } K_{UL} > 0 \end{cases} \tag{4.3}$$

4.4.2 Algorithm Behaviour

To further understand the dynamics of this algorithm, let us consider the case in which the DL direction is chosen over the UL direction. Moreover, let us add a weighting factor α, where $0 \leq \alpha \leq 1$, which allows us to control the

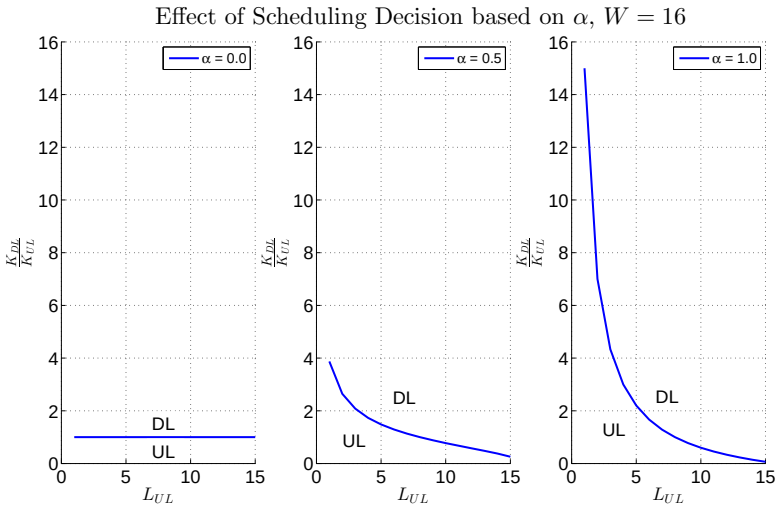


Fig. 4.8: Effect of Alpha on Priority Metric

relationship between the traffic share and previous assigned slots to UL and DL respectively.

$$\begin{aligned}
\left(\frac{K_{DL}}{K_{Total}}\right)\left(\frac{W-L_{DL}}{W}\right)^{\alpha} &> \left(\frac{K_{UL}}{K_{Total}}\right)\left(\frac{W-L_{UL}}{W}\right)^{\alpha} \\
K_{DL}\left(\frac{W-L_{DL}}{W}\right)^{\alpha} &> K_{UL}\left(\frac{W-L_{UL}}{W}\right)^{\alpha} \\
\frac{K_{DL}}{K_{UL}} &> \frac{\left(1-\frac{L_{UL}}{W}\right)^{\alpha}}{\left(1-\frac{L_{DL}}{W}\right)^{\alpha}} \\
\frac{K_{DL}}{K_{UL}} &> \frac{(W-L_{UL})^{\alpha}}{(W-L_{DL})^{\alpha}}
\end{aligned} \tag{4.4}$$

If it is assumed that all slots in the window where utilized, i.e. no muted slots, $L_{DL} = W - L_{UL}$, Eq. 4.4 can be rewritten as,

$$\begin{aligned}
\frac{K_{DL}}{K_{UL}} &> \frac{(W-L_{UL})^{\alpha}}{(L_{UL})^{\alpha}}, \\
&\text{where } L_{UL} > 0
\end{aligned} \tag{4.5}$$

From this formulation, it can be inferred that if $\alpha = 0$, the scheduling decision will simply be based on the traffic share, irrespective of the previous slot

Table 4.2: Simulation Parameters for LF-Based Scheme Evaluation

Traffic Parameters	
Cell Offered Load	600 Mbps
Downlink Offered Load	514 Mbps
Uplink Offered Load	86 Mbps
Approximate Resource Utilization	75%
File Size	2 Mb
Algorithm Parameters	
Alpha	0, 0.5, 1.0

allocations. If $\alpha = 1$, the DL direction will only be chosen if the DL over UL traffic ratio is greater than $\left(\frac{W}{L_{UL}} - 1\right)$. If many UL slots have been previously assigned i.e. large L_{UL} value, a smaller DL over UL traffic ratio needs to be present for DL to be scheduled, and vice versa if few UL slots have been previously assigned. The relationship between the metric, the number of previous assigned slots L, and the parameter value α can be captured in Figure 4.8, where the decision regions for UL and DL are illustrated.

It is therefore expected that for $\alpha = 0$, higher session throughputs should be achievable for the highly loaded traffic direction, at the cost of reduced performance for the sessions contained in lightly loaded buffers. As α grows further, the algorithm will not only consider the instantaneous buffer size ratio between the two links, but it will also give priority to lightly loaded links which did not have the chance to be served over the past TDD slot allocations. This will in turn limit the throughputs of the highly loaded direction, but it will improve the session throughput performance of lightly loaded link direction.

4.4.3 Performance Evaluation

In order to capture and validate the impact of the α weighting factor, system level simulations are conducted. The simulations are parametrized to operate with around 75% resource utilization and with a high asymmetry between DL and UL, with the offered load for DL being six times that for UL. This allows us to increase the probability of instantaneous concurrent UL and DL data, and at the same time stress the condition where one link has persistently more traffic than the other. The parameters used for the the simulations are summarized in Table 4.2.

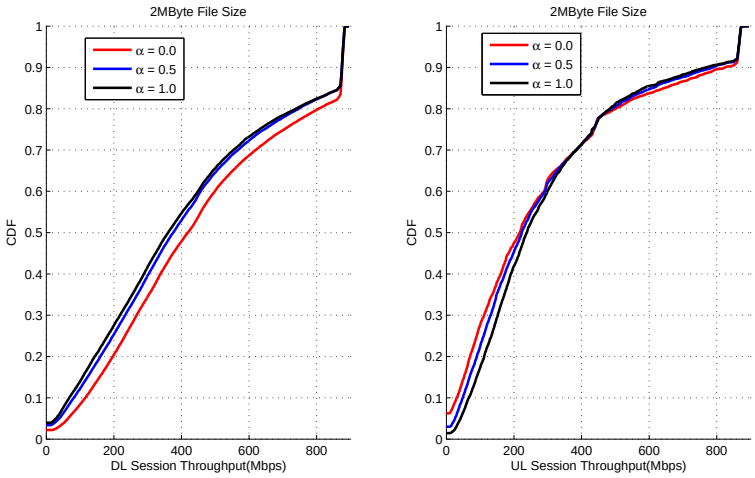


Fig. 4.9: LF-Based Algo with DL:UL 6:1 Asymmetry

The results showing the application layer session throughput in the UL and DL directions are illustrated in Figure 4.9. As expected, on $\alpha = 0$, the highly loaded DL link's performance, benefits at the expense of the lightly loaded UL link. On the other hand, a value of $\alpha = 1$, gives some priority to lightly loaded link, improving the UL session throughput in this case, and degrading the performance of the highly loaded link direction. The value $\alpha = 0.5$ attempts to achieve a subtle balance by prioritizing the DL highly loaded direction in a not so aggressive manner as the $\alpha = 0.0$ case.

4.5 Comparative Analysis

In this section we will compare the DF and LF based algorithms, by inspecting the way they conduct their TDD allocations and by analysing their principle characteristics. This comparison is done via multi-cell system level simulations which also take into account the impact of recovery mechanisms, i.e. HARQ and RLC. Specifically we will investigate how the individual algorithms allocate the available TDD slot resources under low and high load, and with different DL to UL traffic asymmetries. The range of simulation parameters under which the algorithms are compared is summarized in Table 4.3.

[1]Refers to an average value, generated according to an exponential distribution

Table 4.3: Simulation Parameters

Traffic Parameters		
Traffic Model	3GPP FTP Traffic Model 1 [30]	
File Size	2 MB [1]	
	Symmetric Traffic (1:1)	
	Low Load	**High Load**
DL Load	100 Mbps	250 Mbps
UL Load	100 Mbps	250 Mbps
Cell Resource Utilization	25%	75%
	Asymmetric Traffic (6:1)	
	Low Load	**High Load**
DL Load	172 Mbps	429 Mbps
UL Load	28 Mbps	71 Mbps
Cell Resource Utilization	25%	75%
Algorithm Parameters		
DF Based Scheme		
$Th_{HOLDelay}$	4 ms	
$Th_{BufferSize}$	8 MB	
LF Based Scheme		
Window Size	16	
Alpha	1.0	

The results showing the individual UL and DL TDD slot resource utilization for a balanced 1:1 (DL:UL) and unbalanced 6:1 (DL:UL) traffic share, are respectively shown in Figures 4.10 and 4.11.

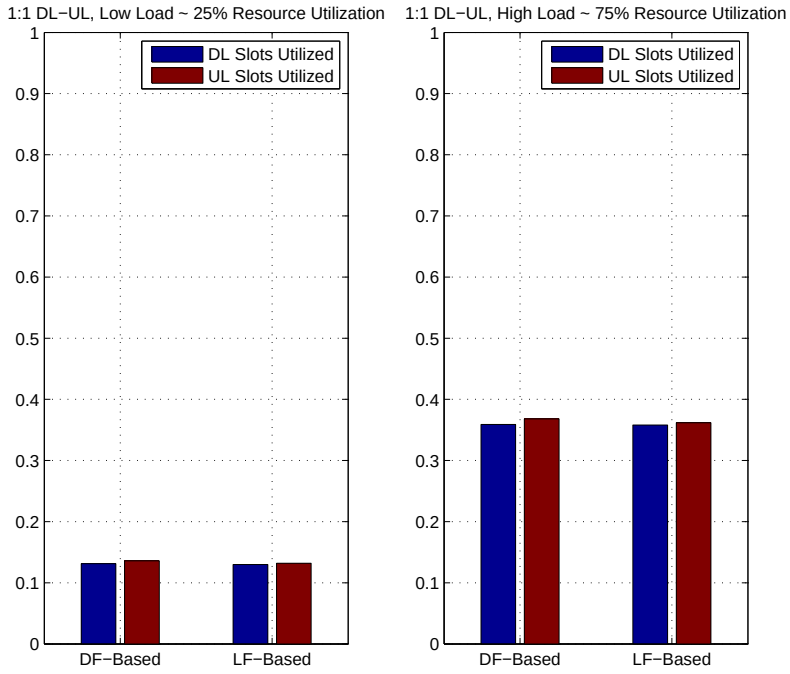


Fig. 4.10: DL/UL Resource Utilization, Symmetric 1:1 Traffic

As seen from Figure 4.10, the algorithms allocate their resources in a similar fashion for a 1:1 traffic profile, ensuring fairness between the two links. This happens at both low and high load.

When traffic asymmetry is introduced, in the presence of both UL and DL data, the DF-Based algorithm tends to converge to a 1:1 allocation, since the buffering time exceeds any reasonably defined threshold, even though in this case the DF-based algorithm is parametrized with a $Th_{HOLDelay}$ threshold of 4 ms. While it is natural that the DL slot utilization is higher in both cases, since most of the traffic is in DL, in Figure 4.11, it must also be said that the LF-based algorithm manages to capture and consider the long term traffic asymmetry better at high load, allocating the available slot resources accordingly. This behaviour occurs only at high load i.e. higher probability of concurrent UL and DL traffic, since in these conditions the individual

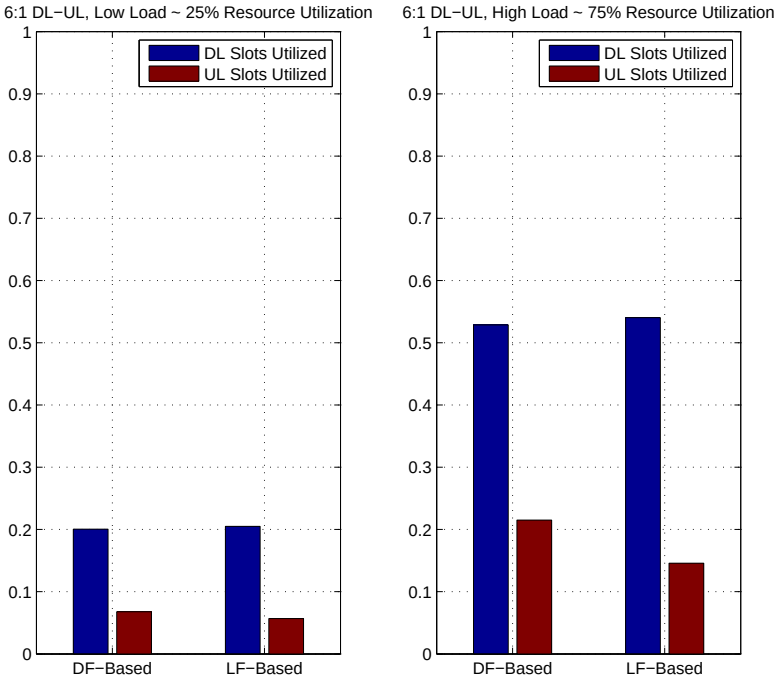


Fig. 4.11: DL/UL Resource Utilization, Asymmetric 6:1 Traffic

algorithms need to decide whether to prioritize the DL or UL traffic.

To understand the implications of the algorithms' TDD slot allocations, we inspect the individual UL and DL application layer session throughput performance, and the overall session throughput performance of all the generated sessions, for the high load case. The CDF's showing these results for the symmetric and asymmetric traffic cases are respectively illustrated in Figure 4.12 and Figure 4.13.

From Figure 4.12, one can observe that the DF based scheme gives slightly better outage performance at the cost of reduced high end session throughput performance when compared to the LF based scheme. In the presence of UL and DL data, the LF-Based scheme will tend to prioritize the direction instantaneously having a lot traffic, while the DF-Based scheme will simply converge to a 1:1 allocation after the $Th_{HOLDelay}$ threshold is exceeded, capturing accurately the symmetric traffic present in the system. The trends in UL and DL session throughput performance are however similar for both

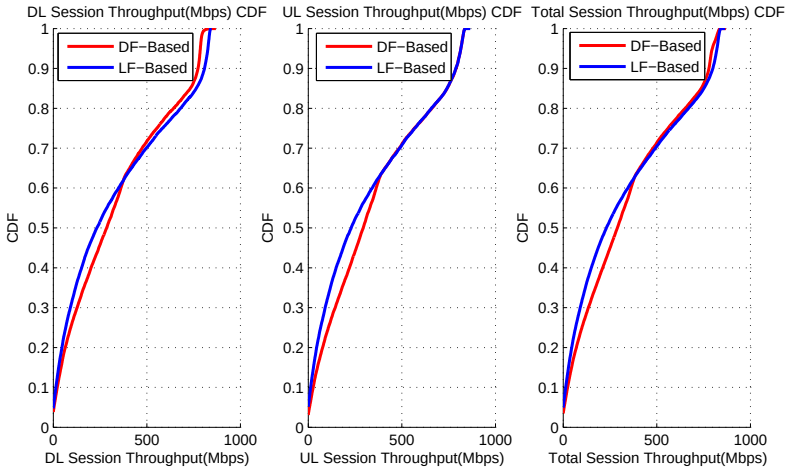


Fig. 4.12: Session Throughput in DL and UL - DF-Based vs. LF-Based Algo, Symmetric 1:1 (DL:UL) Traffic Share Case

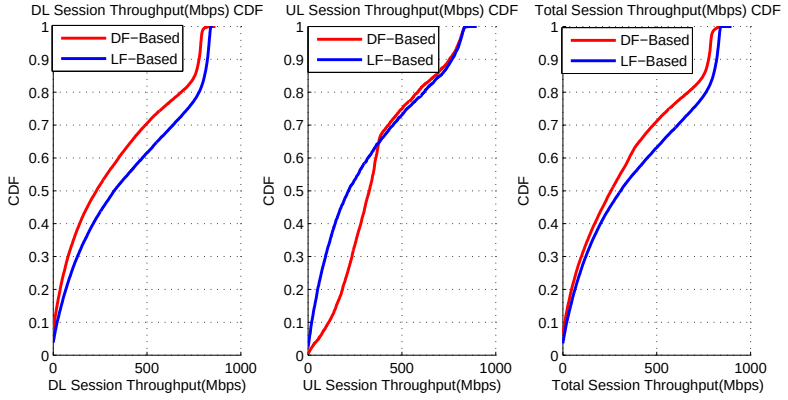


Fig. 4.13: Session Throughput in DL and UL - DF-Based vs. LF-Based Algo, Asymmetric 6:1 (DL:UL) Traffic Share Case

schemes, due to the presence of symmetric traffic.

The asymmetric traffic results in Figure 4.13 confirm that the LF-Based algorithm can capture the traffic asymmetry more accurately offering a superior DL throughput performance. On the other hand the DF-based algorithm converges to an equal UL:DL allocation when both DL and UL buffers are filled with data, hence limiting the DL session throughput, and improving the UL session throughput in the lower end of the CDF curve, when compared to the LF-Based algorithm. However, since in this case most of the sessions are in DL, the LF-Based algorithm manages to offer the best overall session throughput performance, giving a superior performance when one considers all the sessions generated and received through the network.

4.6 Wrap-up

In this chapter we have first explained the link scheduling framework, enabling the flow of information supporting the operation of the proposed F-TDD algorithms. Thereafter, the DF and LF based algorithms were explained, and their characteristics and behaviour were analysed in the light of the different possible parametrizations. Finally a comparative analysis of the two schemes was conducted.

In the next chapter we will propose a distributed rank adaptation scheme, which enables us to use multiple spatial streams for transmission in favourable conditions. Up to this point it has been assumed that all the antenna resources at our disposal were used for interference suppression. As seen from chapter 3, this approach was deemed to be effective in dealing with the additional inter-cell interference variation brought along by F-TDD. The introduction of the rank adaptation will reduce the antenna resources used for interference suppression, and will re-introduce some additional inter-cell interference variation. It is therefore important to see whether with the introduction of this scheme, F-TDD operation still offers gains over fixed S-TDD.

Chapter 5

Rank Adaptation

The availability of multiple transmit and receive antennas has fuelled a lot of research on MIMO. Various topics such as Single User MIMO (SU-MIMO), Multi User MIMO (MU-MIMO), beamforming, transmit/receive diversity, and spatial stream power allocation have in fact been investigated numerous times. The scope of this chapter is limited to the proposal and evaluation of a rank adaptation scheme, which chooses the number of spatial streams used for transmission, assuming the availability of MMSE-IRC receivers capable of suppressing interference.

In the previous chapters we have shown how the proposed frame structure along with the availability of MMSE-IRC receivers and multiple antennas, can significantly alleviate the problem of inter-cell interference along with the inherent channel estimation challenges brought along by F-TDD. This observation is valid if the degrees of freedom brought along by our antennas are used to suppress interference. In principle, for a system with N receive antennas, up to $N-1$ interfering streams can be suppressed effectively.

The envisioned 5G concept presented earlier on, assumes the availability of 4 antennas at both the transmitter and receiver end. The system can be configured to transmit with 1 spatial stream. Assuming the presence of MMSE-IRC receivers, in this configuration, the available degrees of freedom are used to suppress interference, giving us the ability to retain a good outage performance [21], and minimize the experienced inter-cell interference variation introduced by F-TDD. While attractive in highly interfered scenarios, such a transmission configuration limits our ability to reach higher peak throughputs, a highly desirable characteristic in future 5G systems.

There is therefore a classical trade-off to be made which basically limits the amount of transmitted spatial streams to be used, to either retain outage performance or achieve higher peak throughputs. Ideally, the number of spatial streams used for transmission, hereinafter referred to as the transmission rank, needs to be low in highly interfered conditions. Conversely, in low interference conditions, a high transmission rank should be preferred, in order to allow the system to reach high peak throughputs. This technique is known as rank adaptation and is the focus of this chapter.

A fully distributed and opportunistic rank adaptation scheme will inevitably introduce further inter-cell interference variation for both fixed S-TDD and F-TDD systems, causing additional challenges for any scheme requiring stable or slowly varying channel conditions. The sustained usage of high transmission ranks will also remove the ability to suppress interference effectively, causing harm for F-TDD operation by re-introducing the inter-cell interference variation problem. On the other hand, for a given absolute load, F-TDD can service an arriving session more quickly, potentially reducing the probability that two neighbouring cells are simultaneously active. This will in turn allow some nodes to perceive better channel conditions, giving them further incentive to use higher transmission ranks, and thus allowing them to experience higher peak throughputs. Apart from the performance validation of the rank adaptation algorithm itself, it is therefore also of interest to understand the interaction between rank adaptation and F-TDD.

The chapter is structured as follows. We will first give a short overview of the different nature of rank adaptation schemes proposed in literature, outlining also some rank adaptation algorithms which have been proposed for the envisioned 5G concept. Thereafter, we will present the main challenges associated with the design of a rank adaptation algorithm in the envisioned 5G concept. Following that, the proposed rank adaptation algorithm will be presented, and its intended behaviour and performance will be evaluated against a selfish rank adaptation scheme and fixed rank transmission schemes. Finally, we will inspect the performance of the rank adaptation algorithm in fixed S-TDD and F-TDD configurations to understand whether F-TDD operation should still be promoted in more realistic scenarios.

5.1 Rank Adaptation Schemes

The goal of this section is focused on presenting the diverse nature of rank adaptation techniques, highlighting whether the spatial degrees of freedom are used in a selfish or cooperative way. The cooperative approaches can also

be further classified into centralized, semi-distributed and fully distributed approaches, each requiring different degrees of signalling information exchange to apply such cooperation.

An example of a distributed selfish rank adaptation algorithm is presented in [70]. The goal of the algorithm is to choose a rank transmission configuration which maximizes the capacity from an own cell's perspective, irrespective of the potential harm which may be generated to neighbouring cells. In this work, the entities deciding the transmission rank consider the presence of inter-cell interference and simply choose a transmission rank which maximize their own capacity.

On the other hand, co-operative approaches attempt to improve the performance of the whole system, instead of just focusing on the individual gain in one's own cell. An example of a practical cooperative distributed rank coordination scheme aiming at improving the system-wide performance via an interference pricing mechanism is given in [71]. In this scheme the nodes are coordinated via a cyclic master-slave architecture and it is assumed that limited amount of information can be exchanged. Another cooperative scheme is presented in [72], where the authors present an algorithm which assigns precoding vectors which attempt to reach a compromise between the beamforming gain and the mitigation of interference created towards other receivers.

5.1.1 Rank Adaptation in the 5G Small Cell Concept

Some rank adaptation techniques, specifically targeting the 5G small cell concept, have already been presented in [73] and [74]. In these works and in our proposed rank adaptation algorithm, we consider rank adaptation techniques which decide upon the amount of spatial streams to be used, assuming the presence of interference suppression receivers as envisioned in the presented 5G small cell concept described earlier on.

A maximum rank planning algorithm is presented in [73]. The algorithm statically limits the amount of spatial streams which can be used. By limiting the transmission rank, the authors show that given the availability of MMSE IRC receivers, the maximum rank planning technique can serve as an alternative to frequency reuse planning, allowing the preservation of outage performance when inter-cell interference levels are high. The dynamic assignment of the transmission rank based on long-term traffic variations is reserved as future work in this particular study.

A fully distributed taxation-based rank adaptation technique is presented in [74]. Similar to our proposed technique which will be described in the fol-

lowing section, the transmission rank is chosen based on the maximization of a utility function which considers the achievable rate for each rank and a taxation function which is dependent on the transmission rank and the perceived interference levels. In this study, the perceived interference levels are classified into 3 states, low, moderate and high. Based on this classification, a corresponding taxation function is applied. In this work it was however assumed that future interference conditions are known before-hand. Such an assumption does not fit well in the defined 5G framework, especially when one considers the causal aspects of a practical system, and the potential inter-cell interference variation occurring from one frame to the next due to the usage of flexible TDD.

Typically there is a delay from when a rank transmission decision is conducted to when it is applied (in our case, 1 frame, i.e. from Scheduling Grant to transmission). Given the distributed nature of the rank adaptation algorithm and the full freedom in assigning the link direction to either UL or DL, the interference conditions at the point when the rank transmission decision was done and the interference conditions at the point when the transmission is applied might be significantly different.

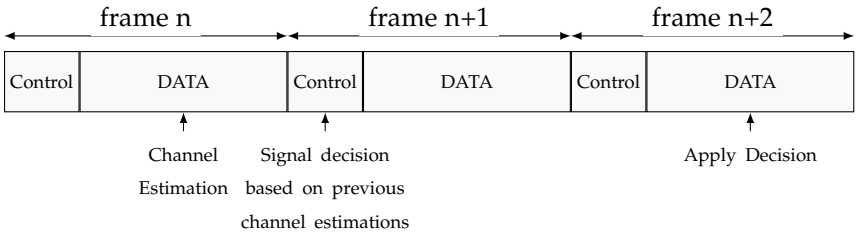


Fig. 5.1: Causality Problem

A pictorial representation of this problem is illustrated in Figure 5.1. In the studied framework, it is assumed that the channel is estimated from the previous receptions occurring in the data part of the frame. Following this channel estimation process at frame n, a decision is signalled in the control part of frame $n+1$, and affected on the following frame $n+2$.

This process is done in a parallel fashion, meaning that all cells are synchronized, and take a decision at the same time. Given the per frame link direction freedom introduced by F-TDD along with the consequential and parallel nature of the process, the decisions taken at frame $n+1$ will be based on potentially outdated conditions.

Such conditions will of course lead us to apply a non-optimal rank selection decision. Even if such a condition is not explicitly addressed, it is of utmost

importance to consider and capture the limitations it might bring along, such that appropriate and realistic conclusions can be drawn when considering the gains and interaction of rank adaptation and F-TDD.

5.2 Taxation Based Rank Adaptation Scheme

The proposed taxation based rank adaptation scheme was published in [75]. In this section we will present the inner workings of the algorithm. The operating principles of the algorithm are inspired by the game-theoretic based channel assignment scheme proposed in [76].

The main philosophy of the taxation based rank adaptation scheme revolves around the concept of a social tax, where one is taxed based on his earnings and the harm generated to the other entities in the society. If a high transmission rank is used, and a high level of outgoing interference is generated, a high tax should be applied. This approach is justifiable, because the choice of a high transmission rank implies increased difficulties for a victim node to suppress interference effectively. It is important to note that the amount of outgoing interference being generated, quantifies the harm inflicted on a set of neighbouring victim nodes. Therefore the applied tax should also consider this factor into account, meaning that the choice for a high transmission rank should not be discouraged via a high taxation if the outgoing interference is low, and the harm being induced is limited.

$$k^* = \arg\max_k \Pi_k \tag{5.1}$$

In essence, the taxation-based rank adaptation scheme, abbreviated as TB-RA hereinafter, aims at choosing a rank k^* which maximizes the utility function Π_k in Equation 5.1, which considers the achievable rate when choosing rank k, and a corresponding taxation term based on the,

- chosen rank k,
- a monotonically increasing weighting vector W_k also based on the rank k, and
- the capacity of the past incoming interference conditions, $C\left(\frac{I}{N}\right)$.

A block diagram showing how a rank transmission decision is made, is shown in Figure 5.2.

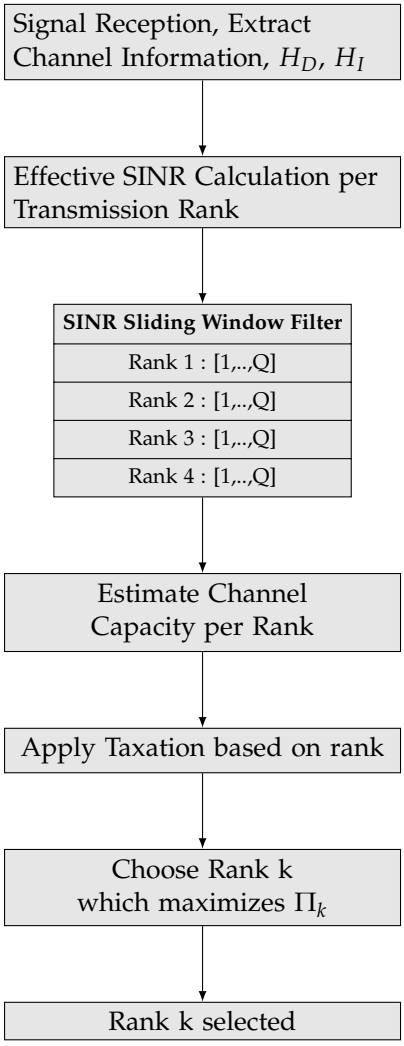


Fig. 5.2: TB-RA Operation

On the reception of a signal, the desired channel matrix H_D and the matrix of the interfering channels H_I, are extracted. The matrix H_D can be acquired over time and is not expected to vary rapidly. On the other hand the value of H_I can potentially vary from one frame to the next due to the usage of F-TDD. However the presence of the DMRS symbol in our frame structure enables us to obtain an updated estimate of the interference channel matrix H_I, subsequently allowing us to compute the interference covariance matrix.

Once this information is retrieved, the effective SINR for each possible rank transmission strategy is estimated, giving us four effective SINR values. The calculated effective SINR values for each rank are then placed in a sliding window filter containing Q samples. The log-averaged SINR is then calculated for each of the ranks, and an estimate of the achievable rate for each of the ranks is obtained.

Once the achievable rate for each of the considered transmission ranks is obtained, a taxation term based on the rank, and the incoming interference conditions is applied. The rank k which maximizes utility function Π_k is then chosen.

