Radio Resource Management for 5G Small Cells in Unpaired Spectrum

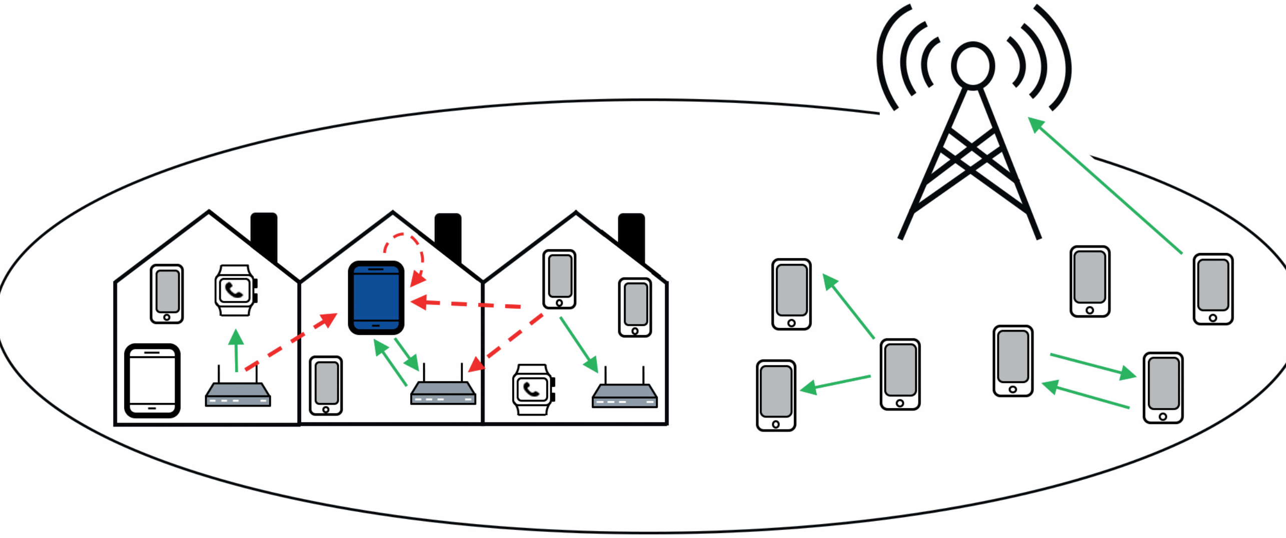
With Focus on Dynamic TDD and Full Duplex Technology

Gatnau, Marta

DOI (link to publication from Publisher):
10.5278/vbn.phd.engsci.00147

Publication date:
2016

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



RADIO RESOURCE MANAGEMENT FOR 5G SMALL CELLS IN UNPAIRED SPECTRUM

– WITH FOCUS ON DYNAMIC TDD AND FULL DUPLEX TECHNOLOGY

BY
MARTA GATNAU SARRET

DISSERTATION SUBMITTED 2016

AALBORG UNIVERSITY
DENMARK

Radio Resource Management for 5G Small Cells in Unpaired Spectrum

- With Focus on Dynamic TDD and Full Duplex Technology

Ph.D. Dissertation

Marta Gatnau Sarret

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

Dissertation submitted: August, 2016

PhD supervisor: Prof. Preben Mogensen
Aalborg University

Assistant PhD supervisors: Assoc. Prof. Gilberto Berardinelli
Aalborg University

Post. Doc. Nurul Huda Mahmood
Aalborg University

PhD committee: Associate Professor Tatiana K. Madsen (chairman)
Aalborg University

Professor Mikko Valkama
Tampere University of Technology

Head of Network Technology Lab dr. Peter Karlsson
Sony Mobile

PhD Series: Faculty of Engineering and Science, Aalborg University

ISSN (online): 2246-1248
ISBN (online): 978-87-7112-779-9

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

Printed in Denmark by Rosendahls, 2016

Abstract

Wireless communication is driving a networked society, where data is exchanged anytime, everywhere, between everyone and everything. The inclusion of smart devices and the explosion of new applications requiring mobile connectivity entails ambitious requirements for the next 5th Generation (5G) system: data rates up to 10 Gbps, latency as low as 1 millisecond and support for new use cases such as Device-to-Device (D2D) Communication.

Several solutions may be considered to achieve higher system capacity: using larger spectrum bands, increasing the number of cells and improving the spectral efficiency. The first approach has limited potential, since the spectrum is scarce and expensive at conventional frequency bands. The strategy of increasing the number of radio cells and reducing their size has proven to be promising. However, small cells serve a small set of users, creating sudden traffic imbalances between uplink and downlink directions. Thus, to optimally accommodate the instantaneous traffic demands, dynamic Time Division Duplex (TDD) is the most appropriate transmission mode. Improving the spectral efficiency is usually achieved by adding more antennas in a device. However, this approach may bring constraints in terms of space and cost. On the other hand, Full Duplex (FD) has been positioned as a potential technology for 5G by allowing a node to transmit and receive simultaneously in the same frequency band, thus theoretically improving the spectral efficiency and the system capacity.

Nevertheless, these approaches have a common drawback: increased and unpredictable interference. Denser networks lead to a larger inter-cell interference (ICI) and FD doubles the amount of interfering streams. Therefore, mechanisms to deal with the interference are essential for the 5G design.

The focus of this Ph.D study is on Radio Resource Management (RRM) for 5G small cells in unpaired spectrum, targeting dynamic TDD and FD nodes. In the first part of the Ph.D. thesis, link adaptation and recovery mechanisms are studied to overcome the drawbacks caused by dynamic TDD. System level simulations show that recovery mechanisms, specially at lower layers, are needed to meet the 5G targets. Moreover, results also show that interference

cancellation receivers highly alleviate the described interference challenges.

In the second part of this work, an RRM framework supporting FD is designed and studied. The joint performance of dynamic TDD and FD is evaluated and compared in terms of throughput gain and delay reduction. Two cases of FD are studied: when both the Base Station (BS) and the User Equipment (UE) are FD capable, and when only the BS can exploit FD. Intensive and detailed system level simulations are carried out, showing that in realistic environments the theoretical FD gain is significantly reduced, mainly due to the traffic profile and the increased ICI.

Finally, motivated by the strict 5G latency target, direct D2D discovery, i.e., the node awareness procedure prior to the communication phase, is studied. FD is considered as an attractive solution to achieve fast discovery and meet the latency target since a node can continuously receive while still broadcasting messages to its neighbors. The study shows that, in order to meet the latency target and get benefits from FD, interference cancellation receivers are a must. In that case, FD can reduce the latency significantly and reach the strict 5G D2D discovery target.

Resumé

Trådløs kommunikation stimulerer et netværkssamfund, hvor der konstant udveksles data alle vegne, imellem alle mennesker og alle ting. Inkluderingen af smarte apparater, og eksplosionen af nye applikationer, som kræver mobilforbindelse, medfører ambitiøse krav til det næste 5. generationssystem: Data hastigheder på op til 10 Gbps, forsinkelse ned til 1 millisekund, og support i form af ”Device-2-Device” (D2D) kommunikation til nye brugsmønstre.

Man kan overveje forskellige muligheder for at opnå højere kapacitet på systemet: Brug af større spektrum bånd, forøge antallet af celler, og forbedre spektre-effektiviteten. Det første tiltag giver et begrænset udbytte, da spektrummet er begrænset og dyrt i de konventionelle frekvensbånd. Det har vist sig at være mere lovende at vælge en strategi, hvor man forøger antallet af radioceller, og formindsker deres størrelse. Men små celler betjener en lille gruppe af brugere, hvilket skaber pludselige trafikale ubalancer imellem uplink og downlink retninger. Så for at imødekomme de øjeblikkelige behov for trafik bedst muligt, er ”Time Division Duplex” (TDD) den bedste måde at overføre på. Muligheden for at forbedre spektre-effektiviteten opnås normalt ved at tilføje flere antenner på apparatet, men denne tilgang medfører begrænsninger i form af volume og pris. På den anden side anses "full duplex" (FD) for at være en mulig teknologi indenfor 5G, ved at tillade at et adgangspunkt sender og modtager samtidigt i det samme frekvensbånd, således at spektre-effektiviteten og systemkapaciteten i teorien forbedres.

Disse tiltag har ikke desto mindre en ulempe: Forøget og uforudsigelig interferens. Tættere netværk fører til en større interferens imellem cellerne (ICI), og FD fordobler antallet af interferens strømninger. Derfor er det vigtigt for designet af 5G netværket, at finde mekanismer til at håndtere interferens. Fokus i dette PhD studie er lagt på ”Radio Ressource Management” (RRM) til små 5G celler i ”unpaired” spektrum, for at opnå dynamisk TDD og FD. I den første del af PhD afhandlingen har man undersøgt tilpasning af forbindelse og gendannelsesmekanismer, for at imødekomme ulemperne forårsaget af dynamisk TDD. Simuleringer på systemlag viser, at gendannelsesmekanismer – specielt på lavere lag – er nødvendige for at imødekomme 5G målene.

Desuden viser resultater også at interferens-elliminerende modtagere, som skal afbryde/modstå interferens, i høj grad dæmper de beskrevne udfordringer med interferens.

I den anden del af denne afhandling har man designet og undersøgt en RRM struktur, som understøtter FD. Man har evalueret og sammenlignet den samlede performance af dynamisk TDD og FD i forhold til forøgelse af data hastigheder og reduktion af forsinkelse. Der er undersøgt 2 scenarier af FD: Når både basestationen (BS) og brugerudstyret (UE) har FD – og når kun BS kan udnytte FD. Når der udføres intensive og detaljerede simuleringer på systemniveau viser det, at den teoretiske FD forøgelse reduceres markant i realistiske omgivelser, mest på grund af trafik-profilen og den forøgede ICI.

Til sidst undersøges proceduren for "adgangspunkt-bevidstheden" forud for kommunikationsfasen, motiveret af det strenge mål for forsinkelse på 5G, og det direkte D2D forskningsresultat. FD anses for at være en attraktiv løsning til at opnå et hurtigt forskningsresultat, og imødekomme målet for forsinkelse, da et adgangspunkt hele tiden kan lytte, alt imens det sender beskeder til sine naboer. Undersøgelsen viser, at for at imødekomme målet for forsinkelse, og for at kunne udnytte fordelene ved FD, skal man bruge modtagere som afviser interferens. I så fald kan FD reducere forsinkelsen betydeligt, og man kan opnå det strenge forskningsresultat for 5G D2D.

Translated by Dorthe Sparre, Aalborg University, Denmark.

Contents

List of Abbreviations

3G 3$^{\text{th}}$ Generation

3GPP 3rd Generation Partnership Project

4G 4$^{\text{th}}$ Generation

5G 5$^{\text{th}}$ Generation

5GPPP 5G Infrastructure Public Private Partnership

AM Acknowledged Mode

AP Access Point

ARQ Automatic Repeat Request

BLER Block Error Rate

BS Base Station

CP Cyclic Prefix

CQI Channel Quality Information

D2D Device-to-Device

DL Downlink

DMRS Demodulation Reference Symbol

eMBB enhanced mobile broadband

FD Full Duplex

FEC Forward Error Correction

FTP File Transfer Protocol

GP Guard Period

GPS Global Positioning System

HARQ Hybrid Automatic Repeat Request

HD Half Duplex

HSPA High Speed Packet Access

ICI inter-cell interference

IRC Interference Rejection Combining

KPI Key Performance Indicator

LTE Long Term Evolution

MAC Medium Access Control

MCS Modulation and Coding Scheme

MIMO Multiple Input Multiple Output

MMSE Minimum Mean Square Error

mMTC massive machine type of communication

MRC Maximum Ratio Combining

OFDM Orthogonal Frequency Division Multiplexing

OLLA Outer Loop Link Adaptation

OSI Open Systems Interconnection

PAPR Peak-to-Average Power Ratio

PHY Physical

PRB Physical Resource Block

QAM Quadrature Amplitude Modulation

RAT Radio Access Technology

RLC Radio Link Control

RRM Radio Resource Management

RTT Round Trip Time

SG scheduling grant

SI Self-Interference

SINR Signal to Noise plus Interference Ratio

SR scheduling request

TCP Transmission Control Protocol

TDD Time Division Duplex

TTI Transmission Time Interval

UDP User Datagram Protocol

UE User Equipment

UL Uplink

UM Unacknowledged Mode

URLLC ultra-reliable low latency communication

V2X Vehicle-to-Anything

Thesis Details

Thesis Title:	Radio Resource Management Solutions for Unpaired 5G Small Cells - With Focus on Dynamic TDD and Full Duplex Technology
Ph.D. Student:	Marta Gatnau Sarret
Supervisors:	Prof. Preben Mogensen, Aalborg University Assoc. Prof. Gilberto Berardinelli, Aalborg University Post. Doc. Nurul Huda Mahmood, Aalborg University

The work conducted in this thesis is the result of three years of research at the Wireless Communication Networks section, Department of Electronic Systems, Aalborg University, Denmark, in close cooperation with Nokia – Bell Labs. The work presented in this thesis was carried out in parallel with the mandatory courses and teaching/working obligations needed to obtain the Ph.D. degree.

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

[A] Marta Gatnau Sarret, Davide Catania, Frank Frederiksen, Andrea Fabio Cattoni, Gilberto Berardinelli, Preben Mogensen, "Improving Link Robustness in 5G Ultra-Dense Small Cells by Hybrid ARQ", *IEEE 11th International Symposium on Wireless Communications Systems (ISWCS), 2014.*

[B] Marta Gatnau Sarret, Davide Catania, Frank Frederiksen, Andrea Fabio Cattoni, Gilberto Berardinelli, Preben Mogensen, "Dynamic Outer Loop Link Adaptation for the 5G Centimeter-Wave Concept", *IEEE 21th European Wireless Conference (EW), 2015.*

[C] Marta Gatnau Sarret, Gilberto Berardinelli, Nurul H. Mahmood, Marko Fleischer, Preben Mogensen, Helmut Heinz, "Analyzing the Potential of Full Duplex in 5G Ultra-Dense Small Cell Networks" *EURASIP Journal on Wireless Communications and Networking - Special issue: Full-Duplex*

Radio: Theory, Design, and Applications. **Submitted.**

[D] Marta Gatnau Sarret, Gilberto Berardinelli, Nurul H. Mahmood, Preben Mogensen, "Can Full Duplex reduce the discovery time in D2D Communication?" *IEEE 13th International Symposium on Wireless Communications Systems (ISWCS), 2016.*

[E] Marta Gatnau Sarret, Gilberto Berardinelli, Nurul H. Mahmood, Preben Mogensen, "Providing Fast Discovery in D2D Communication with Full Duplex Technology" *Springer 9th International Workshop on Multiple Access Communications (MACOM), 2016.* **Submitted.**

In addition to the main papers, the following invention disclosures were submitted in Nokia:

- Marta Gatnau Sarret, Frank Frederiksen, "Channel Quality Indicator (CQI) reporting procedure for 5G systems".
- Marta Gatnau Sarret, Frank Frederiksen, "Interference Cancelling Aware (ICA) Channel State Indicator (CSI) reporting for 5G systems".
- Marta Gatnau Sarret, Frank Frederiksen, "Separation between protected and non-protected data for 5G systems".
- Marta Gatnau Sarret, Beatriz Soret, "Increasing the reliability by inter-cell signaling in full duplex systems".
- Marta Gatnau Sarret, Beatriz Soret, Gilberto Berardinelli, Nurul Huda Mahmood, Frank Frederiksen, "Fast discovery for autonomous devices".

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.

Preface

The research project was financed by Nokia Bell Labs and has been completed under the supervision of Professor Preben Mogensen (Aalborg University and Nokia Bell Labs), Associate Professor Gilberto Berardinelli (Aalborg University) and Post. Doc. Nurul Huda Mahmood (Aalborg University).

First of all, I would like to thank my supervisors, for their patience and for supporting and guiding me during this journey. I have learned something different from each of you, which helped me to grow both personally and as an engineer.

I would also like to express my gratitude to the persons who accepted to be part of my Ph.D. assessment committee, Doctor Peter Karlsson, Professor Mikko Valkamma and Associate Professor Tatiana Kozlova Madsen, for taking the time to read and assess this work.

Thanks to my parents, my brother and my sister-in-law, for always believing in me and making me feel close to you, even though we were at 2.500 kilometers. Your visits and your support was what kept me running.

Thanks to my boyfriend, for his patience and positivism. Your company and your ability to make me laugh makes everything easier.

Last but not least, I would like to thank all my colleagues and friends at Aalborg University and Nokia Bell Labs. I have really enjoyed all these years working with you, chating, laughing, having so much fun in Summer and Christmas parties. The memories that I have will remain always with me. Finally, special thanks my girls, Laura, Isa and Carmen, for being amazing friends and teach me so many things, not engineering related.

Marta Gatnau Sarret
Aalborg University
August 31, 2016

Part I

Introduction

Background and Thesis Overview

1 Mobile Traffic Growth

Over the last few years, a remarkable growth of data-enabled devices has been observed, which has led to an increase of the mobile data traffic. The amount of data carried by mobile networks moved from 10 gigabytes per month in 2010 to 3.7 exabytes per month in 2015 [1]. This outstanding growth has been caused by cellular network advances, moving from the 3$^{\text{th}}$ Generation (3G) to the 4$^{\text{th}}$ Generation (4G) of mobile systems; by the massive increase of smart devices; and by the emergence of new type of applications, such as social networks, whose number of users grows notably day by day. Only from 2014 to 2015, mobile data traffic grew 74%, and even though smart devices with minimum 3G connectivity only represented 36% of the total number of mobile devices and connections, they represented 89% of the mobile data traffic [1].

Such numbers are continuously growing, and the prediction is that global mobile data traffic will increase approximately by eight times between 2015 and 2020, as shown in Figure 1.1. Smart devices are forecasted to generate 98% of the mobile data traffic by 2020 [1], indicating that devices using wireless data connection are becoming more popular year by year.

To increase the network capacity and accommodate the future traffic demands, the strategies that must be considered, shown in Figure 1.2, are:

- Using larger frequency bands
- Enhancing the spectral efficiency
- Increasing the cell density

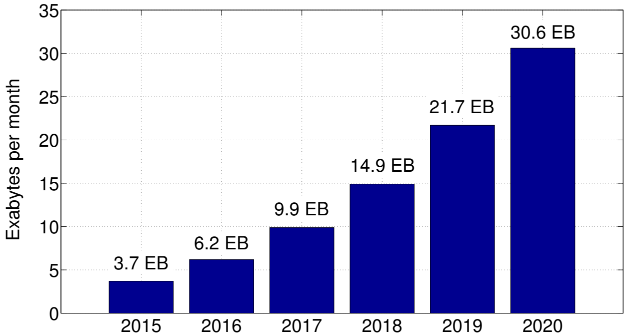


Fig. 1.1: Exabytes per Month of Mobile Data Traffic - Source: Cisco VNI 2016 [1].

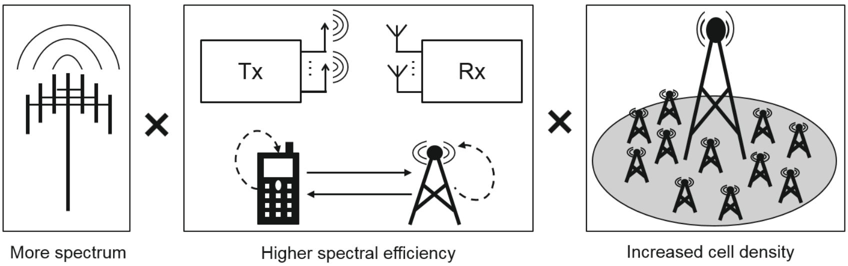


Fig. 1.2: Possible strategies to accommodate the future traffic demands.

The frequencies ranging from 3 to 30 GHz, also known as *centimeter-wave* frequencies, have the most attractive propagation characteristics. However, using this frequency range has limited potential because the spectrum is scarce and expensive, making it insufficient to deal with the future traffic growth [2]. On the other hand, higher frequency bands ranging from 30 up to 300 GHz, also referred as *millimeter-wave* frequencies, have been considered as an option to increase the network capacity [3, 4]. Research shows that there is potential for using such bands. However, given the different propagation conditions compared to the traditional *centimeter-wave* frequencies, large antenna arrays and beamforming techniques are required to overcome the pathloss. Orthogonal Frequency Division Multiplexing (OFDM) might not be the most suitable modulation for *millimeter-wave*. This modulation is very appropriate for bandwidth limited systems to multiplex users in frequency. However, since a large amount of spectrum is available in the *millimeter-wave* frequencies, users can be multiplexed in time. There are already some proposals on a modulation for these high frequencies [5], as well as discussions on channel modeling and beamforming design [3].

To enhance the spectral efficiency, the most common and used approach is to exploit Multiple Input Multiple Output (MIMO) antenna technology. This technique allows to overcome the effects of the signal multipath and fading in systems with limited bandwidth that require high throughput. The main drawback of such technology is that, as the number of device antennas grows, the constraints in terms of space and cost increase. Thus, it is unlikely that future devices will be equipped with a large number of antennas. Another option that could be considered is the emerging Full Duplex (FD) technology. It allows a device to transmit and receive simultaneously in the same frequency band, thus, theoretically, doubling the throughput compared to conventional Half Duplex (HD) systems. Currently, the main concern about FD technology is the cancellation of the Self-Interference (SI) power, which refers to the interference generated by the transmitted signal at the receiver end of the same node. The SI power must be highly attenuated in order to have an operational FD transceiver. Nowadays, the advances on transceiver design show that approximately 110 dB of SI cancellation are possible to be achieved [6], according to a certain bandwidth and power constraints, thus enabling FD as a potential technology to provide better spectral efficiency.