The considered utility function is represented mathematically in equations 5.2 to 5.6.

$$\Pi_k = \underbrace{kC(\overline{SINR_{effective_k}})}_{\text{Estimated Capacity for rank k}} - \underbrace{kW_kC\left(\frac{I}{N}\right)}_{\text{Taxation for rank k}} \tag{5.2}$$

where $C(\overline{SINR_{effective_k}})$ is estimated as,

$$C(\overline{SINR_{effective_k}}) = log_2\left(1 + \overline{SINR_{filtered_k}}\right) \tag{5.3}$$

$\overline{SINR_{filtered_k}}$ represents the log-averaged effective $SINR$ at rank k over the previous Q samples, and is given by,

$$\overline{SINR_{filtered_k}} = \frac{1}{Q}\sum_{i=1}^{Q}\left(\frac{1}{k}\sum_{j=1}^{k} SINR_{ij}\right) \tag{5.4}$$

$$= \frac{1}{kQ}\sum_{i=1}^{Q}\sum_{j=1}^{k} SINR_{ij} \tag{5.5}$$

with $SINR_j$ representing the $SINR$ of stream j.

$C\left(\frac{I}{N}\right)$ is given by,

$$C\left(\frac{I}{N}\right) = log_2\left(1 + \frac{tr(H_I {H_I}^H)}{\sigma_n^2}\right) \tag{5.6}$$

and represents the estimated capacity of the incoming interference over noise ratio, over the previous receptions. When considering the taxation and specifically the interference term, one should ideally consider the outgoing rather than the incoming interference, since this represents the actual harm generated to the other nodes. Obtaining this information in a fully distributed manner is generally difficult and would require some alterations in the frame structure or explicit communication between the nodes. We therefore assume that the incoming interference is equal to the outgoing interference, even if this is not always the case. Moreover, the incoming interference should give us a rough estimate of the currently perceived interference levels.

It is important to note that the rank transmission decision calculation is done internally at both the AP and UE. The UE will simply decide a DL transmission rank based on the calculation of the utility function, considering its locally perceived interference conditions. Such an approach enables a receiver centric operation, relieving the AP from requiring accurate UE channel information and calculating the corresponding DL transmission rank based on this information. After deciding the desired DL transmission rank, the UE simply feeds back this information to the AP via the SR message in the UL control channel. The AP, the final decision maker, will simply utilize this information when instructing which transmission rank to use in DL.

5.3 Behavioural Evaluation

In this section we shall evaluate the behavioural aspects of the TB-RA algorithm, showing how the rank adaptation algorithm adapts in different interference conditions.

The interference conditions will be controlled by first varying the traffic load, showing whether the TB-RA algorithm can adapt in time to different interference conditions. Afterwards, we will vary the deployment ratio to control the intensity of the interference conditions, and see whether the TB-RA algorithm can adapt accordingly in less dense scenarios.

In doing so we will assume a flexible TDD slot allocation based on the DF-Based F-TDD scheme, operating under an equal UL to DL traffic load. We will also benchmark the algorithm against a selfish scheme which applies no taxation, hereinafter denoted as SRA, and fixed rank 1 and 2 transmission schemes. The TB-RA algorithm will also be configured with two different W_k vector parametrizations representing a conservative and aggressive rank transmission selection scheme. The elements of the W_k vectors specify the amount, by which a particular rank transmission strategy should be taxed.

The W_k vectors are chosen to be W_1 = [0, 0.5, 0.66, 0.75], and W_2 = [0, 0.25, 0.66, 0.75]. When comparing W_1 and W_2, one can observe that a rank 1 transmission is not taxed, and rank 3 and 4 transmissions are taxed equally. The only difference lies in how a rank 2 transmission is taxed, with W_1 applying a higher tax for this particular rank configuration. For this sole reason, the W_1 and W_2 configurations are respectively referred to as the conservative (TB-RA Conservative) and aggressive (TB-RA Aggressive) configurations.

The inspected KPI metric by which the performance of the individual schemes are evaluated, is the Average Node Session Throughput at the application layer. This represents the average experienced session throughput over multiple sessions by a particular node during the course of a simulation.

5.3.1 Traffic Load

In order to assess the ability of the rank adaptation algorithm to adapt to different interference conditions, we load the system to approximately use 25% and 75% of the resources, corresponding to an offered load of 100 Mbps and 250 Mbps per node respectively.

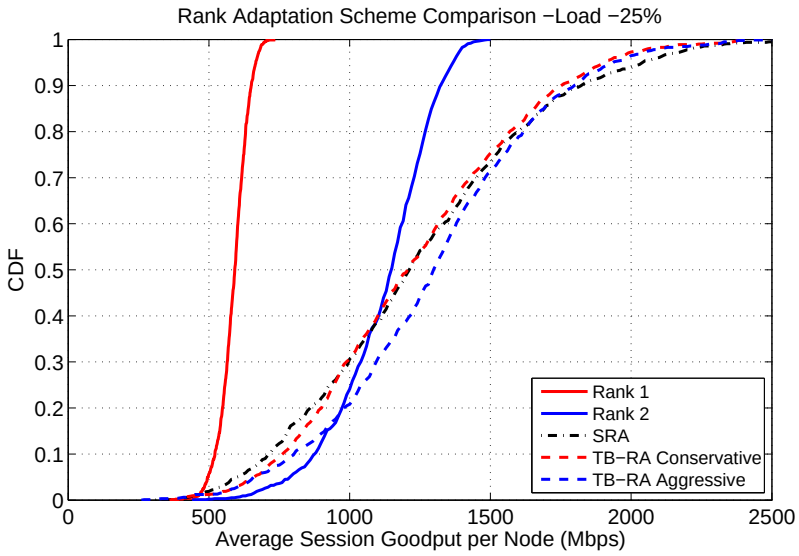


Fig. 5.3: TB-RA at low load, RU 25%

The CDF of the average node session throughput for the individual schemes at around 25% resource utilization is shown in Figure 5.3. In this case, the interference conditions are low, and the fixed rank 1 scheme's throughput is

limited. At low load conditions, the fixed rank 2 scheme offers a superior performance, even on the low end of the CDF curve. The performance of rank 1 was previously shown to provide the best outage performance in the studies carried out in [73], but in that particular study full buffer traffic was assumed. The result shown in Figure 5.3 stresses the importance of testing the performance of a scheme over different traffic loads, and also proves the need for a rank adaptation scheme which can adapt to different interference and traffic conditions.

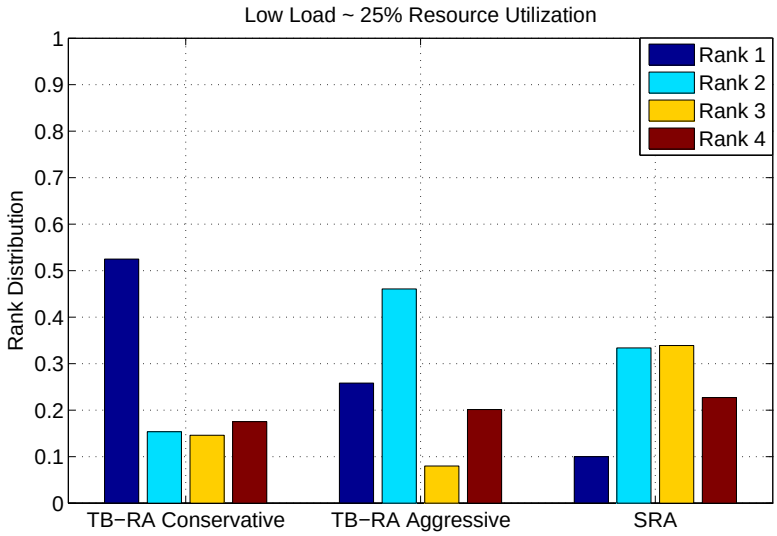


Fig. 5.4: TB-RA at low load, RU 25%

Figure 5.3 also shows that all the rank adaptation schemes can exploit temporary favourable channel conditions, and hence make use of higher transmission ranks, allowing some nodes to experience higher average session throughputs. The provided result also shows that in the low end of the CDF curve, the rank adaptation schemes can almost achieve fixed rank 2 performance. The rank selection distribution of the different schemes is shown in Figure 5.4. From Figures 5.3 and 5.4, one can see that the conservative TB-RA approach tends to choose the transmission ranks conservatively, with the aggressively configured TB-RA scheme giving the best overall performance. The selfish scheme offers a satisfactory performance in this scheme but tends to choose the transmission ranks a bit too aggressively, leading to a slightly inferior performance in the low end of the CDF curve.

The performance results for the rank adaptation schemes at approximately 75% resource utilization are shown in Figure 5.5. The findings conducted in

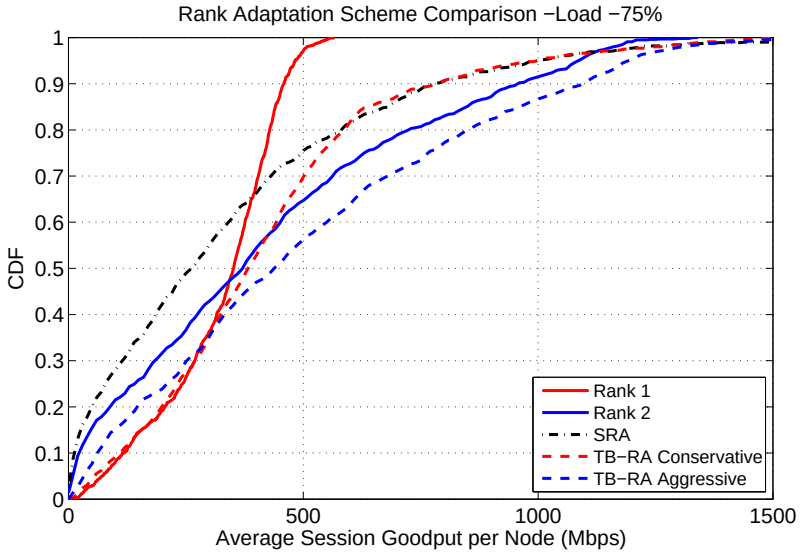


Fig. 5.5: TB-RA at high load, RU 75%

[73] are now in line with our results, since the network is significantly loaded and more severe interference conditions are present, with the fixed rank 1 scheme giving the best outage performance. While the fixed rank 1 scheme offers the best outage performance, it also limits the maximum achievable throughput. The fixed rank 2 scheme has an inferior outage performance in this case, but can reach higher throughputs.

In this case, the different rank adaptation schemes offer a noticeable difference in performance. We once again collectively inspect the average node session throughput CDF, and the rank transmission distribution chosen by the schemes, shown in Figure 5.6. Here we notice that all the rank adaptation schemes lower the chosen transmission ranks as a result of more severe interference conditions. However, the SRA scheme still chooses the ranks aggressively even if there is little gain in doing so. This happens because it designed to maximize its own capacity in a selfish manner, leading to unsatisfactory outage performance, since the MMSE-IRC receiver cannot suppress interference effectively. The conservative TB-RA scheme manages to match fixed rank 1 outage performance, and also exploits the usage of higher transmission ranks where possible. The aggressive TB-RA approach tends to choose slightly higher transmissions ranks than the conservative approach in this case, resulting in higher peak throughputs at the cost of reduced outage performance.

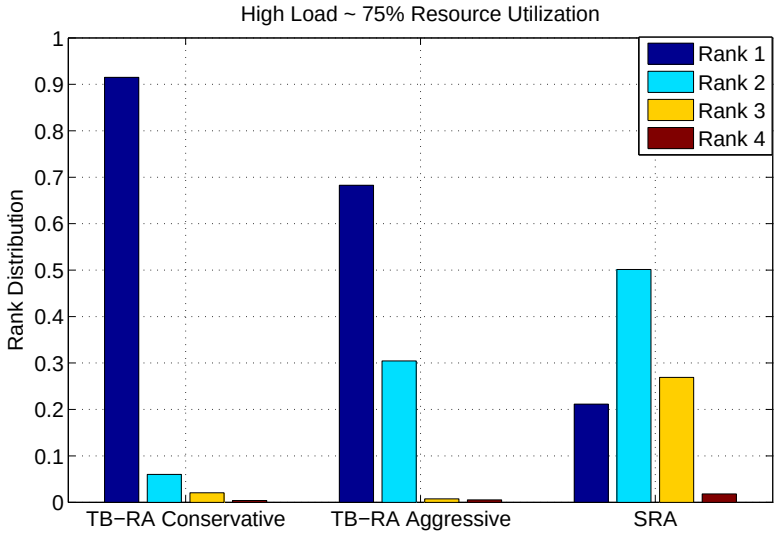


Fig. 5.6: TB-RA at high load, RU 75%

The results illustrated above, show that the proposed TB-RA scheme outperforms the SRA scheme at both low and high loads, and manages to exploit the usage of higher transmission ranks whenever its possible to do so. The taxation term applied in the TB-RA approach limits the transmission rank to be used if high interference conditions are perceived, and avoids choosing a higher transmission rank if there is little gain in doing so. The presented TB-RA algorithm can also be parametrized conservatively to retain outage performance, or aggressively to enjoy higher peak throughputs. The automatic parametrization of the algorithm is left as future work.

5.3.2 Deployment Ratio Analysis

In this section we will study the potential of the presented TB-RA algorithm, by studying its behaviour in more sparse scenarios. The offered load is fixed to represent a resource utilization of around 75% as in the previous high load case. The point of this study is to show whether in more sparse scenarios, the TB-RA scheme can exploit higher transmission ranks.

To conduct this investigation, a 10x2 small cell scenario is considered. The density of the scenario is controlled via a deployment ratio parameter specifying the deployment probability of an AP and UE in the same room. A lower deployment ratio represents a sparser network, while a high deploy-

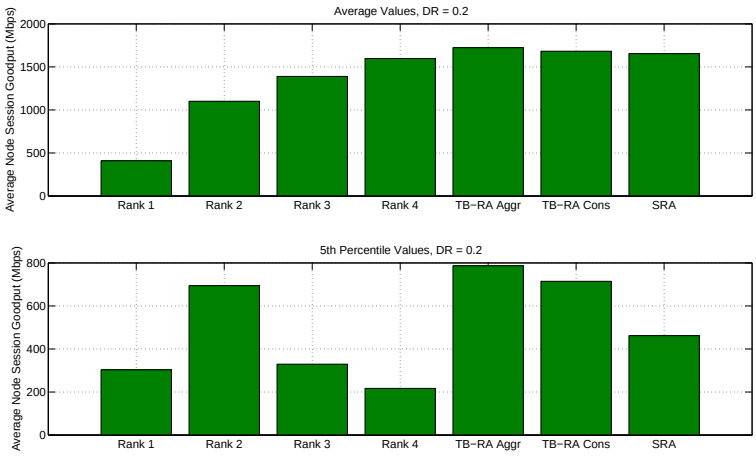


Fig. 5.7: RA Scheme Performance - Deployment Ratio = 0.2

ment ratio represents a dense network.

The average and 5th percentile results of the average node session goodput, for deployment ratios 0.2, 0.6 and 1.0, are shown in Figures 5.7, 5.8 and 5.9 respectively.

A deployment ratio of 0.2 represents a statistically low interference scenario. From Figure 5.7, one can notice that amongst the fixed schemes, the fixed rank 4 scheme offers the best average goodput performance, while significantly degrading the 5th percentile performance. All the benchmarked rank adaptation algorithms show good performance in terms of average throughput, with the SRA scheme significantly losing performance in the 5th percentile performance, due to a too aggressive choice of transmission rank. In this particular scenario we can see that the TB-RA Aggressive scheme offers the best average goodput performance, along with the best 5th percentile performance. A similar result is obtained for the TB-RA Conservative scheme, with slightly inferior performance.

The results for a deployment ratio of 0.6 are shown in Figure 5.8. This scenario produces moderate interference conditions and represents a particularly challenging scenario since any rank adaptation algorithm needs to be conservative and choose lower ranks at times, but also exploit any potential favourable conditions and make use of higher transmission ranks. In this case the TB-RA schemes give the best average performance with the TB-RA Aggressive reaching the highest experienced average throughputs per node. In

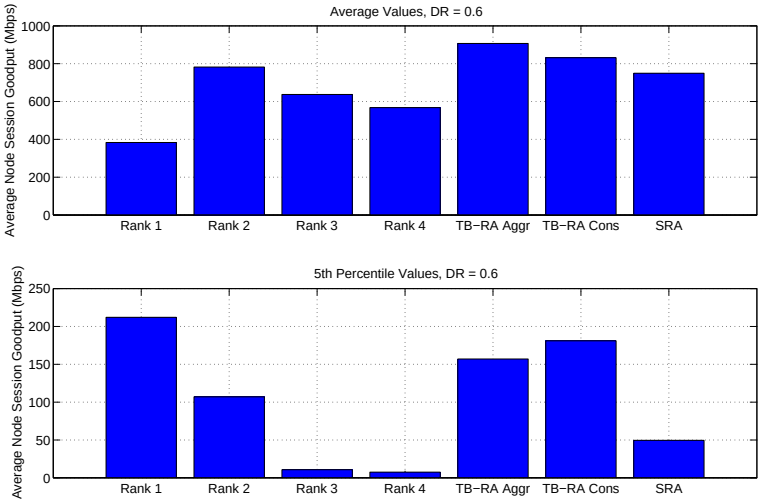


Fig. 5.8: RA Scheme Performance - Deployment Ratio = 0.6

the 5th percentile case the fixed rank 1 scheme offers the best performance, with the TB-RA conservative scheme offering a similar yet slightly inferior performance due to the opportunistic usage of higher ranks generating additional interference which can not always be effectively suppressed.

The results for a deployment ratio of 1.0 are shown in Figure 5.9. This case represents the previously investigated high load case in the preceding subsection. Here we notice again that the TB-RA Aggressive approach offers the best average performance, but offers inferior performance to the TB-RA Conservative approach in the 5th percentile metric, since it tends to choose the transmission ranks to aggressively.

In all these cases it can be seen that the TB-RA approach with one configuration or the other, can reach and even exceed the best average performance, and also approach and sometimes exceed the best 5th percentile performance. These results show that the proposed TB-RA scheme can adapt to different interference conditions automatically while at the same time offering the best possible performance when compared to the fixed rank schemes or the SRA algorithm.

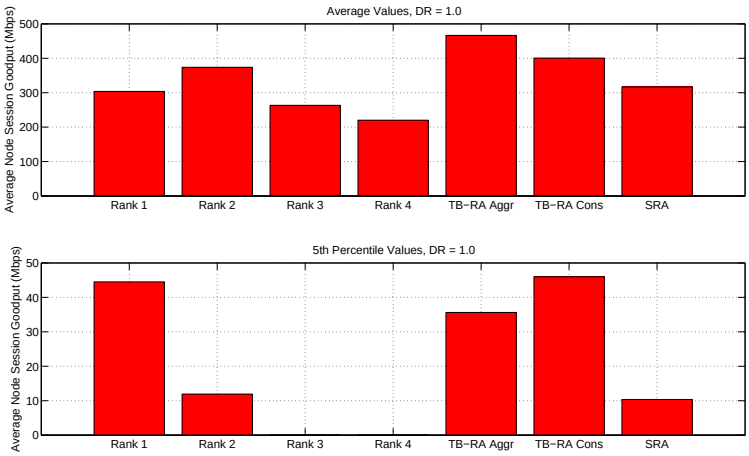


Fig. 5.9: RA Scheme Performance - Deployment Ratio = 1.0

5.4 Rank Adaptation and Flexible TDD

In this section we would like to investigate the performance of F-TDD against fixed S-TDD, when operating over the presented TB-RA scheme. In Chapter 3, the potential demerits of F-TDD related to added inter-cell interference variation were outlined. It was however shown that this problem can be counteracted, if the available degrees of freedom are used to suppress interference with the help of MMSE-IRC receivers. The presented rank adaptation algorithm however allows the usage of higher transmission ranks where possible thus limiting the possibility of interference suppression independently from the source. Moreover it increases the inter-cell interference variation due to the introduced liberty of choosing a transmission rank in an adaptive and varying manner.

The goal of this section is to inspect whether F-TDD still offers noticeable gains in such conditions. In order to verify this, we run system-level simulations, at low and high load scenarios with the proposed TB-RA algorithm, for fixed S-TDD and F-TDD configurations. The results for low and high load are shown in Figures 5.10 and 5.11 respectively.

In low load conditions, the gains of F-TDD over fixed S-TDD are still present even though the TB-RA scheme instructs the nodes to use higher transmission ranks whenever its possible to do so. In this case there is a clear benefit of applying F-TDD, and there is no noticeable difference between the conser-

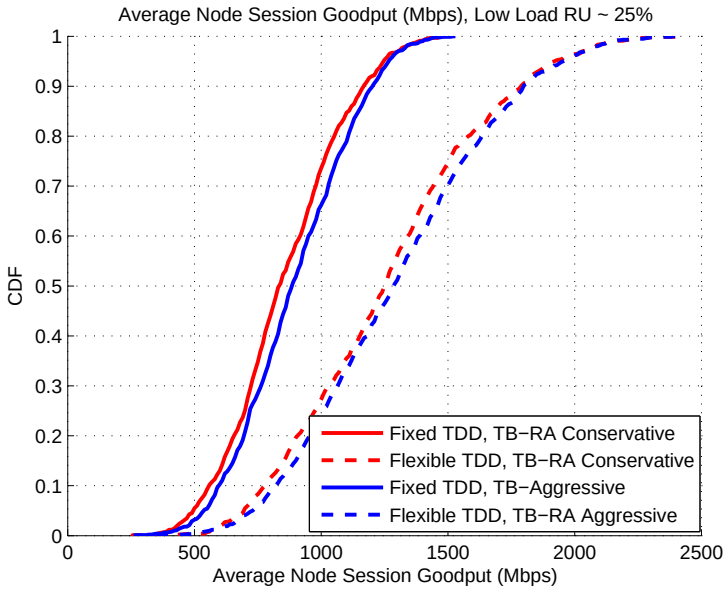


Fig. 5.10: Fixed TDD vs Flexible TDD Performance over TB-RA Scheme - Low Load

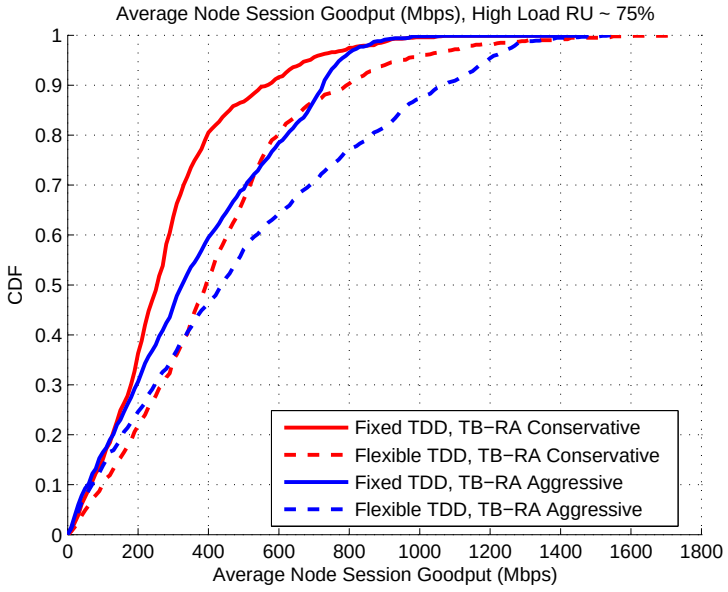


Fig. 5.11: Fixed TDD vs Flexible TDD Performance over TB-RA Scheme - High Load

vative and aggressive parametrizations of the schemes.

In the high load case, there is a difference in performance between the conservative and the aggressive approach. For the F-TDD case, there is a clear benefit in being conservative when choosing the W_k parametrization, especially if outage performance needs to be improved. For the aggressive TB-RA parametrization there is a similar behaviour to the previous results for both the fixed and flexible TDD schemes. With this parametrization, higher transmission ranks are favoured, allowing the system to reach higher peak throughputs. The important thing to note here is that when the interference conditions become severe, the TB-RA scheme has a tendency to use lower ranks for both the fixed S-TDD and F-TDD schemes, reclaiming the benefit of suppressing interference independently from the source, and hence F-TDD still retains a gain over fixed S-TDD, even at 75% resource utilization.

Chapter 6

Performance Evaluation

The goal of this chapter is to provide a complete holistic overview with regards to the performance of Flexible TDD (F-TDD) in the envisioned 5G small cell concept. In order to do so we test the performance of F-TDD over different traffic loads and more realistic traffic asymmetries. The performance of F-TDD is also evaluated in the context of different transport protocols, mainly UDP and TCP, allowing us to capture the interaction between F-TDD and the diverse requirements of these higher layer protocols.

The fully flexible TDD scheme is also benchmarked against semi-static schemes with limited flexibility, similar to what is available in LTE-EIMTA [1]. In doing so, we also attempt to assess the combined merits of these semi-static schemes and the proposed 5G framework by assuming different levels of knowledge on the interference co-variance matrix at the receiver when suppressing interference. All these results are extracted from a detailed implemented system-level simulation model representing the envisioned 5G small cell framework described earlier on. This simulation framework is explained in further detail in Appendix A.

The chapter starts by specifying the local area scenario for which the results are presented, and the corresponding KPI's utilized. After defining these general aspects, we will then proceed to show and discuss the aforementioned results, and final concluding remarks on these results will be supplied at the end of this chapter.

[1] see Section 2.4 for further details

6.1 Indoor Small Cell Scenario

The scenario in which the performance evaluation is conducted is shown in Figure 6.1 and consists of 20 rooms aligned in a 10x2 grid fashion. Each room represents a small cell, with dimensions of 10m x 10m, and contains an access point (AP) and user equipment (UE) deployed in a random fashion. The deployment intensity of these small cells is controlled by the Deployment Ratio parameter and is assumed to be 1.0. Please note that a deployment ratio of 1.0 represents the case, where each and every room contains a small cell deployment.

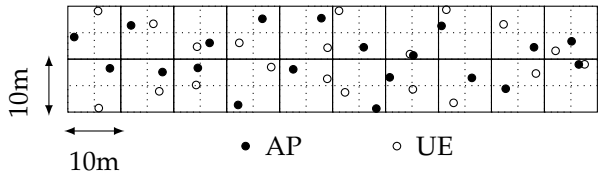


Fig. 6.1: Indoor Scenario 10x2 Grid

A closed-subscriber group AP-UE association model is assumed, such that each UE can only connect to the AP in its room. The rooms are separated by walls, and although a wall loss is considered, due to the random deployment of the nodes, a neighbouring AP may have a stronger signal than a UE's associated AP, creating unfavourable and challenging interference conditions at times.

6.2 Key Performance Indicators

The main key performance indicators that will be considered in the scope of this chapter are the average values of the,

- Application Layer Session Throughput
- Average Node Application Layer Session Throughput

The application layer session throughput results aggregate all the sessions generated by all the nodes in multiple simulation drops. The session throughput represents the experienced throughput for each session generated and is calculated by dividing the session size in bits, and the time taken to transmit the session. A session is made of multiple fragmented packets, and the time

taken to transmit a session represents the difference between the last packet fragment received, t_{end}, and the first packet fragment created, t_{start}. Please note that this time is composed of the waiting time in the buffer, the transmission time over the air, and the minimal required decoding time which is assumed to be instantaneous in our evaluation. For a session S, consisting of N packet fragments with size k bits, the application layer session throughput can be calculated as shown in Eq. 6.1.

$$S = \frac{\sum_{i=1}^{N} k_i}{t_{end} - t_{start}} \tag{6.1}$$

The average node application layer session throughput is calculated in a similar manner, but the statistics are separated on a per node basis to represent the average session throughput experienced by each individual node during a simulation. If we assume that in one simulation drop, a particular node is intended to receive N sessions, each with session throughput S_i, then the average node application layer session throughput S_{node} is calculated as,

$$S_{node} = \frac{\sum_{i=1}^{N} S_i}{N} \tag{6.2}$$

For a better understanding of the behaviour of the individual TDD schemes, we also show results related to the average served versus offered load. The served load is calculated by considering the number of bits received during the entire simulation time, and should ideally be equal to the offered load, indicating that the system can cope well with the offered traffic.

6.3 Benchmarked Schemes

In order to assess the performance of F-TDD, a number of other TDD allocation schemes are used as a benchmark. The F-TDD scheduling scheme we will be showing refers to the LF-based link scheduling algorithm discussed in Chapter 4.

6.3.1 Fixed S-TDD

One baseline scheme that will be used for benchmarking is the fixed S-TDD scheme. The fixed S-TDD scheme is a static TDD configuration that represents the long term average traffic share between UL and DL. For very high

load or full buffer cases, this allocation represents an optimal scheme since the allocated resources match the traffic profile of both links. In less intense finite buffer configurations (low to medium load), this scheme starts suffering from its rigid configuration since it cannot adapt dynamically to match the instantaneous traffic demands.

6.3.2 Upscaled LTE eIMTA

In addition to the fixed S-TDD case, we also benchmark a semi-static TDD case. We consider the LTE eIMTA TDD configuration patterns and apply them directly on our envisioned 5G frame structure, allowing the reconfiguration of the TDD pattern every 5 ms. Every 5 ms, the DL and UL buffer sizes are assessed and the LTE eIMTA TDD configuration which closely represents the instantaneous traffic share is chosen. This benchmark does not fully represent the LTE eIMTA TDD system since it uses the envisioned 5G frame structure. Therefore we refer to it as an upscaled version of LTE eIMTA. Please note that any special subframes in the LTE-TDD configurations are considered as DL timeslot resources. For reference, the LTE eIMTA TDD configurations are included in Section 2.4.

One further demerit of LTE eIMTA relates to the fact that LTE eIMTA cannot estimate the interference co-variance matrix when same-entity interference is present, since LTE uses different frame formats in UL (SC-FDMA) and in DL (OFDMA). We therefore introduce another benchmark where we assume that the system has reduced interference co-variance matrix knowledge. This is achieved by separating the interference sources, same-entity (AP-AP, UE-UE) interference and different entity (AP-UE) interference. It is then assumed that the interference co-variance matrix can accurately acquire the different entity interference, and consider the rest of the interference as noise. This particular scheme will be referred to as *LTE eIMTA Non Ideal* in the results section. Additionally we will also benchmark the case where full interference co-variance matrix knowledge is assumed and refer to it as *LTE eIMTA Ideal*.

The choice for benchmarking both schemes is to allow us to separate the gains arising from the limited flexibility, and LTE eIMTA's interference co-variance matrix estimation demerits.

Table 6.1: Relationship between absolute offered load and Resource Utilization

Offered Absolute Cell Load (Mbps)	%RU UDP	%RU TCP
100	6%	13%
200	16%	37%
300	30%	62%
400	48%	77%
500	68%	84%

6.4 Results and Discussion

In order to evaluate the fully flexible TDD scheme at different loads, the absolute offered load is varied by reducing the average session inter-arrival time (increasing the average session arrival rate λ). This allows us to inspect the regions where flexible TDD gives substantial gain over the other schemes. The session size is assumed to be exponentially distributed with an average size of 2 MBytes. Table 6.1 shows the corresponding average cell Resource Utilization (RU) with different simulated absolute loads, when using a reference fixed S-TDD case, and symmetric traffic, for both UDP and TCP traffic. Please note that the RU is higher for the TCP case due to the congestion window mechanism and TCP ACK feedback requirements of the protocol. Also note, that as the absolute load increases, the relative difference between the UDP's and TCP's RU decreases, as the TCP protocol has more opportunities to piggyback TCP ACK information with data.

The schemes are also compared in the cases where long term average traffic symmetry is present, and also when long term average traffic asymmetry is present. In the symmetric case (1:1) the offered UL and DL traffic is the same, while in the asymmetric case (6:1) the DL offered traffic load is six times larger than the UL. In both cases the offered traffic per cell is equal, and the only thing which varies is the corresponding DL and UL traffic share.

Both UDP and TCP traffic are used to test the schemes, and RLC Acknowledged Mode is used to allow retransmissions for potentially HARQ unrecoverable packets. The rank adaptation scheme described in Section 5.2 is assumed in all cases, and uses a weight parametrization W_k vector of [0.0, 0.5, 0.66, 0.75]. Please note that a frequency reuse factor of 1 is assumed throughout the evaluations, since we rely on the availability of MMSE-IRC receivers to combat inter-cell interference.

Each configuration is simulated for 10 seconds, with 10 simulation drops.

Each drop was randomly generated, and saved in a file, such that the same 10 random identical drops are used for each scheme for a more accurate comparison between the different schemes. A deployment ratio of 1.0 is assumed throughout. A detailed summary of all the relevant simulation parameters is presented in Table 6.2.

6.4.1 UDP Traffic, 1:1 DL/UL Traffic Share

The TDD schemes are first evaluated with an equal traffic share in UL and in DL over the range of offered traffic loads described in Table 6.1. In this first evaluation, UDP traffic is considered.

The average results of the average node session throughput in UL and in DL are shown in Figures 6.2 and 6.3 respectively.

A simple inspection at the average node session throughput results in Figure 6.2, immediately show that the flexible TDD scheme offers superior throughput performance over the fixed TDD scheme, claiming significant absolute gains at low load, and reduced absolute gains at high load. The relative gains between these 2 schemes is however still considerable, even at high load.

The LTE eIMTA Ideal scheme's throughput performance lies between the fully flexible and fixed TDD scheme, highlighting its partial yet limited flexibility in addressing instantaneous traffic requirements.

The LTE eIMTA Non-Ideal scheme's performance is better than the fixed TDD scheme at very low load, due to the low interference scenario and good opportunities to exploit any empty TDD slots and accommodate the instantaneous requirements of the traffic. However, the performance of this scheme quickly degrades when the offered load per cell increases and the interference conditions become more challenging, offering the worst performance out of all the TDD schemes. This is mainly attributed to the scheme's inability to identify and suppress any interference which is classified as same-entity interference (AP-AP, UE-UE). In fact, when inspecting the served versus offered load for each TDD scheme, shown in Figure 6.4, one can notice that this is the only scheme which is not able to serve more than 300 Mbps of cell load.

Similar trends are observed for the UL average node session throughput results. However, in these results, the performance of the LTE eIMTA Ideal scheme degrades significantly. This behaviour is simply attributed to the

[2]Refers to an average value, generated according to an exponential distribution
[3]Ideal & Non-Ideal

Table 6.2: Performance Evaluation Parameters

Traffic Parameters	
Traffic Model	3GPP FTP Traffic Model [30]
File Size	2 MB[2]
Cell Offered Load	100, 200, 300, 400, 500 Mbps
DL to UL Offered Load Ratio	1:1, 6:1
Transport Layer	
Transport Layers	UDP, TCP
TCP Parameters	
Transport Protocol	TCP New Reno [31]
Maximum Segement Size (MSS)	9 kBytes
Minimum Retransmission Timeout	0.1 s
Initial Slow Start Threshold	1 MBytes
Initial Congestion Window	1 MSS
Restart Window	2 MSS
RLC Parameters	
RLC Mode	Acknowledged Mode
$RLC_{ReorderingTimer}$	30 ms
$RLC_{RetransmissionTimeout}$	10 ms
$RLC_{ACKTimer}$	4 ms
$RLC_{ACKCounter}$	8
RRM & PHY Parameters	
Frequency Reuse	1
Rank Adaptation	TB-RA, $W = [0.0, 0.5, 0.66, 0.75]$
Receiver Implementation	MMSE-IRC
AP & UE Tx Power	10 dBm
Channel Parameters	
Propagation Model	Winner II A1 Model [77]
System Bandwidth	200 MHz, f_c = 3.5 GHz
Deployment Parameters	
Scenario	Indoor Small Cell (Fig.6.1)
Room Size	10x10 m
Grid	10x2 (20 Cells)
Deployment Ratio	1.0
UE's per AP	1
Association Mode	Closed Subscriber Group
Simulation Parameters	
Drops	10 (Identically Random)
Simulation Length	10 s (40000 TDD Frames)
TDD Schemes	
Schemes	Fixed, Flexible, LTE eIMTA[3]
Flexible TDD Parameters	
Algorithm	LF-Based, W = 16, Alpha = 1.0

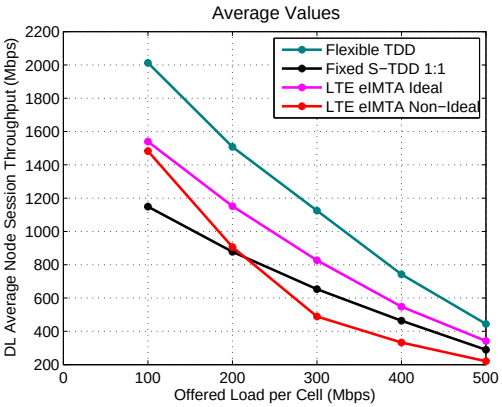


Fig. 6.2: UDP Traffic, DL/UL Traffic Share Configuration (1:1), DL Average Node Session Throughput

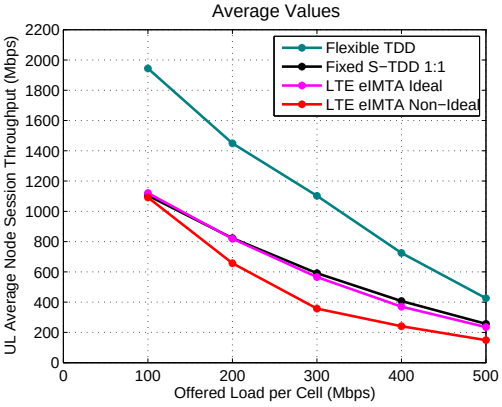


Fig. 6.3: UDP Traffic, DL/UL Traffic Share Configuration (1:1), UL Average Node Session Throughput

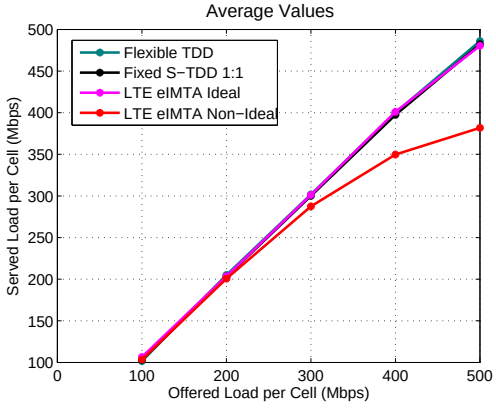


Fig. 6.4: UDP Traffic, DL/UL Traffic Share Configuration (1:1), Served vs. Offered Cell Load

limited availability of LTE eIMTA TDD configurations, which at best can either allocate 90% of the resources to the DL direction, or 60% of the resources to the UL direction. This means that unlike the flexible TDD scheme, if instantaneously only UL traffic is available, 40% of the resources cannot be exploited.

Summarizing, when taking into account the overall session throughput performance gains (in both UL and DL) of flexible TDD over fixed TDD, at low load, a relative gain of 1.8x is achieved. This relative gain diminishes to a factor of 1.6x at high load. This is mainly attributed to the mildly increased inter-cell interference variation re-introduced by the rank adaptation algorithm, and the increased difficulty to opportunistically exploit empty slots due to the high probability of concurrent UL and DL data, especially for those nodes who are in unfavourable conditions and have a very high cell resource utilization.

6.4.2 UDP Traffic, 6:1 DL/UL Traffic Share

In the next evaluation, we consider an asymmetric traffic scenario where the DL offered traffic is six times larger than the UL offered traffic. In essence, this means that for a cell offered load of 100 Mbps, $\frac{6}{7}$ of the traffic is in DL while $\frac{1}{7}$ is in UL, such that the DL offered traffic load is 86 Mbps, while the UL offered traffic load is 14 Mbps. This case represents a more realistic traffic scenario, as predictions state that most of the traffic will be DL dominated in the future [78].

The average results of the average node session throughput in DL and in UL for this traffic configuration, are shown in Figures 6.5 and 6.6 respectively.

The results for this particular traffic scenario show some interesting differences when compared to the previous results. As discussed earlier on, in this scenario, the fixed TDD scheme is configured with a DL to UL allocation of 6:1 to represent the long term average traffic share. When comparing the flexible and fixed TDD schemes it becomes immediately noticeable that there is a small gain in DL as shown in Figure 6.5 and a large gain in UL as shown in Figure 6.6. Again, the absolute gains are pronounced at low load, and diminish at high load.

In DL, the flexible, fixed and LTE eIMTA Ideal schemes offer similar performance, with the LTE eIMTA non-Ideal scheme offering the worst performance, again unable to serve more than 300 Mbps load as shown in Figure 6.7. In UL, the story is reversed with the LTE eIMTA non Ideal scheme offering better session throughput performance than the fixed TDD scheme. This

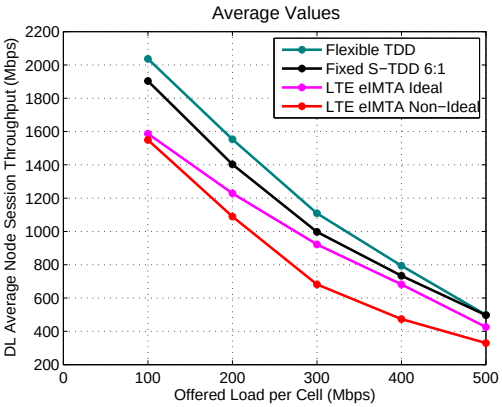


Fig. 6.5: UDP Traffic, DL/UL Traffic Share Configuration (6:1), DL Average Node Session Throughput

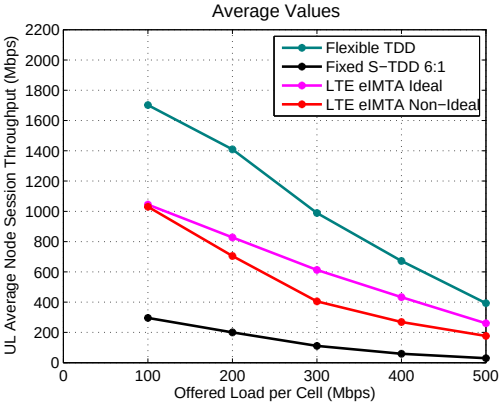


Fig. 6.6: UDP Traffic, DL/UL Traffic Share Configuration (6:1), UL Average Node Session Throughput

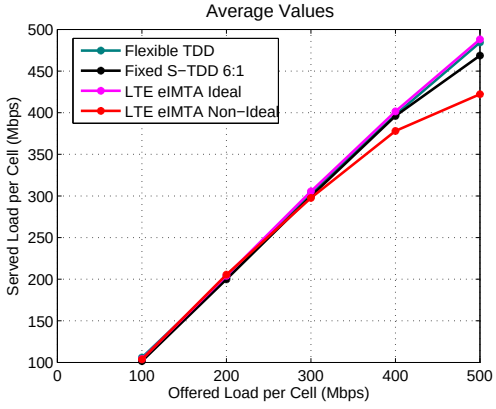


Fig. 6.7: UDP Traffic, DL/UL Traffic Share Configuration (6:1), Served vs. Offered Cell Load

occurs because in UL, the fixed TDD scheme has limited slots to service the UL traffic, and also needs to reserve some TDD slot resources to accommodate RLC ACK's.

In terms of overall session throughput considering both UL and DL traffic, the flexible TDD scheme offers a relative average gain of 1.2x at low load. As the load increases, the relative gain diminishes to 1.1x. In general, the flexible TDD scheme offers better performance in both DL and UL, at low load. At very high load, it offers the same DL performance as in the fixed TDD case, but significantly outperforms the fixed TDD's performance in UL.

6.4.3 TCP Traffic, 1:1 DL/UL Traffic Share

In order to assess the benefits of flexible TDD with more realistic traffic, the previous evaluations are repeated with TCP traffic. Figures 6.8 and 6.9 shows the average node session throughput performance under TCP traffic in DL and in UL respectively.

The trends are similar to what has been obtained under UDP traffic, but in this case, it can be seen that the flexible TDD scheme consistently offers significantly better performance, even at higher loads. It is also to be noted that for TCP traffic, the experienced average node session throughput performance is lower when compared to the UDP traffic case. This is due to the added need of TCP ACK's and the limits imposed by the TCP congestion window mechanism, dictating the amount of data that can be transferred [79]. An inspection at Figure 6.10, showing the served versus the offered load, reinforces this statement, as none of the schemes is able to serve much more than 400 Mbps now.

The work presented in [69] allowed us to make some interesting observations with regards to interaction of different TDD schemes and the requirements of traffic shaped by the TCP protocol. It was seen that in the presence of unidirectional flows (DL or UL traffic only), TCP traffic has different TDD pattern requirements based on the application layer session size.

If the session size is small, a fast TCP ACK reception is typically very beneficial, since it enables a rapid growth of the TCP congestion window, allowing more data to be pushed through. This means that for small session sizes, a fixed 1:1 (DL:UL) TDD scheme is very appropriate. On the other hand, when large session sizes are expected, a more asymmetric TDD configuration would be preferred, since the TCP congestion window is allowed to grow significantly, well before the termination of the transfer. In this case the asymmetry should be exploited, and the transfer of a TCP ACK should

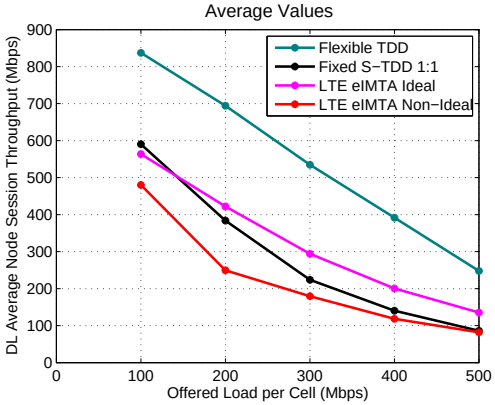


Fig. 6.8: TCP Traffic, DL/UL Traffic Share Configuration (1:1), DL Average Node Session Throughput

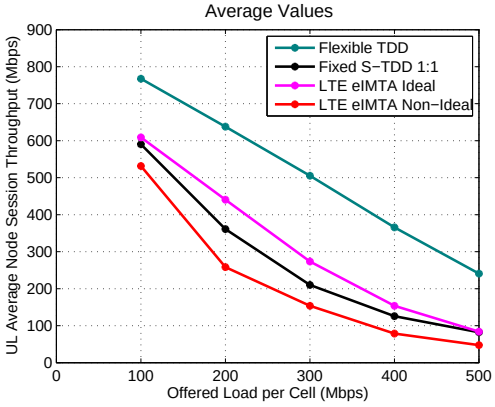


Fig. 6.9: TCP Traffic, DL/UL Traffic Share Configuration (1:1), UL Average Node Session Throughput

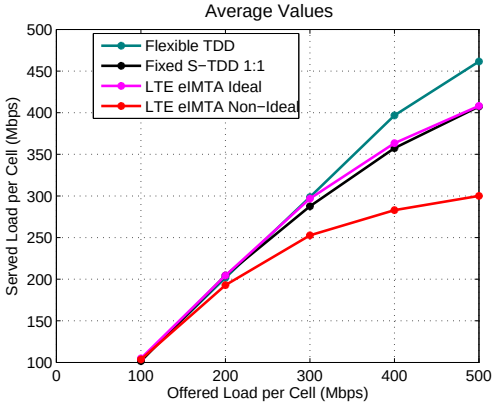


Fig. 6.10: TCP Traffic, DL/UL Traffic Share Configuration (1:1), Served vs. Offered Cell Load

not be over prioritized at the cost of data which can be pushed through the network. Moreover, at this stage of the TCP transfer, multiple TCP ACK's can be aggregated, and a larger amount of data can be acknowledged in a single TDD slot. The study conducted in [69] and attached in Paper B has shown us that an adequately designed flexible TDD scheme, can adapt automatically to any of these conditions.

Since in our evaluations the session size is exponentially distributed with a mean value of 2 MBytes, medium sized sessions are expected. Some samples with small and large session sizes will however also be present. There is therefore a benefit of using flexible TDD over fixed TDD when considering the presence of different session sizes. This is also confirmed in Figures 6.8 and 6.9, with the DL and UL results exhibiting very similar trends, due to the symmetric nature of the offered traffic. The overall relative gains of flexible TDD over fixed TDD are also higher for TCP traffic, and remain consistently high even at more intense traffic loads. From Figure 6.10, one can also notice that flexible TDD is the only scheme that can serve a bit more than 400 Mbps load. In fact, the other TDD schemes already start being congested at an offered load of 300 Mbps, with the LTE eIMTA Non-Ideal scheme only being able to serve up to 200 Mbps load.

It is also to be noted that in these results the LTE eIMTA Ideal TDD scheme performs well in both DL and UL. In the UDP symmetric traffic case, we stated that the LTE TDD eIMTA scheme performs poorly in UL, because the pool of TDD configurations available, do not allow for the allocation of more than 60% of TDD slot resources in UL. However, in the TCP case, the previously unexploited DL slots can now be used for the transmission of a TCP ACK. This behaviour leads to a positive impact in terms of session throughput performance.

6.4.4 TCP Traffic, 6:1 DL/UL Traffic Share

In this subsection, the TCP performance is evaluated, assuming asymmetric traffic loads, with the DL traffic being six times more intensive than the UL traffic, for a given cell offered traffic load. This case represents the most realistic case, since it assumes an expected traffic asymmetry between UL and DL, and also assumes the usage of the most dominant transport layer protocol, TCP [80]. The DL and UL average node session throughput results are respectively shown in Figures 6.11 and 6.12.

In these results, it is clear that the flexible TDD scheme outperforms all the other schemes in both UL and DL throughput performance, even for high

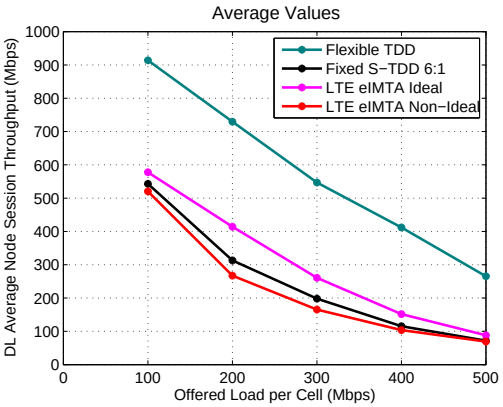


Fig. 6.11: TCP Traffic, DL/UL Traffic Share Configuration (6:1), DL Average Node Session Throughput

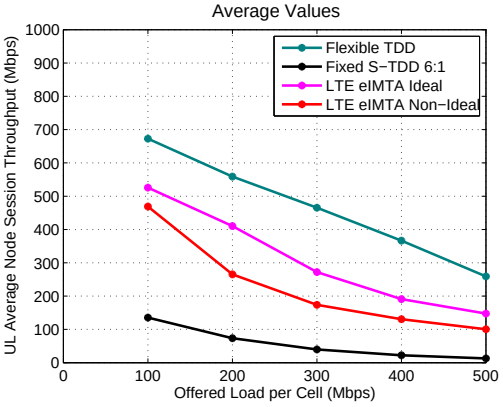


Fig. 6.12: TCP Traffic, DL/UL Traffic Share Configuration (6:1), UL Average Node Session Throughput

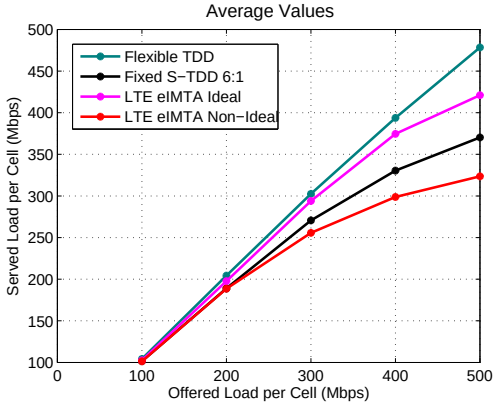


Fig. 6.13: TCP Traffic, DL/UL Traffic Share Configuration (6:1), Served vs. Offered Cell Load

loads. There are multiple factors contributing to the superior performance of the flexible TDD scheme. Apart from being able to accommodate instantaneous traffic demands, the flexible TDD scheme can also accommodate any specific needs arising from higher layer protocols requiring acknowledgements (both at the TCP and the RLC layer).

On the other hand the fixed TDD scheme, configured to a TDD configuration of 6:1 (UL:DL) to match the long term traffic asymmetry, exhibits significantly inferior performance in DL because it cannot account for such higher layer protocol needs, and therefore cannot quickly guarantee the resources for any higher layer TCP acknowledgements. As a result of this, the fixed TDD configuration, has a very slow congestion window growth, limiting its actual performance even though ample time-slot resources for the direction having data are available. This is also witnessed from Figure 6.13, where the fixed TDD case cannot effectively serve more than 200 Mbps in the cell.

The UL performance is also severely compromised in the fixed TDD scheme, since the only UL TDD slot available (over the 6:1 TDD configuration), needs to accommodate not only the data, but also any ACKs coming from both RLC [4] and TCP layers. In principle this means that in realistic traffic scenarios, if a fixed TDD scheme had to be chosen, and the traffic share could be predicted accurately, extra consideration should be taken when selecting the TDD pattern, such that the lightly loaded link is not starved.

When compared to the UDP case, we also observe that the flexible TDD scheme offers increased relative gains throughout all the offered loads, indicating the need of allocating time slot resources not just for the amount of traffic but also for feedback traffic (TCP ACKs in this particular case), since this has a direct impact on the amount of traffic that can be delivered. The flexible TDD scheme can also serve a higher load when compared to the other schemes. Moreover for TCP asymmetric traffic, the limited flexibility available in the LTE eIMTA Ideal benchmark, proves to be beneficial, and as opposed to the UDP asymmetric traffic case, offers better performance than the fixed TDD case in both DL and UL.

6.5 Final Remarks

In this chapter we have presented our simulation framework, and highlighted the main simulation assumptions and KPI's. Following that, the performance

[4]The simulated model does not aggregate RLC ACKs with RLC data, potentially aggravating the situation, see Appendix A for RLC implementation details

evaluation was carried out for a number of TDD schemes over a range of offered traffic loads and configurations, considering both UDP and TCP traffic.

Throughout these evaluations we have confirmed flexible TDD's superior performance over all the benchmarked schemes.

For symmetric UDP traffic, it was observed that flexible TDD offers significant benefits in terms of absolute session throughput. This is clearly noticeable at low loads. While at high loads the absolute gains are less evident, flexible TDD still offers considerable relative gains of around 1.6x over fixed TDD.

In the asymmetric UDP traffic case, the flexible TDD scheme gives a very small gain in the highly loaded direction while offering large gains in the lightly loaded direction. Considering the overall session throughput performance in both links, relative gains in the order of 1.2x are achieved at low load, with the gains diminishing to 1.1x at high load, for this asymmetric case.

In the TCP cases, the flexible TDD scheme offers even better performance since it also adapts well to match the specific ACK feedback requirements of TCP, allowing faster TCP congestion window growth where needed, and appropriate TDD slot resource assignment, once the TCP congestion window size has grown considerably. In fact, it was seen that apart from delivering significant absolute and relative session throughput gains, the flexible TDD scheme can support the highest offered load out of all the investigated TDD schemes.

Chapter 7

Conclusion

This project has mainly focused on the suitability of fully flexible TDD in the context of future 5G small cells. The performance of flexible TDD, as envisioned within the 5G small cell concept, was studied using a detailed system level simulator. Within the simulator, several building blocks were explicitly modelled and implemented, thus allowing us to capture the interacting behaviour of flexible TDD with various elements of the system. Some of these elements include the usage of MMSE-IRC receivers, link and rank adaptation procedures, recovery mechanisms at the HARQ and RLC layers, and transport protocols such as TCP. The conditions at which flexible TDD offers substantial benefits were also evaluated via extensive simulations, always taking into consideration the potential demerits related to increased inter-cell interference variation brought along by this same feature.

In this brief chapter we will outline our main conclusions, provide some general recommendations and finally indicate possible directions to relevant future work.

7.1 Main Conclusions

The main conclusion of this study confirms that within the scope of the envisioned 5G small cell system, fully flexible TDD is a viable solution for indoor small cells.

For this conclusion to be valid one needs to consider the efficacy of the various building blocks available in the envisioned 5G concept.

The role of the frame structure allowing fully flexible TDD is instrumental in the affirmation stated above. In fact, flexible TDD induces added inter-cell interference variation causing potential errors in channel estimation. However, the presence of MMSE-IRC interference suppression receivers along with the designed frame structure enabling just in-time interference suppression, allows a node to perceive relatively stable channel conditions, since an interferer can be suppressed independently from the source. These building blocks are therefore essential in exploiting the benefits of flexible TDD.

Under these conditions, the interaction of rank adaptation with flexible TDD becomes of relative importance. In a practical MIMO system, if the interference conditions are low, multiple spatial streams can be used for transmission rather than for interference suppression. Within this study a distributed taxation based rank adaptation algorithm has therefore been proposed. By being victim-aware, the proposed rank adaptation scheme has been shown to retain both good outage and average session throughput performance, superior to fixed rank transmission and selfish schemes. Moreover, it has been shown that with such an adequate rank adaptation scheme, the flexible TDD scheme still outperforms the fixed S-TDD scheme.

The traffic regions showing the gains of flexible TDD over other semi-static and fixed TDD schemes have also been identified. Intuitively, at low load, flexible TDD offers significant gains over fixed S-TDD. As the load increases, the absolute session throughput gains of flexible TDD decrease, but the relative gains over fixed S-TDD are still considerably attractive.

The benchmarked semi-static TDD scheme, considers a reconfigurable TDD system which periodically assesses the buffer status in UL and DL, and chooses one of the TDD configurations defined in LTE eIMTA, to closely match the current traffic share between the two link directions. As expected the achievable session throughput performance of this semi-static TDD scheme, lies between that of flexible TDD and fixed S-TDD, due to the limited flexibility available. Significant limitations in the achievable UL session throughput performance have also been observed for this scheme, again attributed to the limited number of TDD configurations defined in LTE eIMTA.

When considering asymmetric traffic conditions, the flexible TDD scheme offers marginal benefits in the heavily loaded traffic direction and huge benefits in the lightly loaded conditions at low load, when compared to a fixed S-TDD scheme which can predict the long-term average traffic asymmetry.