Increasing the cell density has been positioned as the other key solution to deal with the expected mobile traffic growth. It is achieved by deploying a large number of low-power base stations in scenarios with high traffic density, while traditional macro cell base stations provide basic service coverage. According to [2, 7], a massive deployment of small cells is fundamental for moving towards the desired increase in network capacity. Nowadays, the number of small cell deployments already exceeds the number of macro-cell base stations [8], and according to [9], most of the traffic is generated indoors. This statements reinforces the necessity of using indoor small cells with short ranges, even though the deployment of outdoor small cells has already been started by some operators to improve the coverage performance of the macro-cells [8].

2 Introduction to 5G

Current mobile systems may not be able to accommodate the expected traffic demands by 2020. Consequently, both the industry and the academia have started developing a concept for the future 5th Generation (5G) Radio Access Technology (RAT). This concept is debated as either an evolution or a revolution, i.e., a 5G system conceived as an improvement of the current 4G system or a disruptive new RAT, respectively. Ericsson believes that 5G will be an improvement of the current 4G system [10]. On the other hand, Huawei is confident that 5G will emerge as a new RAT [11]. The European Project 5G Infrastructure Public Private Partnership (5GPPP), composed of members from both academia and industry, also supports the idea of 5G being a new disruptive system [12]. Independently of the concept, the targets of the system are the same: peak data rates of 10 Gbps, extremely low latencies of 1 millisecond, coverage everywhere, and support for ultra-high reliability and ultra-low latency applications, such as Vehicle-to-Anything (V2X).

The 5G standardization process has already been started, given the fact that the system should be ready for deployment by 2020. Figure 1.3 shows the anticipated time line from research to standardization and commercialization. According to the defined time schedule, there will be a first trial of the basic 5G functionality by 2017 and a second one by 2018.

Fig. 1.3: 5G time line from research to standards - Source: Nokia 2015.

This dissertation is part of a large project in Nokia Bell Labs. The position of Nokia Bell Labs is to design 5G as a new disruptive RAT, since the limits of 4G technologies are being rapidly approached. The 5G vision of the company is depicted in Figure 1.4. The figure shows the use cases that 5G should support and the corresponding system targets. The first use case is enhanced mobile broadband (eMBB). The target is to improve the system capacity to provide peak data rates of 10 Gbps, whereas getting 100 Mbps must be possible whenever and wherever. The second use case is the ultra-reliable low latency communication (URLLC). Very strict requirements are envisioned, such as latencies below 1 millisecond and 99.999% reliability [13]. Finally, to support the third use case, namely massive machine type of communication (mMTC), the system should be optimized to handle a large amount of low cost devices with rigorous energy consumption requirements. It is important to notice that those requirements cannot be fulfilled simultaneously due to fundamental theoretical limits [14], but 5G aims at a flexible design which is able to cope with these requirements on a link basis.

In terms of spectrum band, there might be frequencies that are more suitable for certain type of applications than others. According to [15], frequencies below 6 GHz can provide support for the three use cases, eMBB, URLLC and mMTC. In the range between 3 and 40 GHz, the main focus is on eMBB, whereas URLLC could still be considered. Then, in frequencies above 40 GHz, only the eMBB use case can be addressed.

For the eMBB use case, where the main objective is to maximize the system capacity, an heterogeneous deployment is envisioned [16]. It refers to

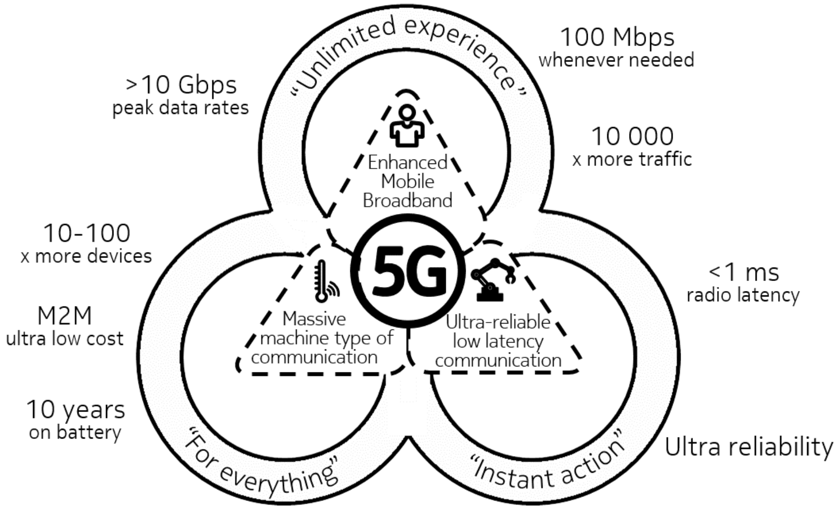


Fig. 1.4: 5G Nokia Vision with the corresponding targets - Source: Nokia 2015.

a combination of macro sites, outdoor micro cells and indoor small cells, as shown in Figure 1.5, which also indicates the ratio of cell density increase among the different radio cells. The target of the macro sites is to provide basic service coverage, to be able to ensure voice and data services. Outdoor micro cells play an important role to achieve high data rate everywhere, e.g., 10 Mbps with 90% coverage. Indoor small cells have the largest amount of traffic share [9], but the signal propagation is more sensitive to attenuations due to the floor and walls, and the transmit power is lowered due to health regulations. Since the indoor coverage may be limited, a large number of cells should be deployed to efficiently accommodate the wireless traffic.

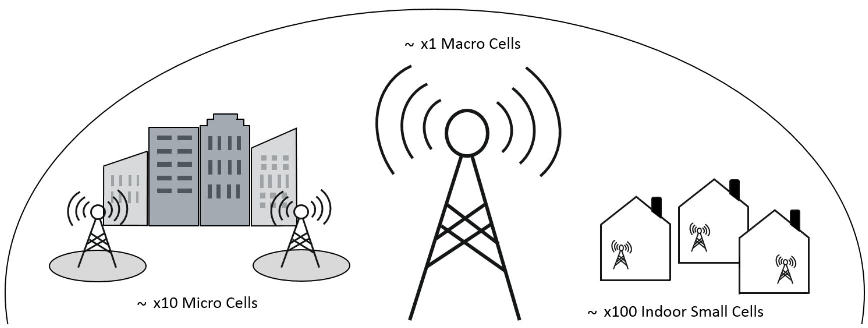


Fig. 1.5: Heterogeneous network, with deployment ratios according to [16].

Figure 1.4 shows that the 5G scope is very broad. This dissertation addresses two topics: the eMBB, in particular the performance of 5G small cells in unpaired spectrum; and the URLLC, more precisely how to provide fast discovery in Device-to-Device (D2D) communication. Next section presents the literature review on these topics.

3 Literature review

The general state of the art of the research areas of the project is divided in three sections: the performance of dynamic Time Division Duplex (TDD) in small cells, the potential of full duplex technology in small cells and how to accelerate the discovery procedure of D2D communication.

Small Cells and Dynamic TDD

The potential of small cells, i.e., of increasing the cell density for boosting the network capacity has already been proven [2, 7, 8]. As the cell size shrinks, the number of users served by the Access Point (AP) is reduced. Therefore, since the level of flow aggregation is low, an unpredictable traffic burtiness is expected. To optimally accommodate such traffic demands, the most appropriate solution would be to provide dynamism in terms of transmission direction, independently in each cell. This technique is commonly known as *dynamic TDD*.

The feasibility of dynamic TDD has been addressed in the literature. In 3rd Generation Partnership Project (3GPP) Release 12, 7 configurations with different Uplink (UL)/Downlink (DL) asymmetries are defined for Long Term Evolution (LTE), named LTE-TDD [17]. The proposed configurations and their corresponding switching times are shown in Figure 1.6. In [18], the evaluation of LTE-TDD with a switching point periodicity of 10 milliseconds is evaluated, showing that providing dynamism in terms of transmission direction gives higher gains than using a static allocation. In addition, the authors conclude that the system performance improves as the switching point periodicity reduces. On the other hand, the study also indicates that interference mitigation techniques are required, since dynamic TDD generates unpredictable interference.

In [19], an uncoordinated and greedy dynamic TDD scheme is evaluated. The study shows the necessity for interference mitigation and management techniques, given the interference variability caused by dynamic TDD. The authors use a beamsteering strategy to reduce the impact of the interference

Uplink-downlink configuration	Downlink-to-Uplink Switch-point periodicity	Subframe number									
		0	1	2	3	4	5	6	7	8	9
0	5 ms	D	S	U	U	U	D	S	U	U	U
1	5 ms	D	S	U	U	D	D	S	U	U	D
2	5 ms	D	S	U	D	D	D	S	U	D	D
3	10 ms	D	S	U	U	U	D	D	D	D	D
4	10 ms	D	S	U	U	D	D	D	D	D	D
5	10 ms	D	S	U	D	D	D	D	D	D	D
6	5 ms	D	S	U	U	U	D	S	U	U	D

Fig. 1.6: LTE-TDD configurations defined by 3GPP in release 12. Source: [17].

without inter-cell coordination. Results show that dynamic TDD outperforms an approach with a static allocation of the transmission direction. The work presented in [20] addresses the use of interference cancellation to mitigate the problem introduced by dynamic TDD. Results show that the use of techniques to reduce the interference are required to benefit from dynamic TDD. In that case, the gains are significant over a non-flexible transmission direction approach.

It is expected that even if advanced techniques to reduce the interference are used, interference cannot be avoided completely. To mitigate the impact of such residual interference, several techniques are investigated, e.g., power control, recovery mechanisms, rank adaptation or interference alignment. In [21], a hybrid solution which combines a dynamic and a static TDD approach with power control is presented. The authors show the benefits of the proposed solution to improve the system capacity. An inter-cell alignment based MIMO transmission scheme is presented in [22]. The proposed strategy combines interference cancellation and interference alignment according to the level of interference. Results show the feasibility of the proposed scheme in improving the system throughput. The well-known Hybrid Automatic Repeat Request (HARQ) recovery mechanism is still considered as an important part for the design of the future RAT. In [23], a novel signaling scheme for HARQ is proposed to deal with the feedback misalignment in the dynamic LTE-TDD system [17].

Other solutions are based on rank adaptation techniques, to improve the system performance by reducing the overall network interference. In [24], a cooperative distributed rank coordination scheme is presented. The goal of the mechanism it to maximize the network utility function instead of the performance of each individual node. The authors assume that cells are coordinated following a master-slave architecture, where nodes exchange limited amount of information. A power control strategy to mitigate the interference caused by dynamic TDD in LTE is presented in [25]. The proposed scheme aims at minimizing the AP-to-AP interference to improve the UL throughput performance. Results show that the proposed scheme can improve the UL direction at expenses of a minor degradation in the DL.

Small Cells and Full Duplex Technology

Given the unpredictable interference caused by dynamic TDD and the advances in the self-interference cancellation techniques [26, 27], FD technology has been positioned as a potential candidate for the future 5G system. For a FD node to be operational, the SI power should be highly attenuated. Recent results show that SI can be reduced approximately by 100 dB [27], which may suffice for considering FD a viable option, according to certain bandwidth and transmit power constraints. Nevertheless, the double throughput gain that FD promises may be compromised by several limitation, namely residual SI, increased interference and traffic asymmetry. The inter-cell interference (ICI) is increased with FD since the amount of interfering streams are doubled. In addition, to exploit FD it is required that data traffic is present on both link directions, i.e., UL and DL.

FD technology is rather expensive and energy consuming, and it could be that for the time being, it would only be implemented in the base stations [28–30]. However, the case where the User Equipments (UEs) are FD capable is also under study [31–33].

In [28], a scheduler that minimizes the UE-to-UE interference is presented, showing that FD can achieve a throughput gain over conventional HD. However, the penetration wall loss assumed in this work mitigates the impact of the inter-cell interference. Note that the penetration wall loss dictates the isolation between cells, thus defining the impact of the inter-cell interference on the system performance. The authors in [29] evaluate the performance of FD in ultra-dense small cells. Several user scheduling techniques alongside an optimal power allocation scheme are provided, showing that FD outperforms HD if a certain level of SI cancellation is achieved, e.g., 70 dB. The area spectral efficiency and the coverage in a small cell network are evaluated with HD and FD in [30]. Results show that FD achieves higher area spectral efficiency than HD, but at the expenses of higher outage, i.e., lower coverage. The authors indicate a compromise in the UL and DL transmit power to achieve the optimal performance.

The study presented in [31] evaluate FD in a dense deployment of small cells. The authors conclude that the SI dominates over the inter-cell interference if the SI cancellation capabilities are below 100 dB. In [32], results comparing the performance of MIMO HD and FD in a small cell scenario are discussed. Results show that under strong interference, HD may outperform FD due to MIMO spatial multiplexing gains. Under low interference conditions, FD can improve both the throughput and delay of the system. Finally, the authors in [33] study the achievable bit rate depending on the residual SI power and the inter-cell interference conditions. The work analyzes the Signal to Noise plus Interference Ratio (SINR) region where FD outperforms

HD. The authors conclude that in highly interfered scenarios, a strategy that switches between FD and HD provides the optimal system performance.

Note that most of the research work on the performance of full duplex assumes full buffer traffic. Therefore, the impact of the increased interference is studied, but the effect of the traffic asymmetry, the high layers protocols or the jointly repercussion of all the mentioned constraints are not addressed.

Device-to-Device Communication

D2D communication is an important topic in current radio research, since allowing devices to exchange data directly among them helps to offload the infrastructure. However, the discovery phase of such type of communication is not extensively. The discovery procedure refers to the detection of peer devices in the surroundings, which is required to establish a unicast or multicast communication. It can be performed with the aid of the infrastructure or autonomously by the devices. The former option generates additional control signaling and thus increases the overhead. For this reason, in this dissertation, the focus is on an autonomous discovery procedure, since allowing devices to exchange control messages directly can benefit the system in reducing the control overhead and the discovery time. According to the latest specifications for the next generation access technologies, the control plane latency target is set to 10 milliseconds [13].

A discovery mechanism for contention-based networks is proposed in [34]. It defines an initiator or group owner and the joiners. The former uses a dedicated channel to send discovery messages, whereas the latter send their requests in the channel specified by the initiator. The study shows that defining two device categories and using a dedicated discovery channel improves the discovery time. The authors in [35] propose a design for the discovery message to compensate for link failures and reduce the control overhead. Such message contains the identifier of the latest k devices that have successfully transmitted the discovery message. In [36], a proposal to dedicate a small portion of the resources to new devices appearing in the network is presented. By using this approach, the discovery message of the newcomers can be transmitted with a shorter delay, thus reducing their discovery time.

Given the fact that FD technology is viable, it can be used in D2D communication to overcome the HD limitation of not being able to listen to discovery messages from neighboring devices while transmitting the own message. The strategy presented in [37] to reduce idle slots and collisions uses FD to detect the activity of other devices. The study assumes that a device stops transmitting the discovery message when it has been discovered. However, this assumption may not be valid when mobility or dynamic (de)activation of devices is considered. In these kind of scenarios, transmitting the discovery message continuously is required. The authors in [38] combine FD with compressed sensing to overcome the drawbacks introduced by HD transmission mode and single packet reception. The presented results show that the discovery phase can be completed in a single time slot. However, the number of considered neighbors is rather limited and the design of the feedback mechanism for the transmission of acknowledgments is not addressed. Finally, [39] evaluates FD with directional antennas. Each device selects a transmission direction randomly at each time slot. This mechanism is effective when two devices coincide in the same direction. However, the time scale is too long to meet the requirements defined for future systems [13], given the proposed time setting and the fact that only two devices can discover each other in each transmission.

It is important to note that, independently of the transmission mode (HD or FD), the conventional approach is to define a fixed transmission probability or periodicity to transmit the discovery message. This approach may be adequate in networks where the control latency is not a limitation. However, in the future 5G RAT, strict requirements are posed in terms of control and data plane latencies. Consequently, an appropriate design for the system to be able to meet the defined requirements is needed.

4 Scope and Objectives of the Thesis

As described in the previous sections, one of the main research areas in mobile communication is the design of a 5G system to accommodate the expected wireless mobile traffic demands by 2020. Since the standardization process is still ongoing and there is not a clear design for the future RAT, this dissertation is based on the 5G Nokia Bell Labs vision. In particular, the eMBB and the URLLC use cases are addressed.

The eMBB use case is studied assuming the 5G indoor small cell concept, optimized for dense local area deployments. The concept is presented in [16, 40, 41] and is described in detail in the chapter entitled *5G Small Cell System Overview*. It was originally proposed as a TDD system, because of its flexibility and the possibility of exploiting unpaired frequency bands, operating at *centimeter-wave* frequencies. Techniques to provide synchronization in time and frequency are considered, as well as MIMO transceivers and interference suppression receivers. A key element to optimally accommodate the traffic and improve the spectral efficiency is to exploit the adaptability of TDD to provide flexibility in the transmission direction slot assignment, i.e., dynamic TDD. However, this strategy brings challenges in terms of interference variability, which affects the signal reception. Another approach to achieve higher spectral efficiency is to use FD technology, given its promise of doubling the throughput of traditional HD systems. Nevertheless, the constraints that compromise the gain that FD can provide over dynamic TDD have to be addressed.

The theoretical potential of FD technology has been studied in terms of throughput. However, such technology could also bring large benefits in terms of latency. For this reason, FD is studied for the URLLC use case, with focus on D2D communication. D2D allows devices to communicate directly among them, thus being an attractive solution for offloading the infrastructure and accommodating the massive increase of mobile devices and the explosion of new services. In D2D, before the data exchange procedure, devices must discover their peers. A critical challenge in this type of communication is how to accelerate the discovery process in order to meet the strict latency target of 10 milliseconds defined for next generation access technologies [13].

To address the described challenges, this thesis presents detailed performance evaluation and Radio Resource Management (RRM) solutions, i.e., all the strategies to efficiently utilize the radio resources, such as scheduling, decision of the transmission parameters or resource allocation. More precisely, these solutions are designed for the proposed 5G small cell system to improve the capacity and accommodate the exponentially growing mobile data traffic, and for direct D2D communication to reduce the discovery time. Figure 1.7 shows the scope of the thesis.

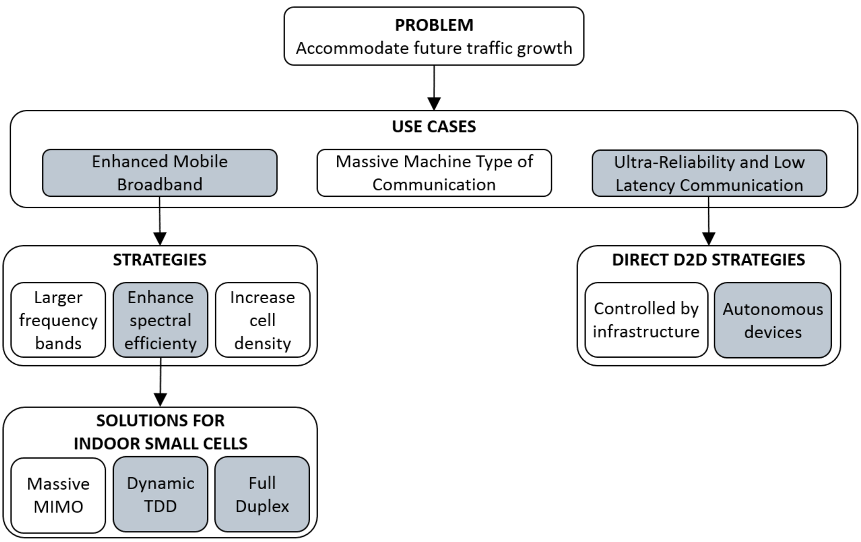


Fig. 1.7: Scope of the thesis. Study cases highlighted in gray.

The hypothesis (H) and research questions (Q) that are addressed in this dissertation are listed below:

H1 Dynamic TDD [42] has the potential of achieving higher system throughput. However, providing flexibility in the transmission direction on a time slot basis leads to larger SINR variability. Since the system performance is affected by the remaining interference that advanced receivers cannot suppress, stabilization mechanisms are required to further enhance the system.

Q1 Can fast recovery and link adaptation mechanisms deal with packet losses caused by such increased interference variability?

H2 Traditional HD systems allow a node either to transmit or receive at a given time instant. By using FD technology, which allows simultaneous transmission and reception in the same frequency band, the system throughput may, theoretically, be doubled compared to HD systems. However, in practice, there might be limitations in achieving the double throughput gain.

Q2 Which are these limitations and how do they impact the FD performance?

H3 Given the limited gains of FD in interference dominant scenarios, new applications where FD can bring larger benefits are studied. In direct D2D communication, devices must discover their neighbors prior to the actual data exchange. This procedure must be completed within the latency requirements, which is set to 10 milliseconds for 5G [13].

Q3 Is it possible to achieve such requirement? How can FD accelerate the device discovery procedure?