The benefits of flexible TDD were also assessed with higher layer protocols such as TCP, and it was shown that the flexible TDD scheme accommodates appropriately the ACK requirements of such protocols, allowing a

faster growth of the TCP congestion window, hence offering a superior session throughput performance over the fixed S-TDD scheme. Moreover, it was shown that flexible TDD can serve a higher load than the other investigated TDD schemes. It was therefore concluded that there are additional benefits of flexible TDD when considered in the context of more realistic TCP traffic.

All these observations confirm the suitability of flexible TDD for the envisioned 5G small cell concept.

For unreliable UDP traffic, it is therefore recommended to enable the use of flexible TDD, given the presence of the 5G small cell frame structure and the availability of MMSE-IRC interference suppression receivers. A similar recommendation can be made for reliable TCP traffic, due to the benefits emerging from the interaction of the higher layer protocols and flexible TDD. Given these conclusions, fully flexible TDD should be strongly considered for inclusion in a future 5G small cell system.

7.2 Future Work

There are numerous opportunities for future work based on this study.

The next logical step is to address the limitation of a single user per cell and assess the performance of the flexible TDD in the presence of multiple users per cell. While in small cells, it is expected that only few users are concurrently active at the same time instance, the presence of multiple users will introduce some interesting scheduling problems, especially when confronted with the interaction of flexible TDD and the specific needs of different classes of devices. The interaction of packet and user/device scheduling in a flexible TDD framework is therefore an interesting topic for future work.

Another possible line of study relates to the comparison of flexible TDD and full duplex. While from a traffic perspective, the benefits of flexible TDD can mainly be exploited whenever there is data in either UL or DL, the benefits of full duplex can only be exploited whenever there is data in both UL and DL. It is therefore expected that the full duplex gain at low load is quite limited, while at high load there are several possibilities to obtain performance benefits which are unobtainable with flexible TDD. While attractive, the usage of full duplex at high load will induce more severe interference conditions, potentially making it unusable in the region in which it is best exploited. The suitability of such a technology should therefore be benchmarked against flexible TDD by also considering any practical imperfections in the technology and the cost associated with enabling such a feature.

While it was concluded that flexible TDD is a suitable technology for small cells, its suitability in larger cell scenarios should also be assessed. In macro cell scenarios, there is typically a huge unbalance in the transmit powers employed by AP's and UE's. The wider distribution of distances between the desired and interference signals, and the differences in DL and UL transmit power will entail some dynamic range problems when considering the operation of MMSE-IRC receivers, significantly reducing their efficacy. Moreover the flexibility introduced by flexible TDD may introduce harmful AP-AP interference, a significantly serious problem in outdoor cells, where access points are tightly coupled due to the presence of line of sight conditions between the two entities. Moreover, macro cells typically serve a much larger number of UE's and their benefit from fully flexible TDD might be limited, since the traffic asymmetry changes on a much slower basis. These implications necessitate a more careful study with regards to the suitability of flexible TDD in larger cells.

References

[1] Cisco. "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2014-2019". In: *White Paper* (Feb. 2015).

[2] Ericsson. *Ericsson Mobility Report*. http://www.ericsson.com/res/docs/2015/ericsson-mobility-report-feb-2015-interim.pdf. 2015.

[3] G. Arunabha et al. *Fundamentals of LTE, Chapter 1 - Evolution of Cellular Technologies*. 1. ed. Prentice Hall, 2010.

[4] *FP7 European Project 317669 METIS (Mobile and Wireless Communications Enablers for the Twenty-Twenty Information Society)*. https://www.metis2020.com. 2012.

[5] *FP7 European Project 318555 5G NOW (5th Generation Non-Orthogonal Waveforms for Asynchronous Signalling)*. https://www.5gnow.eu. 2012.

[6] J.G. Andrews et al. "What Will 5G Be?" In: *Selected Areas in Communications, IEEE Journal on* 32.6 (June 2014), pp. 1065–1082. ISSN: 0733-8716. DOI: 10.1109/JSAC.2014.2328098.

[7] P. Mogensen et al. "B4G local area: High level requirements and system design". In: *Globecom Workshops (GC Wkshps), 2012 IEEE*. Dec. 2012, pp. 613–617. DOI: 10.1109/GLOCOMW.2012.6477644.

[8] Preben Mogensen et al. "Centimeter-Wave Concept for 5G Ultra-Dense Small Cells". In: *Vehicular Technology Conference (VTC Spring), 2014 IEEE 79th*. May 2014, pp. 1–6. DOI: 10.1109/VTCSpring.2014.7023157.

[9] P. Mogensen et al. "5G small cell optimized radio design". In: *Globecom Workshops (GC Wkshps), 2013 IEEE*. Dec. 2013, pp. 111–116. DOI: 10.1109/GLOCOMW.2013.6824971.

[10] A. Osseiran et al. "Scenarios for 5G mobile and wireless communications: the vision of the METIS project". In: *Communications Magazine, IEEE* 52.5 (May 2014), pp. 26–35. ISSN: 0163-6804. DOI: 10.1109/MCOM.2014.6815890.

[11] Huawei. *5G: A Technology Vision*. www.huawei.com/5gwhitepaper/. 2013.

[12] Erik Dahlman et al. "5G Radio Access". In: *Ericsson Review* (2014).

[13] E. Hossain et al. "Evolution toward 5G multi-tier cellular wireless networks: An interference management perspective". In: *Wireless Communications, IEEE* 21.3 (June 2014), pp. 118–127. ISSN: 1536-1284. DOI: 10.1109/MWC.2014.6845056.

[14] E. Dahlman et al. "5G wireless access: requirements and realization". In: *Communications Magazine, IEEE* 52.12 (Dec. 2014), pp. 42–47. ISSN: 0163-6804. DOI: 10.1109/MCOM.2014.6979985.

[15] M. Matinmikko et al. "Cognitive Radio Trial Environment: First Live Authorized Shared Access-Based Spectrum-Sharing Demonstration". In: *Vehicular Technology Magazine, IEEE* 8.3 (Sept. 2013), pp. 30–37. ISSN: 1556-6072. DOI: 10.1109/MVT.2013.2269033.

[16] Jamshid Khun-Jush et al. "Licensed shared access as complementary approach to meet spectrum demands: Benefits for next generation cellular systems". In: *ETSI WORKSHOP ON RECONFIGURABLE RADIO SYSTEMS* (Dec. 2012).

[17] R. Berry, M.L. Honig, and R. Vohra. "Spectrum markets: motivation, challenges, and implications". In: *Communications Magazine, IEEE* 48.11 (Nov. 2010), pp. 146–155. ISSN: 0163-6804. DOI: 10.1109/MCOM.2010.5621982.

[18] PCAST. *Annual Report on Intellectual Property Enforcement, Realizing the Full Potential of Government-Held Spectrum to Spur Economic Growth*. July 2012.

[19] A. Ghosh et al. "Millimeter-Wave Enhanced Local Area Systems: A High-Data-Rate Approach for Future Wireless Networks". In: *Selected Areas in Communications, IEEE Journal on* 32.6 (June 2014), pp. 1152–1163. ISSN: 0733-8716. DOI: 10.1109/JSAC.2014.2328111.

[20] S.G. Larew et al. "Air interface design and ray tracing study for 5G millimeter wave communications". In: *Globecom Workshops (GC Wkshps), 2013 IEEE*. Dec. 2013, pp. 117–122. DOI: 10.1109/GLOCOMW.2013.6824972.

[21] F.M.L. Tavares et al. "On the Potential of Interference Rejection Combining in B4G Networks". In: *VTC Fall, 2013 IEEE 78th*. Sept. 2013, pp. 1–5. DOI: 10.1109/VTCFall.2013.6692318.

[22] M. Dohler et al. "Is the PHY layer dead?" In: *Communications Magazine, IEEE* 49.4 (Apr. 2011), pp. 159–165. ISSN: 0163-6804. DOI: 10.1109/MCOM.2011.5741160.

[23] V. Chandrasekhar, J.G. Andrews, and Alan Gatherer. "Femtocell networks: a survey". In: *Communications Magazine, IEEE* 46.9 (Sept. 2008), pp. 59–67. ISSN: 0163-6804. DOI: 10.1109/MCOM.2008.4623708.

[24] M.-S. Alouini and A.J. Goldsmith. "Area spectral efficiency of cellular mobile radio systems". In: *Vehicular Technology, IEEE Transactions on* 48.4 (July 1999), pp. 1047–1066. ISSN: 0018-9545. DOI: 10.1109/25.775355.

[25] Monica Paolini. *Mobile Data Move Indoors*. http://www.senzafiliconsulting.com/Blog/tabid/64/articleType/ArticleView/articleId/59/Mobile-data-move-indoors.aspx. 2011.

[26] J.G. Andrews et al. "Femtocells: Past, Present, and Future". In: *Selected Areas in Communications, IEEE Journal on* 30.3 (Apr. 2012), pp. 497–508. ISSN: 0733-8716. DOI: 10.1109/JSAC.2012.120401.

[27] N. Golrezaei et al. "FemtoCaching: Wireless video content delivery through distributed caching helpers". In: *INFOCOM, 2012 Proceedings IEEE*. Mar. 2012, pp. 1107–1115. DOI: 10.1109/INFCOM.2012.6195469.

[28] A. Pingyod and Y. Somchit. "Rank-based content updating method in FemtoCaching". In: *TENCON 2014 - 2014 IEEE Region 10 Conference*. Oct. 2014, pp. 1–6. DOI: 10.1109/TENCON.2014.7022376.

[29] Preben Mogensen et al. "Centimeter-wave concept for 5G ultra-dense small cells". In: *IEEE VTS Vehicular Technology Conference. Proceedings* (2014). ISSN: 1550-2252.

[30] 3GPP TR 36.814 V9.0.0. *3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects (Release 9)*. Mar. 2010.

[31] S. Floyd, T. Henderson, and A. Gurtov. *The NewReno Modification to TCP's Fast Recovery Algorithm*. RFC 3782 (Proposed Standard). Obsoleted by RFC 6582. Internet Engineering Task Force, Apr. 2004. URL: http://www.ietf.org/rfc/rfc3782.txt.

[32] Fernando M. L. Tavares. "Interference-robust Air Interface for 5G Small Cells: Managing inter-cell interference with advanced receivers and rank adaptation". PhD thesis. Aalborg University, 2015.

[33] E. Lähetkangas et al. "On the TDD subframe structure for beyond 4G radio access network". In: *Future Network and Mobile Summit (FutureNetworkSummit), 2013*. July 2013, pp. 1–10.

[34] Ian Poole. *TDD FDD Duplex Schemes*. http://www.radio-electronics.com/info/cellulartelecomms/cellular_concepts/tdd-fdd-time-frequency-division-duplex.php.

[35] G. Berardinelli et al. “Distributed synchronization for beyond 4G indoor femtocells”. In: *Telecommunications (ICT), 2013 20th International Conference on*. May 2013, pp. 1–5. DOI: 10.1109/ICTEL.2013.6632081.

[36] G. Berardinelli et al. “Distributed Initial Synchronization for 5G Small Cells”. In: *Vehicular Technology Conference (VTC Spring), 2014 IEEE 79th*. May 2014, pp. 1–5. DOI: 10.1109/VTCSpring.2014.7022884.

[37] G. Berardinelli et al. “On the Potential of OFDM Enhancements as 5G Waveforms”. In: *Vehicular Technology Conference (VTC Spring), 2014 IEEE 79th*. May 2014, pp. 1–5. DOI: 10.1109/VTCSpring.2014.7023019.

[38] B. Farhang-Boroujeny. “OFDM Versus Filter Bank Multicarrier”. In: *Signal Processing Magazine, IEEE* 28.3 (May 2011), pp. 92–112. ISSN: 1053-5888. DOI: 10.1109/MSP.2011.940267.

[39] IEEE. *802.16-2004 IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems, 01 October 2004*. http://www.wimaxforum.org/index.htm. 2004.

[40] Zukang Shen et al. “Dynamic uplink-downlink configuration and interference management in TD-LTE”. In: *Communications Magazine, IEEE* 50.11 (Nov. 2012), pp. 51–59. ISSN: 0163-6804. DOI: 10.1109/MCOM.2012.6353682.

[41] P. Mogensen et al. “5G small cell optimized radio design”. In: *Globecom Workshops (GC Wkshps), 2013 IEEE*. Dec. 2013, pp. 111–116. DOI: 10.1109/GLOCOMW.2013.6824971.

[42] CATT, Ericsson, and ST-Ericsson. *3GPP TSG RAN WG1 Meeting 51, R1-110450, New study item proposal for Further Enhancements to LTE TDD for DL-UL Interference Management and Traffic Adaptation*. Kansas City, USA, Mar. 2010.

[43] Samsung. *3GPP TSG RAN WG1 Meeting 68, R1-120196, Performance evaluation of dynamic TDD reconfiguration*. Dresden, Germany, Feb. 2012.

[44] CATT. *3GPP TSG RAN WG1 Meeting 68, R1-120118, Evaluation on TDD UL-DL reconfiguration for isolated Pico scenario*. Dresden, Germany, Feb. 2012.

[45] Huawei and HiSilicon. *3GPP TSG RAN WG1 Meeting 68, R1-120059, Evaluation of TDD traffic adaptive DL-UL reconfiguration in isolated cell scenario*. Dresden, Germany, Feb. 2012.

[46] *LTE; Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN)*. Tech. rep. TS 36.300 version 12.5.0 Release 12. 3rd Generation Partnership Project, Apr. 2015.

[47] Yuancao Li et al. “Performance Evaluation of a Resource Allocation Scheme for Mixed Traffic in Dynamic-TDD Network”. In: *Vehicular Technology Conference (VTC Fall), 2014 IEEE 80th*. Sept. 2014, pp. 1–5. DOI: 10.1109/VTCFall.2014.6966121.

[48] Li Chen et al. “Utility-based resource allocation for mixed traffic in wireless networks”. In: *Computer Communications Workshops (INFOCOM WKSHPS), 2011 IEEE Conference on*. Apr. 2011, pp. 91–96. DOI: 10.1109/INFCOMW.2011.5928944.

[49] H. Haas, P.K. Jain, and B. Wegmann. “Capacity improvement through random timeslot opposing (RTO) algorithm in cellular TDD systems with asymmetric channel utilisation”. In: *Personal, Indoor and Mobile Radio Communications, 2003. PIMRC 2003. 14th IEEE Proceedings on*. Vol. 2. Sept. 2003, 1790–1794 vol.2. DOI: 10.1109/PIMRC.2003.1260423.

[50] P. Omiyi, H. Haas, and G. Auer. “Analysis of Intercellular Timeslot Allocation in Self-Organising Wireless Networks”. In: *Personal, Indoor and Mobile Radio Communications, 2006 IEEE 17th International Symposium on*. Sept. 2006, pp. 1–5. DOI: 10.1109/PIMRC.2006.254252.

[51] P. Omiyi, H. Haas, and G. Auer. “Analysis of TDD Cellular Interference Mitigation Using Busy-Bursts”. In: *Wireless Communications, IEEE Transactions on* 6.7 (July 2007), pp. 2721–2731. ISSN: 1536-1276. DOI: 10.1109/TWC.2007.05975.

[52] P. Janis, V. Koivunen, and C.B. Ribeiro. “On the performance of flexible UL-DL switching point in TDD wireless networks”. In: *GLOBECOM Workshops (GC Wkshps), 2011 IEEE*. Dec. 2011, pp. 225–230. DOI: 10.1109/GLOCOMW.2011.6162442.

[53] Pekka Janis, Visa Koivunen, and CassioB Ribeiro. “Interference-Aware Radio Resource Management for Local Area Wireless Networks”. In: *EURASIP Journal on Wireless Communications and Networking* 2011.1 (2011), p. 921623. ISSN: 1687-1499. DOI: 10.1155/2011/921623. URL: http://jwcn.eurasipjournals.com/content/2011/1/921623.

[54] M.S. ElBamby et al. “Dynamic uplink-downlink optimization in TDD-based small cell networks”. In: *Wireless Communications Systems (ISWCS), 2014 11th International Symposium on*. Aug. 2014, pp. 939–944. DOI: 10.1109/ISWCS.2014.6933488.

[55] M. Al-Rawi and R. Jantti. “A dynamic TDD inter-cell interference coordination scheme for Long Term Evolution networks”. In: *Personal Indoor and Mobile Radio Communications (PIMRC), 2011 IEEE 22nd International Symposium on*. Sept. 2011, pp. 1590–1594. DOI: 10.1109/PIMRC.2011.6139772.

[56] Alexis A. Dowhuszko et al. “A decentralized cooperative Uplink/Downlink adaptation scheme for TDD Small Cell Networks”. In: *Personal Indoor and Mobile Radio Communications (PIMRC), 2013 IEEE 24th International Symposium on*. Sept. 2013, pp. 1682–1687. DOI: 10.1109/PIMRC.2013.6666413.

[57] A.M. El Hajj and Z. Dawy. “Dynamic Joint Switching Point Configuration and Resource Allocation in TDD-OFDMA Wireless Networks”. In: *Global Telecommunications Conference (GLOBECOM 2011), 2011 IEEE*. Dec. 2011, pp. 1–6. DOI: 10.1109/GLOCOM.2011.6133553.

[58] Lei Jiang, Ming Lei, and Jun Du. “Cross-Subframe Co-Channel Interference Mitigation Scheme for LTE-Advanced Dynamic TDD System”. In: *Vehicular Technology Conference (VTC Spring), 2013 IEEE 77th*. June 2013, pp. 1–5. DOI: 10.1109/VTCSpring.2013.6692497.

[59] Chongning Na, Xiaolin Hou, and Huiling Jiang. “Interference alignment based dynamic TDD for small cells”. In: *Globecom Workshops (GC Wkshps), 2014*. Dec. 2014, pp. 700–705. DOI: 10.1109/GLOCOMW.2014.7063514.

[60] WunCheol Jeong and Mohsen Kavehrad. “Cochannel interference reduction in dynamic-TDD fixed wireless applications, using time slot allocation algorithms”. In: *Communications, IEEE Transactions on* 50.10 (Oct. 2002), pp. 1627–1636. ISSN: 0090-6778. DOI: 10.1109/TCOMM.2002.803991.

[61] Bo Yu, H. Ishii, and Liuqing Yang. “System Level Performance Evaluation of Dynamic TDD and Interference Coordination in Enhanced Local Area Architecture”. In: *Vehicular Technology Conference (VTC Spring), 2013 IEEE 77th*. June 2013, pp. 1–6. DOI: 10.1109/VTCSpring.2013.6692777.

[62] Bo Yu et al. “Dynamic TDD support in the LTE-B enhanced Local Area architecture”. In: *Globecom Workshops (GC Wkshps), 2012 IEEE*. Dec. 2012, pp. 585–591. DOI: 10.1109/GLOCOMW.2012.6477639.

[63] Young Choi, Illsoo Sohn, and Kwang Bok Lee. “A novel decentralized time slot allocation algorithm in dynamic TDD system”. In: *Consumer Communications and Networking Conference, 2006. CCNC 2006. 3rd IEEE*. Vol. 2. Jan. 2006, pp. 1268–1272. DOI: 10.1109/CCNC.2006.1593242.

[64] H. Ishii, Y. Kishiyama, and H. Takahashi. “A novel architecture for LTE-B :C-plane/U-plane split and Phantom Cell concept”. In: *Globecom Workshops (GC Wkshps), 2012 IEEE*. Dec. 2012, pp. 624–630. DOI: 10.1109/GLOCOMW.2012.6477646.

[65] 3GPP Workshop on Release 12 NTT DOCOMO Inc. and Onwards. *Requirements Candidate Solutions, and Technology Roadmap for LTE Rel. 12 Onward*. http://www.3gpp.org/ftp/workshop/2012-06-11_12_RAN_REL12/Docs/RWS-120010.zip. Ljubljana, Slovenia, June 2012.

[66] Chunyi Wang et al. "HARQ Signalling Design for Dynamic TDD System". In: *Vehicular Technology Conference (VTC Fall), 2014 IEEE 80th*. Sept. 2014, pp. 1–5. DOI: 10.1109/VTCFall.2014.6965995.

[67] Leonard Kleinrock. *Queueing Systems. Volume 1: Theory*. Wiley-Interscience, 1975.

[68] J. Virtamo. *38.3143 Queueing Theory / The M/G/1/ queue*. https://www.netlab.tkk.fi/opetus/s383143/kalvot/E_mg1jono.pdf.

[69] Davide Catania et al. "Flexible UL/DL in Small Cell TDD Systems: A Performance Study with TCP Traffic". In: *2015 IEEE 81st Vehicular Technology Conference: VTC2015-Spring*. 2015.

[70] Alexandra Oborina et al. "MIMO performance evaluation in UTRAN Long Term Evolution downlink". In: *Information Sciences and Systems, 2008. CISS 2008. 42nd Annual Conference on*. Mar. 2008, pp. 1179–1183. DOI: 10.1109/CISS.2008.4558697.

[71] B. Clerckx et al. "A Practical Cooperative Multicell MIMO-OFDMA Network Based on Rank Coordination". In: *Wireless Communications, IEEE Transactions on* 12.4 (Apr. 2013), pp. 1481–1491. ISSN: 1536-1276. DOI: 10.1109/TWC.2013.013013.112100.

[72] Z.K.M. Ho and D. Gesbert. "Balancing Egoism and Altruism on Interference Channel: The MIMO Case". In: *Communications (ICC), 2010 IEEE International Conference on*. May 2010, pp. 1–5. DOI: 10.1109/ICC.2010.5501882.

[73] F.M.L. Tavares et al. "Inter-cell interference management using maximum rank planning in 5G small cell networks". In: *Wireless Communications Systems (ISWCS), 2014 11th International Symposium on*. Aug. 2014, pp. 628–632. DOI: 10.1109/ISWCS.2014.6933430.

[74] N.H. Mahmood et al. "A distributed interference-aware rank adaptation algorithm for local area MIMO systems with MMSE receivers". In: *ISWCS*. Aug. 2014, pp. 697–701. DOI: 10.1109/ISWCS.2014.6933443.

[75] Davide Catania et al. "A Distributed Taxation Based Rank Adaptation Scheme for 5G Small Cells". In: *2015 IEEE 81st Vehicular Technology Conference: VTC2015-Spring*. 2015.

[76] G.W.O. da Costa et al. “A Scalable Spectrum-Sharing Mechanism for Local Area Network Deployment”. In: *Vehicular Technology, IEEE Transactions on* 59.4 (May 2010), pp. 1630–1645. ISSN: 0018-9545. DOI: 10.1109/TVT.2009.2039361.

[77] IST-WINNER, II. *Deliverable 1.1.2 v.1.2, "WINNER II Channel Models"*. Tech. rep. Tech. Rep., 2008, 2008. URL: http://projects.celtic-initiative.org/winner+/deliverables.html.

[78] Qualcomm. *Wireless Broadband Future and Challenges*. 2010.

[79] V. Jacobson. “Congestion Avoidance and Control”. In: *Symposium Proceedings on Communications Architectures and Protocols*. SIGCOMM ’88. Stanford, California, USA: ACM, 1988, pp. 314–329. ISBN: 0-89791-279-9. DOI: 10.1145/52324.52356. URL: http://doi.acm.org/10.1145/52324.52356.

[80] W. John and S. Tafvelin. “Heuristics to Classify Internet Backbone Traffic based on Connection Patterns”. In: *Information Networking, 2008. ICOIN 2008. International Conference on*. Jan. 2008, pp. 1–5.

Part II

Appendix

Appendix A

Simulation Framework

A.1 Simulation Framework

In order to achieve realistic results, several levels of the protocol stack were modelled. The protocol layers modelled are shown in Figure A.1. In the upper part of the protocol stack, the application, transport, and internet protocol layer are modelled. The lower part of the protocol stack considers the radio layers of the stack, and the Radio Link Control (RLC), Medium Access Control (MAC) and Physical (PHY) layers are modelled explicitly to represent the envisioned 5G small cell concept described in Chapter 2.

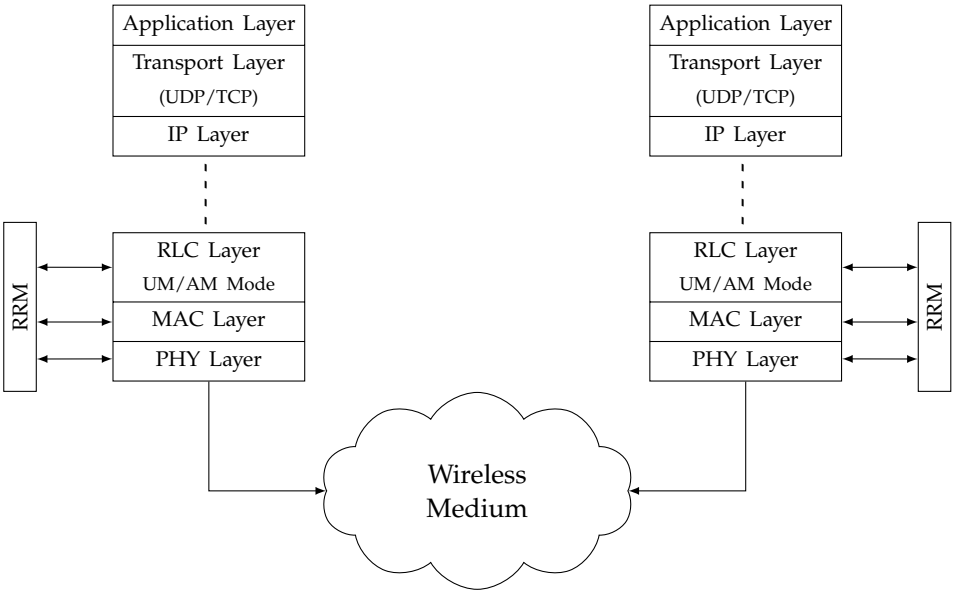


Fig. A.1: Simulation Model & Protocol Stack

A.1.1 Upper Protocol Stack

The application layer consists of a finite buffer traffic model as specified in the FTP Traffic Model 1 in [1]. In this model application layer payload (file) bursts of a certain size S are generated at an inter-arrival time t_{ia}. The size of the file S can be fixed or generated according to an exponential distribution with a specified average value. The absolute offered traffic load is controlled via the parametrization of the S and t_{ia} parameters. The arrival rate λ, a parameter which depends on the average exponentially distributed t_{ia}, specifies the average number of payload arrivals per second. The actual resource

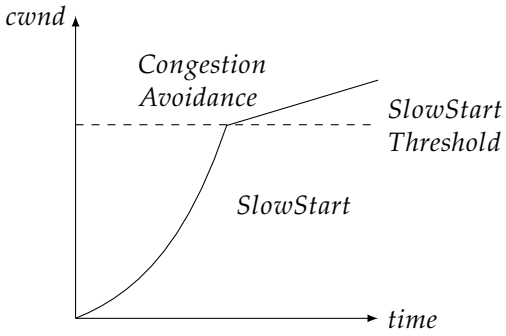


Fig. A.2: TCP Congestion Control Mechanism

utilization is dependent on the performance of a specific configuration, and is typically approximated to represent a certain value in most of the results.

In this specific application layer model and in our assumptions, unlimited payload buffering is allowed. Each arriving payload, referred to also as a session, is tagged with a session number such that individual session statistics can be extracted. It is assumed that a session consists of multiple fragmented packets, each having a size equal to the minimum transport block size, whose value depends on the lowest rank and modulation and coding scheme allowed in the system.

The transport layer can be configured to operate in UDP or TCP mode [2]. In the UDP mode, packet losses are allowed, while the TCP protocol does not allow any packet losses and follows the typical TCP congestion control dynamics, limiting the amount of data that can be sent, based on the timely reception of a corresponding acknowledgement which acknowledges that data. The corresponding protocol overhead are modelled in both cases.

The TCP model does not include any initial 3-way handshaking and connection termination procedures available in the TCP protocol, but models the congestion control mechanism shown in Figure A.2 in detail. The TCP congestion control mechanism dictates how many TCP segments are allowed to be transmitted before a TCP acknowledgement (TCP ACK) is required. It was first proposed by Van Jacobson in [3] and its role proved to be instrumental in avoiding the collapse of the Internet in the late 1980's. The task of TCP congestion control is to limit and avoid transmitting more TCP segments in the case where network congestion can be deduced due to the missing receptions of TCP ACK's. The most relevant aspect in our studies relate to the fact that during the initial TCP transfer phase, a small amount of TCP segments can be transmitted, and a larger amount of TCP segments can only be sent if

Table A.1: TCP Parameters

Parameter	Value
Transport Protocol	TCP New Reno [8]
Maximum Segement Size (MSS)	9 kBytes
Minimum Retransmission Timeout	0.1 s
Initial Slow Start Threshold	1 MBytes
Initial Congestion Window	1 MSS
Restart Window	2 MSS

TCP ACK's are received in a timely manner, allowing the congestion window size shown in Figure A.2 to grow accordingly. This mechanism will of course degrade the achievable application layer session throughput, but is essential in discovering whether the network is begin congested. The main motivation for explicitly modelling these dynamics stem from two main facts,

- TCP is used by a lot of important end-user applications which do not tolerate packet loss
- TCP represents most of the Internet traffic share [4]

The modelled TCP version is TCP New Reno [5] and the parameters shown in Table A.1 are assumed throughout all the corresponding TCP results. It is to be noted that newer versions of the TCP protocol dealing with more advanced congestion control mechanisms catering for issues present in large bandwidth-delay product networks have been released and are used as the default implementation in many operating systems. The two most prominent implementations are Compound TCP [6], and TCP Cubic [7], both of which attempt to increase the congestion window size more quickly in the congestion avoidance stage.