These research questions are respectively addressed in parts II, III and IV of the dissertation. Part II and III of the dissertation focus on investigating mechanisms to enhance the throughput and the delay of indoor ultra-dense small cell networks. In Part II, recovery mechanisms and link adaptation techniques have been studied to mitigate the drawback introduced by dynamic TDD, i.e., the increased interference variability [42]. Mechanisms such as HARQ and Outer Loop Link Adaptation (OLLA) are evaluated in the proposed 5G TDD system. In Part III, the emerging FD technology is considered as a potential candidate for indoor small cells. The performance of FD is studied considering dynamic TDD as the baseline system performance. Two types of FD are studied: when both the Base Station (BS) and the UE can exploit simultaneous transmission and reception (bidirectional FD), and only when the BS is FD capable (BS FD). The constraints that prohibit to achieve the theoretical FD gain are investigated and analyzed, individually and jointly. Part IV of the dissertation studies how to provide fast discovery in the direct D2D communication. The discovery procedure can be performed autonomously by the devices or with the involvement of the infrastructure. This work focuses on how to provide autonomous fast discovery for ad-hoc type of networks, where the infrastructure is not involved in the process. Further details about the system design will be given in Part IV of the dissertation. The potential of FD technology in meeting the latency requirements defined in [13] is investigated.

5 Research Methodology

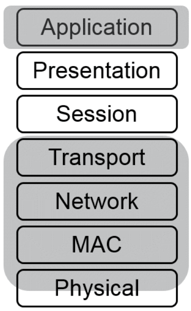
During the studies, a scientific methodology was employed to carry out the studies. Such methodology aims at dividing the research process into several phases:

1. Identify the research question to answer.
2. Literature review to get acquainted with the research area and related state of the art.
3. Design of solutions aiming at solving the identified problem.
4. Implementation and testing of the feature/mechanism through the simulation tool.
5. Collection of the simulation results and its corresponding dissemination in the form of a scientific article.

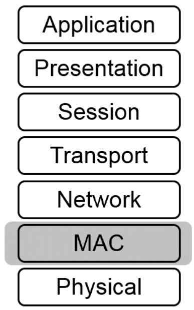
The target of this work is to provide a meaningful insight into how 5G would perform in 2020, assuming realistic conditions and a complete system, taking into account some of the envisioned technology components. Given the complexity of the work, it becomes unfeasible to perform the studies from a theoretical point of view. Consequently, the considered approach to evaluate the work of the thesis is through Monte Carlo system level simulations [43]. By using this methodology, meaningful results via comprehensive simulation campaigns can be extracted.

To evaluate the performance of the 5G system in ultra-dense small cells, a C++ event-driven simulator is used. It includes the implementation of the Open Systems Interconnection (OSI) protocol stack and the proposed 5G system design [41]. Figure 1.8a shows the layers which are implemented in the simulator. In the application layer, several traffic models are available, e.g., File Transfer Protocol (FTP) [44] and full buffer; in the transport layer, both User Datagram Protocol (UDP) and Transmission Control Protocol (TCP) are implemented; at the network level, the Acknowledged Mode (AM) and Unacknowledged Mode (UM) modes defined for Radio Link Control (RLC) [45] are included; in the Medium Access Control (MAC) layer, mechanism such as HARQ, scheduling and link and rank adaptation are available; finally, at the physical layer, two types of receivers are implemented, the Minimum Mean Square Error (MMSE)-Maximum Ratio Combining (MRC) receiver and the MMSE-Interference Rejection Combining (IRC) receiver. The modeling of the physical layer is out of the scope of this thesis, and both receiver models are extracted from [46]. SI cancellation is assumed ideal, and its implementation is not included in the simulator. The work carried out in this studies mainly focuses on schemes for the MAC layer, considering the interaction with the already existing transport and network layers. For more details regarding the simulator, please refer to the Appendices of [42].

The evaluation of the D2D study case is done using a Matlab simulator which uses some of the 5G concepts proposed in [41]. The focus of this study is on designing novel MAC protocols, as shown in Figure 1.8b. The simulator includes the ideal modeling of two receivers: a *basic* receiver that treats interference as noise, and one that ideally suppress a certain amount of interfering streams, based on the degrees of freedom in the MIMO antenna domain. Finally, the implementation of a SI cancellation model is not implemented since it is considered ideal.



(a) Ultra-dense small cells.



(b) D2D communication.

Fig. 1.8: Layers from the OSI protocol stack where the study focus on.

6 Contributions and publications

The main contributions of this study can be summarized as follows:

1. **Evaluating the potential of recovery mechanisms and link adaptation on the dynamic TDD 5G system.**

   Dynamic TDD allows to optimally accommodate the traffic, but also brings challenges in terms of interference variability, since the interfering source may change at each time slot. Although the use of interference suppression receivers has been proven to mitigate the drawback introduced by dynamic TDD, other mechanisms are required to cope with the unstable interference. HARQ and OLLA are introduced to the 5G system design to evaluate their potential in dealing with the unpredictable interference. Results show that recovery mechanisms at lower layers are beneficial for the system. On the other hand, OLLA shows limited gain but it does not come at the expenses of increased overhead or delay.

2. **Introducing a RRM module to provide support for FD technology into the envisioned 5G system.**

   Given the recent advances in the transceiver design, FD communication is positioned as a potential technology component for 5G systems. There are three different types of FD communication: when both the BS and the UE are FD capable (bidirectional FD); when only the BS is able to exploit such technology for data transmission (BS FD); and when only the BS is FD capable and uses the technology for relaying user data (relay FD). The module introduced allows for the evaluation of the first two FD types. It is composed by two submodules. Firstly, the optimal transmission direction is extracted per each node. Secondly, the decision of which user or pair of users is going to be scheduled is obtained. By separating functionalities, the number of operations to perform can be reduced and hence the processing time and the system complexity.

3. **Identifying the constraints that limit the achievable gain of FD in 5G small cell networks.**

   The performance of bidirectional FD and BS FD is evaluated in the 5G ultra-dense small cell system, considering dynamic TDD as the baseline system performance. An exhaustive analysis is carried out, identifying the constraints that affect the gain that FD can provide over dynamic TDD. In addition, an evaluation of the interaction of FD with higher protocols, such as TCP, is provided. The overall analysis is carried out assuming ideal self-interference cancellation, thus providing an upper bound of the achievable FD gain in indoor small cell networks.

The evaluation of the potential of FD in 5G indoor small cell networks shows that this technology brings limited gains over dynamic TDD. Consequently, identifying other applications where FD could provide larger improvements is required.

4. **Evaluating the potential of FD to accelerate the discovery procedure in D2D communication.**

   Direct D2D communication is becoming popular given its potential to offload traffic from the infrastructure and to allow for a faster communication among devices. However, prior to the exchange of data, devices must discover their peers. This procedure, known as device discovery, should be completed in 10 milliseconds to meet the requirements imposed by 3GPP [13]. The potential of FD technology to speed up such process and to satisfy the latency requirements is evaluated. The analysis shows that the use of interference cancellation receivers is a requirement to achieve low latency, and in that case, FD technology is able to further reduce the discovery time.

5. **Designing a protocol supporting FD technology to provide fast discovery in D2D communication.**

   Given the potential of FD in accelerating the discovery phase in D2D communication, a protocol to perform such procedure is proposed. A technique to estimate the number of neighboring devices alongside a signaling exchange mechanism to reduce the network interference are introduced. Furthermore, several approaches on how to use the signaled information are proposed and evaluated. The study compares the performance of HD transmission mode with FD communication, proving that FD technology achieves lower latency, achieving the 3GPP requirements in most of the evaluated scenarios.

In addition to these contributions, exhaustive development of the simulators was performed during the entire period of the studies. The C++ event-driven simulator used to extract the results for Part II and III of the dissertation was originally implementing the LTE system, and it was modified to include the envisioned 5G design [41]. The main contributions to the simulator development are listed below:

- **5G processing delay**
  The control and data transmission timing, processing delay and the proposed frame structure defined in the concept [41] were implemented in the simulator to extract realistic results.

- **HARQ and link adaptation**
  An asynchronous and non-adaptive implementation of the HARQ mechanism was implemented in the simulator. In terms of link adaptation, an scheme to obtain the Modulation and Coding Scheme (MCS) for transmission based on the average of the latest k SINR samples in logarithmic scale was implemented, as well as a static and a dynamic OLLA mechanism.

- **Full duplex transmission**
  The procedure for the signal transmission and reception in the simulator was modified to allow for both HD and FD transmission modes. The modification was included in both control and data parts, although in this dissertation such functionality is only exploited in the data part.

- **Multi-user functionality**
  The limitation of a single user per small cell did not allow for evaluating the FD case where the AP is the only node which can exploit simultaneous transmission and reception. The multi-user functionality affected the entire simulator, e.g., the generation of data flows or the RLC data aggregation and acknowledgment transmission.

- **User and transmission direction scheduler**
  Introducing the multi-user functionality required a mechanism to decide the transmission mode, the transmission direction in case of HD and the scheduled user(s).

- **Ideal RLC control transmission**
  The RLC acknowledgment transmission occupied a single Transmission Time Interval (TTI) and the whole bandwidth. Since the generated overhead was unrealistic, an ideal RLC acknowledgment transmission through the control channel was implemented.

The post-processing of the data obtained from the simulator was performed with Matlab to extract and present the results. Finally, Part IV of the thesis is carried out with an own Matlab simulator, developed entirely and only for the purpose of the last study in D2D communication.

The following publications were authored or co-authored in relation with this study:

Part I:

- 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. IEEE 11th International Symposium on Wireless Communications Systems (ISWCS), 2014.

- Dynamic Outer Loop Link Adaptation for the 5G Centimeter-Wave Concept. *Gatnau, Marta*; Catania, Davide; Frederiksen, Frank; Cattoni, Andrea Fabio; Berardinelli, Gilberto; Mogensen, Preben. IEEE 21th European Wireless Conference, 2015.

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

- 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. IEEE 81st Vehicular Technology Conference (VTC) Spring, 2015.

Part II:

- 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) Fall, 2015.

- Can Full Duplex Boost Throughput and Delay of 5G Ultra-Dense Small Cell Networks?. *Gatnau, Marta*; Berardinelli, Gilberto; Mahmood, Nurul Huda; Mogensen, Preben. IEEE 83rd Vehicular Technology Conference (VTC) Spring, 2016.

- Impact of Transport Control Protocol on Full Duplex Performance in 5G Networks. *Gatnau, Marta*; Berardinelli, Gilberto; Mahmood, Nurul Huda; Mogensen, Preben. IEEE 83rd Vehicular Technology Conference (VTC) Spring, 2016.

- On the Potential of Full Duplex Performance in 5G Ultra-Dense Small Cell Networks. *Gatnau, Marta*; Fleischer, Marko; Berardinelli, Gilberto; Mahmood, Nurul Huda; Mogensen, Preben; Heinz, Helmut. IEEE European Signal Processing Conference (EUSIPCO), 2016.

- Analyzing the Potential of Full Duplex in 5G Ultra-Dense Small Cell Networks. *Gatnau, Marta*; Berardinelli, Gilberto; Mahmood, Nurul Huda;

Fleischer, Marko; Mogensen, Preben; Heinz, Helmut. EURASIP Journal on Wireless Communications and Networking - Special issue: Full-Duplex Radio: Theory, Design, and Applications, 2016. **Status:** submitted.

- Evaluating Full Duplex Potential in Dense Small Cells from Channel Measurements. Berardinelli, Gilberto; Assefa, Dereje; Mahmood, Nurul Huda; *Gatnau, Marta*; Sørensen, Troels Bundgaard; Mogensen, Preben Elgaard. IEEE Vehicular Technology Conference (VTC) Spring, 2016.

Part III:

- Can Full Duplex reduce the discovery time in D2D Communication?. *Gatnau, Marta*; Berardinelli, Gilberto; Mahmood, Nurul Huda; Soret, Beatriz; Mogensen, Preben. IEEE 13th International Symposium on Wireless Communications Systems (ISWCS), 2016.
- Providing Fast Discovery in D2D Communication with Full Duplex Technology. *Gatnau, Marta*; Berardinelli, Gilberto; Mahmood, Nurul Huda; Soret, Beatriz; Mogensen, Preben. Springer 9th International Workshop on Multiple Access Communications (MACOM), 2016. **Status:** submitted.
- Radio resource management for V2V discovery. Soret, Beatriz; *Gatnau, Marta*; Kovács, István Z.; Martín-Vega, Francisco Javier; Berardinelli, Gilberto; Mahmood, Nurul Huda. IEEE Vehicular Technology Magazine, 2016. **Status:** submitted.

7 Thesis Outline

This Ph.D. dissertation is composed as a collection of papers. For this reason, the contributions and outcomes extracted from this work are presented in the form of accepted or submitted conference and journal articles, included in Parts II, III and IV of the thesis. Each of these parts consist of a general overview, i.e., the problem definition, the assumptions and the main findings. Part I of the dissertation provides an introduction to the study and a description of assumed the 5G system design. Part V exposes the main conclusions and future research work. Finally, Part VI corresponds to the Appendix, which includes the conference articles not included in the main body of the dissertation. The outline of the thesis is as follows:

Part I. Includes the current chapter and a detailed description of the proposed 5G indoor small cell system.

Part II. Analyses the benefits of using recovery mechanisms and link adaptation schemes to mitigate the drawbacks introduced by dynamic TDD. This part is composed by papers A and B.

Part III. Studies the potential of FD technology in improving the 5G system performance, in terms of throughput and delay. For the purpose of the reader, only paper C is included in this part, since it provides a complete overview of the four accepted conference papers included in Part VI Appendix.

Part IV. Examines the requirements to provide fast discovery in direct D2D communication and achieve the strict latency target requirements. Papers D and E compose this part.

Part V. Concludes the Ph.D. dissertation and provides recommendations for future research work.

Part VI. Includes additional first-author papers complementing the dissertation.

A number of abbreviations are used on this thesis, spelled out in their first appearance. We recommend the reader to use the List of Abbreviations included after the Table of Contents while reading the thesis. A reference list is included at the end of each chapter and paper. Note that the references which are cited within each chapter are not necessarily represented by the same number in all chapters.

References

[1] Cisco, "White paper: Cisco visual networking index: Global mobile data traffic forecast update, 2015-2020," Feb 2016.

[2] J. G. Andrews, S. Singh, Q. Ye, X. Lin, and H. Dhillon, "An overview of load balancing in hetnets: old myths and open problems," *IEEE Wireless Communications*, vol. 21, no. 2, pp. 18–25, Apr 2014.

[3] A. Ghosh *et al.*, "Millimeter-wave enhanced local area systems: A high-data-rate approach for future wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1152–1163, Jun 2014.

[4] S. Larew, T. A. Thomas, M. Cudak, and A. Ghosh, "Air interface design and ray tracing study for 5G millimeter wave communications," in *IEEE Globecom Workshops (GC Wkshps)*, Dec 2013, pp. 117–122.

[5] M. Cudak, A. Ghosh, T. Kovarik, R. Ratasuk, T. A. Thomas, F. W. Vook, and P. Moorut, "Moving towards mmwave-based beyond-4G (B-4G) technology," in *IEEE 77th Vehicular Technology Conference (VTC Spring)*, Jun 2013, pp. 1–5.

[6] M. Heino *et al.*, "Recent advances in antenna design and interference cancellation algorithms for in-band full duplex relays," *IEEE Communications Magazine*, vol. 53, no. 5, pp. 91–101, 2015.

[7] A. G. Gotsis, S. Stefanatos, and A. Alexiou, "Ultra dense networks: The new wireless frontier for enabling 5G access," *CoRR*, vol. abs/1510.05938, 2015. [Online]. Available: http://arxiv.org/abs/1510.05938

[8] L. Lopez-Perez, M. Ding, H. Claussen, and A. H. Jafari, "Towards 1 Gbps/UE in cellular systems: Understanding ultra-dense small cell deployments," *IEEE Communications Surveys Tutorials*, vol. 17, no. 4, pp. 2078–2101, Fourthquarter 2015.

[9] M. Paolini. Mobile data move indoors. [Online]. Available: http://www.senzafiliconsulting.com/Blog/tabid/64/articleType/ArticleView/articleId/59/Mobile-data-move-indoors.aspx

[10] Ericsson, "White paper: 5G radio access," Feb 2015.

[11] Huawei, "White paper: 5G: New air interface and radio access virtualization," Apr 2015.

[12] 5G Infrastructure Public Private Partnership, "White paper: 5G vision," Feb 2015.

[13] 3rd Generation Partnership Project TR 38.913 V0.3.0, "Study on scenarios and requirements for next generation access technologies," Mar 2016.

[14] B. Soret, P. Mogensen, K. I. Pedersen, and M. C. Aguayo-Torres, "Fundamental tradeoffs among reliability, latency and throughput in cellular networks," in *IEEE Globecom Workshops (GC Wkshps)*, Dec 2014, pp. 1391–1396.

[15] 3rd Generation Partnership Project TSG-RAN WG1 #85 R1-165011, "OFDM based waveform for 5G new radio interface," May 2016.

[16] P. Mogensen *et al.*, "B4G local area: high level requirements and system design," in *IEEE Globecom Workshops (GC Wkshps)*. IEEE Press, 2012.

[17] 3rd Generation Partnership Project TS 36.211 V12.8.0, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation (Release 12)," Dec 2015.

[18] J. Liu, R. Fan, H. Wang, J. Liu, and F. Wang, "Dynamic TDD Testbed and Field Measurements," in *IEEE 83rd Vehicular Technology Conference (VTC Spring)*, May 2016, pp. 1–5.

[19] H. Celik and K. W. Sung, "On the Feasibility of Blind Dynamic TDD in Ultra-Dense Wireless Networks," in *IEEE 81st Vehicular Technology Conference (VTC Spring)*, May 2015, pp. 1–5.

[20] A. K. Gupta, M. N. Kulkarni, E. Visotsky, F. W. Vook, A. Ghosh, J. G. Andrews, and R. W. Heath, "Rate analysis and feasibility of dynamic TDD in 5G cellular systems," in *IEEE International Conference on Communications (ICC)*, May 2016, pp. 1–6.

[21] H. Lee and D. H. Cho, "Combination of Dynamic-TDD and Static-TDD Based on Adaptive Power Control," in *IEEE 68th Vehicular Technology Conference (VTC Fall)*, Sep 2008, pp. 1–5.

[22] C. Na, X. Hou, and H. Jiang, "Interference alignment based dynamic TDD for small cells," in *IEEE Globecom Workshops (GC Wkshps)*, Dec 2014, pp. 700–705.

[23] C. Wang, X. Hou, A. Harada, S. Yasukawa, and H. Jiang, "HARQ Signalling Design for Dynamic TDD System," in *IEEE 80th Vehicular Technology Conference (VTC Fall)*, Sep 2014, pp. 1–5.

[24] B. Clerckx, H. Lee, Y. J. Hong, and G. Kim, "A Practical Cooperative Multicell MIMO-OFDMA Network Based on Rank Coordination," *IEEE Transactions on Wireless Communications*, vol. 12, no. 4, pp. 1481–1491, Apr 2013.

[25] H. Takahashi, K. Yokomakura, and K. Imamura, "A Transmit Power Control Based Interference Mitigation Scheme for Small Cell Networks Using Dynamic TDD in LTE-Advanced Systems," in *IEEE 79th Vehicular Technology Conference (VTC Spring)*, May 2014, pp. 1–5.

[26] S. Hong *et al.*, "Applications of self-interference cancellation in 5G and beyond," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 114–121, February 2014.

[27] M. Heino *et al.*, "Recent advances in antenna design and interference cancellation algorithms for in-band full duplex relays," *IEEE Communications Magazine*, vol. 53, no. 5, pp. 91–101, May 2015.

[28] S. Goyal, P.Liu, S. S. Panwar, R. A. DiFazio, R. Yang, and E. Bala, "Full duplex cellular systems: will doubling interference prevent doubling capacity?" *IEEE Communications Magazine*, vol. 53, no. 5, pp. 121–127, May 2015.

[29] I. Atzeni, M. Kountouris, and G. C. Alexandropoulos, "Performance evaluation of user scheduling for full-duplex small cells in ultra-dense networks," in *22th European Wireless Conference (EW)*, May 2016, pp. 1–6.

[30] S. Goyal, C. Galiotto, N. Marchetti, and S. Panwar, "Throughput and coverage for a mixed full and half duplex small cell network," in *IEEE International Conference on Communications (ICC)*, May 2016, pp. 1–7.

[31] H. Alves, C. H. M. de Lima, P. H. J. Nardelli, R. D. Souza, and M. Latva-aho, "On the average spectral efficiency of interference-limited full-duplex networks," in *9th International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*, Jun 2014, pp. 550–554.

[32] W. Zhou and K. Srinivasan, "Sim+: A simulator for full duplex communications," in *International Conference on Signal Processing and Communications (SPCOM)*, July 2014, pp. 1–6.

[33] A. C. Cirik, R. Wang, and Y. Hua, "Weighted-sum-rate maximization for bidirectional full-duplex MIMO systems," in *Conference on Signals, Systems and Computers (ASILOMAR)*, Nov 2013, pp. 1632–1636.

[34] H. B. Li and R. Miura, "Discovery protocol for peer aware communication networks," in *IEEE 82nd Vehicular Technology Conference (VTC Fall)*, Sep 2015, pp. 1–5.

[35] Y. Kang and X. Xu, “A fast ad-hoc neighbor discovery algorithm based on a friend trust mechanism,” in *IEEE 12th International Conference on Networking, Sensing and Control (ICNSC)*, Apr 2015, pp. 69–74.

[36] S. Park and S. Choi, “Expediting D2D discovery by using temporary discovery resource,” in *IEEE Global Communications Conference (GLOBECOM)*, Dec 2014, pp. 4839–4844.

[37] G. Sun, F. Wu, X. Gao, and G. Chen, “PHED: Pre-Handshaking Neighbor Discovery Protocols in full duplex wireless ad hoc networks,” in *IEEE Global Communications Conference (GLOBECOM)*, Dec 2012, pp. 584–590.

[38] X. Yang, X. Wang, R. Yang, and J. Zhang, “Full-duplex and compressed sensing based neighbor discovery for wireless ad-hoc network,” in *IEEE Wireless Communications and Networking Conference (WCNC)*, Mar 2015, pp. 1643–1647.

[39] R. Xu, J. Li, L. Peng, and Y. Ye, “A neighbor discovery algorithm for full duplex ad hoc networks with directional antennas,” in *The 27th Chinese Control and Decision Conference (2015 CCDC)*, May 2015, pp. 5235–5240.

[40] P. Mogensen *et al.*, “5G small cell optimized radio design,” in *IEEE Conference and Exhibition Globecom*. IEEE, 2013, pp. 111–116.

[41] ——, “Centimeter-wave concept for 5G ultra-dense small cells,” in *IEEE 79th Vehicular Technology Conference (VTC Spring)*, May 2014.

[42] D. Catania, “Performance of 5G small cells using flexible TDD,” Ph.D. dissertation, Department of Electronic Systems, Aalborg University, Oct 2015. [Online]. Available: http://vbn.aau.dk/files/225798349/PhD_Thesis_Davide_Catania_Performance_of_5G_Small_Cells_using_Flexible_TDD.pdf

[43] R. Coates, G. Janacek, and K. Lever, “Monte Carlo Simulation and Random Number Generation,” *IEEE Journal on Selected Areas in Communications*, vol. 6, no. 1, pp. 58–66, Jan 1988.

[44] 3rd Generation Partnership Project TR 36.814 V9.0.0, “Further advancements for E-UTRA physical layer aspects (Release 9),” Mar 2010.

[45] 3rd Generation Partnership Project TR 36.322 V8.0.0, “Technical specification group radio access network; evolved universal terrestrial radio access (E-UTRA) radio link control (RLC) protocol specification,” Dec 2007.

[46] F. Tavares, “Interference-robust air interface for 5G small cells: Managing intercell interference with advanced receivers and rank adaption,” Ph.D. dissertation, Department of Electronic Systems, Aalborg University, Apr 2015. [Online]. Available: http://vbn.aau.dk/files/225797874/PhD_Thesis_Fernando_Tavares.pdf

5G Small Cell System Overview

This chapter provides a detailed description of the envisioned 5G small cell concept, since it is the reference system used in this dissertation to design features for the next generation system. Such concept has been under research study since 2012. Therefore, some aspects of the system described in the following sections may differ from standardization, but they are still assumed in this dissertation.

Several articles providing a general description of the envisioned 5G small cell system are available [1–3]. Nokia Bell Labs' concept includes technology components such as an optimized frame structure, interference suppression receivers, MIMO antenna technology and dynamic UL and DL transmission scheduling.

1 Physical layer overview

The envisioned concept, tailored for indoor small cells, is applied in the *centimeter-wave* frequency band, ranging from 3.4 to 4.9 GHz, and assumes a system bandwidth of 200 MHz. The preferred waveform is OFDM, for both UL and DL. The reasons for the modulation choice are its cost-effectiveness for coping with channel multipath and its straightforward extension to MIMO [4, 5]. OFDM presents some drawbacks, such as a large Peak-to-Average Power Ratio (PAPR) and out-of-band emissions, sensitivity to hardware impairments, and the overhead generated by the Cyclic Prefix (CP). The CP should, ideally, match the channel delay spread, and it should also compensate for propagation delays if timing advance techniques are not used. In the targeted local area small cell scenario, the CP can be very short given the shorter expected delay spread and propagation delay, thus mitigating one of the OFDM drawbacks. For the other shortcomings of OFDM, several solutions have been proposed by the research community [4, 6]. For the purpose of this dissertation, we assume that the above limitations of OFDM can be effectively addressed. The proposed duplexing operational mode is TDD, given its cost-effectiveness, flexibility and possibility of exploiting unpaired frequency bands. Table 2.1 provides a summary of the 5G numerology assumed in this dissertation.

Table 2.1: Assumed 5G numerology according to [2]

Parameter	Value
Carrier Bandwidth [MHz]	200
Subcarrier spacing [kHz]	60
Symbol length [μs]	16.67
OFDM symbols per frame	14
Frame duration [ms]	0.25
CP duration [μs]	1
Guard Period (GP) duration [μs]	0.89
Number of GPs	3
Overhead (CP+GP) [%]	6.67
HARQ Processes	4

An important aspect to address is the time synchronization at OFDM symbol level among the APs. It allows to support efficient interference coordination schemes, enable advanced transmission techniques to boost the system throughput, and ease the support for advanced receivers, among others. The level of accuracy should be in the order of a fraction of μs [3]. A Global Positioning System (GPS) reference for the APs could achieve such level of accuracy. However, for indoor scenarios it is not the most appropriate choice, since the wall penetration loss worsens its performance. An attractive option to provide sufficient accuracy in indoor scenarios is to use distributed algorithms. A distributed synchronization solution is proposed in [7, 8], based on a periodically exchange of beacon messages. The concrete strategy to provide synchronization for 5G systems is a topic for further studies. In this dissertation, ideal time and frequency synchronization is assumed.

2 Optimized frame structure

The introduction of a new frame structure is intended for:

- Achieving low latencies and a shorter Round Trip Time (RTT), to meet the

1 millisecond latency target of 5G

- Supporting advanced interference management techniques and interference suppression receivers
- Enabling pipeline-processing at the receiver
- Reducing the energy consumption, specially at the UE
- Reducing the necessity for big data storage blocks, for cost-effectiveness

The proposed 5G small cell frame structure is shown in Figure 2.1. Each TTI has a duration of 0.25 milliseconds. The control and data planes are separated in time to enable cost-effective pipeline-processing at the receiver and reduce the UE power consumption. Each TTI is composed of 14 OFDM symbols. The first two are allocated to the DL and UL control information, respectively. The remaining symbols are for the data part, which is dedicated entirely to DL or UL. The first symbol within the data part contains the Demodulation Reference Symbol (DMRS). This symbol is used for channel estimation. It allows to estimate simultaneous channel responses from multiple interfering APs and/or UEs in the data part, assuming the use of orthogonal reference sequences. Having the DMRS symbol allocated at the beginning of the data part, and the data part entirely being DL or UL, makes the interference stable within a frame. This allows for an effective usage of interference suppression receivers. The GP is inserted at each potential switch of the communication link direction to adjust to the on-off power transition. Finally, in the frequency domain, a certain number of Physical Resource Blocks (PRBs) divide the system bandwidth, in order to support frequency coordination and/or reuse techniques.

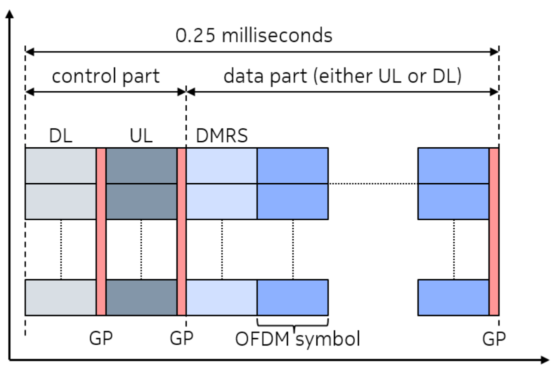
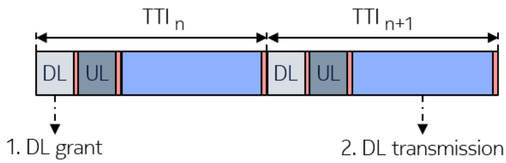


Fig. 2.1: Proposed 5G frame structure

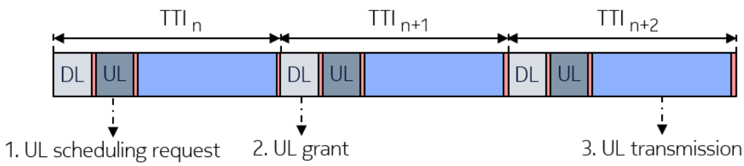
3 Enhanced scheduling and HARQ

The introduction of a novel frame structure allows to provide enhancements in terms of transmission scheduling and HARQ.

Figure 2.2a shows the DL scheduling procedure. In the DL control symbol, the AP sends a scheduling grant (SG), indicating the necessary parameters for the transmission that will occur in the next TTI. The UL scheduling procedure is shown in Figure 2.2b. In the UL control symbol, the UE sends a scheduling request (SR), containing information related to buffer size, HARQ feedback, Channel Quality Information (CQI) and transmission rank. Note that in case of UL, an extra TTI is needed compared to DL, since the UE has to send a request for transmission to the AP. There is always one TTI delay between the scheduling and the corresponding transmission, assuming that the processing time does not exceed 0.25 milliseconds.



(a) DL scheduling procedure



(b) UL scheduling procedure

Fig. 2.2: Scheduling procedures in DL and UL. Frequency domain and DMRS symbol not depicted for simplicity.

The latency target of 1 millisecond can be achieved using the proposed frame structure, the described scheduling and an appropriate HARQ design. Figure 2.3 shows a representation of the DL RTT, which requires 0.75 milliseconds. Such RTT is fixed because the control part in each radio frame offers the possibility of transmitting the HARQ feedback. In addition, the figure shows the maximum number of parallel HARQ processes, which has been set to 4, as specified in Table 2.1. The decision on the low number of HARQ processes is because the 5G small cell concept is designed to be cost-effective, and the most expensive component in the baseband chip is the buffer.

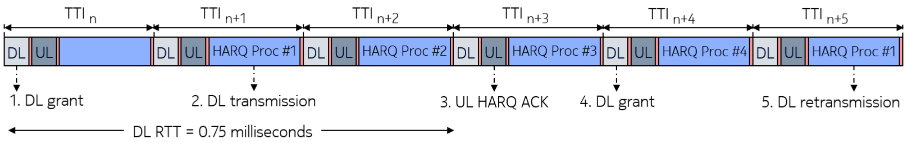


Fig. 2.3: RTT definition and DL HARQ procedure. Frequency domain and DMRS symbol not depicted for simplicity.

4 Radio resource management design

RRM refers to all the strategies to efficiently utilize the radio resources. In this section we describe several RRM techniques that have been studied for the design of our system.

As described in the introductory chapter *Background and Thesis Overview*, one approach to increase the system capacity is to consider MIMO antenna technology. The proposed concept assumes that all the nodes will be equipped with 4×4 MIMO transceivers and interference suppression receivers.

The design of the RRM should be differentiated for the control and the data planes, since the requirements are different. In the control plane, multiple APs and UEs can simultaneously transmit control information, thus posing challenges on how to provide high probability of successful decoding. Therefore, the transmissions should be robust enough to allow APs and UEs to decode all the received packets, for example by using a fixed and robust MCS and a low transmission rank. Techniques such as frequency reuse should also be considered to improve the decoding probability. The design of the control channel for the envisioned 5G small cell concept is still under study. In this dissertation, the control plane is assumed ideal, i.e., the control messages are always decoded.

In the data part, the design of the RRM can be more aggressive. The system can be more permissive with interference, allowing for bounded link failures since HARQ works as an insurance for unsuccessful packet decoding. Consequently, it can be worth to have higher payoffs in system throughput at the expenses of loosing a few packets. The RRM design should consider the use of fast link adaptation, adaptive transmission rank, dynamic transmission scheduling and advanced receivers [9, 10].

Dynamic TDD

In local area small cells, a large and unpredictable traffic burstiness is expected, since only a few users are served by the AP resulting in a low level of flow aggregation. Consequently, providing cell independent dynamism in terms of transmission direction would be the most adequate solution. In that sense, a dynamic TDD approach is considered for the 5G small cell design [9]. This means that the transmission direction can change at every TTI, so the traffic can flexible accommodated. In [9], two algorithms to decide the optimal transmission per TTI are presented. The first scheme is the *delay fairness based*, which uses the head of line delay and the buffer size information to optimally allocate the data. The second proposed algorithm is the *load fairness based*, which decides the optimal transmission direction based on the instantaneous traffic share between DL and UL, and the previous slot allocation decisions.

Even though the interference is stable at each TTI, providing such dynamism creates challenges in terms of SINR variability, given the short TTI duration and the possibility of changing the transmission direction at each TTI. To exemplify the interference challenge, Figure 2.4 presents a SINR trace of a static TDD scheme (1 DL slot followed than 1 UL slot) and a flexible one (50% random probability of selecting DL or UL). The scenario is a 10×2 small cell grid, with 1 AP and 1 UE randomly deployed in each cell, both equipped with 4×4 MIMO transceivers and IRC receivers. The transmission rank is fixed to 1, i.e., each node has 3 degrees of freedom in the antenna domain to suppress interference. A finite buffer traffic model is used [11], and the offered load leads to a channel occupancy of 60%, approximately. From the figure we can observe that the flexible scheme shows a larger SINR variability than the static scheme. In particular, the SINR range with the flexible scheme is doubled compared to the static. Since the flexible TDD approach is random, it would correspond to an upper limit of the expected SINR variability from the proposed small cell system. This unpredictable interference caused by dynamic TDD poses challenges on the design of the link and rank adaptation mechanisms. Given the fact that interference can change at each TTI, it could be that the available estimation of the MCS and transmission rank is not the most appropriate for the TTI where they are applied to. Consequently, it may lead to link failures and HARQ retransmissions. Figure 2.5 depicts the considerations that should be taken into account for the RRM design in order to mitigate the impact of the SINR variability, which are described in the following sections.

Interference variability protections: advanced receivers and HARQ

The first and essential protection is to use advanced receivers with interference suppression capabilities. The 5G small cell scenario is interference-limited, and one of the important challenges is to overcome the inter-cell interference. An interference-robust air interface is presented in [12], using advanced receivers as the main interference mitigation technique. Our pro-

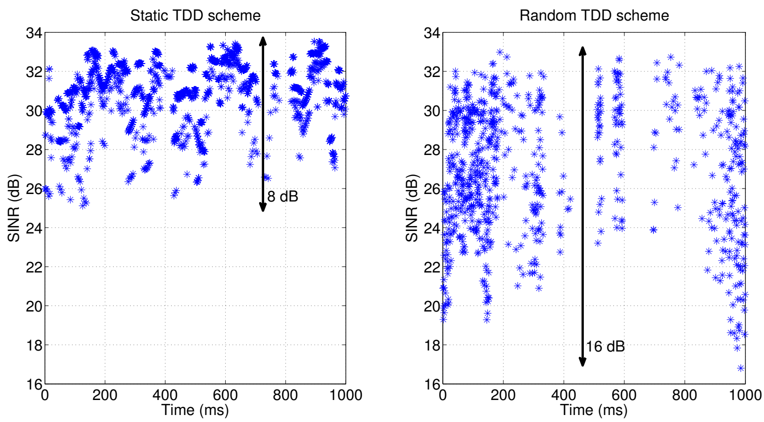


Fig. 2.4: Example of the SINR variability introduced by a static and a flexible TDD scheme.

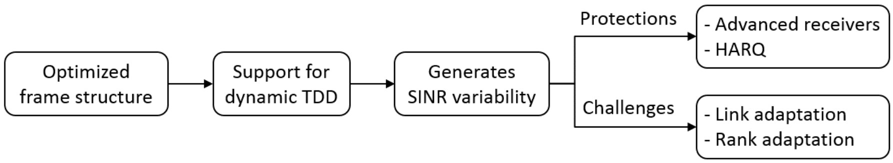


Fig. 2.5: From the optimized frame structure to the consideration for the RRM design.

posed frame structure is very convenient for supporting advanced receivers. APs and UEs will send their reference sequences within the DMRS symbol, located right before the data part, allowing the receiver to estimate cross-link interference and improve the quality of the received signal. The IRC receiver [13] exploits the degrees of freedom in the antenna domain to suppress interference. Therefore, its suppression capabilities will be larger when fewer desired streams are received. Assuming a node with N receive antennas and M desired streams, the IRC receiver is able to suppress at most N - M interfering streams. The other receiver which is considered is successive interference cancellation [14]. This type of receiver aims at decoding both desired and interfering streams, so the latter can be subsequently canceled. The main drawback of successive interference cancellation is its higher complexity compared to IRC, and the fact that it requires transmission information, such as MCS and transmission rank, to successfully decode the interfering signals. To obtain this information, a node should read the control channel or use blind decoding, which increases the complexity. Nevertheless, successive interference cancellation may be a convenient solution to suppress intra-stream interference, since the required information is available, whereas IRC takes care of suppressing the inter-cell interference [12]. The assumption in this dissertation is the use of IRC receivers without intra-stream interference cancellation.

The second protection technique is HARQ. This mechanism, performed between Physical (PHY) and MAC, was first introduced in High Speed Packet Access (HSPA) [15] and is still a effective retransmission technique to provide fast recovery. System level simulation results presented in Part II of the dissertation show that HARQ is able to improve the 5G system throughput, in particular the outage performance (5^{th} percentile). More details about the HARQ performance in 5G ultra-dense small cells is given in Part II of the dissertation.

Interference variability challenges: fast link and rank adaptation

In terms of link adaptation, we foresee the inclusion of the 256-Quadrature Amplitude Modulation (QAM) modulation and a fast mechanism that extracts the MCS and the transmission rank according to updated channel measurements. Such measurements are obtained in every frame that a node is receiving data. In that sense, the proposed frame structure is very efficient, since in each TTI, the UE channel information can be sent within the UL control symbol. Figure 2.6 shows a representation of the link adaptation procedure in UL. The figure shows the shortest delay, which is three TTIs, between the estimation of the MCS and rank and its corresponding application. Note that it represents the shortest delay because the channel information used could be estimated in previous TTIs if the possibilities for data reception are low. The figure shows the UL direction because the UE has to send the channel information to the AP, thus adding extra delay compared to DL.

A technique that complements fast link adaptation is OLLA. It follows the same principle as the outer loop power control mechanism introduced in HSPA [16]. OLLA is intended for coping with systematic errors, e.g., from the CQI report. If such errors are too high, OLLA brings limited gains to the system [17]. Part II of the dissertation presents system level results of the OLLA performance in the 5G indoor small cell concept. Simulations show that OLLA is able to improve the system performance when the SINR variability is low. Further details of the proposed OLLA technique for the 5G small cell concept are discussed in Part II of this dissertation.

To exploit MIMO antenna technology efficiently, an adaptive scheme that selects the most appropriate transmission rank must be studied. In [12], two rank adaptation techniques are presented, the *selfish rank adaptation* and the *victim-aware rank adaptation*. The former algorithm tries to maximize individually the Shannon throughput, and it is performed independently in each node. In the latter scheme, each node selects the most appropriate transmission rank taking into account the amount of interference that it may generate to the neighboring cells. The principle of the algorithm is to apply higher taxation to the higher ranks in order to reduce the overall network interference [9].

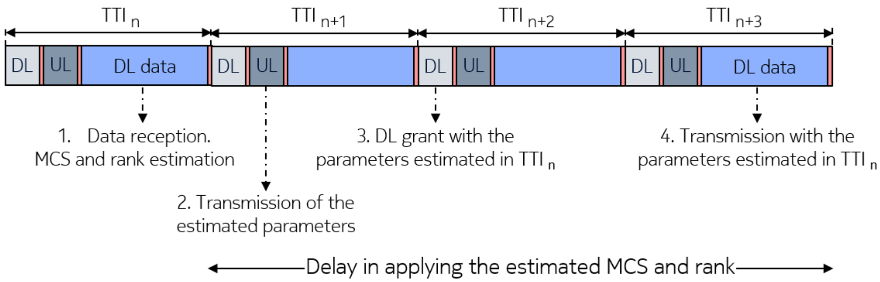


Fig. 2.6: Description of the delay between the estimation of the transmission parameters and its corresponding application.

5 Energy consumption

Last but not least, the UE battery life will be an important Key Performance Indicator (KPI) for 5G, since it will affect the user satisfaction. As previously mentioned, our concept has also been designed to improve the energy consumption [18]. TDD does not require the use of a duplexer, being able to lower the consumed power. In addition, the optimized frame structure allows a device to reduce the active time given its short duration and the separation of the control and the data planes. Consequently, the sleeping time, i.e., the period of time where the UE powers OFF unused components to save energy, can be increased. According to [19], the sleep mode consumes at least 40 times less power than the active mode. The study is based on LTE measurements, but the results are expected to be similar for 5G UEs. Finally, reducing the number of parallel HARQ processes also allows for saving energy, since the buffer is a highly energy consuming element of the baseband chip. Detailed energy consumption studies of 5G small cells is out of the scope of this dissertation.

References

[1] P. Mogensen *et al.*, "B4G local area: high level requirements and system design," in *IEEE Globecom Workshops (GC Wkshps)*. IEEE Press, 2012.

[2] ——, *5G small cell optimized radio design*. IEEE, 2013, pp. 111–116.

[3] ——, "Centimeter-wave concept for 5G ultra-dense small cells," in *IEEE 79th Vehicular Technology Conference (VTC Spring)*, May 2014.