The IP layer routing protocols are not explicitly modelled since the evaluations carried out are mainly focused on the radio access network part of the network. Therefore, with regards to the IP Layer, only the relevant IP packet overhead is considered.

A.1.2 Lower Radio Protocol Stack

In order to obtain a realistic evaluation of the performance benefits of F-TDD in our envisioned 5G concept, a detailed model was implemented for the

lower radio protocol stack. The RLC, MAC and physical layers were implemented. All these entities are responsible for exchanging information with the RRM block, which in turn is responsible for collecting data from these layers, and re-routing it to provide valuable information to the appropriate layer.

The RLC layer has multiple responsibilities. These include retrieving packets from the upper layer, and storing them into appropriate separate buffer queues based on the user and flow. These packets, which once stored in the RLC layer are referred to as RLC Service Data Unit (SDU)'s, are typically encapsulated within an RLC Protocol Data Unit (PDU) when handed over to the lower layer. For the sake of efficiency, in this particular implementation, the RLC layer aggregates multiple PDU's intended for the same user/flow pair, when a transmission opportunity is received from the lower layer. Once this aggregated RLC PDU is formed, the RLC hands over this block to the lower layer, in our case the MAC layer. The RLC layer also communicates with the RRM entity to provide relevant information with regards to the current buffer status and current HOL delay statistics for each queue. This information, as seen from the previous chapters, is used to convey information to the DL/UL link scheduling decision maker.

On the reception side, the RLC layer is also responsible for removing any duplicate data and attempting to deliver in-order packets to the higher layer. Due to the varying nature of the wireless channel, along with the interaction of temporary interfering entities, it is expected that some receptions are received in error, and are therefore required to be retransmitted. Any lower layer retransmissions, can induce out-of-order RLC PDU's at the reception side. The RLC layer introduces the notion of a reordering timer, $RLC_{ReorderingTimer}$, which is triggered once an out-of-order RLC PDU is received. The $RLC_{ReorderingTimer}$ timer will wait for a configurable amount of time, in the attempt of recovering any potentially lost RLC PDU's (either through active RLC transmission, or via Hybrid Automatic Repeat Request (HARQ) retransmissions). If these RLC PDU's are not received within this predefined amount of time, the RLC layer will simply give up on these RLC PDU's and simply extract the corresponding RLC SDU's and pass them to the higher layer.

The RLC layer can be configured to operate in either Unacknowledged Mode (UM) or Acknowledged Mode (AM). In the UM no acknowledgement feedback is required, but in this mode the RLC layer will not attempt to recover any potentially lost packets which can or not be recoverable by the lower HARQ layer retransmission mechanism. In this configuration, no signalling overhead associated with RLC acknowledgements is present.

Table A.2: RLC Parameters

	RLC UM	RLC AM
$RLC_{ReorderingTimer}$	5 ms	30 ms
$RLC_{RetransmissionTimeout}$	x	10 ms
$RLC_{ACKTimer}$	x	4 ms
$RLC_{ACKCounter}$	x	8

The RLC AM introduces some RLC feedback overhead but is aimed to reduce the amount of lost RLC PDU's. This is done by the receiving entity by sending an RLC ACK, acknowledging the number of aggregated RLC PDU's received. This feedback is sent on every $RLC_{ACKCounter}$ aggregated RLC PDU's received, or whenever a $RLC_{ACKTimer}$ expires. Please note that in this simplified implementation, any RLC ACK's are prioritized over data. The RLC ACK's are also not aggregated with data, even if there is the possibility to do so at times.

The RLC retransmission mechanism is also simplified for the sake of these studies, and is conducted as follows. A $RLC_{RetransmissionTimeout}$ is introduced in order to retransmit any potential lost RLC data. If a missing set of RLC PDUs are detected before the expiry of this timer, these RLC PDU's are selectively retransmitted. If the timer expires, a go-back-N approach is assumed, and the RLC layer retransmits the RLC PDU's, starting from the highest in-order RLC PDU received. The number of times this retransmission is allowed is governed by the relationship between the $RLC_{RetransmissionTimeout}$ and $RLC_{ReorderingTimer}$ timers. For instance, if the $RLC_{ReorderingTimer}$ is 3.5 times greater than the $RLC_{RetransmissionTimeout}$, then up to 3 retransmission attempts should be allowed.

Since the detailed design of the RLC layer is outside of the scope of this thesis, this simplified model was adopted. It should however be noted that in this simplified design, there is several room for improvement. The main parameters used for the RLC UM and RLC AM modes are summarized in Table A.2.

The MAC and PHY layers are tightly coupled with each other and serve a myriad of purposes. The appropriate timing related to the frame structure is controlled in these layers, and any formed MAC SDU's are put into coding blocks which are eventually synthesized to corresponding signals in the physical layer.

The actual scheduling and transmitter parameter decisions are taken in the RRM. These decisions are influenced from any remote CQI reports and in-

formation related to the signal quality in previous receptions which are forwarded to the RRM block from the MAC and PHY layers. Based on these decisions, such information is passed to the MAC layer which constructs appropriate scheduling grant and scheduling requests messages and embeds them in the control channel of the frame. The MAC layer is also responsible for creating transmit opportunity messages based on these decisions which are eventually forwarded to the upper layer, which in turn delivers an appropriate RLC PDU which can then be encapsulated in the MAC SDUs.

The receiver and link-to-system models used are discussed in Appendices B and C respectively.

A.2 Propagation Model

The implemented propagation model is based on the Winner II A1 indoor model specified in [9], and models the large scale fading (path loss, and slow fading due to shadowing), and small scale fading which is time varying and frequency selective.

The Line Of Sight (LOS) and Non-Line of Sight (NLOS) path loss models, including a corresponding log normal shadowing standard deviation parameter are calculated as follows,

$$PathLoss(dB) = Alog_{10}(d(m)) + B + Clog_{10}\left(\frac{f_c(GHz)}{5.0}\right) + X \qquad \text{(A.1)}$$

Table A.3 shows the corresponding parameters for the path loss formula shown in equation A.1. The NLOS condition represents the room-to-room scenario.

Please note that a carrier frequency f_c of 3.5 GHz is assumed in these evaluations, and the specified path loss model is only valid from frequencies ranging from 2 to 6 GHz.

The fast fading time-varying and frequency selective effects were also extracted from [9], assuming a 3 km/hr mobility to account for moving objects and people in the indoor scenario. The model generated in [10] is used in these evaluations. This model was generated using a fading sample generator script, which was a deliverable of the Winner II project [9], and was configured to generate channel samples between four transmit and four receive antennas for every tenth frame (2.5 ms). The script was adjusted to represent

Table A.3: Channel Parameters

	LOS	NLOS
A	18.7	20
B	46.8	46.4
C	20	20
X	0	5(n_{walls}-1)
Shadow Fading Std. Deviation (dB)	3	6
Minimum Coupling Loss (dB)	45	45

the frequency domain signal characteristics of the envisioned 5G concept explained in Chapter 2, with a frequency resolution of 900 kHz representing 15 sub-carriers. The model assumes that each device consists of a uniform linear array with four dipole elements, with half wave-length distance between the adjacent elements.

Appendix B

Receiver Model

B.1 Receiver Model

B.1.1 Signal Model

Throughout our investigations we have assumed a 4x4 MIMO system, with 4 transmit and 4 receive antennas. In this section, for the sake of generality, we assume a system where each node has N_T transmit antennas and N_R receive antennas, and the amount of streams utilized for transmission varies between $1 \leq N_s \leq min(N_T, N_R)$.

The received signal **r** is then composed of four components, the desired signal, the inter-stream interference, the inter-cell interference and noise. From the receiver's perspective, the inter-stream interference represents the interference originating from the additional spatial streams transmitted by the intended transmitter and the inter-cell interference represents the set of interfering streams originating from each of the neighbouring interferers.

Assuming that we are interested in the received signal $r_{i,j}$ from transmitter i at stream j, and the number of streams generated by the intended transmitter i is represented by N_s^i, while the number of interfering streams generated by an interfering transmitter K is represented by N_s^K, then for N_I interfering nodes, the received signal $r_{i,j}$ in a particular OFDM sub carrier and symbol can be represented as,

$$r_{i,j} = \overbrace{h_{i,j}s_{i,j}}^{\text{Desired Signal}} + \underbrace{\sum_{\substack{l=1 \\ l \neq j}}^{N_s^i} h_{i,l}s_{i,l}}_{\text{Inter-stream Interference}} + \overbrace{\sum_{\substack{k=1 \\ k \neq i}}^{N_I} \sum_{l=1}^{N_S^K} h_{k,l}s_{k,l}}^{\text{Inter-cell Interference}} + n \tag{B.1}$$

Please note that **s** represents the transmitted signal with zero mean, and unity power, and it assumed that the transmitted signals **s** from the different nodes are uncorrelated. Each element of vector **n** represents the noise with power σ_n^2 at the receive antennas, and the individual elements are assumed to be uncorrelated. The complex vector **h** represents the channel gain from the transmitting to the receiving stream, and incorporates any additional pre-coding, transmit and/or receive antenna gains along with the small and large scale fading and propagation effects. A block fading model is assumed, meaning that the complex channel vector **h** is constant and non-varying within a specified time block.

The expression defined in equation B.1 can be written in succinct matrix form as shown in Equation B.2, where $H \in \mathbb{C}^{N_R \times N_T}$ represents the effective $N_R \times N_T$ channel matrix, and subscripts D and I represent the desired and interference signals respectively.

$$r = H_D s_D + H_I s_I + n \tag{B.2}$$

B.1.2 MMSE estimation and Combining

In order to recover the transmitted signal which has been altered by the channel and the noise, an estimate of the transmitted signal needs to be obtained. We assume a minimum mean square error estimator, which is a simple estimation technique that does not result in noise enhancement as in the zero-forcing estimator.

The generic MMSE estimate is given by [11],

$$W = (H_D H_D^H + R_N)^{-1} H_D \tag{B.3}$$

We assume that the acquisition of H_I is considered to be possible due to the assumed frame structure which includes a DMRS symbol, allowing just in-time interference covariance matrix estimation. In principle, the acquisition of this information is non-ideal due to the some receiver imperfections, but studies carried out in [10] show that the impact of such imperfections is minimal on the performance of this scheme.

For the model of the MMSE-IRC receiver we then consider R_N to be,

$$R_N = H_I H_I^H + \sigma_n^2 I \tag{B.4}$$

while for the model of the MMSE-MRC receiver we consider R_N to be,

$$R_N = diag(H_I H_I^H) + \sigma_n^2 I \tag{B.5}$$

The desired s_D is then estimated as,

$$s_D = W^H r \tag{B.6}$$

Unlike the modelled MMSE-IRC receiver, the modelled MMSE-MRC receiver assumes that the sum of all the interference originating from neighbouring

interfering transmitters, can be modelled as additive noise, assuming no correlation between the signals received by the receive antennas [10].

B.1.3 SINR Estimation

The SINR estimation is extracted from the model used in [10] and for each stream j, the SINR can be given by,

$$SINR_j = \frac{W_j^H H_{Dj} H_{Dj}^H W_j}{W_j^H (\overline{H}_{Dj} \overline{H}_{Dj}^H + H_{Ij} H_{Ij}^H + \sigma^2_{N_{Rx}} I) W_j} \tag{B.7}$$

where X_j represents the j-th column of matrix X, while $\overline{X_j}$ represents the matrix X with the omission of the j-th column.

The SINR at stream j is therefore given as ratio of the combined signal power, over the combined inter-stream interference, inter-cell interference and noise.

Appendix C

Link To System Mapping

C.1 L2S Mapping

C.1.1 Effective SINR Mapping

In order to run the simulations in an adequate time frame, a method to extract a single SINR value from each reception was introduced using an elementary effective SINR mapping.

For a single PRB, the SINR perceived on each sub-carrier is different. For a system consisting of K sub-carriers per PRB, and L PRB's, an effective SINR mapping is introduced. The simplified SINR mapping considered simply log-averages the SINR across all subcarriers and PRB's, as follows.

$$SINR_{eff}(dB) = \frac{\sum_{i=1}^{L}\sum_{n=1}^{K} 10log_{10}SINR_{i,n}}{KL} \tag{C.1}$$

The same approach is extended to compress the SINR values perceived at each stream j where if a single reception is spread over M streams, where $1 \leq M \leq N_{rx}$, the effective SINR is simply calculated in a similar manner.

It is acknowledged that more accurate models such as the EESM model [12–14] are available, but these models require a β scaling value parameter of the SINR based on the MCS used for transmission, and the unavailability of detailed link-level simulations tailored for the envisioned 5G concept led us to consider the simplified model specified above.

References

[1] 3GPP TR 36.814 V9.0.0. *3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects (Release 9)*. Mar. 2010.

[2] J. Postel. *Transmission Control Protocol*. RFC 793 (INTERNET STANDARD). Updated by RFCs 1122, 3168, 6093, 6528. Internet Engineering Task Force, Sept. 1981. URL: http://www.ietf.org/rfc/rfc793.txt.

[3] V. Jacobson. "Congestion Avoidance and Control". In: *Symposium Proceedings on Communications Architectures and Protocols*. SIGCOMM '88. Stanford, California, USA: ACM, 1988, pp. 314–329. ISBN: 0-89791-279-9. DOI: 10.1145/52324.52356. URL: http://doi.acm.org/10.1145/52324.52356.

[4] W. John and S. Tafvelin. "Heuristics to Classify Internet Backbone Traffic based on Connection Patterns". In: *Information Networking, 2008. ICOIN 2008. International Conference on*. Jan. 2008, pp. 1–5.

[5] M. Allman, V. Paxson, and W. Stevens. *TCP Congestion Control*. RFC 2581 (Proposed Standard). Obsoleted by RFC 5681, updated by RFC 3390. Internet Engineering Task Force, Apr. 1999. URL: http://www.ietf.org/rfc/rfc2581.txt.

[6] Murali Sridharan et al. *Compound TCP: A New TCP Congestion Control for High-Speed and Long Distance Networks*. Internet-Draft draft-sridharan-tcpm-ctcp-02. IETF Secretariat, Nov. 2008. URL: https://tools.ietf.org/html/draft-sridharan-tcpm-ctcp-02.

[7] Sangtae Ha, Injong Rhee, and Lisong Xu. "CUBIC: A New TCP-friendly High-speed TCP Variant". In: *SIGOPS Oper. Syst. Rev.* 42.5 (July 2008), pp. 64–74. ISSN: 0163-5980. DOI: 10.1145/1400097.1400105. URL: http://doi.acm.org/10.1145/1400097.1400105.

[8] S. Floyd, T. Henderson, and A. Gurtov. *The NewReno Modification to TCP's Fast Recovery Algorithm*. RFC 3782 (Proposed Standard). Obsoleted by RFC 6582. Internet Engineering Task Force, Apr. 2004. URL: http://www.ietf.org/rfc/rfc3782.txt.

[9] IST-WINNER, II. *Deliverable 1.1.2 v.1.2, "WINNER II Channel Models"*. Tech. rep. Tech. Rep., 2008, 2008. URL: http://projects.celtic-initiative.org/winner+/deliverables.html.

[10] Fernando M. L. Tavares. "Interference-robust Air Interface for 5G Small Cells: Managing inter-cell interference with advanced receivers and rank adaptation". PhD thesis. Aalborg University, 2015.

[11] Jinho Choi. *Optimal Combining and Detection, Statistical Signal Processing for Communications*. Cambridge, 2010.

[12] 3GPP TSG RAN 1 Meeting 37. *OFDM-HSPDA System level simulator Calibration (R1-040500)*. May 2004.

[13] 3GPP TSG RAN 1, TS RAN 4 Ad Hoc. *System-Level Performance Evaluation for OFDM and WCDMA in UTRAN (R1-04-0090)*. Jan. 2004.

[14] 3GPP TSG RAN WG1 36. *New Results on Realistic OFDM Interference, (R1-040189)*. Feb. 2004.

Part III

Papers

Paper A

The Potential of Flexible UL/DL Slot Assignment in 5G Systems

Davide Catania[1], Marta Gatnau Sarret[1], Andrea F. Cattoni[1], Frank Frederiksen[2], Gilberto Berardinelli[1], Preben Mogensen[1,2]

[1]Department of Electronic Systems, Aalborg University, Denmark, [2]Nokia, Denmark

dac@es.aau.dk, mgs@es.aau.dk, afc@es.aau.dk, frank.frederiksen@nsn.com, gb@es.aau.dk, pm@es.aau.dk

The paper has been published in the *Vehicular Technology Conference (VTC Fall), 2014 IEEE 80th*

The layout has been revised.

Abstract

5th Generation (5G) small cells are expected to satisfy the increasing demand for wireless data traffic. In the presence of large scale dense and randomly deployed cells, autonomous and distributed configuration mechanisms are highly desirable. However, small cells typically serve a small number of users, such that sudden traffic imbalances between downlink (DL) and uplink (UL) are expected in the new 5G system. We exploit the flexibility of time-division duplex (TDD) to deal with such imbalances by adapting swiftly to instantaneously varying traffic needs. In this paper we propose a distributed algorithm to deal with these varying traffic requirements. We also exploit the availability of interference rejection capable receivers. Simulation results show that in the presence of the aforementioned features, we can approximately double the session throughput and halve the packet delay in a large number of cases.

A.1 Introduction

The demand for high speed wireless data services is always on the rise. An exponential increase in traffic growth is expected in the coming years [1], paving the way for a future 5G system. One way of handling this traffic growth is via cell densification, and further spectral reuse [2]. We therefore expect the proliferation of a new wireless local area 5G system. While the exact requirements of a future 5G system might still be unclear, some key performance indicators to take into account include service availability, peak throughput, and latency.

Some envisioned characteristics include interference rejection capable receivers, and a redesigned frame structure. A direct impact on minimizing latency is to reduce the frame duration significantly as presented in [3]. Another envisioned feature is the usage of Time Division Duplex (TDD) mode, with complete freedom of assigning each frame as uplink or downlink. As opposed to frequency division duplex (FDD), TDD can offer advantages in terms of cost, possibilities of exploiting unpaired bands and coping with unbalanced uplink/downlink (UL/DL) traffic. Such a feature is particularly useful in local area scenarios, where the amount of active users in a cell is typically small. In this scenario, a traffic burst in one particular direction, UL or DL, can significantly unbalance the UL and DL traffic demands of a cell. This is in contrast with cellular macro cells, where the amount of active uplink users is typically larger and their traffic fluctuations have a minimal impact on the instantaneous UL/DL traffic needs of a cell. Long Term Evolution (LTE) [4] and WiMAX [5] have already included the concept of a switching point to

deal with the aforementioned needs. In WiMAX, the switching point within a frame can be set dynamically to assign the appropriate amount of UL and DL timeslots. In LTE, a set of configurations are defined, and each cell can select one of these configurations allowing a downlink assignment varying from 40 to 90 percent of the available transmission time intervals (TTI) [4]. The reconfiguration time is limited by a timer [6] where studies in [7] show that lowering down such reconfiguration time can provide benefits to the system in terms of throughput, especially at lower loads. In our envisioned 5G system, we remove any switching point restrictions, such that each slot can be arbitrarily set to either uplink or downlink [3].

The key contribution of this paper lies in presenting a simple distributed UL/DL slot selection scheme showing the potential gains in session throughput and reduced packet delays when having complete control and freedom in assigning each slot as either UL or DL, as opposed to a fixed slot strategy, in the context of the envisioned 5G system. We also show that interference rejection combining (IRC) capable receivers are an effective way to suppress interference variations introduced by the proposed flexible UL/DL slot allocation scheme. The paper is organized as follows. In Section II we give a short overview on our envisioned 5G concept. In Section III we then proceed to introduce our proposed UL/DL selection scheme and its associated parameters. System level simulation results are shown in Section IV, and Section V finally concludes the paper and states the future work.

A.2 Envisioned 5G Concept

In this section we present the most relevant concepts of our envisioned 5G concept related to this study. In the first subsection we will describe our envisioned frame structure, its direct impact on latency and the flexibility it provides in assigning each slot arbitrarily as UL or DL. Thereafter we will proceed to describe the potential of IRC receivers in the context of our problem. A general overview of the whole concept can be found in [8].

A.2.1 Frame Structure

One key characteristic feature of the envisioned 5G frame format is its short 0.25ms duration. The main goal of using such a short frame is to reduce latency. This simplified frame structure is shown in Figure A.1.

The frame consists of a downlink and uplink control part, a demodulation

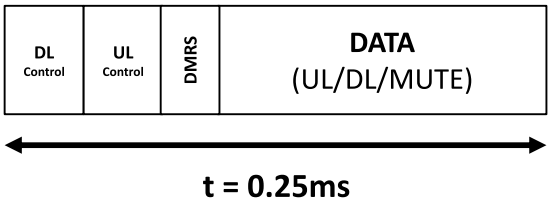


Fig. A.1: 5G Frame Structure

reference symbol (DMRS) and a data slot part. The data slot can be set to UL, DL or MUTE, giving us full flexibility to assign each slot arbitrarily every 0.25ms. Further details related to the frame design and the switching costs incurred from having such a short frame structure can be found in [3].

The DL control part is used by the access point (AP) to signal grant messages. A grant is essentially a transmission opportunity indication along with associated transmission parameters. It indicates whether the corresponding transmission is in DL or UL, and related information such as the Modulation and Coding Scheme (MCS), rank indicator (number of transmission streams), and allocated channels to be used.

A complete DL transmission procedure operates as follows. Let us consider an AP that decides to schedule a DL transmission towards a particular user equipment (UE). A grant is sent in the DL control part. On the following frame, the DL transmission towards the UE occurs. The one frame delay gives sufficient time to the UE receiving the grant to decode and process such information.

In the case of an UL transmission the following occurs. A scheduling request (SR) is sent by a UE to the AP in the UL control part, indicating its desire to be scheduled in UL. The AP decodes this SR and decides to grant the UE an UL transmission. Such a grant is sent on the subsequent frame, and the associated UL transmission occurs in data part, one frame after the grant is sent.

The DMRS symbol is used to enhance IRC operation, by allowing the receiver to estimate the interference covariance matrix that will be present in the data part [9].

A.2.2 Reducing interference variation via IRC

Another relevant key feature of our envisioned 5G system lies in the availability of interference rejection via IRC. Multiple input multiple output (MIMO)

antenna systems offer extra degrees of freedom in the spatial domain. The antennas can be exploited by using multiple transmission streams, or by using the antennas to suppress interferers. In a system with more than two antennas a balance between the two is also possible.

In a fully flexible TDD UL/DL system, where each cell can switch its transmission direction independently from the neighbouring cells at each frame, the signal to noise plus interference ratio (SINR) experienced by a node is expected to vary quite wildly making it problematic for link and rank adaptation as well as for interference coordination schemes to converge. While the focus of this paper is on the analysis of the potential gains of flexible UL/DL slot selection rather than on rank adaptation, in this section we show that IRC receivers can be an extremely effective antidote to mitigate such a problem. IRC relies on an estimate of the interference covariance matrix for projecting the significant interferers onto an orthogonal subspace with respect to the desired signals. The system design in [8] enables an instantaneous estimate of the interference covariance matrix at each frame via the DMRS symbol, thus enabling the possibility of rejecting the active interferers from neighbouring cells regardless of their transmission direction.

Two important metrics that can severely impact system performance are the experienced average SINR, and the SINR variance. Intuitively we would like to increase the average SINR, and minimize the SINR variance. By minimizing the SINR variance, we ensure that the interference conditions felt by a particular node are independent of a neighbouring cell's varying UL/DL slot selection. We simulate a set of 10 randomly deployed dense networks positioned in a 5 by 2 grid fashion, where each network consists of a single AP and a single UE. The simulation is repeated for different rank transmission schemes, e.g. Rank 1 transmission scheme - all nodes use a transmission rank of 1. We then analyse the SINR variation experienced by the different nodes when using IRC. The DL and UL traffic share is set to be equal, such that at a particular frame, each cell can equally go for either UL, or DL based on which direction has the most data to transmit.

A concise representation showing the SINR variation behaviour in the different schemes is shown in figure A.2. Here we show the average SINR versus the average SINR variance for all nodes. The average SINR variance is extracted by computing the SINR variance for each node, and taking the average of this variance across all nodes for each deployment. It is quite clear that at rank 1, IRC can be effective in mitigating interference, by providing a high average SINR and low average SINR variance. The price to pay here is the inability to fully utilize all the resources via higher order spatial multiplexing, as all the available degrees of freedom are used to reject interference.

On increasing the rank, the average SINR decreases, since fewer antennas are used to reject interference. At rank 2 and 3 the SINR variance grows significantly, since a neighbouring cell containing an interferer we cannot reject might be active or inactive based on the neighbour's decision to schedule it or not. In summary, figure A.2 shows that if needed, at rank 1, IRC is able to mitigate interference variation effectively.

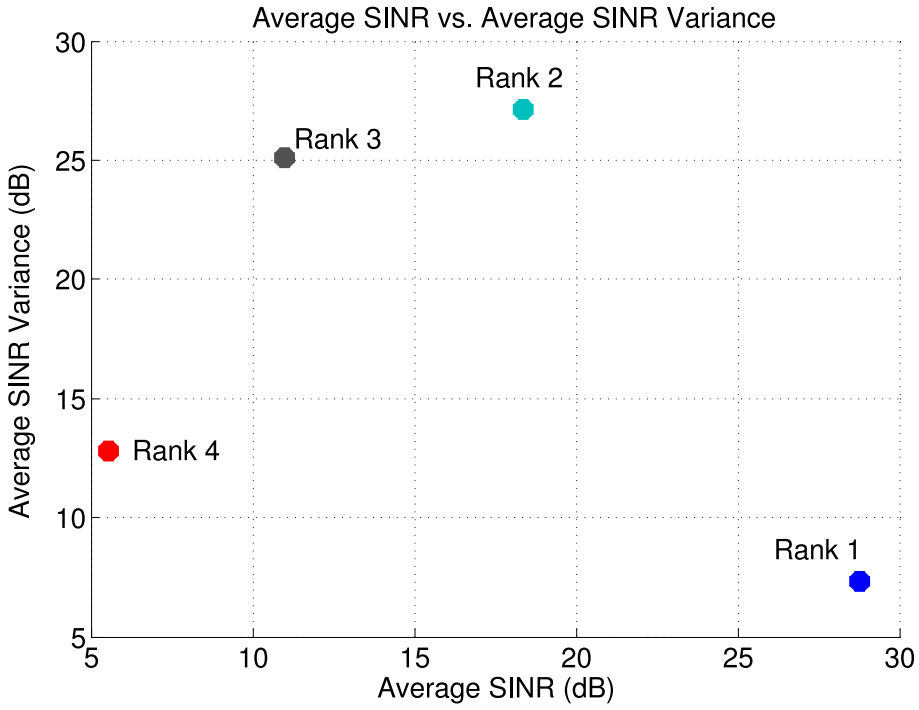


Fig. A.2: Average SINR vs. Average SINR Variance

A.3 UL/DL Selection Algorithm

Most of the previous work related to flexible UL/DL slot allocation, focuses on new types of generated interference when employing unsynchronized UL/DL time slot allocations between cells. In traditional TDD and FDD systems, the UL and DL periods of all cells coincide with each other such that in UL, the AP's experience interference from the UE's, while in DL the UE's experience interference from the AP's. In a fully flexible system, one cell could be in DL while its neighbouring cell is in UL, such that a UE can experience interference from another UE, and an AP experiences interference from another AP.

In [10], Haas et. al study this interaction and show that statistically this scenario could be seen as an advantage due to the statistical randomness of the system, such that if a user is always in outage due to high perceived interference in previous UL/DL synchronized systems, allowing randomization of a time slot can statistically remove that particular user from outage at times. In [11] Omiyi et. al introduce the concept of a busy burst signal sent at the end of each transmission by the receiver to enable interference aware scheduling. A similar enhanced concept is proposed in [12], this time in an LTE oriented context. In this study more information in the busy burst signal is provided, such as signal power, interference and receiver throughput. Here the focus lies in studying the possible throughput gains when having a fully flexible system, where the authors claim a large increase in throughput in lowly loaded systems and even marginal gains at high loads when using an interference aware scheduler. In [13], Duwhuszko et. al study the potential gains of having a distributed and cooperative scheme for selecting UL/DL adaptation schemes by interference pricing exchange messages.

Unlike previous studies, in this paper we rely on IRC to deal with the interference variations brought along with the freedom of assigning each slot as UL or DL arbitrarily. The main advantage of a flexible UL/DL allocation scheme lies in its fast adaptiveness to deal with sudden bursts of traffic in a particular direction making it a highly desired feature in local area networks where such conditions are expected to occur. In this study we devise an algorithm that decides the transmission direction of a cell (UL or DL), with the scope of showing the benefits of having this flexibility from a higher layer point of view.

Our proposed algorithm works as follows. We introduce two main parameters which impact the direction of transmission. These are a buffer size threshold and a Head-of-Line (HOL) delay threshold, denoted as $Th_{BufferSize}$ and $Th_{HOLDelay}$ respectively. The main role of $Th_{BufferSize}$ is to avoid the buffer from overflowing once the buffer size starts growing excessively, while the role of $Th_{HOLDelay}$ lies in bounding the delay experienced by a packet.

The AP is considered to be the decision maker, since it is the entity sending the grants. Internally the AP has updated information related to the buffer size and the current HOL delay in DL, but it needs to be informed of such metrics in UL. We assume that such UL information is ideally embedded in the SR sent from the UE to the AP. This allows the AP to have updated information related to buffer status and HOL delays in both UL and DL, allowing it to take a sensible decision.

The algorithm initializes its current direction to downlink by default. It then waits until it is requested a direction decision, on the start of a frame. Once

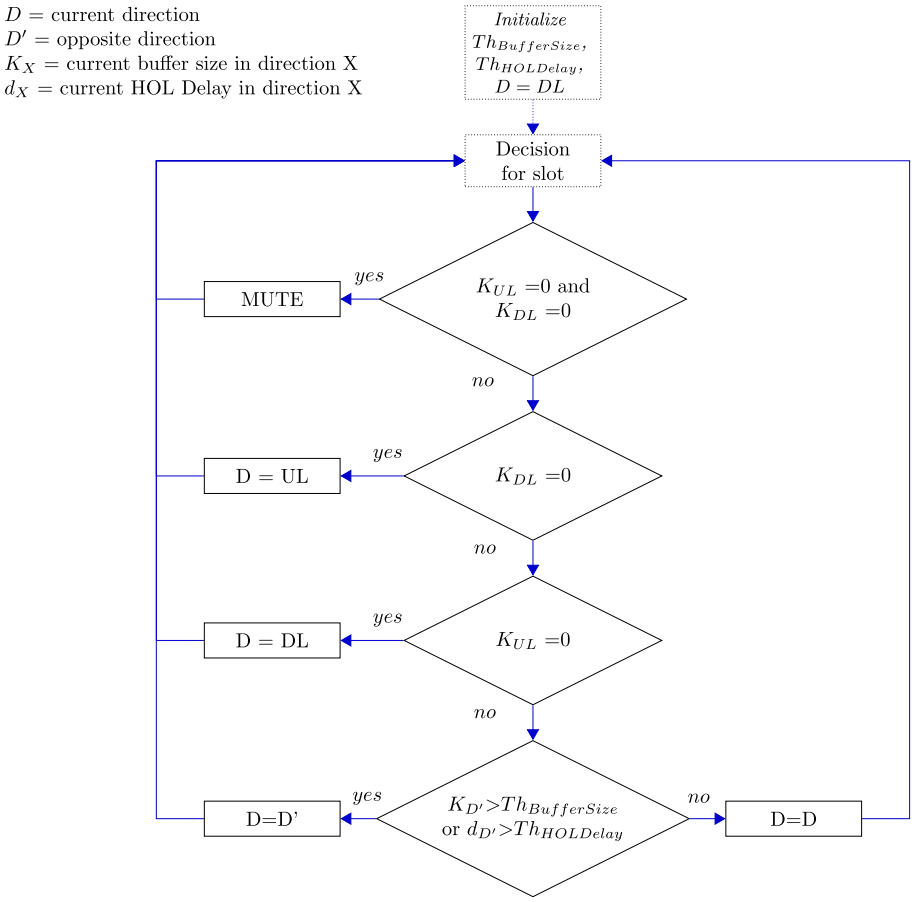


Fig. A.3: UL/DL Selection Algorithm

requested, it checks whether there is any data in the UL or DL direction. If no data is to be transmitted in any of the two directions, the frame is muted. If data is available in only one direction, the algorithm will schedule the direction having data. If data is available in both buffers, the algorithm will inspect the buffer size and HOL delay of the direction not being utilized. If any of the buffer size or HOL delay exceed the predefined thresholds, it switches the direction, otherwise it keeps the current direction. This operation is illustrated in figure B.3.

A.4 Performance Evaluation

In this section we present the results of the presented algorithm. To conduct our analysis, we use a custom discrete event based system level simulator. The most relevant simulation parameters are shown in Table C.1. Our scenario consists of a 10x2 grid consisting of 20 apartments each having a size of 10m x 10m as shown in figure C.1. The apartments are separated by walls and in each apartment we deploy an AP and a single UE terminal, both of which are randomly deployed. The UE's affiliate to an AP according to a closed subscriber group policy, meaning that a UE connects to the AP located in the same apartment.

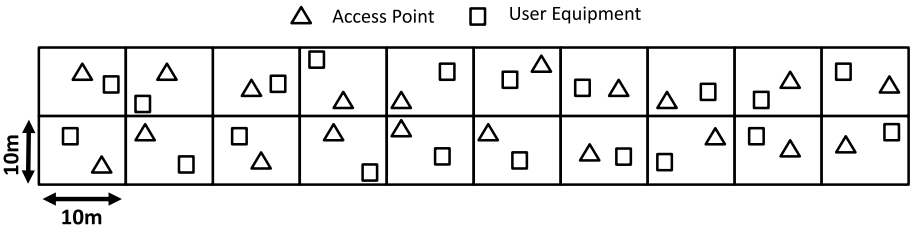


Fig. A.4: 10x2 Scenario

Based on the analysis carried out in Section A.2.2, in order to lower the interference variation, a rank 1 configuration is used. This allows us to limit the inherent convergence problems of link and rank adaptation algorithms, allowing us to solely focus on the benefits of a flexible UL/DL allocation.

The packets generated during the on period of the bursty traffic model, are eventually fragmented into 1500 byte packets, representing the very common Ethernet v2 maximum transmission unit (MTU). All links use the same on-off traffic model specified in Table C.1. The instantaneous packet arrival will temporarily overload the network but the radio link control (RLC) layer will buffer the packets. On average the traffic load will represent 30% load in one case and 70% in the second case. We also use a simplified RLC model where we simply aggregate and send as many packets as we can during a frame transmission opportunity. If the transmission opportunity provided by the MAC is much larger than the data available in the RLC buffers we simply use data padding. The buffer size was set to be infinite. While this is ideal and unrealistic in practical systems, such an assumption allows us to keep our focus on the merits of the algorithm, neglecting any external behaviours occurring due to buffer overflows.

Given our traffic model, we have found out that the performance of the algorithm is insensitive to the $Th_{BufferSize}$ threshold, since the algorithm was in

most cases performing a switch in direction based on the $Th_{HOLDelay}$ threshold. To explain such behaviour, let us consider an arriving burst of data in UL, and an active burst of data in DL currently being served. In principle, setting a $Th_{BufferSize}$ threshold close to the average packet size, will allow the arriving burst of data to be served immediately, but the switch in the other direction will also happen quickly, since there is a high probability that the current experienced buffering delay already exceeds the defined $Th_{HOLDelay}$. Eventually, the direction which was triggered because of $Th_{BufferSize}$, will also experience buffering delay, and will subsequently switch because of the $Th_{HOLDelay}$ threshold. Setting a $Th_{BufferSize}$ threshold well below the average packet size expected in the system, will eventually make the algorithm switch direction on every slot, for any defined $Th_{HOLDelay}$. Such behaviour occurs because the algorithm will switch direction continuously until the buffer size goes under the defined threshold. However, when that point is reached, the buffering delay would have grown above any reasonably defined $Th_{HOLDelay}$, and hence a fast switch will occur independently of the defined $Th_{HOLDelay}$. On the other hand setting a large $Th_{BufferSize}$ will make the system switch direction based on the $Th_{HOLDelay}$ threshold. This proves that the $Th_{BufferSize}$ threshold does not significantly affect the algorithm outcome. Given our traffic model, instantaneously high loads were expected and all schemes ended up giving a similar performance when sweeping through different $Th_{HOLDelay}$ values. Similar performance was observed because at high bursty loads, the delay of a packet will eventually grow due to buffering delay and exceed any reasonably defined $Th_{HOLDelay}$. At this point, the algorithm will then converge to switch on every slot independently of the chosen $Th_{HOLDelay}$ parameter. The $Th_{HOLDelay}$ parameter was hence chosen to be very short to switch as fast as possible to reduce the individual packet delay.

As a baseline result we consider two different schemes. Firstly we assume a simple fixed one UL to one DL slot allocation, referred to as a Fixed Slot 1:1 strategy hereinafter. We also consider a traffic based scheme which simply allocates the transmission to the direction instantaneously having more data to transmit. If no data in either direction is present the frame is muted for both schemes.

In Figures A.5, and A.6 we compare the Fixed Slot 1:1, as well as the traffic based scheme to our proposed algorithm in terms of session delay and session throughput, respectively, when the system is loaded at 30% of the system's maximum capacity. A session is defined as a burst of data occurring during the on period of the traffic generator. The session throughput represents the amount of bits received from that session divided by the time taken to receive all the fragmented packets in the session, starting from the

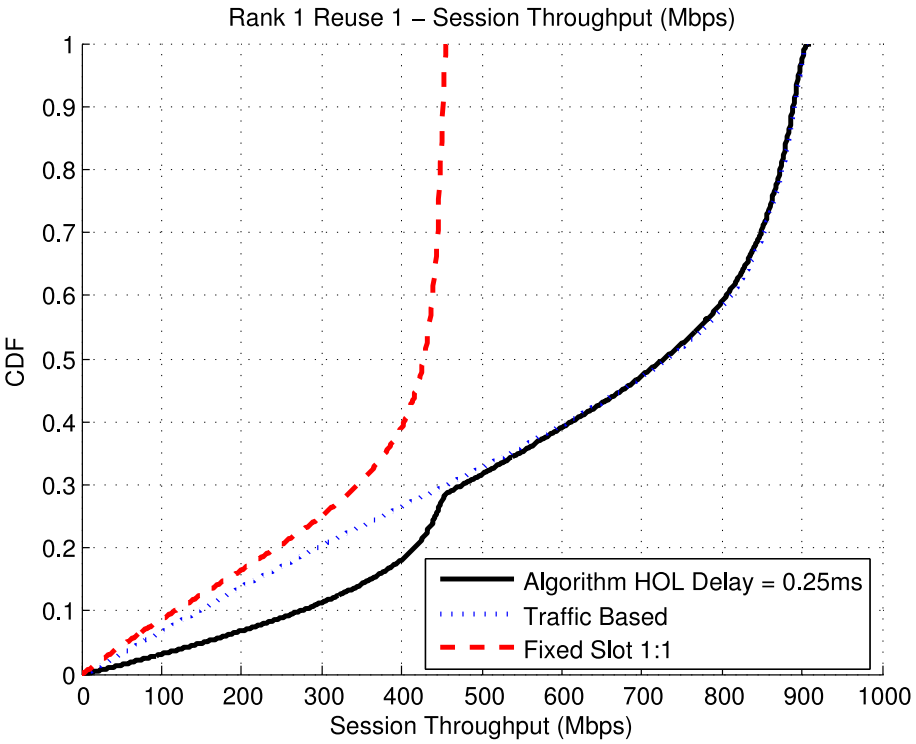


Fig. A.5: Session Throughput (Mbps) at 30% load

first received fragmented packet since the session is created. The session delay measures the amount of time taken to receive the last fragmented packet of the session, starting from the first received fragmented packet since the session is created. Additional delays incurred due to packet processing are neglected and excluded from the session delay.

As expected and witnessed by Figures A.5 and A.6 we observe gains in both session throughput and delay from our algorithm compared to the Fixed Slot 1:1 scheme. This is because whenever a sudden burst in traffic in one direction exceeding half of the cell's capacity occurs, a rigid fixed slot allocation scheme can only accommodate half of the cell's capacity while our algorithm can exploit the full cell capacity if the other direction is inactive. Similar benefits from the traffic based scheme are also observed in the session throughput in figure A.5. The traffic based scheme is efficient in terms of exploiting the available capacity in the channel as it tends to maximize the link efficiency by sending as much data as it can, hence it is also able to reach the maximum system capacity in terms of session throughput. The problem with the traffic based scheme occurs when both UL and DL directions have data to transmit,

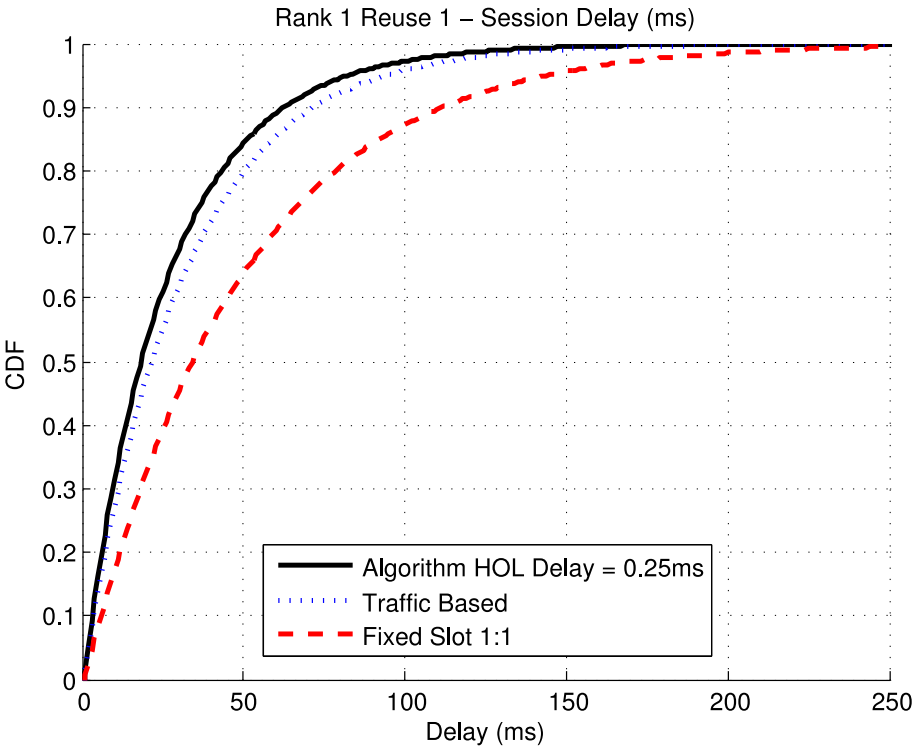


Fig. A.6: Delay (ms) at 30% load

and one direction is more loaded than the other. The less loaded direction, i.e. the direction having smaller amounts of data in its buffer, can be momentarily starved. This increases the packet delay as shown in figure A.6 and also lowers the session throughput for that link as shown in the lower percentiles of the CDF in figure A.5. The defined $Th_{HOLDelay}$ in our algorithm ensures a degree of fairness and helps us bound the delay without letting it grow excessively unlike the traffic based scheme.

Tables A.2 and A.3 show some numerical results at 30% and 70% average load for the session delay, session throughput and final throughput. The final throughput is defined as the total amount of bits received over the total simulation time.

Ideally if the bursty nature of traffic allowed us to always instantaneously have only one direction active we would expect to double the session throughput and halve the packet delay, when compared to a Fixed Slot 1:1 scheme. However, since this is not always the case the observed gains from Tables A.2 and A.3 are slightly reduced to 49% for the delay and to 41% for the session

throughput at 30% load for half of the cases. On increasing the load to 70%, and therefore increasing the probability of having two active links at a time, the gains of our proposed algorithm reduce to 41% in terms of delay for half of the cases.

In a multi-cell scenario there are cases where we can increase the gains even more than 50%. Even though IRC is effective at suppressing interference, there will be some nodes who are interfered by an amount of nodes that exceed the number of degrees of freedom dedicated to interference suppression. Since our algorithm services packet bursts more quickly, it can statistically reduce the interference levels over time, hence improving the conditions for these interfered nodes, allowing them to increase their session throughput significantly. This can be observed in the 5th percentile of the session throughput at 30% load.

Finally, having full UL/DL flexibility, allows us to also reach a higher peak final throughput when loading each cell with more than 50% load. This can be observed from Table A.3 where the proposed algorithm gets an improvement of 8% in final throughput at 70% load compared to the Fixed Slot 1:1 scheme. This happens because the provided flexibility, unlike a rigid Fixed Slot 1:1 scheme, does not lock the UL and DL capacity to half the cell's capacity.

A.5 Conclusions & Future Work

In this paper we have presented a simple and effective algorithm that exploits flexible UL/DL slot allocation in our envisioned 5G system. The algorithm is capable to automatically track traffic variations, and adapt accordingly to the instantaneous needs of the cell. The flexibility of the envisioned frame structure gives us the possibility to adapt to sudden traffic imbalances between UL and DL and allows us to exploit the full capacity of our system rather than being constrained by a rigid fixed allocation scheme. Compared to a fixed slot scheme with one slot in uplink and the other in downlink we are capable of almost doubling the experienced session throughput and reducing the experienced delay by half in approximately 50% of the cases. We are also able to reduce the delay compared to a traffic based scheme. This is achieved by introducing a head of line delay parameter threshold. We also analyse and confirm that interference rejection via IRC is an effective tool in increasing robustness from the source of interference and can help us in stabilizing the experienced interference, hence counteracting and minimizing previous challenges experienced with flexible UL/DL slot allocation.

Our future work will focus on exploiting spatial multiplexing gains while

limiting SINR variations via appropriate link and rank adaptation algorithms. The impact of hybrid automatic repeat request (HARQ) on reducing packet losses will also be studied and its impact on the delay will be investigated. The behaviour of such an algorithm in the presence of multiple users requires further investigation and the impact of reduced signalling and the compression of feedback reports needs to be quantified. Finally, the behaviour of the algorithm with different parametrizations, in the presence of higher layer protocols such as the Transmission Control Protocol (TCP), also needs to be assessed, as such protocols provide the possibility of testing the algorithm in instantaneously high asymmetric conditions due to the presence of TCP acknowledgements.

References

[1] Cisco. "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2011–2016". In: *White Paper* (Feb. 2012).

[2] V. Chandrasekhar, J.G. Andrews, and Alan Gatherer. "Femtocell Networks: A Survey". In: *Communications Magazine, IEEE* 46.9 (Sept. 2008), pp. 59–67. ISSN: 0163-6804. DOI: 10.1109/MCOM.2008.4623708.

[3] E. Lähetkangas et al. "On the TDD subframe structure for beyond 4G radio access network". In: *Future Network and Mobile Summit (FutureNetworkSummit), 2013*. July 2013, pp. 1–10.

[4] 3GPP. *Long Term Evolution*. http://www.3gpp.org/article/lte. 2009.

[5] IEEE. *802.16-2004 IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems, 01 October 2004*. http://www.wimaxforum.org/index.htm. 2004.

[6] 3GPP TS 36.331 v10.3.0. *3rd Generation Partnership Project; Technical Specification Group Radio Access network; Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC); Protocol Specification (Release 10)*.

[7] Zukang Shen et al. "Dynamic uplink-downlink configuration and interference management in TD-LTE". In: *Communications Magazine, IEEE* 50.11 (Nov. 2012), pp. 51–59. ISSN: 0163-6804. DOI: 10.1109/MCOM.2012.6353682.

[8] "Preben Mogensen et al. "5G Small Cell Optimized Radio Design". In: *Globecom. IEEE Conference and Exhibition* (2013). ISSN: 1930-529X.

[9] F.M.L. Tavares et al. "On the Potential of Interference Rejection Combining in B4G Networks". In: *VTC Fall, 2013 IEEE 78th*. Sept. 2013, pp. 1–5. DOI: 10.1109/VTCFall.2013.6692318.

[10] H. Haas, P.K. Jain, and B. Wegmann. "Capacity improvement through random timeslot opposing (RTO) algorithm in cellular TDD systems with asymmetric channel utilisation". In: *Personal, Indoor and Mobile Radio Communications, 2003. PIMRC 2003. 14th IEEE Proceedings on*. Vol. 2. Sept. 2003, 1790–1794 vol.2. DOI: 10.1109/PIMRC.2003.1260423.

[11] P. Omiyi, H. Haas, and G. Auer. "Analysis of Intercellular Timeslot Allocation in Self-Organising Wireless Networks". In: *Personal, Indoor and Mobile Radio Communications, 2006 IEEE 17th International Symposium on*. Sept. 2006, pp. 1–5. DOI: 10.1109/PIMRC.2006.254252.

[12] P. Jänis, V. Koivunen, and C.B. Ribeiro. "On the performance of flexible UL-DL switching point in TDD wireless networks". In: *GLOBECOM Workshops (GC Wkshps), 2011 IEEE*. Dec. 2011, pp. 225–230. DOI: 10.1109/GLOCOMW.2011.6162442.

[13] Alexis A. Dowhuszko et al. "A decentralized cooperative Uplink/Downlink adaptation scheme for TDD Small Cell Networks". In: *Personal Indoor and Mobile Radio Communications (PIMRC), 2013 IEEE 24th International Symposium on*. Sept. 2013, pp. 1682–1687. DOI: 10.1109/PIMRC.2013.6666413.

[14] IST-WINNER, II. *Deliverable 1.1.2 v.1.2, "WINNER II Channel Models"*. Tech. rep. Tech. Rep., 2008, 2008. URL: http://projects.celtic-initiative.org/winner+/deliverables.html.

Table A.1: Simulation Parameters

Parameter	**Value**
System Parameters	Bandwidth 200MHz; f_c = 3.5GHz
Antenna Configuration	4x4
Rank, Frequency Reuse	Rank 1, Reuse 1
Propagation Model	WINNER II A1 w/fast fading[14]
Transport Protocol	UDP
Traffic Model	Bursty, On-Off Generator
ON Period Neg. Exponential Mean Interarrival Time	0.25ms
OFF Period Neg. Exponential Mean Interarrival Time	110ms (average 30% load), 47ms (average 70% load)
Average Packet Size (ON Period)	2MBytes
Theoretical Cell Capacity	968Mbps, at highest MCS
UL/DL Strategies	Fixed Slot 1:1, Traffic Based, Algorithm based on HOL Delay
$Th_{HOLDelay}$	0.25ms
$Th_{BufferSize}$	2MBytes
Simulation Time	1 second
Simulation Drops	50

Table A.2: Results for different Loads in Percentiles

Delay (ms)	Fixed Slot			Algorithm		
	5%	50%	95%	5%	50%	95%
30% Load	3.5	35	144	2	18	82
70% Load	6	70	282	3.5	41	211
Session Throughput (Mbps)	Fixed Slot			Algorithm		
	5%	50%	95%	5%	50%	95%
30% Load	57	428	453	154	729	898
70% Load	13	219	451	23	354	870
Final Throughput (Mbps)	Fixed Slot			Algorithm		
	5%	50%	95%	5%	50%	95%
30% Load	59	146	267	57	148	273
70% Load	179	305	429	173	304	468

Table A.3: Algorithm Gains over Fixed Slot Strategy

Delay (ms)	Gains in %		
	5%	50%	95%
30% Load	42.9	48.6	43.1
70% Load	41.7	41.4	25.2
Session Throughput (Mbps)	Gains in %		
	5%	50%	95%
30% Load	63.0	41.3	49.6
70% Load	43.5	38.1	48.2
Final Throughput (Mbps)	Gains in %		
	5%	50%	95%
30% Load	-3.5	1.4	2.2
70% Load	-3.5	-0.3	8.3

Paper B

Flexible UL/DL in Small Cell TDD Systems: A Performance Study with TCP Traffic

Davide Catania[1], Marta Gatnau Sarret[1], Andrea F. Cattoni[1], Frank Frederiksen[2], Gilberto Berardinelli[1], Preben Mogensen[1,2]

[1]Department of Electronic Systems, Aalborg University, Denmark, [2]Nokia Networks, Denmark

dac@es.aau.dk, mgs@es.aau.dk, afc@es.aau.dk, frank.frederiksen@nsn.com, gb@es.aau.dk, pm@es.aau.dk

The paper has been published in the *Vehicular Technology Conference (VTC Spring), 2015 IEEE 81th*

The layout has been revised.

Abstract

Time division duplex (TDD) systems offer a substantial amount of freedom to deal with downlink (DL) and uplink (UL) traffic asymmetries. Most TDD-based systems define either multiple static configurations or adaptive approaches to deal with such asymmetries. Our envisioned 5G concept embraces the flexibility brought along by TDD, and allows us to switch the link direction on a slot by slot basis. In this paper we study the interaction of Transmission Control Protocol (TCP) traffic, with a fully flexible UL/DL TDD allocation scheme. We show that flexibility is not only beneficial for exploiting the different instantaneous UL and DL traffic variations, but also performs well with TCP traffic, where the protocol behaviour plays an important role in throughput performance. The advantages of full flexibility compared to fixed static allocations for TCP traffic are reported for both small and large payloads, and for multi-cell scenarios where both DL and UL traffic are present.

B.1 Introduction

The demand for higher speed wireless access systems is constantly on the rise. Measurements carried out by Cisco [1] forecast an exponential increase in traffic in the coming years. To deal with such growth, further cell densification and smaller cells are expected. Such trends indicate that a 5th generation (5G) radio access system for small cells will be inevitable in the future. A future 5G system will also spawn a myriad of new services and applications, highlighting the importance of studying the interaction of such a system with higher layer protocols such as the Transmission Control Protocol (TCP).

In order to accommodate the introduction of a new 5G system, a number of key radio technology components are expected. In our previous work [2], we have presented our envisioned 5G concept, describing the features which we believe will be part of this future technology. Some foreseen key characteristics which are particularly relevant to this study include the design of a new short frame structure, the presence of Interference Rejection Combining (IRC) receivers and most importantly the possibility of having a fully flexible Time-Division Duplex (TDD) system where each time slot can be arbitrarily allocated to operate in either downlink (DL) or uplink (UL). The proposed short frame structure has a duration of 0.25 ms, and consists of a DL and UL control part, a Demodulation Reference Symbol (DMRS) aiding the operation of interference rejection, and a data part which can be arbitrarily set to DL or UL on a per frame (time slot) basis [3]. Such a feature is desirable in small cell scenarios where the number of users per cell is typically smaller than

in traditional macro cell deployments. In such conditions the probability of sudden traffic imbalance between DL and UL increases. Another important key feature is the presence of IRC receivers where studies carried out in [4] show that IRC receivers have the ability to suppress high interference levels in ultra dense small cells. In [5], we have shown the benefits of having a fully flexible TDD system where each time slot can be assigned as UL or DL, by presenting a simple flexible UL/DL allocation algorithm which exploits temporal UL/DL bursty traffic variations. The introduced dynamic UL/DL behaviour will potentially introduce varying interference conditions, providing serious challenges for interference coordination or link adaptation schemes to converge. However we have seen that the usage of IRC receivers, is an effective way to limit the experienced varying interference conditions [5]. The reason for this is that the DMRS symbol located on every frame allows for the estimation of the interference co-channel matrix, giving the IRC receiver the ability to suppress interferers independently from their origin [4]. This is an important characteristic in our system, since it gives us an extra level of protection against high interference variation, allowing us to exploit TDD flexibility without experiencing highly variable interference conditions due to neighbouring cells switching their transmission mode from UL to DL or vice versa.

In our previous investigation [5], on the benefits of flexible UL/DL allocation, we have considered bursty traffic using a simplified Poisson process traffic model operating over the User Datagram Protocol (UDP). However, studies reveal that most traffic is carried over TCP [6], stressing the importance of studying the behaviour of our flexible UL/DL allocation algorithm with TCP traffic. TCP is typically the protocol of choice for any application requiring a reliable communication that cannot tolerate packet loss.

The focus of this work lies in studying the gains of flexible UL/DL allocation with TCP traffic, where acknowledgement feedback channels and protocol dynamics play an important role in determining suitable UL/DL TDD configurations. In this paper we highlight the benefits of having a flexible TDD scheme over a static fixed TDD configuration when dealing with this type of traffic. The scope of this work is limited to local area small cells, where the originating traffic is located close to the access point, such that the delay between the server and the access point is relatively small or negligible.

The paper is structured as follows. In the next section we give a short background on TCP and a brief summary on the state of the art with regards to TCP traffic behaviour over wireless TDD systems. In section B.3, we present the 5G framework and system model used to carry out our analysis. In Section IV, we show our performance evaluation and discuss the results obtained

and finally, Section C.5 concludes the paper.

B.2 State of the Art

B.2.1 TCP Overview

TCP [7] is an end-to-end transport protocol used for the reliable transfer of data over the Internet Protocol (IP). One of the most important mechanisms of TCP is the congestion control mechanism [8]. The basic operation of congestion control defines the notion of a congestion window, which limits the amount of data a transmitting host can send. On reception of this data, the receiving host will generate an acknowledgement (ACK) stating the amount of bytes it has correctly received. On receiving this ACK, the transmitting side can increase its congestion window, such that it is allowed to transmit a larger amount of data. This means that the reception of ACKs increases the congestion window size, allowing the transmitting host to send more data at once before requiring an ACK. Therefore a short round trip time (RTT) between the transmission of data and the acknowledgement of data is highly desirable since it allows the congestion window size to grow, especially in the initial TCP transfer phase. As shown in figure B.1, the growth of the congestion window is governed by two stages, a slow start phase where the congestion window increases exponentially on the reception of each ACK, and a congestion avoidance stage, where the congestion window increases linearly on reception of each ACK. Since the initial congestion window size is typically very small, it is expected that during the initial transfer period severe network underutilization might be experienced. While the congestion control mechanism might initially limit the maximum attainable system performance, it serves the indispensable role of protecting the network from congestion, disallowing a transmitting host from immediately sending large amounts of data in a potentially already congested network.