[4] H. Holma and A. Toskala, *LTE for UMTS - OFDMA and SC-FDMA Based Radio Access*. Wiley, 2009.

[5] L. Hanzo, M. Munster, B. Choi, and T. Keller, *OFDM and MC-CDMA for Broadband MultiUser Communications, WLANs and Broadcasting*. John Wiley - IEEE Press, 2003.

[6] G. Berardinelli, F. M. L. Tavares, T. B. Sørensen, P. Mogensen, and K. Pajukoski, *Zero-tail DFT-spread-OFDM signals*. IEEE Press, 2013, pp. 229 – 234.

[7] G. Berardinelli, F. M. L. Tavares, N. H. Mahmood, O. Tonelli, A. F. Cattoni, T. B. Sørensen, and P. Mogensen, *Distributed synchronization for Beyond 4G Indoor Femtocells*. IEEE, 2013, pp. 1–5.

[8] G. Berardinelli, F. M. L. Tavares, O. Tirkkonen, T. B. Sørensen, and P. Mogensen, "Distributed initial synchronization for 5G small cells," in *IEEE 79th Vehicular Technology Conference (VTC Spring),*. IEEE, 2014, pp. 1–5.

[9] D. Catania, "Performance of 5G small cells using flexible TDD," Ph.D. dissertation, Department of Electronic Systems, Aalborg University, October 2015.

[10] N. H. Mahmood, D. Catania, M. Lauridsen, G. Berardinelli, P. E. Mogensen, F. M. L. Tavares, and K. Pajukoski, *A Novel Centimeter-Wave Concept for 5G Small Cells*. C R C Press LLC, 4 2016.

[11] 3rd Generation Partnership Project TR 36.814 V9.0.0, "Further advancements for E-UTRA physical layer aspects (Release 9)," March 2010.

[12] F. M. L. Tavares, "Interference-robust air interface for 5G small cells: Managing inter-cell interference with advanced receivers and rank adaption," Ph.D. dissertation, Department of Electronic Systems, Aalborg University, April 2015.

[13] J. Choi, *Optimal combining and detection: Statistical signal processing for communications*. Cambridge University Press, 2010.

[14] S. Verdu, *Multiuser Detection, 1st edn.* Cambridge University Press, 1998.

[15] 3rd Generation Partnership Project TS 21.101 V5.14.0, "Technical specifications and technical reports for a UTRAN-based 3GPP system," December 2009.

[16] 3rd Generation Partnership Project TS 25.427 V6.8.0, "UTRAN iub/iur interface user plane protocol for DCH data streams (release 6)," December 2006.

[17] C. Rosa, D. L. Villa, C. U. Castellanos, F. D. Calabrese, P. H. Michaelsen, K. I. Pedersen, and P. Skov, "Performance of fast AMC in E-UTRAN uplink," in *IEEE International Conference on Communications (ICC)*, May 2008, pp. 4973–4977.

[18] M. Lauridsen, G. Berardinelli, F. M. L. Tavares, F. Frederiksen, and P. E. Mogensen, "Sleep modes for enhanced battery life of 5G mobile terminals," in *IEEE 83rd Vehicular Technology Conference (VTC Spring)*, 2016.

[19] M. Lauridsen, L. Noël, T. B. Sørensen, and P. Mogensen, "An empirical lte smartphone power model with a view to energy efficiency evolution," *Intel Technology Journal*, vol. 18, no. 1, pp. 172–193, March 2014.

Part II

Dynamic TDD in 5G Small Cells

Overview

1 Problem Description and Assumptions

Dynamic TDD refers to the strategy that provides flexibility, on a time slot basis, in terms of transmission direction. This part of the dissertation focuses on improving the capacity of the proposed 5G indoor small cell concept [1], assuming the use of dynamic TDD. Research work shows that increasing the cell density and providing flexible allocation of the transmission direction has the potential of improving the system throughput [2–4]. However, dynamic TDD comes at the expense of larger interference variability [4]. The first protection against interference is the use of advanced receivers, such as IRC receivers, which enhance the system performance by suppressing a part of the interference [5].

Nevertheless, there may still be remaining interference affecting the signal reception. Under unpredictable interference, the link adaption mechanism may not be able to match the appropriate MCS, causing unsuccessful decoding. Consequently, there is the need for stabilization mechanisms to cope with the residual interference and the link failures that interference variability may cause.

In order to improve the link quality and overcome the drawback of dynamic TDD, the following stabilization mechanisms should be considered:

- **Interference suppression receivers**. As explained in Part I of the dissertation, two types of receiver are considered, MRC and IRC. The latter has been proved to have potential in stabilizing the network interference by using the degrees of freedom in the antenna domain to cancel interference [5]. Therefore, there is a trade-off between the number of desired streams that have to be decoded and the interference suppression capabilities.
- **Interference management techniques**. It can be performed by considering a distributed or centralized controller, or employing techniques such as

frequency reuse or rank adaptation. The former case generates larger overhead since extra information needs to be shared. Consequently, the delay is increased. Using frequency reuse techniques reduces the transmission rate and requires planning. The rank adaptation technique presented in [4] shows potential on increasing the system throughput without the need of information exchange or planning.

- **Recovery mechanisms**. This category includes HARQ at the PHY-MAC layer, Automatic Repeat Request (ARQ) at the RLC layer and TCP at the transport layer. This part of the dissertation focuses on HARQ, since it is the fastest recovery mechanism used to improve the physical layer robustness. Its first appearance was in HSPA [6] and it has also been adopted by LTE [7]. It retransmits erroneous packets, combining ARQ and Forward Error Correction (FEC). On top of HARQ, the other mentioned mechanisms could also be considered.

- **Link adaptation schemes**. Optimizing the link adaptation procedure to anticipate the interference pattern would help at reducing the link failures. This work studies the potential of OLLA in improving the system. OLLA is a technique that aims at compensating for systematic errors based on a defined Block Error Rate (BLER) target and the received HARQ feedback. Upon the reception of positive or negative acknowledgments, an offset is updated. This offset is applied in the MCS domain. In this way, if a negative acknowledgment is received, OLLA selects a more robust MCS at the expenses of a lower amount of transmitted data, and vice versa in case of a positive acknowledgment.

Those mechanisms are not exclusive but cumulative. In this work, the potential of HARQ in improving the system is firstly studied. Afterwards, OLLA is added on top to analyze the benefits that this technique may bring in terms of throughput. In this work, the goal of the OLLA mechanism is shifted from incorrect reports to variable interference.

The HARQ and OLLA mechanisms for the 5G indoor small cell concept [1] are studied in this part of the dissertation under the following assumptions:

- **Ultra-dense indoor small cell network**
  The studied scenario corresponds to a 10×2 grid of 10×10 m^2 small cells, containing one AP and one UE deployed randomly. The wall penetration loss is set to 5 dB, according to the Winner II model [8]. The transmit power of both the AP and the UE is set to 10 dBm, given health regulation constraints and the small cell size [1].

- **Full buffer traffic with random DL/UL slot allocation**
  This assumption creates defiant situations to HARQ and OLLA. A random slot allocation, with 50% probability of selecting one of the two link directions, DL or UL, represents a challenging interference variability situation, since the transmission direction is totally unpredictable, and hence the interference. In addition, having full buffer traffic means that the interference is persistent. It represents a challenge for HARQ because there is always new data to be transmitted, while retransmissions may also be waiting in the buffer.

- **Fast link adaptation**
  The mechanism to extract the MCS is based on the average in logarithmic value of the latest 5 post-IRC SINR samples. The MCS is extracted from a SINR-to-MCS mapping table, according to a BLER target of 10%. In terms of transmission rank, Paper A uses a fixed approach, whereas Paper B assumes the use of the rank adaptation technique described in [4]. Note that, in each TTI, the UE sends the latest estimate of MCS and transmission rank within the UL control channel, even if such UE does not get a scheduling grant.

- **4$\times$4 MIMO transceiver with IRC receiver**
  The IRC receiver uses the degrees of freedom in the antenna domain to suppress interference [5]. Therefore, the interference suppression capabilities of the receiver, and consequently the system performance, depends on the number of desired streams. The impact of varying the transmission rank is addressed in Paper A and Paper B.

The results presented in the included articles compare the system performance with and without the considered mechanisms. The focus is mainly on the final throughput and the number of retransmissions required to successfully decode a data packet.

2 Main Findings

As indicated in the introduction, HARQ is the second protection tier to deal with the increased interference variability caused by dynamic TDD. System level results show that HARQ can increase the system throughput, including the 5th percentile or outage users. The benefits that this mechanism brings depend on the interference level, which is related to the transmission rank, i.e., the number of desired streams. The larger the number of desired streams, the lower are the capabilities of the IRC receiver to suppress interference. Thus, higher gains are obtained from HARQ.

At high SINR variability, triggering HARQ is the most appropriate option because slower recovery mechanisms, such as ARQ at the RLC layer or TCP at the transport layer, would lead to a decrease in throughput and an increase in delay.

The study reported in Paper A assumes full buffer traffic, as mentioned previously. Therefore, the delay statistics are highly impacted by the buffering time, i.e., the time that data needs to wait in the buffer to be transmitted. For this reason, Figure 3.1 shows the session throughput and the packet delay with FTP traffic [9] for fixed rank 1 and 2, since in the studied dense small cell network, those are the most selected ranks [4]. In this case, the statistics are obtained using the *delay fairness based* [4] scheme with an the average channel occupation is 65%. From the figure we can observe that both throughput and delay are improved with HARQ. The gain from HARQ is larger with rank 2 because the interference is stronger since IRC has one degree of freedom less compared to rank 1 to suppress interference.

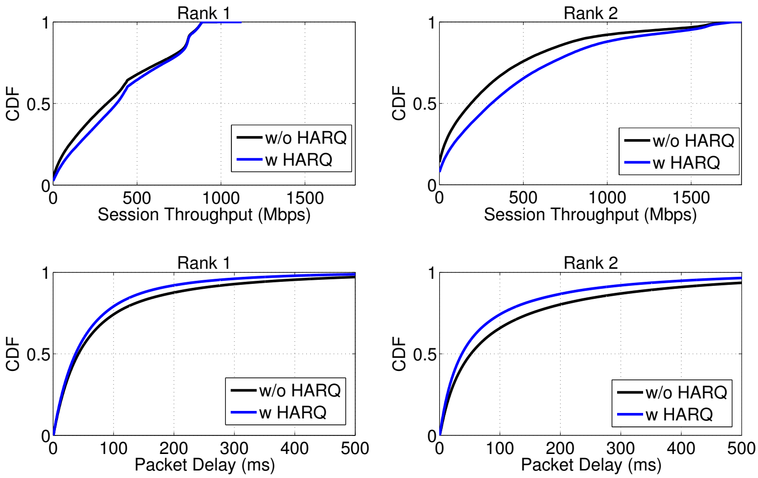


Fig. 3.1: Throughput comparison with and without HARQ, with FTP traffic and for fixed rank 1 and 2.

OLLA provides limited gain because the interference variability in the targeted study is high. System level results show that the most appropriate approach is to use a low BLER target in the OLLA algorithm, in the order of 5% to 10%. By using such conservative approach, a failure in the transmission of a packet (negative HARQ acknowledgment) is more penalized than a successful transmission (positive HARQ acknowledgment), in terms of MCS selection. Thus, the number of retransmissions can be reduced at the expense of an eventual throughput reduction.

Paper B shows that selecting a low transmission rank provides better system results because the post-IRC interference and thus the interference variability are lower. Since OLLA was designed to cope with systematic errors, such as delay and quantization of measurements, having unpredictable interference prohibits this technique to provide large gains. Finally, the paper also shows that providing adaptability in the OLLA technique in terms of BLER target, according to the SINR conditions, brings benefits to the system.

3 Included articles

The following two articles compose the main body of this part of the thesis.

Paper A: Improving Link Robustness in 5G Ultra-Dense Small Cells by Hybrid ARQ

This article investigates the most suitable operational mode for HARQ given the requirements of the envisioned 5G system. The proposed mode is an asynchronous and adaptive HARQ procedure to have full flexibility on handling retransmissions. The performance of HARQ is evaluated in a non-adaptive manner due to simulator constraints, using two dynamic TDD algorithms as well as varying the transmission rank. System level results show that HARQ is able to improve the final throughput, as well as the performance of the outage users. The impact of OLLA is not considered in this work.

Paper B: Dynamic Outer Loop Link Adaptation for the 5G Centimeter-Wave Concept

This article focuses on the performance of OLLA in the 5G small cell system, where HARQ is enabled. A novel algorithm based on the SINR characteristics is presented and compared against an OLLA scheme from the literature. An analysis of the SINR variability as a function of the number of desired streams is also included. System level results show that the gain that OLLA provides is limited, since this mechanism requires stable interference to compensate only for systematic errors [10]. However, this is not the situation in the considered scenario. It is important to note that even if the gain is limited, it does not come at the expense of an increased system complexity.

References

[1] P. Mogensen *et al.*, "Centimeter-wave concept for 5G ultra-dense small cells," in *IEEE 79th Vehicular Technology Conference (VTC Spring)*, May 2014.

[2] J. G. Andrews, S. Singh, Q. Ye, X. Lin, and H. Dhillon, "An overview of load balancing in hetnets: old myths and open problems," *IEEE Wireless Communications*, vol. 21, no. 2, pp. 18–25, April 2014.

[3] A. G. Gotsis, S. Stefanatos, and A. Alexiou, "Ultra dense networks: The new wireless frontier for enabling 5G access," *CoRR*, vol. abs/1510.05938, 2015. [Online]. Available: http://arxiv.org/abs/1510.05938

[4] D. Catania, "Performance of 5G small cells using flexible TDD," Ph.D. dissertation, Department of Electronic Systems, Aalborg University, October 2015.

[5] F. Tavares, "Interference-robust air interface for 5G small cells: Managing inter-cell interference with advanced receivers and rank adaption," Ph.D. dissertation, Department of Electronic Systems, Aalborg University, April 2015.

[6] 3rd Generation Partnership Project TS 25.321, "Technical specification group radio access network; medium access control (MAC) protocol specification," September 2008.

[7] 3rd Generation Partnership Project TS 36.321, "Technical specification group radio access network; evolved universal terrestrial radio access (E-UTRA); medium access control (MAC) protocol specification," December 2008.

[8] W. W. I. N. Radio, "WINNER II channel models," February 2008.

[9] 3rd Generation Partnership Project TR 36.814 V9.0.0, "Further advancements for E-UTRA physical layer aspects (Release 9)," March 2010.

[10] C. Rosa, D. L. Villa, C. U. Castellanos, F. D. Calabrese, P. H. Michaelsen, K. I. Pedersen, and P. Skov, "Performance of fast AMC in E-UTRAN uplink," in *IEEE International Conference on Communications (ICC)*, May 2008, pp. 4973–4977.

Paper A

Improving Link Robustness in 5G Ultra-Dense Small Cells by Hybrid ARQ

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

The paper has been published in the
IEEE 11th International Symposium on Wireless Communications Systems (ISWCS), 2014.

Abstract

A new 5^{th} generation (5G) radio access technology is expected to cope with an estimated factor of $\sim$x1000 growth in mobile data traffic in the upcoming years. Such system will be optimized for a massive uncoordinated deployment of small cells, where autonomous operation of the individual nodes may bring unpredictable and fast varying link quality. In this paper, Hybrid Automatic Repeat Request (HARQ) is studied as a solution to cope with such unpredictability. An operational mode of HARQ for our 5G system definition is proposed, and its performance is evaluated for two different scheduling options. Simulation results confirm the capability of HARQ of improving the final throughput and solving outage problems, with limited impact on the end delay.

1 Introduction

According to the latest forecasts, mobile data traffic is expected to grow by a factor of $\sim$x1000 by 2020 [1]. Current radio standards such as Long Term Evolution (LTE)/LTE-Advanced (LTE-A) are not designed to cope with such traffic growth, and eventual upgrades may prohibitively increase their complexity. This justifies the design of a novel and disruptive 5th Generation (5G) Radio Access Technology (RAT). In [2] [3] we have presented our vision on such 5G system, optimized for local area, where a massive and uncoordinated deployment of small cells with limited coverage is foreseen. Our envisioned 5G system aims at significantly outperforming LTE/LTE-A, reducing at the same time its complexity. Time Division Duplexing (TDD) is foreseen as the preferred operational mode, since a high traffic with variable data rates in Uplink (UL) and Downlink (DL) is expected; in that sense, a flexible UL/DL time slot assignment would provide an agile support to accommodate such type of traffic [3].

Though the instantaneous decision on the transmission direction is intended to boost the performance of the high layers of the protocol chain, it introduces uncertainty in terms of expected Signal-to-Noise plus Interference Ratio (SINR). This may negatively affect the link adaptation and cause incorrect reception. The transmission rank, i.e., number of Multiple Input Multiple Output (MIMO) spatial streams, may also affect the network performance since it may generate interference signatures among neighbour cells that cannot be suppressed by computationally feasible receivers. Furthermore, errors coming from measurements, quantization or signalling have also a negative impact on the link quality, and likewise the outdated measurements due to the fast varying interference conditions. Hence, the use of a stabilization mechanism to improve the link quality is required. Such stabilization can rely on coordinated operation among neighbour cells, or to simple retransmission mechanisms.

Hybrid Automatic Repeat and Request (HARQ) has extensively been adopted in previous RATs [4] [5] to improve the physical layer robustness [6]. It retransmits erroneous packets, combining Automatic Repeat Request (ARQ) and Forward Error Correction (FEC). It was first introduced in High Speed Packet Access (HSPA) systems and has been then adapted to next generation RATs. In [7], HARQ performance is improved for HSPA by considering the Channel Quality Indicator (CQI) from retransmissions as well as from new transmissions. In [8], an algorithm which optimizes HARQ for MIMO is proposed, targeting HSPA$^+$. HARQ has also been adopted in LTE. An implementation to meet Quality of Service (QoS) requirements is proposed in [9], while its throughput under QoS constraints is analysed in [10].

In this paper, the usage of HARQ as a solution to deal with the SINR uncertainties in 5G systems is analyzed. We propose a preferred HARQ operational mode given our 5G design. The effective potential of HARQ in 5G is addressed through system level simulations, providing throughput and delay results.

The paper is structured as follows. Section II describes our main design choices for 5G systems, also motivating the necessity of a HARQ mechanism. Section III presents our envisioned HARQ mechanism for 5G. Section IV describes the simulator used for this study and discusses the performance results. Finally, Section V concludes the paper and describes the future work.

2 Envisioned 5G concept

In [2], a clean slate approach for a 5G RAT optimized for local area is proposed. The targets should be superior than the LTE/LTE-A ones: peak data rates in order of 10 Gbps, Round Trip Time (RTT) of 1 ms, and wake-up time from "inactive" to "active" of 10 ms.

Optimization of a number of technology components would be needed. The 1 ms RTT forces the definition of a new frame structure of 0.25 ms, shown in figure A.1, defined as the transmission time interval (TTI). Control and data planes are separated in time, to enable efficient pipeline processing at the receiver. The first Orthogonal Frequency Division Multiplexing (OFDM) symbol is dedicated to DL control, the second one to UL control, and the re-

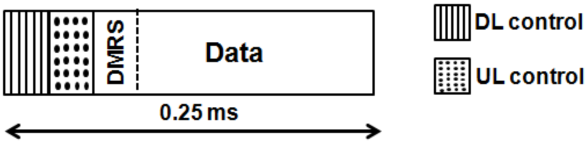


Fig. A.1: Proposed 5G frame structure

maining are dedicated to data, including the Demodulation Reference Signal (DMRS) for enabling channel estimation at the receiver. The data part can carry either UL or DL traffic, and the type of carried data can vary at every frame. This means that every 0.25 ms, a new and independent decision is taken to determine which type of data is going to be sent in the next TTI.

TDD is considered as the preferred operational mode due to its cost-effectiveness and the possibility of exploiting unpaired frequency bands. Further, it allows to maximize the similarity of all links (traditional UL and DL, self-backhauling and device-to-device). To achieve the desired network capacity, the use of MIMO antenna technology is assumed, being 4x4 the default configuration.

A massive deployment of indoor cells working on a dedicated spectrum is foreseen as a solution to cope with the aforementioned target. In such type of networks, the number of User Equipments (UEs) is typically small, leading to a low level of flow aggregation and imbalances between UL and DL traffic. To deal with such traffic imbalances, and exploiting the adaptability of TDD, we envision a totally flexible UL/DL slot assignment, breaking with the transmission patterns of LTE/LTE-A [6]. This means that the data direction in slot TTI_n is independent from both $\mathrm{TTI}_{n\text{-}1}$ and TTI_{n+1}. As a starting point, the decision will be taken autonomously by each access point (AP). Moreover, we assume that all nodes in the network (UEs and APs) are synchronized and time and frequency [11].

The UL/DL Decision Maker (DM) is defined as the entity which takes care of selecting the transmission direction at each TTI. The decision may depend on higher layer reports such as buffer status or Head of Line (HOL) delay [3]. This is meant to improve the performance of the higher layers. Nevertheless, such flexibility comes at the expenses of SINR unpredictability, since the type of interference that a node perceives in each TTI is unknown. Advanced receivers such as Interference Rejection Combining (IRC) [12] and Successive Interference Cancellation (SIC) [13] aim at reducing the SINR variance by suppressing the most significant interfering streams. Nonetheless, in case of a large number of interfering streams with significant power, their use would not be sufficient to cope with the unpredictable SINR. Moreover,

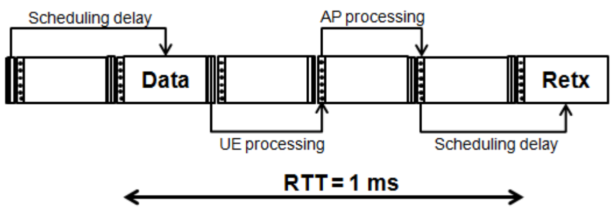


Fig. A.2: Round trip time for a downlink transmission

measurement inaccuracies and signalling errors impact negatively the system too. Measurements delays also affect the link quality, since outdated information may not be valid at another time instant if the interference conditions change rapidly.

Stabilization mechanisms are then needed to improve the link robustness. Collaborative decisions among neighbour cells may improve the overall network performance, but they require a significant signalling overhead. A simpler approach is based on the usage of HARQ and/or Radio Link Control (RLC) recovery mechanisms. In this paper, the HARQ mechanism is considered. It needs to be adapted to the described 5G design in order to improve the overall system performance. Next section presents our recommendations on the design of HARQ for the envisioned 5G concept.

3 HARQ in 5G

As discussed in section 2, the link quality is negatively impacted by several factors. In this first study, we consider the effects of the UL/DL switching, measurement errors and delays, and the chosen approach to deal with such impairments is the use of HARQ with an acknowledged mode RLC [14] on top. Our envisioned HARQ mechanism for 5G is presented in this section.

Figure A.2 shows the timeline of a DL HARQ retransmission [2] [3]. The AP grants a time slot to the UE in TTI_n to receive data in TTI_{n+1}. Once the UE receives the packet, it requires one TTI to process this information and in TTI_{n+3}, the positive or negative acknowledge (ACK/NACK) feedback is sent in the UL control symbol. The AP requires another TTI to process such control information. If a NACK is received, a new time slot is granted to the UE (in TTI_{n+4} in case of synchronous HARQ operation) to receive the retransmission (in TTI_{n+5}). Therefore, in the best case, there is 1 ms delay between a transmission and its corresponding retransmission. If an asynchronous operation is used, the delay will be 1 ms or more, depending

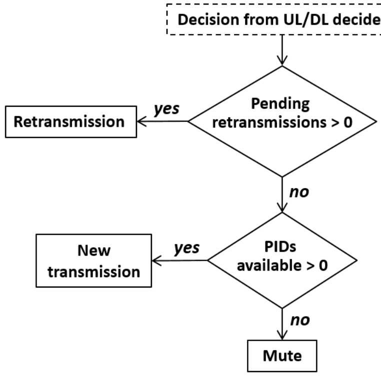


Fig. A.3: HARQ scheme dependent of the UL/DL decision maker

on when the retransmission is scheduled.

In case of UL, the procedure is similar. The difference is that the AP does not need to send the ACK/NACK feedback since the AP is the entity gathering the scheduling information. It is important to notice that the minimum UL delay between the start of a new transmission and its corresponding retransmission is 0.75 ms, one TTI less than in DL.

It would be beneficial to adopt a HARQ operational mode preserving full flexibility and adaptability in both time and frequency domains. Therefore, the preferred approach is to define the two link directions as adaptive and asynchronous, which means that there is not a predefined time where retransmissions will be scheduled, and allocation and transmission parameters can be different at each retransmission. In this way, the UL/DL dynamicity is not being compromised since the asynchronism does not force the system to retransmit at a certain specified time. Moreover, the adaptability can ensure a successful retransmission if the correct resource allocation and Modulation and Coding Scheme (MCS) are chosen. Finally, the number of HARQ processes would be limited to 4, since this is the number of parallel processes that can run during the RTT [3].

Figure A.3 shows the interaction between the UL/DL decision maker and the HARQ manager that has been used for this study. The UL/DL DM performs its decisions regardless of the presence of pending retransmissions in UL or DL. Once the decision is made, the HARQ manager will check if there are pending retransmissions in such direction. In this case, the oldest retransmission will be scheduled for the next TTI. Otherwise, if the maximum number of HARQ processes is not reached, a new transmission with

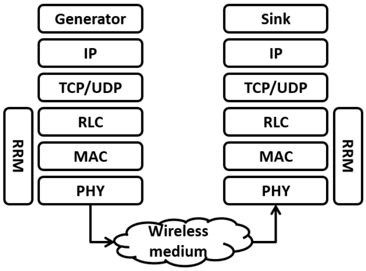


Fig. A.4: Simulator protocol stack

an unused Process Identifier (PID) will be scheduled for the next TTI. If the maximum number of HARQ processes is reached, no transmission will be scheduled for the next TTI. It is important to notice that priority is always given to retransmissions.

4 Performance Evaluation

This section presents an initial performance evaluation of HARQ in our envisioned 5G concept.

The results are extracted from our event driven based system level simulator. It implements the physical (PHY), medium access control (MAC), RLC, transmission control protocol (TCP), user datagram protocol (UDP) and internet protocol (IP) layers. It also features a vertical Radio Resource Management (RRM) layer which interacts with the PHY, the MAC and the RLC. Figure A.4 shows the simulator structure.

At the transmitter side, the data generator is characterized by a traffic model, such as constant bit rate (CBR) or exponential on/off. It creates data units (DUs) according to the defined parameters (packet size, generation period, on and off time). Afterwards, the DUs are passed to the lower layers (IP and TCP/UDP) that add the headers (modelled as overhead). The DUs are then buffered in the RLC layer, until the RRM entity located in the AP decides that a new data transmission will be scheduled. A scheduling decision is taken according to the channel state, SINR conditions and CQI information from UE to AP of the previous TTIs. Then, a set of DUs that fits the transmission bandwidth are aggregated and passed to the MAC, that adds its header and passes the resulting protocol data unit (PDU) to the PHY layer. Afterwards, a coding block, whose size depends on the selected MCS and rank, is

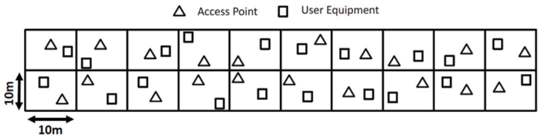


Fig. A.5: 10x2 grid scenario

created once the corresponding PHY header has been added. Finally, a signal is transmitted over the wireless medium.

At the receiver side, both desired and interfering streams arrive to the antennas. The IRC receiver computes the effective SINR, which is passed to the decoding module that decides, according to a block error rate target of 10%, whether the packet is decodable. In case of failure, the HARQ manager takes care of notifying the RRM that a retransmission is required. If the packet is decodable, it is passed to the higher layers up to the sink, where the delay and throughput statistics are computed. The throughput is calculated as the total number of received bits divided by the total simulation time, and the delay is computed from the time that the DU reaches the transmission buffer to the instant that it arrives at the sink.

We define an SINR soft combining model used by the receiver to get the effective SINR upon retransmissions. Soft combining keeps memory of previous transmission of the same packet to achieve SINR gain and then improve the probability of correct detection [6]. The model can be expressed as follows:

$$SINR_{effective} = \sum_{i=1}^{n} SINR_i \bullet \eta^{n-1} \tag{A.1}$$

where n is the transmission number, $SINR_i$ is the SINR for the i^{th} transmission/retransmission of the same packet, and η is the combining efficiency, used to model the non-ideality of the combining process. It is set to 1.0 for simplicity.

The simulated scenario is a 10x2 grid, shown in figure A.5. One AP and one UE are randomly deployed per room. Results are collected over 100 simulation drops; at each drop, a UE is always affiliated to the AP which is in the same room (closed subscriber group). Each node is equipped with 4 antennas and an IRC receiver, and all of them use a fixed rank (1,2,3 or 4) during the whole simulation. Simulation parameters are summarized in Table A.1.

Table A.1: Simulation parameters

Parameter	Value/State/Type
System parameters	BW = 200MHz; f_c = 3.5GHz
Frequency reuse	1 (whole band)
Propagation model	WINNER II A1 w/fast fading [16]
Antenna configuration	4x4
Scenario	10x2
Rank (fixed)	1, 2, 3 and 4
UL/DL decision maker	HOL and random
Max HARQ processes	4
Max retransmission counter	4
RLC mode	Acknowledged
Transport protocol	UDP
Simulation time per drop	1 second
Number of drops	100

This first study shows the performance of an asynchronous and non-adaptive HARQ. Frequency reuse 1 is used, i.e., each transmission occupies the whole band, and the link adaptation decides the MCS to transmit with according to an average of the latest 5 SINR measurements. This is meant to mitigate the impact of measurement errors. The RLC layer operates in acknowledged mode, which includes a retransmission recovery procedure. If an RLC PDU is not acknowledged within a given timeout, a retransmission is triggered. Moreover, the RLC layer also aims at delivering RLC PDUs in-sequence. In case eventual missing PDUs are not recovered within a predefined time, they are considered dropped. Finally, neither QoS nor priority are here considered.

The used traffic model is CBR. Note that this represents a challenging situation for HARQ since there is always new data in the buffer to transmit, and this data has to be delayed if there are pending retransmissions. In addition, two UL/DL DM schemes are evaluated: one based on the buffer size and the Head-of-Line (HOL) delay [3], where the direction is chosen to avoid the buffer from overflowing and to bound the delay experienced by a packet, and a random one, meaning that the direction for each TTI is randomly chosen, with 50% probability for each link to be selected. Finally, we compare the performance of the system when the nodes transmit with fixed rank from 1 to 4.

The usage of rank 1 allows the IRC receiver to use the degrees of freedom of MIMO for suppressing the significant interfering streams [15], and then the gain from HARQ is expected to be limited since no many retransmissions are

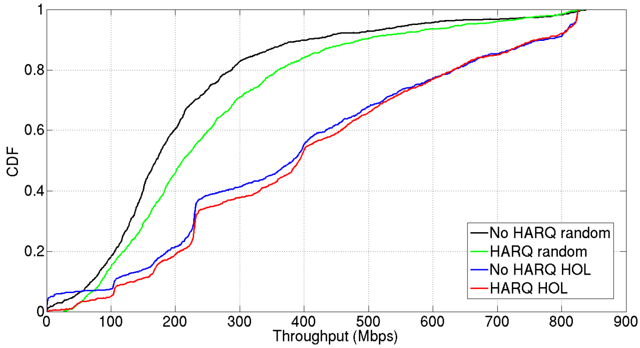


Fig. A.6: Throughput with rank 2

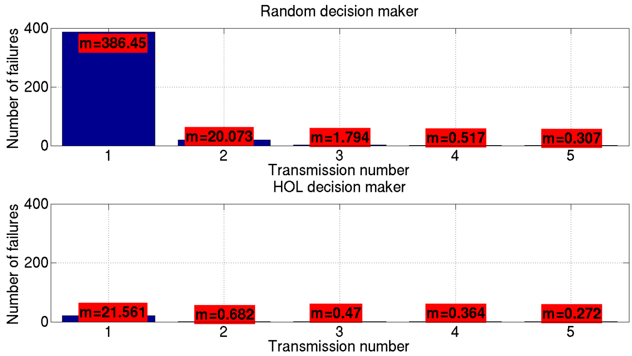


Fig. A.7: Number of failures per attempt

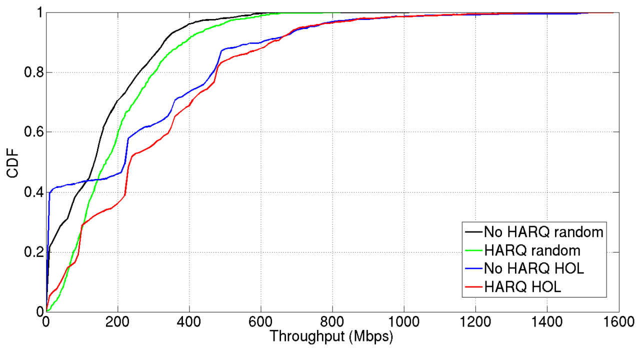


Fig. A.8: Throughput with rank 4

likely to occur. Nevertheless, the interference suppression capability of the IRC receiver diminishes for higher ranks.

Figure A.6 shows the cumulative distribution function (CDF) of the cell throughput when nodes operate with rank 2, for both DM schemes. We can observe that the gain of HARQ when operating with the HOL algorithm is smaller than in the case of random scheduling. This is due to the fact that, with the considered CBR traffic model, the HOL algorithm converges to a fixed downlink:uplink pattern because both buffers have the same amount of data. It is worth to notice the presence of some nodes in outage in the lower part of the CDF, i.e., nodes getting no throughput, a situation solved when HARQ is enabled.

Another interesting aspect is the lower throughput of the random DM compared to the HOL algorithm. This behaviour is due to the uncorrelated decisions on the link direction at each TTI, leading to larger SINR variance. Therefore, a higher impact of HARQ is expected. To prove such behaviour, figure A.7 shows the total number of failures per transmission attempt. The first bar indicates the number of first retransmissions, the second bar denotes the number of second retransmissions, and so on. The last bar indicates the number of dropped packets, since the maximum attempts to retransmit a packet is set to 4. As we can see in the graphic, the number of first retransmissions in the random case is significantly higher than the HOL algorithm case (x8 times). This is due to the fact that with the random DM we spend more time retransmitting, and therefore we reduce the chance of transmitting new data.

Figure A.8 shows the CDF of the throughput when the nodes operate with rank 4. In this case, there are 40% of the nodes in outage when the HOL algorithm is used and 22% in the random case. Such percentages are reduced to 5% and 0% when HARQ is enabled, respectively. It is interesting to notice that the random DM case is less affected than the HOL algorithm case by the high transmission rank, since the random DM leads to unpredictable interference in all rank configurations, and therefore increasing the number of streams has a lower impact on the system performance.

Table A.2 shows the throughput results in percentiles for all the possible ranks, in Mbps. We can see that in all cases HARQ gives a gain, as already visible from inspection from the plots in figure A.6 and figure A.8. The most significant gain is obtained for higher rank transmissions and in the 5% percentile.

Finally, table A.3 shows the delay reduction per rank in percentiles when using HARQ. We can notice that significant delay reduction is obtained for the random DM case; this is due to the faster recovery mechanism provided by HARQ with respect to RLC. Nevertheless, such delay reduction is not visible in the HOL DM case (negative delay detection), since HOL DM does not

Table A.2: Throughput results per rank per percentiles (Mbps)

		Random DM			HOL DM		
		5%	50%	95%	5%	50%	95%
R1	**w/oHARQ**	92.7	280.2	404.5	181.5	389.8	402.2
	w/HARQ	99.3	325.0	405.5	201.0	394.7	402.3
R2	**w/oHARQ**	43.8	170.7	568.5	7.5	383.5	815.0
	w/HARQ	62.5	212.0	651.0	100.5	392.5	819.2
R3	**w/oHARQ**	0	152.9	432.5	0	318.6	973.8
	w/HARQ	50.2	185.3	523.0	37.8	328.3	988.3
R4	**w/oHARQ**	0	139.6	390.7	0	225.7	740.0
	w/HARQ	35.5	167.6	462.5	11.3	240.0	740.0

Table A.3: Delay reduction (%) per rank per percentiles

	Random DM			HOL DM		
	5%	50%	95%	5%	50%	95%
R1	7.5	5.4	13.3	3.2	2.7	-0.4
R2	5.9	12.0	6.6	-2.0	-2.7	-0.5
R3	12.9	11.3	0.9	-11.1	-4.4	-2.9
R4	0	5.6	0.7	-21.8	-4.3	-4.5

take into account HARQ information, and therefore new data transmissions can be prioritized over retransmissions. The delay increase is however rather minor.

5 Conclusions and future work

In this paper, we have evaluated HARQ as a solution for improving the link robustness to the large SINR variance experienced in our envisioned 5G small cells networks. Two approaches for the selection of the instantaneous transmission direction (random and HOL based) are also tested. System level performance results confirm the capabilities of HARQ to improve the link quality in the considered scenarios, especially for the cases of high transmission rank and random UL/DL direction, with limited impact on the end delay.

Future work will be focused on the design of Outer Loop Link Adaptation (OLLA) and dynamic rank adaptation solutions based on HARQ feedback. The impact of different traffic models on the HARQ performance will be also investigated.

References

[1] Cisco; *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2011–2016*, White Paper, February 2012

[2] P. Mogensen et al.; *Beyond 4G local area: high level requirements and system design*, IEEE Globecom Workshops (GC Wkshps), 2012

[3] P. Mogensen et al.; *Centimeter-wave concept for 5G ultra-dense small cells*, IEEE Vehicular Technology Conference, Spring 2014

[4] 3GPP TS 25.321, Technical Specification Group Radio Access Network; Medium Access Control (MAC) protocol specification

[5] 3GPP TS 36.321, Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) protocol specification

[6] H. Holma and A. Toskala; *LTE for UMTS - OFDMA and SC-FDMA Based Radio Access*, Wiley April, 2009

[7] J. F. Cheng, Y.P.E. Wang and S. Parkvall; *Adaptive incremental redundancy [WCDMA systems]*, IEEE Vehicular Technology Conference, Fall 2003

[8] W. C. Chung, Y. Y. Chen and C.J. Chang; *HARQ Control Scheme by Fuzzy Q-Learning for HSPA+*, IEEE Vehicular Technology Conference, Spring 2011

[9] M. Kim, S. Kim and Y. Lim; *An implementation of downlink asynchronous HARQ for LTE TDD system*, IEEE Radio and Wireless Symposium, 2012

[10] Y. Li, M. C. Gursoy and S. Velipasalar; *On the Throughput of Hybrid-ARQ under QoS Constraints*, Computing Research Repository, 2013

[11] G. Berardinelli, F. M. L. Tavares, N. H. Mahmood, O. Tonelli, A. F. Cattoni, T. B. Sørensen, P. Mogensen; *Distributed Synchronization for Beyond 4G Indoor Femtocells*, International Conference on Telecommunications, May 2013

[12] 3GPP TR 36.829, Technical Specification Group Radio Access Network; Enhanced performance requirement for LTE User Equipment (UE)

[13] C. N. Manchon, L. Deneire, P. Mogensen and T. B. Sørensen; *On the Design of a MIMO-SIC Receiver for LTE Downlink*, IEEE Vehicular Technology Conference, Fall 2008

[14] 3GPP TS 25.322, Technical Specification Group Radio Access Network; Radio Link Control (RLC) protocol specification

[15] F. M. L. Tavares, G. Berardinelli, N. H. Mahmood, T. B. Sørensen and P. Mogensen; *On the Potential of Interference Rejection Combining in B4G Networks*, IEEE Vehicular Technology Conference, Fall 2013

[16] I. IST-WINNER; *"Deliverable 1.1.2 v.1.2, WINNER II channel models, IST-WINNER2"*, Tech. Rep., 2007

Paper B

Dynamic Outer Loop Link Adaptation for the 5G Centimeter-Wave Concept

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

The paper has been published in the
IEEE 21th European Wireless Conference (EW), 2015.

Abstract

A 5th generation (5G) of wireless communication systems is expected to be introduced around 2020 to cope with a rapid increase of mobile data traffic. One of the main challenges of our envisioned 5G centimeter-wave concept is a large signal to interference plus noise ratio (SINR) variability, due to a flexible uplink/downlink (UL/DL) scheduler that aims at dynamically accommodating the traffic while reducing the latency. The use of advanced receivers and recovery mechanisms has been proved not to be sufficient to deal with such variability. In this paper, the potential of outer loop link adaptation (OLLA) to cope with the aforementioned problem is studied, and a dynamic OLLA (d-OLLA) algorithm is proposed. System level simulations show that d-OLLA provides an average throughput gain of up to 23% and 16.6% when using fixed rank 1 and rank 2, respectively, and 20.7% gain with rank adaptation, without adding extra complexity in the system.

1 Introduction

Latest forecasts show an exponential growth in the mobile data traffic in the next years, reaching a factor of $\sim$x1000 by 2020 (with reference to 2010) [1]. Given the intrinsic limitation of radio standards such as Long Term Evolution (LTE)/LTE-Advanced (LTE-A) [2], a novel 5$^{\text{th}}$ generation (5G) radio access technology (RAT) is expected to cope with such traffic increase. We have presented our vision on a 5G system in [3]. Such system is optimized for local area, where an ultra-dense deployment of small cells with limited coverage is foreseen. Inter-cell interference is the main limiting factor in such deployments when operating in the centimeter-wave spectrum region. Time Division Duplex (TDD) is foreseen as the preferred operational mode, where a flexible uplink/downlink (UL/DL) time slot assignment is a key feature to provide dynamicity and accommodate the traffic optimally [3]. Nevertheless, such flexibility may increase the inter-cell interference variability, which leads to a large instability in the measured signal to interference plus noise ratio (SINR), since the interference patterns may change at each time slot. Such SINR instability may negatively affect the link performance since it leads to incorrect link adaptation decisions.

Advanced receivers such as Interference Rejection Combining (IRC) and Successive Interference Cancellation (SIC) have the potential of suppressing the most significant interfering streams, thus reducing the SINR variability [3]. Nonetheless, their contribution may be insufficient for stabilizing the SINR, and further mechanisms are required. One solution is to use coordination among cells. The drawbacks of such approach is an increase in system complexity and latency. Other solutions are based on the usage of recovery mechanisms, such as Hybrid Automatic Repeat Request (HARQ) or ARQ. The potential of HARQ has already been studied in [4].