B.2.2 TDD Performance over TCP Traffic

The possibility of dynamic TDD is currently under evaluation in Long Term Evolution (LTE) [9]. This allows the system to switch between a set of LTE predefined TDD configurations. The periodicity of the switch is governed by a specified reconfiguration time. Studies in LTE TDD [10] show the benefits of flexibly switching between the specified TDD configurations. The reports claim that the faster the reconfiguration time, the bigger the through-

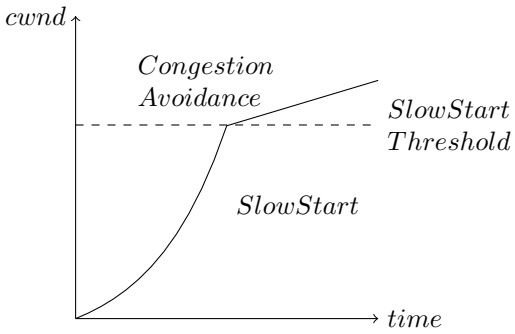
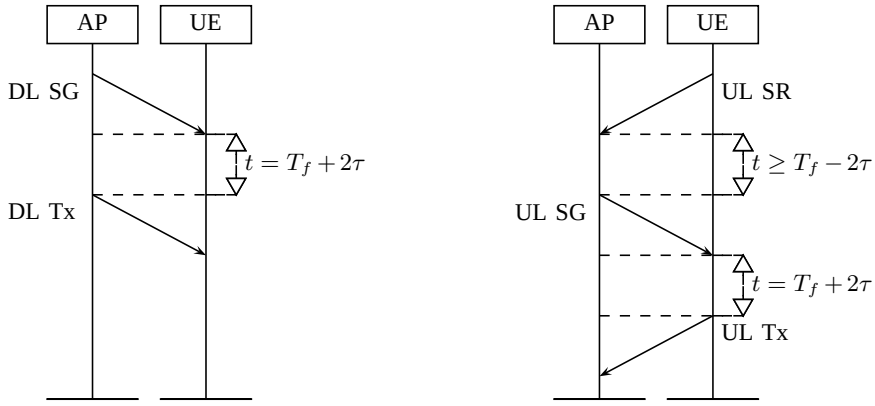


Fig. B.1: TCP Congestion Window (cwnd)

put gains. However such studies are limited to bursty traffic and do not consider any shaping of the traffic from higher layer protocols such as TCP.

In [11], the authors consider the performance of TCP over an LTE TDD system with different DL to UL configurations, and different payload sizes, by considering the download throughput from a single user. In this study, dynamic TDD is not considered, and the different DL to UL LTE TDD configurations are not reconfigured during the course of the simulations. Focus is instead put on the behaviour of different TDD configurations on TCP sessions consisting of payloads having different sizes. The authors conclude that for small payload sizes the fast reception of TCP ACKs is important to increase the throughput since it stimulates the growth of the congestion window. Therefore for small payload sizes, a balanced DL to UL configuration is necessary. The benefit of highly asymmetrical links starts appearing for large payload sizes, where most of the time is spent in transferring the data, rather than waiting for TCP's initial phase responsible for increasing the congestion window to grow with the reception of ACKs. We shall present a related analysis in Section IV and show a similar behaviour for our envisioned 5G concept.

WiMAX is another system allowing adaptive DL/UL bandwidth allocation via TDD. TCP traffic studies in WiMAX [12], show that in multi-user DL only transfers, it is beneficial to increase the DL to UL asymmetry up to a certain point. Beyond that point, ACK's in the UL channel will be starved and will not have enough capacity to be transmitted, hence degrading the DL throughput. In [13], the authors adapt the DL to UL ratio based on the number of long-lived file transfer protocol (FTP) flows in the system and show that their adaptive bandwidth allocation scheme outperforms static schemes.



(a) DL Scheduling Procedure **(b)** UL Scheduling Procedure

Fig. B.2: Downlink and Uplink Scheduling Procedures (frame duration $T_f = 0.25ms$, symbol time $\tau = 0.0192ms$)

B.3 System Model

In this section we will present details about our system's DL and UL transmission procedures, the proposed flexible UL/DL algorithm and the simulation framework used.

B.3.1 Downlink and Uplink Transmission Procedures

Figures B.2a and B.2b show the respective DL and UL scheduling procedures. As mentioned earlier on, the frame consists of a DL and UL control channel part each occupying one symbol, a DMRS symbol, and a data part which can be arbitrarily set to operate in DL or UL. The timing follows the frame structure defined in [3]. For a frame to be scheduled in DL, the access point must send a DL scheduling grant (SG) in the DL control channel. A DL data transmission can then occur in the data part of the subsequent frame, allowing suitable time for control channel message processing. The UL procedure is initiated by a user via a scheduling request (SR) in the UL control part. The access point can then decide to schedule an UL transmission by transmitting an UL SG in the DL control channel, resulting in an UL transmission on the subsequent frame.

B.3.2 Flexible UL/DL Algorithm

Figure B.3 illustrates the basic operation of the algorithm. In essence, the algorithm checks whether there is data in the DL or UL direction. Since the direction decision is taken by the access point, any DL traffic information will be available locally, while similar UL information will be received via the SR in the UL control channel. In the absence of data, the algorithm mutes the frame, otherwise it schedules the direction having data. If both UL and DL directions have data to be transmitted, a decision to schedule a particular link direction needs to be taken. The algorithm first proceeds to inspect the currently unscheduled direction's buffer head-of-line (HOL) delay, representing the delay of the oldest packet in its buffer. If this delay exceeds a predefined HOL delay threshold a switch in direction is carried out, otherwise the direction remains the same. This whole procedure is repeated on each of the following frames, where a decision is taken on the same criteria. Increasing the HOL delay threshold will delay the initial scheduling time of the currently unscheduled link to momentarily favour the currently scheduled link. The packets in each of the individual queues are scheduled in a first in first out (FIFO) manner. Further details related to the algorithm's behaviour and performance can be found in [5].

B.3.3 Simulation Framework

A custom discrete event system level simulator, modelling the various layers of the protocol stack was used to carry out our analysis. The simulator models the application, and internetworking layers, i.e. TCP and IP overhead, and the radio access layer. The radio access layer consists of the radio link control (RLC) layer, a radio resource management (RRM) module, a medium access control (MAC) layer and a physical layer built according to our envisioned 5G concept.

The main key performance indicator (KPI) used to conduct our analysis is the session throughput. The session throughput represents the throughput experienced when downloading a file generated at the application layer. We compare this KPI for various static TDD allocations and the proposed flexible UL/DL algorithm with different parametrizations.

Table C.1 shows the most relevant parameters used to conduct our simulations.

We carry out both single cell, and multi cell simulations. In the single cell scenarios we offer DL only data traffic while in the multi cell case we offer

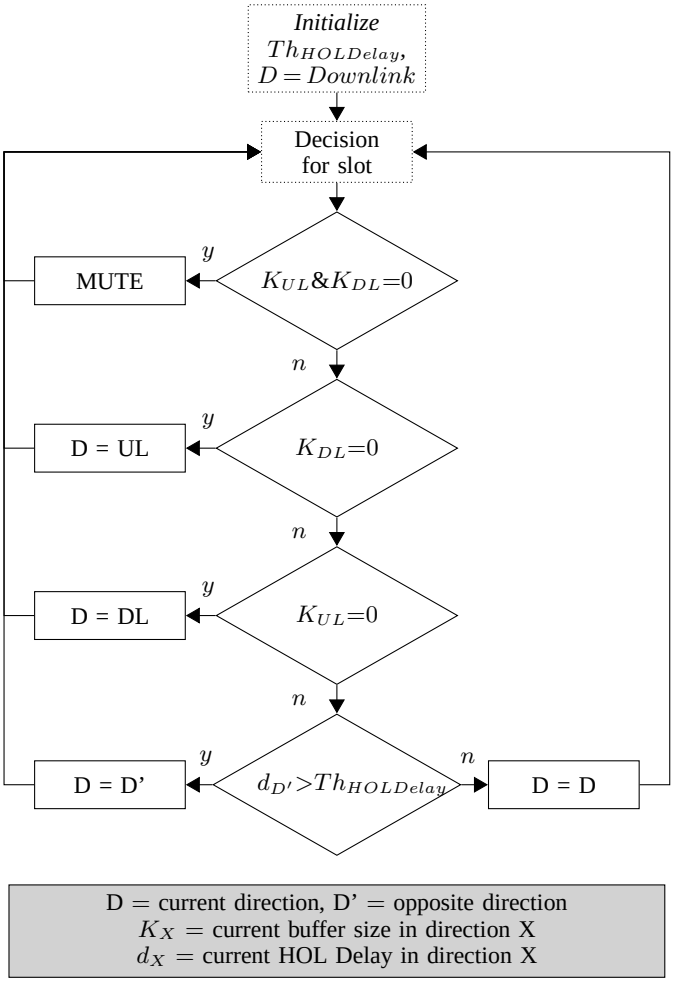


Fig. B.3: UL/DL Selection Algorithm

Table B.1: Simulation Parameters

Parameter	Value
Traffic Model	
Single Cell	Fixed On/Off
Multi Cell	Negative Exponential On/Off
Transport Protocol	TCP New Reno [14]
TCP	
Maximum Segement Size (MSS)	9kB
Minimum Retransmission Timeout	0.1s
Initial Slow Start Threshold	1MB
Initial Congestion Window	1MSS
Restart Window	2MSS
RLC	
Single Cell	Unacknowledged Mode
Multi Cell	Acknowledged Mode
MAC Packet Processing Delay	0.05ms
HARQ Maximum Retransmissions	4
Antenna Configuration	4x4
Rank, Frequency Reuse	Rank 1, Reuse 1
Propagation Model	WINNER II A1 [15]
System Parameters	Bandwidth = 200MHz
	f_c = 3.5GHz
UL/DL Strategies	
Fixed Slot (DL:UL)	1:1, 2:1, 4:1, 10:1
Flexible UL/DL($Th_{HOLDelay}$)	0.25, 1, 2, 4ms
Simulation Parameters	
Single Cell	
Simulation Drops	10
Simulation Time	5s
Multi Cell	
Simulation Drops	10
Simulation Time	30s

both DL and UL data traffic. The motivation for carrying out single cell DL only simulations is to limit our analysis to TCP's throughput performance with fixed-sized small (100 kB) and large (10 MB) payloads. Investigating these simple scenarios enables us to look further into the details with regards to TCP ACK generation to transmission opportunity delay. We will refer to this delay, as the TCP ACK waiting time. This waiting time is correlated with the congestion window growth and throughput performance for small payloads.

For our single cell simulations, we set the RLC protocol to operate in unacknowledged mode (UM), to observe the aforementioned delay statistics as a sole function of the slot configuration and the algorithm parametrization, removing any extra delay generated due to RLC acknowledgements. The multi cell simulations are then conducted to give a more holistic overview of the situation, where the cells have traffic in both UL and DL, and the payload size is not fixed but is generated according to a negative exponential distribution with an average size of 2MB. For our multi cell simulations, where interference is present, we set the RLC to operate in acknowledged mode (AM) to recover from any losses that the Hybrid Automatic Repeat Request (HARQ) module cannot cope with.

B.4 Performance Evaluation

In this section we shall report our findings, by first taking a look at some single cell results with DL only traffic, followed by an analysis of a 10x2 multi-cell scenario with both DL and UL traffic.

B.4.1 Single Cell with DL Traffic

To confirm the findings in [11], we set up a single cell scenario with an access point and a user, with DL traffic. We set the payload file size to 100 kB and 10 MB respectively and inspect the session throughput and the TCP ACK waiting time, with different static TDD and flexible UL/DL configurations. The session throughput S is defined in Eq. (B.1) as the payload size P divided by time T. P is fragmented into N packets of size b, each consisting of 1500 bytes. T is defined as the time from when the first packet of P was generated until the last packet N of P was received, i.e.

$$S_P = \frac{\sum_{i=1}^{N} b_i}{T} \tag{B.1}$$

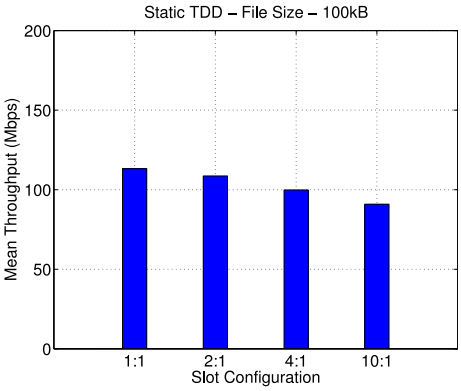
Figures B.4 and B.5 show these results with 100 kB and 10 MB payloads, for different slot configurations, and algorithm parametrizations. The slot configurations represent the DL to UL ratio as DL:UL.

From Figure B.4a, we see that for small payload sizes, as the DL to UL configuration ratio increases, the throughput decreases and the corresponding ACK waiting time shown in Figure B.4c increases, limiting the growth of the congestion window and blocking DL transfer of data. From Figure B.5a, we observe that when the payload size is large, we benefit from having a large DL to UL configuration ratio since we spend most of the time transmitting data, rather than in TCP's initial congestion window growth phase. When the congestion window size has grown sufficiently to fill multiple frames with data, it is thus beneficial to delay the acknowledgement and use the extra DL slots for data transfer, as witnessed in Figures B.5a and B.5c. This avoids limiting the system to half of the cell's capacity. These findings are consistent with the results provided in [11].

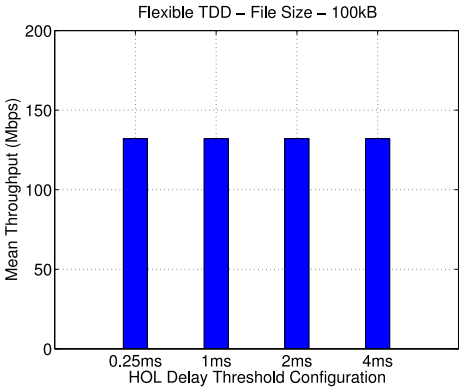
When comparing the throughput results with the ones obtained with the flexible UL/DL algorithm located in Figures B.4b and B.5b, we notice that the flexible approach provides comparable or better results. From Figure B.4b, we observe that for small payloads the performance of the flexible UL/DL algorithm is independent of the chosen HOL delay threshold. The reason is that with small payloads, we spend most of the time in TCP's initial phase, where focus should be given in increasing the congestion window size. The algorithm will assign traffic to the direction having data, such that when no further data can be pushed in DL due to the congestion window limit, the algorithm will immediately allocate resources to UL where the TCP ACK traffic is present. When the ACK is allowed to be transmitted, data in DL can then flow.

In Figure B.5b we notice that on increasing the payload size to 10 MB, there is a benefit in delaying the switch and increasing the HOL delay threshold for the flexible scheme. Once the congestion window size has grown, it starts becoming wasteful to over prioritize the TCP ACK by parametrizing the algorithm with a short HOL delay, as this will lead the system to oscillate between DL and UL modes, limiting the data direction's capacity to half the available capacity. Such a condition indicates that as soon as the congestion window size has grown sufficiently to span multiple slots of data the impact of the RTT on the throughput diminishes, and increased focus should be given on allocating slots to the direction having the data.

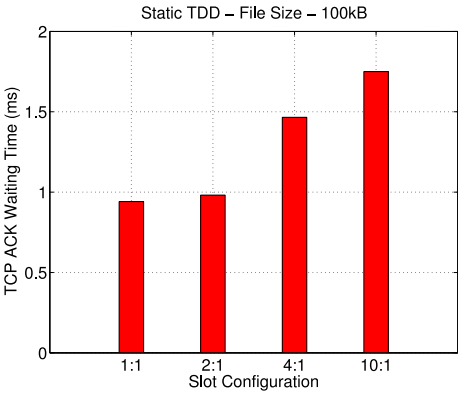
Therefore the flexible UL/DL algorithm presented in Section B.3.1 should be parametrized with a HOL delay threshold of 2 to 4 ms for TCP traffic, since the impact of such a parameter should only kick in when the congestion



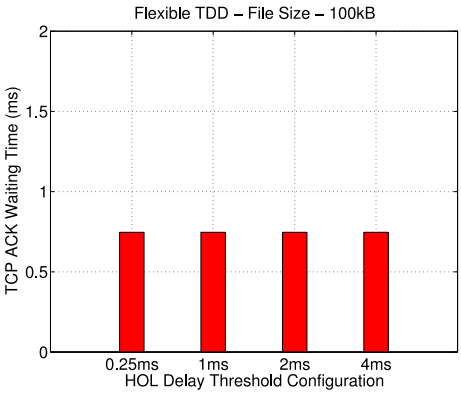
(a) Session Throughput - Fixed



(b) Session Throughput - Flexible

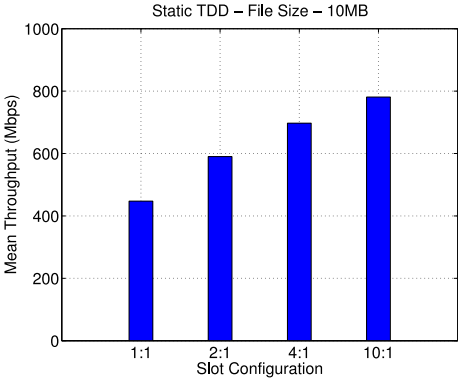


(c) ACK Waiting Time - Fixed

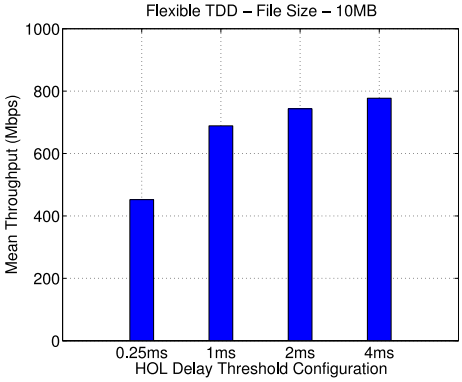


(d) ACK Waiting Time - Flexible

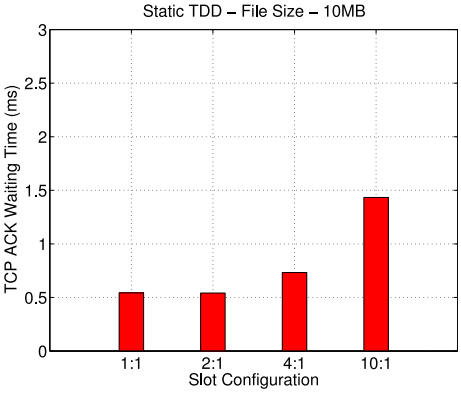
Fig. B.4: Single Cell Results for File Size of 100kB



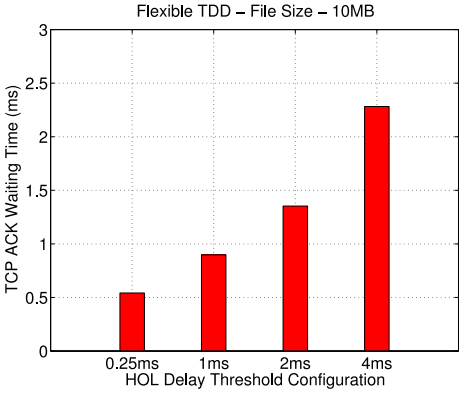
(a) Session Throughput - Fixed



(b) Session Throughput - Flexible



(c) ACK Waiting Time - Fixed



(d) ACK Waiting Time - Flexible

Fig. B.5: Single Cell Results for File Size of 10MB

window size has grown sufficiently and both DL data and UL ACK data are concurrently present. For small payload transfers, we would most of the time be operating in TCP's initial slow start phase. In this phase, as soon as the congestion window size limit has been reached, data arriving from DL will be blocked until an UL TCP ACK is received. This means that the algorithm will only detect data in UL and it will swiftly schedule an UL TCP ACK as required, independently of any parametrization. This is witnessed by Figure B.4d where the ACK waiting time is equal in all cases.

The benefit of the flexible approach is that it adapts to the protocol behaviour of TCP. This can also be observed in Figures B.4d and B.5d, where the TCP ACK waiting time is low for small payloads and increases for large payloads on large HOL delay thresholds. This allows DL capacity to be fully exploited when a UL TCP ACK is not urgently required.

A suggested HOL delay threshold of 2 to 4 ms goes in contrast with our previous investigation of this algorithm with UDP traffic [5]. UDP does not require a transport layer ACK feedback channel. This means that whenever there is both UL and DL traffic present in the system, independently of any HOL delay threshold chosen, the buffering delay would eventually exceed any reasonably defined HOL delay threshold resulting in an alternating UL and DL switch. This behaviour ends up delivering similar throughput performance independently from the HOL delay threshold chosen.

B.4.2 Multi Cell with DL and UL Traffic

To complete our analysis we repeat the study with mixed DL and UL traffic in a 10x2 multi cell scenario, resulting in 20 cells, each consisting of a single access point and a single user. Here we compare our flexible algorithm against a static fixed 1:1 DL to UL configuration. While instantaneously the links can be heavily loaded with either DL or UL traffic, on average, in this specific investigation, the two links are loaded with an equal traffic load, such that a fixed 1:1 DL to UL TDD configuration represents the optimal long term average DL to UL traffic share.

In the application layer, we generate payload bursts having an average size of 2MB, with a negative exponential average inter-arrival time of 0.5s. The payload burst size is generated according to a negative exponential distribution and can be smaller or larger than the average payload size of 2MB. This allows us to capture the merits and demerits with different settings, since we have seen that the performance is dependent on the payload size. We also capture different instantaneous conditions such as DL only, UL only, and both

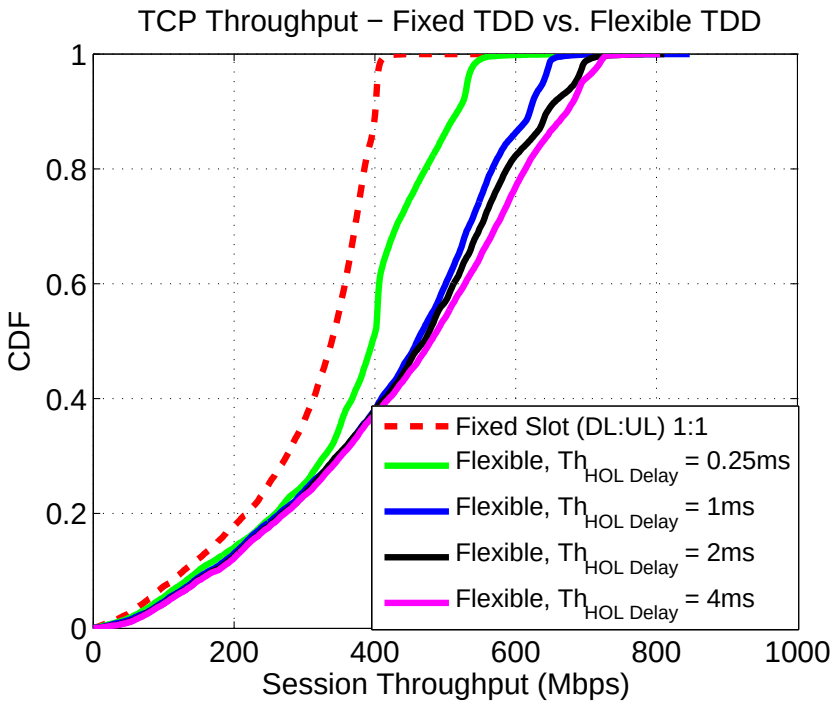


Fig. B.6: Multi cell Results, Session Throughput, Fixed 1:1 vs Flexible Schemes

DL and UL traffic where in this case the TCP ACK is piggybacked with the data. The cumulative distribution function (CDF) of the session throughput results for the different schemes is shown in Figure B.6.

The fixed scheme offers an inferior throughput against any flexible scheme. The best performing schemes for the flexible algorithm occur when the switching HOL delay threshold is delayed. The fixed scheme suffers in various cases. It will suffer whenever there is an average to large payload size, with instantaneous DL or UL traffic only. The flexible approach can exploit such scenarios and can also adapt to different payload sizes when the HOL delay threshold increases. For large payloads, the fixed scheme is limited to half of the cell's capacity, and a similar behaviour is observed for the flexible scheme with very short HOL delay thresholds, implying prioritization of the TCP ACKs even when it is not needed. On the other hand the other flexible schemes can achieve a significantly higher throughput when only DL or UL traffic is instantaneously present in the system.

B.5 Conclusion

In this paper we have studied the advantages of flexible UL/DL TDD for TCP traffic. The benefits of flexibility are not only exploitable when instantaneously there is only DL or UL traffic. The flexible approach also adapts well to TCP protocol behaviour. This is done by allocating TCP ACK's promptly, whenever the data transmission side is blocked due to limited TCP congested window sizes making the flexible approach suitable for small payloads, independently of any parametrization. For large payloads, it is important to delay TCP ACK's with suitable HOL delay threshold parameters. This avoids unnecessary switching when the congestion window has grown sufficiently such that data can still be transmitted without requiring a TCP ACK. In most cases, with an appropriate HOL delay threshold parametrization of 2-4 ms, the flexible scheme outperforms the static TDD configurations, and automatically adapts to different traffic conditions and payload sizes. Such a feature therefore offers an attractive and elegant solution to autonomously and automatically tune the TDD slot configuration in uncoordinated small cell deployments.

In this study we have assumed that the network delay from the server to the AP is negligible. This condition holds for data originating close to the AP. However, for more general cases, such a condition might not be valid, and the TCP throughput performance of flexible TDD operating over network delays spanning multiple TDD slots needs to be evaluated in future work.

References

[1] Cisco. "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2013–2018". In: *White Paper* (Feb. 2014).

[2] Preben Mogensen et al. "Centimeter-wave concept for 5G ultra-dense small cells". In: *IEEE VTS Vehicular Technology Conference. Proceedings* (2014). ISSN: 1550-2252.

[3] E. Lähetkangas et al. "On the TDD subframe structure for beyond 4G radio access network". In: *Future Network and Mobile Summit (FutureNetworkSummit), 2013*. July 2013, pp. 1–10.

[4] F.M.L. Tavares et al. "On the Potential of Interference Rejection Combining in B4G Networks". In: *VTC Fall, 2013 IEEE 78th*. Sept. 2013, pp. 1–5. DOI: 10.1109/VTCFall.2013.6692318.

[5] Davide Catania et al. “The Potential of Flexible UL/DL Slot Assignment in 5G Systems”. In: *Vehicular Technology Conference (VTC Fall), 2014 IEEE 80th*. 2014.

[6] W. John and S. Tafvelin. “Heuristics to Classify Internet Backbone Traffic based on Connection Patterns”. In: *Information Networking, 2008. ICOIN 2008. International Conference on*. Jan. 2008, pp. 1–5.

[7] J. Postel. *Transmission Control Protocol*. RFC 793 (INTERNET STANDARD). Updated by RFCs 1122, 3168, 6093, 6528. Internet Engineering Task Force, Sept. 1981. URL: http://www.ietf.org/rfc/rfc793.txt.

[8] M. Allman, V. Paxson, and W. Stevens. *TCP Congestion Control*. RFC 2581 (Proposed Standard). Obsoleted by RFC 5681, updated by RFC 3390. Internet Engineering Task Force, Apr. 1999. URL: http://www.ietf.org/rfc/rfc2581.txt.

[9] 3GPP. *Long Term Evolution*. http://www.3gpp.org/article/lte. 2009.

[10] *Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Further enhancements to LTE Time Division Duplex (TDD) for Downlink-Uplink (DL-UL) interference management and traffic adaptation (Release 11)*.

[11] R. Susitaival et al. “Internet Access Performance in LTE TDD”. In: *Vehicular Technology Conference (VTC 2010-Spring), 2010 IEEE 71st*. May 2010, pp. 1–5. DOI: 10.1109/VETECS.2010.5493986.

[12] A Eshete et al. “Impact of WiMAX Network Asymmetry on TCP”. In: *Wireless Communications and Networking Conference, 2009. WCNC 2009. IEEE*. Apr. 2009, pp. 1–6. DOI: 10.1109/WCNC.2009.4917743.

[13] Chih-He Chiang, Wanjiun Liao, and Tehuang Liu. “Adaptive Downlink/Uplink Bandwidth Allocation in IEEE 802.16 (WiMAX) Wireless Networks: A Cross-Layer Approach”. In: *Global Telecommunications Conference, 2007. GLOBECOM ’07. IEEE*. Nov. 2007, pp. 4775–4779. DOI: 10.1109/GLOCOM.2007.906.

[14] S. Floyd, T. Henderson, and A. Gurtov. *The NewReno Modification to TCP’s Fast Recovery Algorithm*. RFC 3782 (Proposed Standard). Obsoleted by RFC 6582. Internet Engineering Task Force, Apr. 2004. URL: http://www.ietf.org/rfc/rfc3782.txt.

[15] IST-WINNER, II. *Deliverable 1.1.2 v.1.2, "WINNER II Channel Models"*. Tech. rep. Tech. Rep., 2008, 2008. URL: http://projects.celtic-initiative.org/winner+/deliverables.html.