The usage of an outer loop link adaptation (OLLA) mechanism is another option for dealing with the residual SINR variability. OLLA is based on Outer Loop Power Control (OLPC) [5] and has already been adopted in previous RATs. OLLA was originally designed to improve the robustness of the system to eventual systematic errors coming from Channel Quality Indicator (CQI) reports, delays, quantization and interference variability. Such mechanism is based on an offset, adjusted according to positive/negative acknowledgements (ACK/NACKs) from past transmissions. Such offset is added to the CQI or to the estimated SINR, with the aim of correcting the current Modulation and Coding Scheme (MCS).

OLLA was first used in High Speed Packet Access (HSPA). An algorithm was proposed in [6], following the OLPC principle [5] and working on the MCS domain, so the OLLA offset was extracted directly from the MCS index. In LTE/LTE-A, outer loop link adaptation is implementation specific and it is not mandatory. In [7], an OLLA algorithm based on OLPC [5] is used, but operating on the SINR domain rather than in the MCS domain, as proposed for HSPA [6]. An analysis of how OLLA reacts depending on the predefined block error rate (BLER) target and its tolerable error is done in [8] [9]. The authors conclude that a BLER target higher than 20-30% would have a negative impact on the delay, and that OLLA can cope with errors of approximately 2 dB. Finally, [10] proposes an optimized OLLA algorithm for Multiple-Input Multiple-Output (MIMO) systems.

In this paper, the potential of OLLA in our envisioned 5G system is analysed. A baseline system level performance analysis using the algorithm presented in [7] is provided. Furthermore, a dynamic OLLA (d-OLLA) algorithm is proposed, considering the mean and the standard deviation of the SINR to get the optimal BLER target.

The paper is structured as follows. Section II presents our envisioned 5G design and justifies the necessity of an OLLA algorithm. Section III defines the 5G OLLA framework. Section IV includes the simulation setup and the performance evaluation. Finally, Section V presents the conclusions and states the future work.

2 5G Systems: Design and Challenges

In order to cope with the future mobile data traffic demands, 5G is expected to significantly boost the LTE/LTE-A performance. The peak data rate target is set to 10 Gbps, the Round Trip Time (RTT) to 1 ms, and wake-up time from "inactive" to "active" to 10 ms. We believe that a clean slate approach for the design of a 5G RAT is required for achieving such targets. Our foreseen design for such system is presented in [3], where TDD has been chosen as the operational mode.

A new frame structure of 0.25 ms, specified as the transmission time interval (TTI), is defined [3]. Control and data planes are separated in time, being the control part located before data part, as shown in Figure B.1. The data part, which includes the demodulation reference signal (DMRS) for enabling channel estimation at the receiver, can carry either DL or UL traffic. In that sense, an independent scheduling decision will be taken for every TTI. It is important to remark that the scheduling decision taken in TTI_n will be applied in TTI_{n+1}, i.e., there is one TTI delay between the time instant when the decision is taken until such decision is applied. An UL/DL scheduler, as

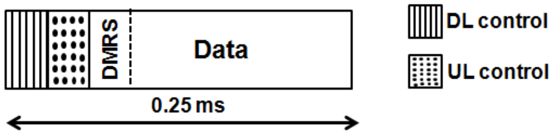


Fig. B.1: Envisioned 5G frame structure

proposed in [3], takes care of deciding the link direction at each frame, depending on the amount of pending data in both UL and DL buffers and the head-of-line delay, i.e., the amount of time that the next packet to be transmitted has been waiting. In this way, the traffic in the two link directions can be optimally accommodated. Nevertheless, as mentioned in the introduction, such flexibility comes at the price of an increased inter-cell interference variability, since the transmission direction is extracted independently in each TTI, instead of using predefined transmission patterns like in TDD LTE [11]. Then, the type of perceived interference in each node per TTI is unknown.

MIMO antenna technology is required to achieve the desired network capacity, being 4x4 the default configuration. In addition, IRC receivers are assumed. These receivers operate on the spatial domain to suppress a number of interfering streams at the expense of reducing the number of desired transmission streams. In that sense, they provide a first tier of protection towards interference variability. In addition, mechanisms such as HARQ in the physical layer and ARQ in the Radio Link Control (RLC) layer are also included in our design, to be able to recover from link failures and to add robustness to the system [4].

Nonetheless, the use of advanced receivers and recovery mechanisms would not be enough to cope with the unpredictable SINR. Figure B.2 shows the mean and standard deviation of the SINR variability, depending on the number of used spatial streams, often referred as the transmission *rank*, and

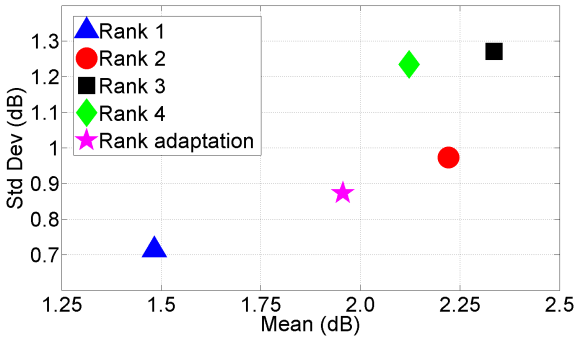


Fig. B.2: SINR variability (σ_{SINR}) characteristics

also when a rank adaptation algorithm [12] is used. These results have been extracted from a 10x2 grid scenario [4] [12], i.e., 20 rooms, each of them with 1 AP and 1 UE randomly deployed (see Figure B.3). The traffic model is full buffer, and in each TTI, the transmission direction is decided randomly (equal probability for DL and UL)[1]. Such SINR variability (σ_{SINR}) is defined as the standard deviation of the post-IRC SINR in dB, i.e.

$$\begin{aligned}\sigma_{SINR} &= \sqrt{\mathrm{E}\{(SINR - \mu_{SINR})^2\}} \\ &= \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(SINR_i - \mu_{SINR})^2} \qquad \text{(B.1)} \\ \mu_{SINR} &= \mathrm{E}\{SINR\} = \frac{1}{N}\sum_{i=1}^{N} SINR_i\end{aligned}$$

where $\mathrm{E}\{\cdot\}$ refers to the expected value, N is a sliding window of size 16 subframes or TTIs and μ_{SINR} is the mean of the post-IRC SINR. From Figure B.2 we can see that, even using IRC receivers and HARQ, the mean of the SINR variability ranges from 1.5 to 2.3 dB. More precisely, Figure B.2 shows that reducing the degrees of freedom to suppress interference increases the SINR variability, making it more challenging for the link adaptation mechanism to make a correct decision. To deal with the residual SINR variability and further improve the system performance, outer loop link adaptation is studied. Next section describes the OLLA framework according to the described 5G system and it introduces the design of a new dynamic OLLA algorithm.

[1]Section IV will further describe the simulated scenario

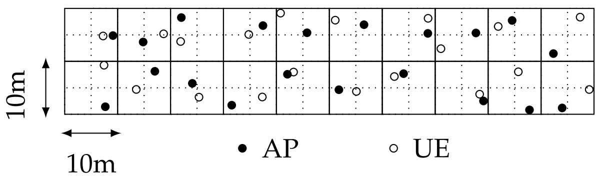


Fig. B.3: Simulated scenario

3 5G OLLA Framework

OLLA was originally designed to compensate for systematic error sources such as delays, quantization errors, inaccuracies in the CQI report or interference variability, in order to maintain a predefined BLER target. Such mechanism adjusts an additive offset, based on HARQ ACK/NACK feedback from past transmissions, that is applied in the SINR domain to extract the most appropriate MCS for transmission [7].

In the previous section, the SINR variability in our system has been characterized. Furthermore, Figure B.4 shows the time line of a DL transmission when OLLA is enabled, according to our 5G concept. In TTI_n the user equipment (UE) measures the instantaneous SINR from the DMRS symbol. This SINR, which is post-IRC, is translated into a CQI, which will be transmitted in TTI_{n+1} in the UL control part. It is important to notice that the CQI is transmitted in each TTI and it is instantaneous, meaning that it is extracted from the latest SINR sample, therefore no filtering over previous samples of the SINR is being applied. Then, in TTI_{n+2}, the UE will send the HARQ ACK/NACK feedback in the UL control channel. Once the access point (AP) has received such information, it will send a grant in TTI_{n+3}. Such grant will include a scheduling decision for TTI_{n+4}, which will be using the latest CQI report (from TTI_n) compensated by the OLLA offset according to the feedback received in TTI_{n+2}. Therefore, in the best case, the delay between the instant when a measurement is done and the instant when it is applied will be 1 ms. Hence, the MCS that is going to be used would most probably not be the one that better matches the current channel conditions. From Figures B.2 and B.4, we can conclude that OLLA needs to compensate for a SINR variability in the order of 1-4 dB and large measurement delays compared to the TTI length. These conditions are more challenging than the ones from HSPA or LTE/LTE-A, where OLLA was also used. Unlike such systems, our envisioned 5G concept introduces fully flexible TDD, causing an increase in interference variation and thus a larger SINR variation. For this reason, as a baseline, the proposed algorithm in [7] has been analysed within our OLLA framework. To the authors knowledge, this is the most relevant OLLA algo-

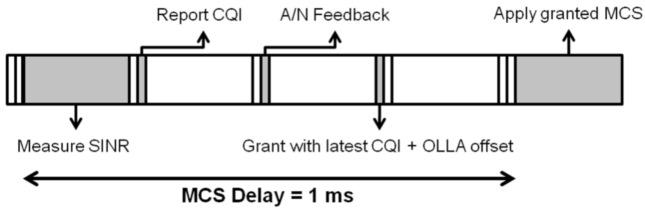


Fig. B.4: Measurement delay time line

rithm in literature.

The OLLA algorithm [7] uses three variables: *stepUp*, *stepDown* and BLER target ($BLER_T$). These variables are fixed, meaning that once the $BLER_T$ and the *stepUp* are predefined, the *stepDown* is extracted as follows:

$$stepDown = \frac{stepUp}{\frac{1}{BLER_T} - 1} \tag{B.2}$$

The working principle of the OLLA algorithm, shown in the pseudo code 1, is very simple and intuitive. When an ACK is received, the OLLA offset, which is initially set to 0 dB, is decreased by *stepDown*. On the other hand, if a NACK is received, the offset is increased by *stepUp*. Once the offset is calculated, it is subtracted from the current SINR measurement and the corresponding MCS index is obtained from a lookup table *SINR-MCS_Index*. When *stepDown* is subtracted from the offset, the SINR is increased and then a more aggressive selection of the MCS is made. With the same principle, a more conservative MCS selection is performed in case of reception of a NACK, since summing *stepUp* to the OLLA offset leads to a SINR decrease.

This algorithm [7] works when the channel conditions and measurement errors are stationary and within a certain limited range. Nevertheless, a large SINR variability is expected in 5G small cells, as discussed in Section II. For this reason, we propose a dynamic OLLA algorithm, d-OLLA, that takes as input the mean and the standard deviation of the post-IRC SINR in order to characterize the perceived interference conditions, and then extract the optimal BLER target. Such algorithm does not add extra complexity into the system, and it is intended to adapt the OLLA behaviour to the current channel conditions and improve the outage performance, i.e., to reduce the number of users getting no throughput. We believe that, depending on the interference conditions that a node is perceiving, it should be more conservative or more aggressive on the MCS. The working principle is the same as the one shown in pseudo code 1, but the definition of the *stepDown* parameter is changed, as shown in equation B.3, since it includes the mean SINR (μ_{SINR}) and the

Algorithm 1 Outer Loop Link Adaptation Algorithm

$fb \leftarrow$ ACK/NACK ▷ HARQ feedback
$sU \leftarrow$ Step Up ▷ Value = 0.5 dB
$sD \leftarrow$ Step Down
$offset \leftarrow$ OLLA Offset ▷ Initial value = 0 dB
{One outer loop per link}
if *fb* is ACK **then**
    {Positive feedback $\rightarrow$ Be aggressive}
    $offset \leftarrow offset - sD$
else
    {Negative feedback $\rightarrow$ Be conservative}
    $offset \leftarrow offset + sU$
end if
$SINR_{OLLA} = SINR - offset$
$MCS_{index} = MCS[SINR_{OLLA}]$

SINR standard deviation (σ_{SINR}). The *stepUp* parameter remains fixed.

$$stepDown = \frac{stepUp}{\frac{1}{BLER_T(\mu_{SINR}, \sigma_{SINR})} - 1} \tag{B.3}$$

When an ACK is received, the *stepDown* is recalculated with the BLER target extracted according to μ_{SINR} and σ_{SINR}. In that sense, a BLER function under different channel conditions needs to be defined.

Next section describes the proposed BLER target $BLER_T(\mu_{SINR}, \sigma_{SINR})$ function, the simulation set-up and the performance evaluation of both algorithms.

4 Performance Evaluation

The provided results are extracted from an event-driven based system level simulator. It implements several layers of the Open Systems Interconnection (OSI) protocol stack. It features the physical (PHY), medium access control (MAC) and RLC layers according to the presented 5G design. The transmission control protocol (TCP) and the user datagram protocol (UDP) layers are fully modelled, while the internet protocol (IP) is modelled in terms of overhead. Moreover, it includes a vertical Radio Resource Management (RRM) layer in the AP side, that gathers information from the PHY, MAC and RLC layers to make the most appropriate decision for a link. Such information includes channel state, SINR conditions, HARQ retransmissions and infor-

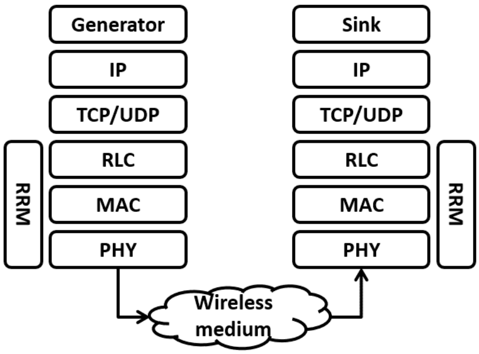


Fig. B.5: Simulator protocol stack

mation provided by the UE to the AP. Figure B.5 shows the protocol stack of such simulator. The simulated scenario is a 10x2 grid [4] [12], resulting in 20 rooms (i.e., 20 small cells), shown in Figure B.3. Results are collected over 50 simulation drops. Each drop creates a deployment, with one AP and one UE randomly deployed per room, and the UE is always affiliated to the AP which is in the same room (closed subscriber group). The same set of deployments is used for all the simulations, to provide a fair comparison between the schemes.

Each node is equipped with 4 transmit and receive antennas and an IRC receiver. The used rank is the same for all nodes during the whole simulation, except when a taxation-based rank adaptation algorithm is used [12]. Such algorithm is interference-aware, and it aims at using a lower rank in high interfered scenarios to guarantee a good outage performance, while providing the benefit of higher average throughputs. For all configurations, a single codeword is transmitted independently of the rank, so the AP will execute an independent OLLA offset per link, i.e., one in DL and one in UL. Moreover, the *SINR-MCS_Index* table, defined according to a BLER target of 10%, includes 31 MCSs, from quadrature phase-shift keying (QPSK) to 256-quadrature amplitude modulation (256-QAM), with several coding rates. Notice that such large number of MCSs offers high granularity for coping with the channel conditions.

The SINR combining model used by the receivers in the physical layer refers to an ideal Chase Combining (CC) [13], shown in equation B.4, where n is the transmission number, $SINR_i$ is the post-IRC SINR in the transmission attempt i and η is the combining efficiency, which is set to 1. The maximum number of allowed HARQ retransmissions is set to 4.

$$SINR_{\text{effective}} = \sum_{i=1}^{n} SINR_i \cdot \eta^{n-1} \tag{B.4}$$

Table B.1: Used parameters to run the simulations

Parameter	Value/State/Type
System parameters	BW = 200MHz; f_c = 3.5GHz
Frequency reuse	1 (whole band)
Propagation model	WINNER II A1 w/fast fading [14]
Antenna configuration	4x4
Rank	Fixed (1, 2, 3 and 4) and adaptive
Receiver type	IRC
UL/DL decider	Random (50% UL, 50% DL)
Max HARQ processes	4
Max retransmission counter	4
OLLA offset range	{-10,3} dB
Step up	0.5 dB
Packet size	12000 bits
RLC mode	Acknowledged
Transport protocol	UDP
Simulation time per drop	1 second
Number of drops	50

The traffic model is full buffer in both link directions. The UL/DL decider is random, with equal probability (50%) of scheduling a DL or an UL transmission. The reason of choosing such UL/DL decider and traffic model is to create a challenging situation in terms of SINR variability. Finally, the OLLA offset range has been extracted empirically. The remaining simulation parameters are summarized in Table B.1. The results are presented in terms of final throughput and number of transmissions/retransmissions. In addition, when analysing the fixed rank case, we focus on the results with rank 1 and rank 2. The reason is because we are considering full buffer traffic, with 100% deployment ratio and frequency reuse 1. In this particular case, since all the cells are always active, the interference level is very high. Therefore, a network configuration set-up in which the maximum rank is limited to 1 or 2 provides the best trade-off between spatial multiplexing gains and inter-cell interference protection (due to the use of IRC), leading to the best overall network performance [15].

The first set of results shows the performance of the OLLA algorithm proposed in [7] for different BLER targets.

Figure B.6 shows the cumulative distribution function (CDF) of the average throughput per link with fixed rank 1. In this case, OLLA targeting a 5% BLER gives the best network performance, providing a gain of 12.7% in the 5th percentile and 23% in the average throughput. The 10% BLER target configuration also provides significant gains (16.4% in the 5th per-

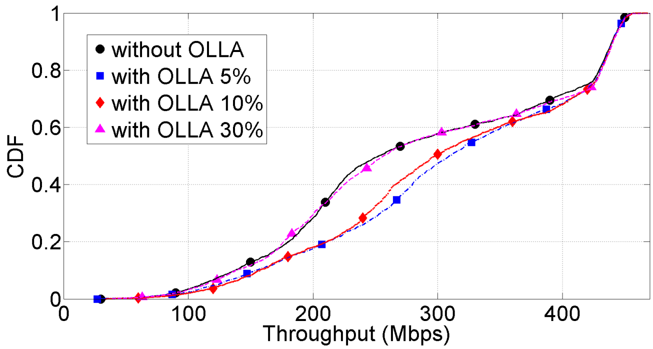


Fig. B.6: Final throughput with rank 1

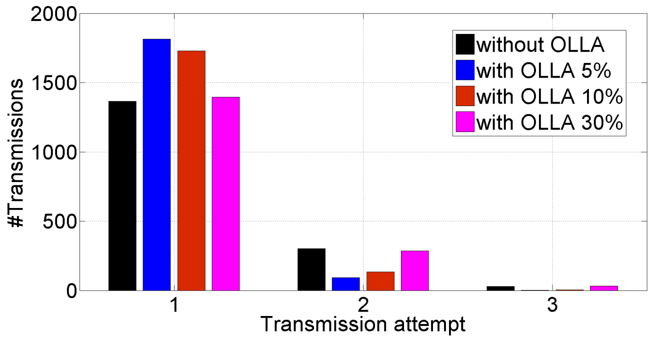


Fig. B.7: Transmission attempts with rank 1

centile and 17.9% in average), which means that it is also a good solution. Therefore, these results show the opposite behaviour from previous studies [7] [8] [9] [10], since 30% BLER target does not provide any gains or even a loss. This is because, even if the interference conditions are more stable since the IRC receiver suppresses up to three interfering streams, the SINR variability is still significant (mean of 1.5 dB and standard deviation of 0.7 dB, according to Figure B.2), meaning that OLLA has to compensate also for the remaining interference that IRC cannot suppress. Hence, it is better to be conservative in the MCS selection by using a low BLER target and reduce the number of retransmissions. This reduction in retransmissions is shown in Figure B.7, where the first column refers to first transmissions, the second one to first retransmissions and the third one to the second retransmissions. Note that in this case, up to two HARQ retransmissions are required to get the packets through. From this figure, we observe that a BLER target of 5% instead of 30% reduces the first retransmissions by a factor of 3, while improving the throughput. The throughput with fixed rank 2 is shown in Figure B.8. In this case, since IRC can only suppress up to two interfering streams,

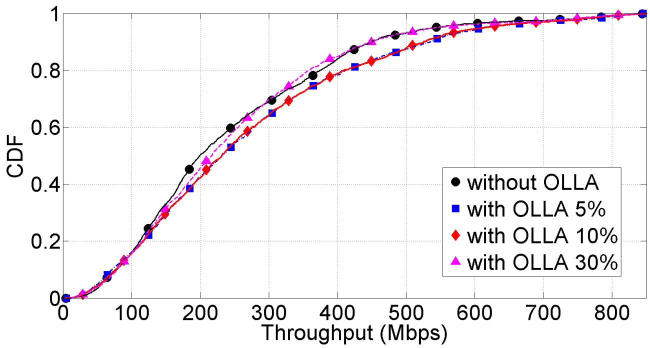


Fig. B.8: Final throughput with rank 2

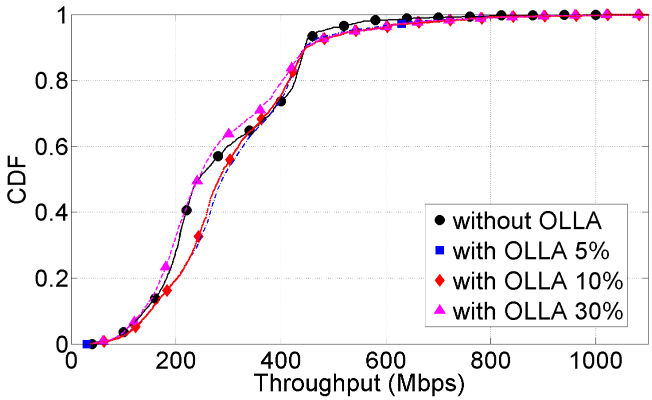


Fig. B.9: Final throughput using rank adaptation

the interference conditions become worse and the SINR variability increases (see Figure B.2). Then, OLLA is not able to compensate for such larger variability, and with 5%BLER target (best case, like with fixed rank 1), the UEs in outage (5^{th} percentile) are negatively impacted. On the other hand, OLLA provides an average gain of 16.6% and 13.3% in the 95^{th} percentile. In this case, the first retransmissions are reduced by 2 due to worse interference conditions, and up to 4 retransmissions are needed to get the packets through.
Finally, we evaluate how the system performs when using the taxation based rank adaptation algorithm [12]. In this case, the throughput, shown in Figure B.9, lays between the results of fixed rank 1 and rank 2, with a gain of 11.7% in the 5^{th} percentile, 19.4% in average and 10.1% in the 95^{th} percentile. These results confirm the previous statement regarding the trade-off between spatial multiplexing gain and inter-cell interference protection. Even though the choice is between four ranks, rank 1 and rank 2 are only selected in practice due to the strong interference conditions. The best solution is still to target a

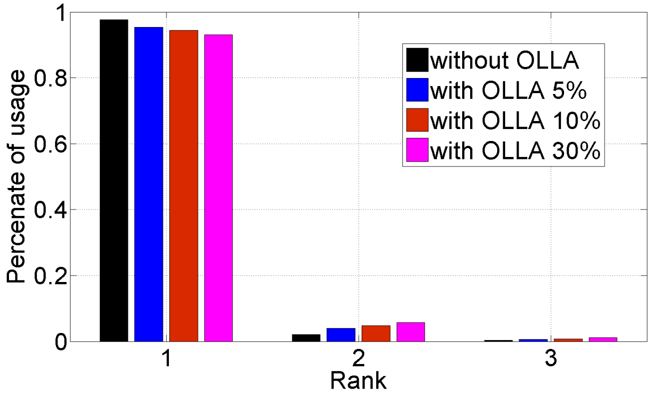


Fig. B.10: Rank distribution when using rank adaptation

low BLER (5-10%). The number of required retransmissions increases slightly compared to the fixed rank 1 approach, because higher ranks are also used. Figure B.10 describes the rank distribution, showing that, without OLLA, the system cannot adapt to the interference conditions and a conservative approach (rank 1) is chosen. Nevertheless, using a conservative OLLA allows the system to choose a higher rank and therefore improve its performance. Notice that increasing the BLER target makes the system more aggressive, providing worse results compared to being conservative.