Paper C

A Distributed Taxation Based Rank Adaptation Scheme for 5G Small Cells

Davide Catania[1], Andrea F. Cattoni[1], Nurul H. Mahmood[1], Gilberto Berardinelli[1], Frank Frederiksen[2], Preben Mogensen[1,2]

[1]Department of Electronic Systems, Aalborg University, Denmark, [2]Nokia Networks, Denmark

dac@es.aau.dk, afc@es.aau.dk, nhm@es.aau.dk, gb@es.aau.dk, frank.frederiksen@nsn.com, pm@es.aau.dk

The paper has been published in the *Vehicular Technology Conference (VTC Spring), 2015 IEEE 81th*

The layout has been revised.

Abstract

The further densification of small cells impose high and undesirable levels of inter-cell interference. Multiple Input Multiple Output (MIMO) systems along with advanced receiver techniques provide us with extra degrees of freedom to combat such a problem. With such tools, rank adaptation algorithms allow us to use our antenna resources to either exploit multiple spatial streams in low interference scenarios, or suppress interference in high interference scenarios. In this paper, we propose and evaluate an interference-aware distributed rank adaptation algorithm. The concept behind our approach is to discourage the choice of transmitting with multiple spatial streams in highly interfered scenarios, and to exploit and encourage the usage of multiple spatial transmission streams in low interference scenarios. We show that our proposed algorithm can be adjusted to preserve and guarantee a good outage performance, while providing the benefit of higher average throughputs, in both low and highly interfered scenarios, when compared to fixed rank configurations, and distributed selfish schemes.

C.1 Introduction

The unrelenting growth of data traffic, the scarcity of available spectrum, and the further densification of cells pose significant challenges to small dense wireless networks operating in centimetre-wave frequencies. Future systems such as the envisioned 5th generation (5G) concept specified in [1] must ensure the delivery of high peak data rates, whilst dealing with severe inter-cell interference problems inherent in ultra dense small cells. One key technology component that can help the envisioned 5G concept address both requirements, is the availability of Multiple Input Multiple Output (MIMO) transceivers.

MIMO systems introduce spatial degrees of freedom, which can be used to increase the system throughput via spatial multiplexing. The number of spatial streams used for transmission is also referred to as the transmission *rank*. Such a feature is however only exploitable in low interference scenarios. In interference limited scenarios, MIMO systems along with Interference Rejection Combining (IRC) receivers can be used to suppress interference [2]. IRC operates by projecting the interfering streams over an orthogonal subspace with respect to the desired ones. The price to pay is a reduction of the number of desired streams since part of the degrees of freedom of MIMO are used for interference suppression. In order to get the best throughput performance, one needs to adjust the configuration of the system based on the experienced

interference scenarios. In low interference conditions a high rank should be favoured, allowing the nodes to reach peak data rates. In high interference conditions, lower ranks should be favoured since this improves the intended receiver's interference suppression capabilities, allowing the system to guarantee a minimum throughput performance. This rank selection technique is known as rank adaptation and is the focus of this paper.

In this paper we propose a novel distributed rank adaptation algorithm. The proposed algorithm is implicitly cooperative and utilizes the concept of interference pricing, by introducing a taxation mechanism which is a function of the transmission rank and the experienced interference conditions. This algorithm is studied within the framework of the envisioned 5G concept [1] via system-level simulations taking into account the causal aspects of a realistic system, and avoiding assumptions which are deemed to be impractical in realistic implementations.

The rest of the paper is structured as follows. In Section II, an overview of the different nature of rank adaptation algorithms is presented. Section III is dedicated to explain our baseline scheme and the proposed rank adaptation algorithm in further detail. In Section IV, the performance evaluation is carried out with system level simulations, and finally Section V concludes the paper.

C.2 Classification of Rank Adaptation Algorithms

This section aims at giving a brief overview of the diverse nature of distributed rank adaptation algorithms. It is out of the scope of this study to give an exhaustive list of all the distributed rank adaptation algorithms available in literature, and therefore a restricted set of works which highlight the different approaches are selected.

One set of distributed rank adaptation algorithms are selfish distributed schemes, which attempt to chose a transmission rank that maximizes the benefit from an own cell's perspective, disregarding the overall system performance and the harm generated to other cells. An example of this is given in [3], where a system-level evaluation of a capacity-based rank adaptation scheme is studied in the context of the Long Term Evolution (LTE) radio access technology. In this scheme, given the presence of inter-cell interference, each node is instructed to choose a transmission rank which maximizes its own MIMO channel capacity.

Another approach to the rank adaptation problem is to cooperate in an at-

tempt to improve the throughput performance of the whole system, rather than just maximizing the benefit from one's own cell perspective. The cooperation can be done explicitly with varying degrees of information exchange. An example of this is given in [4], where a practical cooperative distributed rank coordination scheme motivated by an interference pricing mechanism is proposed with the aim of maximizing a network utility function rather than the individual node performance. In this study it assumed that the cells are coordinated by a cyclic master-slave architecture and can exchange limited amounts of information. Alternatively, the cooperation can be implicit where no information between the cells is exchanged. An example is given in [5], where an algorithm which assigns precoding vectors that achieve a compromise between the beamforming gain and the mitigation of interference created towards other receivers, is proposed.

Particularly relevant to this work is the cooperative distributed rank adaptation algorithm proposed in [6]. The algorithm is evaluated in the context of the envisioned 5G concept in [1]. Implicit cooperation between the cells is achieved by classifying the interference levels into three states. Based on this classification, a corresponding multiplicative taxation is applied to each rank, discouraging the use of higher ranks if high interference levels are detected. In this study, it is however assumed that knowledge of future outgoing inter-cell interference conditions is available. Given that a key feature of the envisioned 5G concept [1] is the usage of Time Division Duplex (TDD) and the full flexibility to assign each frame to operate in uplink (UL) or downlink (DL), it is expected that inter-cell interference conditions can vary rapidly from one frame to the next, making the aforementioned assumption impractical.

C.3 Rank Adaptation Schemes

The goal of this section is to explain the operation of our proposed rank adaptation algorithm, which aims at providing a implicitly cooperative distributed rank adaptation scheme overcoming the impractical assumptions taken in [6]. A benchmarked selfish scheme is also briefly explained. Both schemes revolve around the idea of choosing a rank k^* which maximizes a scheme specific utility function Π_k as shown in Eq. (C.1). k represents the transmission rank, where $1 \leq k \leq min(N_t, N_r)$. N_t and N_r represent the number of antennas at the transmitter and receiver side respectively.

$$k^* = \arg\max_k \Pi_k \tag{C.1}$$

C.3.1 Selfish Scheme

The selfish scheme refers to a rank adaptation algorithm in which each node will simply try to choose a rank which maximizes its capacity, independently of the amount of interference it might generate to its neighbours. This is done by maximizing the utility function Π_k presented in Eq. (C.2).

$$\Pi_k = \sum_{j=1}^{k} C(\overline{SINR_j}) \tag{C.2}$$

where $C(.)$ represents an estimate of the achievable rate, assumed to be the Shannon formula in this study. $\overline{SINR_j}$ represents the log averaged post IRC signal to noise plus interference ratio (SINR) at stream j, of the latest Q, with $Q \geq 1$, $SINR_j$ estimates. For the sake of brevity, the $SINR_j$ expression for an IRC receiver is omitted, but the same model as the one described in [2] is used for this study.

In order to estimate the post IRC SINR and achievable rate for each rank k, the desired channel matrix information H_D related to desired streams is required. The frame structure in the envisioned 5G concept [1] operates in TDD mode, with full flexibility to assign each frame as UL or DL. It includes a demodulation reference symbol (DMRS) on every frame, allowing the acquisition of H_D. It should also be noted, that the calculation of the utility function specified in Eq. C.2, is done on the reception of every frame, and can only be applied to future transmissions, hence taking into account the causality aspects of a practical implementation. The utility function is computed at each node, and while the access point (AP) is the entity taking the final transmission rank decision, it is assumed that the user equipments (UE's) will feedback a preferred rank indication to the AP's in the control channel.

This selfish scheme is a simple proposal to the rank adaptation problem and is expected to perform well in lightly loaded interference scenarios. However, in dense and highly active small cell deployments, there is a tendency that the scheme will instruct its nodes to use a high rank even if there is little gain in doing so, creating undesirable system interference scenarios that can not be suppressed effectively. In general, if the amount of interfering streams becomes larger than the degrees of freedom available for interference suppression, a node will be a potential victim of interference which can not be effectively suppressed. It is expected that such a selfish scheme will start showing its weaknesses on high load conditions.

C.3.2 Proposed Interference Aware Scheme

In order to deal with the aforementioned problems, we propose an interference aware rank adaptation scheme inspired from the interference coordination channel selection scheme proposed in [7]. While our proposed approach operates on a similar philosophy, there are some subtle yet important differences to note. In this study, the resources refer to transmission streams and not to channel resources. It is important to also note that using fewer transmission streams gives us the possibility to suppress interference more effectively, since the unused transmission streams can be utilized to improve interference resilience. Our proposed algorithm instructs each node to independently choose a rank k which maximizes the utility function shown below in Eq. (C.3),

$$\Pi_k = \underbrace{\sum_{j=1}^{k} C(\overline{SINR_j})}_{\text{Estimated Capacity for rank k}} - \underbrace{kW_k C\left(\frac{I}{N}\right)}_{\text{Taxation for rank k}} \tag{C.3}$$

where $C\left(\frac{I}{N}\right)$ is given by,

$$C\left(\frac{I}{N}\right) = log_2\left(1 + \frac{tr(H_I {H_I}^H)}{\sigma_n^2}\right) \tag{C.4}$$

and $tr(H_I {H_I}^H)$ represents the trace of the interference covariance matrix (ICM), and σ_n^2 represents the noise power. H_I refers to the concatenation of all interference matrices, where $H_I \in \mathbf{C}^{N_r \times L}$, and L represents the number of interfering streams coming from neighbouring cells.

On every reception, a node will calculate the achievable rate for each rank as in the selfish scheme, and for each rank it will apply a taxation function which depends on the interference over noise ratio, the rank k, and a weighting function W, which is a monotonically increasing function of rank k. The net effect of this utility function is that it discourages the usage of higher ranks if interference conditions are high. The level of discouragement is quantified via the defined weights W_k. As a boundary condition please note that if all the weights W_k were set to 0, the utility function would degenerate to the one defined in the selfish scheme.

It is assumed that the ICM can be reliably estimated on a per frame basis due to the designed frame structure proposed in [1], allowing effective in-

terference suppression, and updated knowledge with regards to interference conditions. In the envisioned 5G concept in [1] the same frame format is used in UL and DL. The nodes are also synchronized, allowing the estimation of cross-link interference. However, unlike the work proposed in [6], it can not be assumed that future interference conditions are known, due to the potential inter-cell interference variation experienced from one frame to the next. This variation is expected given the flexible UL/DL allocation feature envisioned by the 5G concept described in [1]. When assuming the causality aspects of a practical implementation, it also means that the choice of a transmission rank might not be taking accurately into account the appropriate interference considerations, at the time it is applied. Such a demerit of the system should be captured when analysing the realistic gains of an evaluated rank adaptation algorithm.

In order to maintain a realistic framework and deal with such a problem, no future interference conditions are assumed, and updated incoming interference conditions are applied in the utility function described in Eq. (C.1). To combat further this potential frame-to-frame inter-cell interference variation, instead of considering the instantaneous SINR, we consider the log-averaged SINR, $\overline{SINR}$, for the latest Q SINR samples in Eqs. (C.2) and (C.3), when estimating the achievable rate for a particular rank. In an ideal interference pricing scheme, one would consider outgoing interference conditions rather than incoming interference conditions, since the outgoing interference reflects the harm which is being generated to the neighbouring nodes. However, the knowledge availability of the outgoing interference is also considered unrealistic in our scenario, and hence it is considered that the interference component in Eq. (C.4) represents the actual incoming interference. While this is non-ideal and the incoming interference is not always equal to the outgoing interference, it gives a rough approximation of the current experienced interference levels. The goal of this utility function is to provide a distributed rank adaptation scheme which given its limited knowledge with regards to future interference conditions, attempts to improve the outage performance of highly interfered nodes, in dense and highly active small cell deployments. This is done by discouraging the use of higher transmission ranks, when high interference levels are detected.

C.4 Performance Evaluation

The potential of our proposed algorithm is evaluated via system level simulations, using the scenario shown in Figure C.1. The deployment scenario consists of 20 rooms each having a dimension of 10m x 10m. A room consists

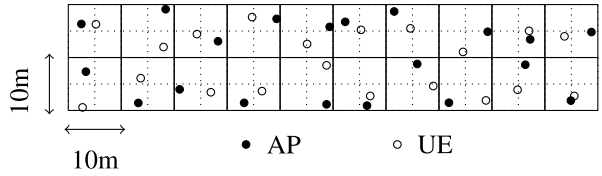


Fig. C.1: 10x2 Scenario

of a randomly deployed AP and UE, each featuring a 4x4 MIMO transceiver, allowing a maximum rank transmission of 4 streams. It is also assumed that both AP's and UE's transmit with the same transmission power. The system level simulator used to evaluate the proposed rank adaptation scheme models the physical layer and frame structure as described in [1]. Flexible UL/DL allocation is allowed on a per frame basis where the link direction decision is based on the algorithm proposed in [8]. This flexible UL/DL algorithm considers the buffer size and the experienced head-of-line delay of both UL and DL links, and effects a switch in direction if any of these exceed a set of predefined thresholds. Receptions are decoded or discarded based on the modulation coding scheme (MCS) selected for the transmission of a frame and the instantaneous SINR conditions of the channel. When selecting from a set of available MCS's a 10% blocking error rate target is assumed, based on the measured SINR of previous receptions. Unsuccessful transmissions are retransmitted via hybrid automatic repeat request (HARQ), and at the radio link control (RLC) layer. In the application layer, we use a finite buffer traffic model, which generates payloads with an average size of 2MB. These payloads are referred to as sessions and are generated every inter-arrival period, $t_{inter-arrival}$ seconds.

The proposed algorithm is simulated with two sets of weights W_1=[0, 0.5, 0.66, 0.75] and W_2=[0, 0.25, 0.66, 0.75], where kW_1=[0, 1, 2, 3] and kW_2=[0, 0.5, 2, 3]. W_1 is a more conservative approach which favours low ranks, while W_2 is a more aggressive approach which favours slightly higher ranks. In both cases a rank 1 transmission is not penalized with any tax. For benchmarking purposes, the proposed rank adaptation scheme is compared against a fixed rank scheme, where all nodes in the networks use a fixed transmission rank, and against the described selfish scheme. A summary of the relevant parameters is shown in Table C.1.

The metric used to compare the results is the average session goodput per node. During a simulation, a node is intended to receive L payloads at its application layer. For each of these payloads, there will be an experienced session goodput associated with it, which we denote as S_i. It is therefore expected that during a simulation a node will finally get L values of S_i, where

the average of these values, represents the average session goodput per node.

Table C.1: Simulation Parameters

Parameter	Value
Traffic Model	Bursty, Negative Exponential On/Off
RLC	Acknowledged Mode
UL/DL Strategy	Flexible UL/DL [8]
Rank Adaptation Schemes	Fixed Ranks 1, 2, Selfish, Proposed Algorithm with W_1, W_2
Log-Averaged Q SINR Samples	5
HARQ Max. Retransmissions	4
Antenna Configuration	4x4
Frequency Reuse	Reuse 1
Propagation Model	WINNER II A1
System Parameters	Bandwidth 200MHz, f_c = 3.5GHz
Scenario	10x2 Stripe (see figure C.1)
Simulation Parameters	
Simulation Drops	25
Simulation Time	5s

C.4.1 Low Load

The rank adaptation schemes are first simulated in a low load environment. The network is loaded with a traffic load representing approximately 25% utilization of the resources available. In such conditions, with a finite buffer bursty traffic model, the probability that two neighbouring cells are simultaneously active is low. This means that statistically, the perceived interference levels are also low and higher ranks should be favoured. The cumulative distribution function (CDF) results for the average node session goodput achieved with the different schemes, are shown in Figure C.2. Interestingly enough, at low load, a fixed rank 2 configuration gives better performance than a rank 1 configuration, even in outage. The reason is that in low interference scenarios, a higher rank gives a higher throughput, allowing the network to get rid of data more quickly, hence allowing the temporal reduction

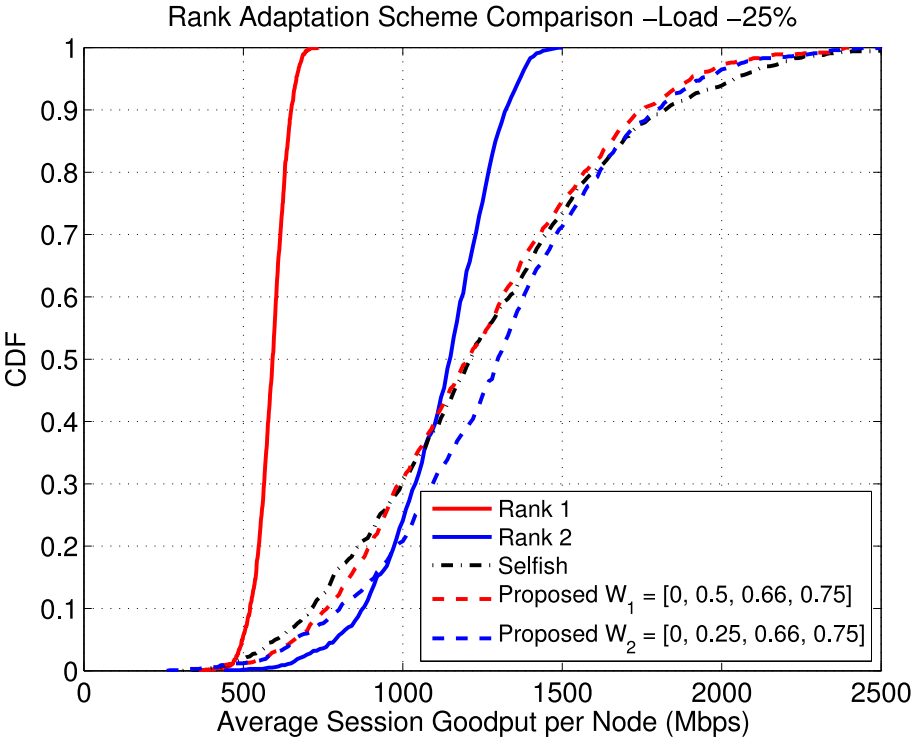


Fig. C.2: Average Session Goodput per Node, Resource Utilization 25%

of interference. All the rank adaptation algorithms behave similarly, reaching close rank 2 outage performance, and exploiting higher ranks and achieving higher speeds for around 60% of the cases. The proposed algorithm with a weighting vector W_2 performs the best, since it closely approaches the outage performance of fixed rank 2, and manages to achieve the highest throughput against all schemes. The distribution of ranks chosen according to the simulated schemes can be seen in Figure C.3. It can be observed that the selfish schemes seem to favour the higher ranks a bit too aggressively, with the W_1 weighting vector, being too conservative with a high share given to rank 1 transmissions. The proposed algorithm with W_2 offers a good compromise by choosing rank 2, for most transmissions. However it is to be noted that the performance gap between the rank adaptation schemes is minimal in this case, and all schemes allow the exploitation of higher ranks when possible, unlike the fixed schemes.

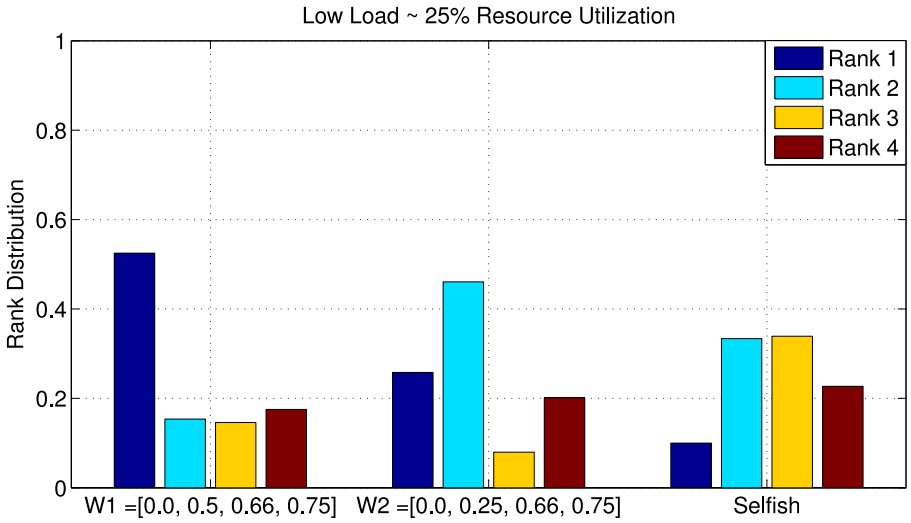


Fig. C.3: Ranks chosen, Resource Utilization 25%

C.4.2 High Load

In high load conditions, stronger interference conditions are expected, since there is a higher probability that neighbouring nodes are simultaneously active. It is therefore expected that lower ranks would usually be preferred in this case. To confirm this hypothesis, the simulation is repeated, with 75% resource utilization. The CDF of the node's average session goodput is shown in Figure C.4. From Figure C.4 we immediately notice that a fixed rank 1 configuration gives the best outage performance but is unable to exploit the higher peak throughputs available with fixed rank 2 configuration. A quick look at Figure C.5 showing the rank distributions chosen by the schemes, indicates that all schemes tend to lower the chosen ranks when compared to the previous low load scenario. This is also seen for the selfish scheme which can deduce that it can achieve a higher capacity with a lower rank at times. However we also observe that the selfish scheme still tends to favour high ranks, creating unfavourable interference conditions which can not be always suppressed. This is also reflected by the bad outage performance achieved by the selfish scheme in Figure C.4. For our proposed scheme we also see that the W_1 configuration gives the best outage performance, since it tends to choose rank 1 most of the times. While guaranteeing the maximum outage performance, it also achieves higher throughputs than a fixed rank 1 configuration for 50% of the cases, giving an ideal balance between outage performance and higher average throughputs. We can also observe that with weighting vector W_2, the algorithm can achieve even higher through-

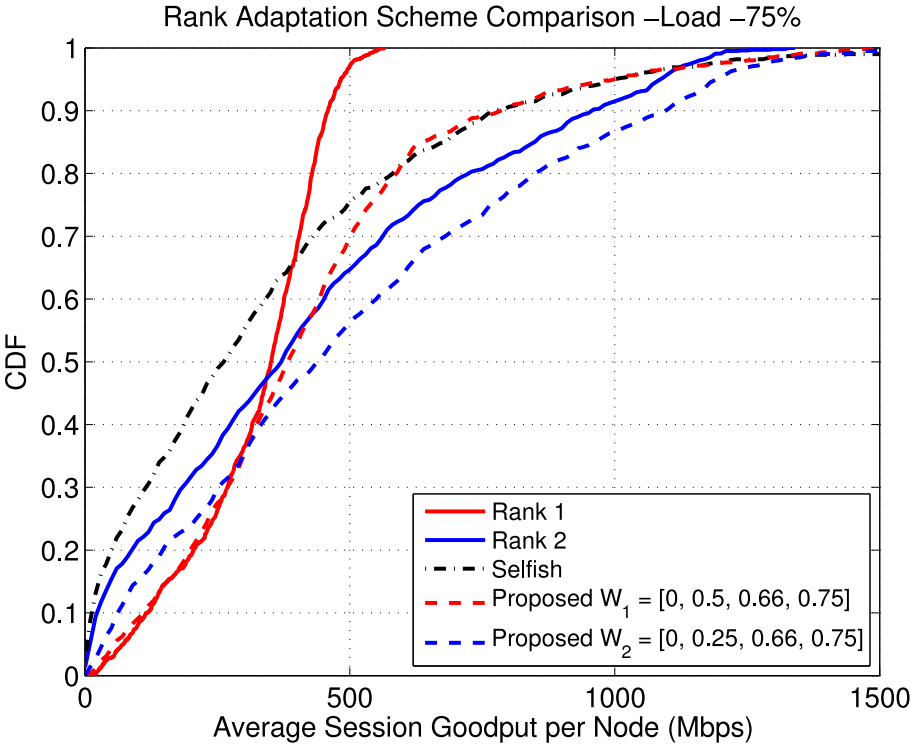


Fig. C.4: Average Session Goodput per Node, Resource Utilization 75%

puts, at the cost of degraded outage performance. For both configurations, there is a trade-off between outage and average throughput performance. A conservative approach will tend to protect the outage performance of potentially highly interfered users at the cost of limiting the attainable maximum throughput performance of users located in more favourable channel conditions which might however be generating undesirable levels of interference to victim nodes. Conversely, a more aggressive approach will tend to give less considerations towards potentially interfered nodes, allowing nodes in favourable conditions to enjoy higher throughput performance.

C.5 Conclusion

In this paper we have proposed and evaluated the performance of a distributed rank adaptation algorithm via system-level simulations. The algorithm does not require any information exchange between the nodes, and operates with no knowledge on future outgoing interference conditions. In-

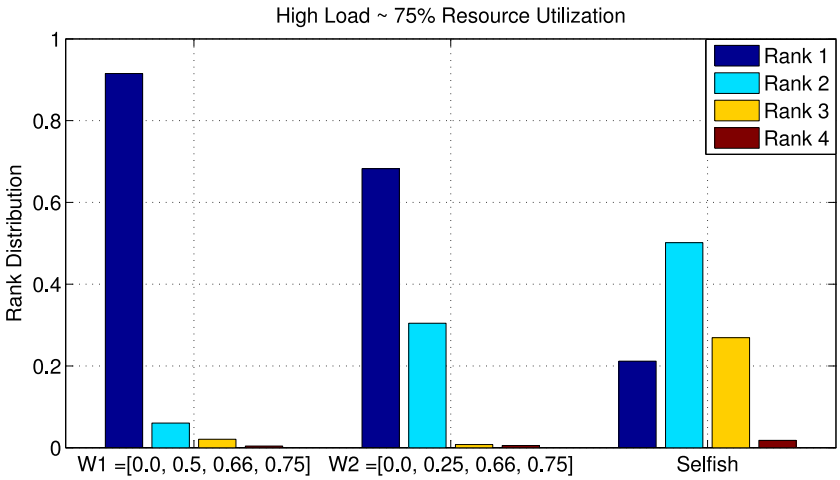


Fig. C.5: Ranks chosen, Resource Utilization 75%

stead we attempt to detect such conditions from past incoming interference measurements. Our proposed scheme, can achieve good outage performance, and enjoy higher average peak throughputs, at both low and high loads, with appropriate weight configuration. It is flexible enough to match and exceed the desired performance of fixed rank configurations in different load conditions. The proposed algorithm is also aware of interference conditions and outperforms the benchmarked selfish scheme.

In its form, the performance of the proposed algorithm is dependent on the parametrization of the weight configuration. Future work will therefore be dedicated to improve the algorithm by having adaptive weights based on the overall perceived network interference and the load of the network. Comparison with algorithms exploiting realistic exchangeable information should also be conducted, to evaluate the potential gains of such additional information.

References

[1] P. Mogensen et al. “Centimeter-wave concept for 5G ultra-dense small cells”. In: *IEEE VTC Spring Proceedings* (2014). ISSN: 1550-2252.

[2] F.M.L. Tavares et al. “On the Potential of Interference Rejection Combining in B4G Networks”. In: *VTC Fall, 2013 IEEE 78th*. Sept. 2013, pp. 1–5. DOI: 10.1109/VTCFall.2013.6692318.

[3] Alexandra Oborina et al. “MIMO performance evaluation in UTRAN Long Term Evolution downlink”. In: *Information Sciences and Systems, 2008. CISS 2008. 42nd Annual Conference on*. Mar. 2008, pp. 1179–1183. DOI: 10.1109/CISS.2008.4558697.

[4] B. Clerckx et al. “A Practical Cooperative Multicell MIMO-OFDMA Network Based on Rank Coordination”. In: *Wireless Communications, IEEE Transactions on* 12.4 (Apr. 2013), pp. 1481–1491. ISSN: 1536-1276. DOI: 10.1109/TWC.2013.013013.112100.

[5] Z.K.M. Ho and D. Gesbert. “Balancing Egoism and Altruism on Interference Channel: The MIMO Case”. In: *Communications (ICC), 2010 IEEE International Conference on*. May 2010, pp. 1–5. DOI: 10.1109/ICC.2010.5501882.

[6] N.H. Mahmood et al. “A distributed interference-aware rank adaptation algorithm for local area MIMO systems with MMSE receivers”. In: *ISWCS*. Aug. 2014, pp. 697–701. DOI: 10.1109/ISWCS.2014.6933443.

[7] G.W.O. da Costa et al. “A Scalable Spectrum-Sharing Mechanism for Local Area Network Deployment”. In: *Vehicular Technology, IEEE Transactions on* 59.4 (May 2010), pp. 1630–1645. ISSN: 0018-9545. DOI: 10.1109/TVT.2009.2039361.

[8] Davide Catania et al. “The Potential of Flexible UL/DL Slot Assignment in 5G Systems”. In: *VTC Fall, IEEE 80th*. 2014.