After a detailed analysis of the system behaviour when using fixed and adaptive rank, we notice that the interference conditions do have an impact in the performance of the network. For this reason, the performance of the proposed d-OLLA algorithm, which aims at further stabilizing the SINR variability, is evaluated.

For our d-OLLA algorithm, the BLER target depends on the estimated μ_{SINR} and σ_{SINR}, as described in Section III. The BLER target mapping table used for the current evaluation is visually depicted in Figure B.11. Notice that, for this first study, an empirical definition of $BLER_T(\mu_{SINR}, \sigma_{SINR})$ has been used. Three regions are distinguished depending on μ_{SINR} and σ_{SINR}. The first region corresponds to very low SINR, i.e, below the minimum MCS decoding threshold. In this case, independently of σ_{SINR}, the nodes are retransmitting all the packets and relying on the CC gain to get them through. Then, it is beneficial to target a higher BLER, since retransmissions may happen anyway. The second region is critical, because a small SINR variation may cause retransmissions. Therefore, σ_{SINR} is the parameter which dictates the most appropriate BLER target. As σ_{SINR} increases, the probability of guessing the correct MCS decreases, meaning that retransmissions may most probably occur given the critical SINR conditions. Therefore, the OLLA algorithm should be more aggressive. The third region corresponds to good

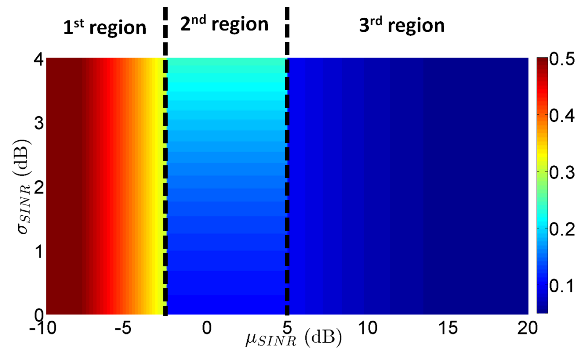


Fig. B.11: BLER Target depending on μ_{SINR} and σ_{SINR}

Table B.2: Throughput gain of 5% OLLA and d-OLLA versus no using OLLA

	5th percentile		50th percentile		95th percentile	
	5% B_T	d-OLLA	5% B_T	d-OLLA	5% B_T	d-OLLA
Rank 1	12.7%	**16.4%**	23.0%	23.0%	0	0
Rank 2	-8.9%	**7.0%**	16.6%	16.6%	13.3%	**16.4%**
Rank 3	0	0	0	**2.4%**	11.1%	**13.0%**
Rank 4	0	0	-9.9%	**-5.5%**	13.0%	**17.0%**
RA	11.7%	**14.2%**	19.4%	**20.7%**	10.1%	**12.6%**

SINR conditions. In this case, from previous results (Figure B.6, Figure B.8 and Figure B.9), we conclude that a conservative solution is the one providing the best performance. Finally, the mathematical function which describes all regions is a parabola, depending on μ_{SINR} for the first and third regions, and on σ_{SINR} for the second one. Table B.2 provides the gain of the baseline algorithm [7] with 5% BLER target (5% B_T) and the d-OLLA algorithm versus no using OLLA. From the results we observe that, as the SINR variability increases, the OLLA gain decreases, even adding a loss for the baseline case when considering fixed rank. Nevertheless, we can see that d-OLLA improves the performance of the system. Even if the gain ranges from 2% to 15.9%, d-OLLA solves the outage problem with fixed rank 2 and improves the one with fixed rank 1. In addition, when using rank adaptation, d-OLLA always performs better than [7], providing an average gain of 20.7% over the no use of OLLA. As future work, we have planned to design a more accurate BLER target mapping table according to μ_{SINR} and σ_{SINR} ($BLER_T(\mu_{SINR}, \sigma_{SINR})$), which also depends on the rank and the load in the system. Nevertheless, it is important to remark that the gain of the proposed algorithm does not come at the expenses of extra complexity in the system.

5 Conclusions and future work

In this paper, the potential of OLLA in improving the link robustness to the SINR variability in our envisioned 5G centimeter-wave concept is analysed. A dynamic algorithm d-OLLA is proposed, which provides gain over a baseline algorithm without adding any extra complexity to the system. System level simulations confirm the capabilities of d-OLLA to improve the link quality in the considered scenarios, leading to a higher final throughput.

The design of different BLER target functions is left for future work. Also, the impact of different traffic models on the OLLA performance will be investigated.

References

[1] Cisco, "Cisco visual networking index: Global mobile data traffic forecast update, 2014-2019," 2015.

[2] H. Holma and A. Toskala, *LTE for UMTS - OFDMA and SC-FDMA Based Radio Access*. Wiley, 2009.

[3] P. Mogensen *et al.*, "Centimeter-wave concept for 5G ultra-dense small cells," in *IEEE 79th Vehicular Technology Conference (VTC Spring)*, May 2014.

[4] M. Sarret *et al.*, "Improving link robustness in 5G ultra-dense small cells by hybrid ARQ," in *IEEE 11th International Symposium on Wireless Communications Systems (ISWCS)*, August 2014.

[5] A. Sampath, P. S. Kumar, and J. Holtzman, "On setting reverse link target SIR in a CDMA system," in *IEEE 47th Vehicular Technology Conference (VTC)*, May 1997.

[6] K. Pedersen *et al.*, "Performance of high-speed downlink packet access in coexistence with dedicated channels," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 3, pp. 1262–1271, May 2007.

[7] K. I. Pedersen *et al.*, "Frequency domain scheduling for OFDMA with limited and noisy channel feedback," in *IEEE 66th Vehicular Technology Conference (VTC)*, 2007.

[8] k. Aho, O. Alanen, and J. Kaikkonen, "CQI reporting imperfections and their consequences in LTE networks," in *IEEE 10th International Conference on Networks (ICN)*, 2011.

[9] C. Rosa *et al.*, "Performance of fast AMC in E-UTRAN uplink," in *IEEE International Conference on Communications (ICC)*, 2008.

[10] I. Kovács *et al.*, "Effects of non-ideal channel feedback on dual-stream MIMO-OFDMA system performance," in *IEEE 66th Vehicular Technology Conference (VTC)*, 2007.

[11] 3rd Generation Partnership Project, "Further enhancements to LTE time division duplex (TDD) for downlink-uplink (DL-UL) interference management and traffic adaptation," 2012.

[12] D. Catania *et al.*, "A distributed taxation based rank adaptation scheme for 5G small cells," in *IEEE 81st Vehicular Technology Conference (VTC Spring)*, May 2015, accepted.

[13] 3rd Generation Partnership Project, "Physical layer aspects of UTRA high speed downlink packet access (release 2000)," 2001.

[14] W. W. I. N. Radio, "WINNER II channel models," 2008.

[15] F. Tavares *et al.*, "Inter-cell interference management using maximum rank planning in 5G small cell networks," in *IEEE 11th International Symposium on Wireless Communications Systems (ISWCS)*, 2014.

Part III

Full Duplex in 5G Small Cells

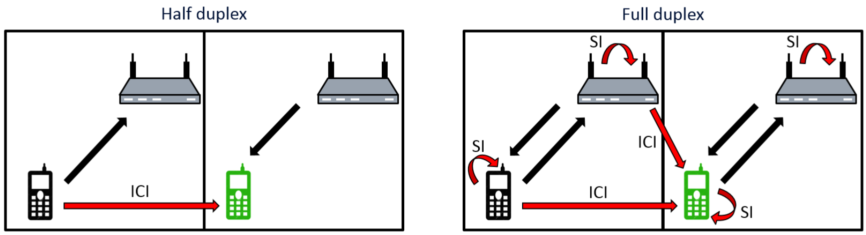
Overview

1 Problem Description and Assumptions

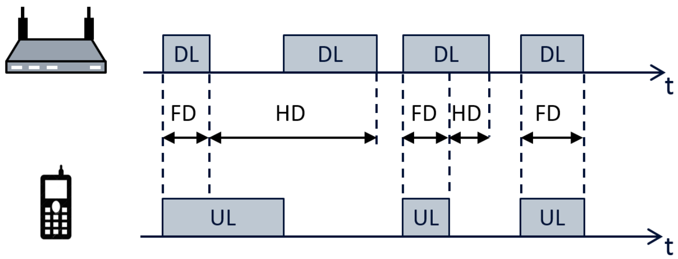
A strategy to enhance the spectral efficiency to increase the system capacity is the usage of MIMO technology with a large number of antennas. However, this strategy brings constraints in terms of space and cost, thus becoming infeasible. In this part of the dissertation, we analyse the role of FD technology in improving the system capacity of a 5G indoor small cell system [1]. FD allows for simultaneous transmission and reception in the same frequency band, which can theoretically double the system throughput over conventional HD systems.

The key limitation in building an operational FD transceiver is the self-interference (SI), i.e., the interference generated by the transmitted signal at the receiver end of the same node. Recent results in the SI cancellation technology have shown that $\sim$110 dB of attenuation of the SI are already achievable [2], according to certain constraints in bandwidth and transmit power. Such results indicate that building an operational FD transceiver is already feasible.

FD promises 100% throughput gain over HD transmission mode. However, there are several limitations that may prohibit FD to accomplish that promise. Figure 4.12 shows the three constraints that impact the FD gain. The first two, namely the SI and the ICI, are depicted in Figure 4.12a, with focus on the green device. The red lines represent interference, whereas the black ones correspond to the desired signals. FD technology requires a high level of attenuation of the transmitted signal to avoid saturating the receiver end of the same node. Even though the recent SI cancellation results indicate that building an operational FD transceiver could be already feasible, the residual SI impacts the performance of a FD node. Furthermore, the increase of ICI caused by FD by doubling the number of interfering streams compared to HD, has a negative effect on the FD performance. The third constraint is the traffic profile, shown in Figure 4.12b. Exploiting FD is only possible when there is data in both ends. Realistic networks are represented by bursty and DL heavy traffic, and such traffic is usually not symmetric between DL and UL, being the former much heavier than the latter [3].



(a) Self-interference and inter-cell interference. Comparison between HD and FD in a two cells scenario.



(b) Traffic profile.

Fig. 4.12: Constraints which limit the gain that FD can provide over HD.

Two FD applications are considered, bidirectional FD and BS FD. The former refers to the case where both the AP and the UE are FD capable. In the latter case, only the BS or AP is able to exploit simultaneous transmission and reception in the same frequency band.

The research community has been actively studying FD technology given its potential. Most of the works are focused on SI cancellation techniques, since as explained earlier, FD requires a high level of attenuation of the transmitted signal to be operational. The evaluation of FD in realistic scenarios is scarce. Many research works assume full buffer traffic model and isolated cells, which means that two of the main constraints that limit the FD gain are not considered. The main target in these works is the evaluation of the residual SI impact. Then, the throughput obtained with FD is approximately doubled compared to the HD throughput. This work takes a different approach, by assuming ideal SI cancellation and evaluating the potential of FD technology in indoor small cell networks considering the occurrence of ICI and realistic traffic profiles. Furthermore, the main KPI is not only throughput but also packet delay. The following assumptions are considered in this part of the dissertation:

- **Ultra-dense indoor small cell network.**
  A 10×2 grid of small cells is considered, with a size of 10×10 m^2 and a wall penetration loss of 5 dB [4]. Each indoor cell contains one AP and 4 UEs, all randomly deployed. The users are affiliated to the AP in the same cell (closed subscriber group).

- **Finite buffer traffic model.**
  As previously described, the traffic asymmetry between UL and DL has an impact on the FD performance. The considered traffic model is the one defined by 3GPP [5]. This model generates payloads of 2 megabytes in average. Both the payloads and their inter-arrival time are extracted according to a negative exponential distribution. Two ratios of the offered load are considered, DL:UL = 1:1 and DL:UL = 6:1, in order to capture the effect of the traffic asymmetry on the FD gain.

- **Dynamic TDD system for comparison.**
  The *load fairness based* scheme [6] is the baseline choice for this study, because it shows better performance than the *delay fairness based* scheme [6] when the traffic is asymmetric, while performing well with symmetric traffic.

- **4×4 MIMO transceiver with IRC receiver.**
  The IRC is combined with a taxation-based rank adaptation algorithm that penalizes the use of high transmission ranks to control the network interference level [6].

- **Recovery mechanisms and link adaptation.**
  The considered recovery mechanisms are HARQ, RLC AM and TCP. A fast link adaptation that uses the latest 5 channel measurements to extract the most appropriate MCS and transmission rank is used. The MCS is extracted from a SINR-to-MCS mapping table according to a BLER target of 10%. OLLA is not considered.

- **Ideal SI cancellation.**
  Given the low transmit power (10 dBm) and the short distance among nodes (10×10 m^2 cell size), taking this assumptions becomes appropriate. This consideration allows for evaluating the upper bound of the gain that FD can provide over a fully dynamic TDD system. Therefore, if the evaluation shows that such gain is not significant, it may indicates that FD may not be the most suitable technology for interference-limited scenarios. Finally, the modeling of SI cancellation is out of the scope of this work.

- **Random UE paring and round robin scheduler.**
  For the BS FD case, the pairing of UEs is done randomly. Therefore, the results presented in this work may slightly improve if a smarter UE pairing algorithm is used. However, the difference in performance would not be significant given the small cell size and hence the marginal UE coupling loss.

2 Main Findings

The analysis of the conditions under which the theoretical FD throughput gain is obtained is provided. Such gain is possible to be achieved under specific conditions, namely ideal ICI, full buffer traffic (100% probability of exploiting FD) and scheduling of same UE in UL and DL to avoid intra-cell interference, i.e., UE-to-UE interference. In that specific case, a delay reduction of 50% is demonstrated, thus proving also the potential of FD in improving the latency of the system.

The traffic asymmetry dictates the probability that FD can be exploited, and thus the gain that FD can provide. Then, when the traffic allows to simultaneously transmit and receive, the number of interfering streams are doubled. Consequently, the ICI is increased and the FD performance is negatively affected. This situation is described in Figure 4.13 for the bidirectional FD case. In the BS FD scenario there is the impact of the intra-cell interference, which is also present in an isolated cell and impacts the FD gain.

The performance of the bidirectional FD and the BS FD is different. Consequently, the main findings for each FD type are presented separately. In the first place, the BS FD scenario suffers from intra-cell interference, which means that the DL user is highly impacted by the transmission of the user scheduled in UL. In the second place, the schemes used to decide the most appropriate transmission direction are different for bidirectional FD and BS FD. The former exploits FD with a single user every time there is data at both ends. The latter decides the optimal transmission direction of each UE according to the *load fairness based* scheme [6] and then decides which users can be scheduled to exploit FD.

Part of the analysis is dedicated to the interaction between TCP and FD. TCP is a well-known protocol to provide a reliable communication by limiting the amount of transmitted data based on TCP acknowledgments [7]. The drawback of this protocol is a reduction in throughput and an increase of the latency. FD shows potential to speed-up the protocol since the acknowledgments can be transmitted without delay, by simultaneously transmitting and receiving. Results presented in Paper C show that FD is able to accelerate the TCP protocol and mitigate its drawbacks when the ICI is not the main

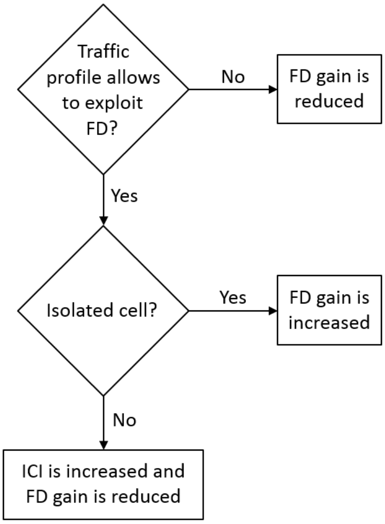


Fig. 4.13: Flow chart describing how the FD gain behaves in the bidirectional FD case.

limiting factor. However, under strong ICI, the benefits of FD in speeding-up the TCP protocol are hidden by the increased interference, since it limits the amount of transmitted data with the selection of a lower MCS and transmission rank.

Bidirectional FD

In an isolated cell only traffic profile affects the FD performance, since ICI is not present. Results show that an increase of the offered load leads to a higher probability of exploiting FD. Consequently, the FD gain in both throughput and delay also increases with the offered load. However, such gain is always below the theoretical one. In case of asymmetric traffic, assuming a traffic ratio of 6DL:1UL, the UL (lightly loaded link) gets higher gains than the DL (heavily loaded link), both in terms of throughput and delay. This behavior comes from the fact that with HD, the lightly loaded link gets 1 out of 7 opportunities to be scheduled, on average, assuming a traffic ratio of 6DL:1UL. On the other hand, with FD, the lightly loaded link gets 7 out of 7 opportunities to be scheduled, in average. This means that the transmission opportunities for the lightly loaded link are 6 times more with FD. On the other hand, the heavily loaded link gets an extra transmission opportunity with FD, on average. This is the reason why the achieved gain is lower than the lightly loaded link one.

In a multi-cell scenario, both ICI and traffic profile affect the FD performance. Paper C first analyzes the isolated effect of the ICI. Simulation results show that the 100% throughput gain is achievable in case of an isolated cell and full buffer traffic, but then, as the interference level increases, the FD gain is reduced fast. For example, for the 5$^{\text{th}}$ percentile of the throughput, which represents the users in outage, the FD gain is approximately 10% when the interference conditions are nearly the same as in an open space scenario.

Then, when the FD performance is affected by the joint impact of the ICI and the traffic profile, results show a trade-off between the possibility of exploiting FD, and the gain that such technology can provide over dynamic TDD. If simultaneous transmission and reception can be exploited in almost each TTI, dynamic TDD may perform better than FD. Consequently, this results indicate that a smart scheme that combines HD and FD transmission mode may be the most appropriate solution for improving the capacity in indoor small cells.

Base station FD

While in the bidirectional FD case the DL and UL performance is nearly the same in case of symmetric traffic, the intra-cell interference in the BS FD scenario leads to different performance in the two links, even assuming symmetric traffic. Furthermore, the used distributed rank adaptation algorithm [6] extracts the transmission rank according to the incoming interference. Consequently, the UL direction behaves more aggressively than the DL, thus increasing the harm to the DL performance.

In an isolated cell, the gain of BS FD is upper bounded by the bidirectional FD result, as shown in Figure 4.14. If the two scheduled UEs can be paired with infinite coupling loss, i.e., at infinite distance between them, the intra-cell interference would be zero. Therefore, both scenarios would be equivalent.

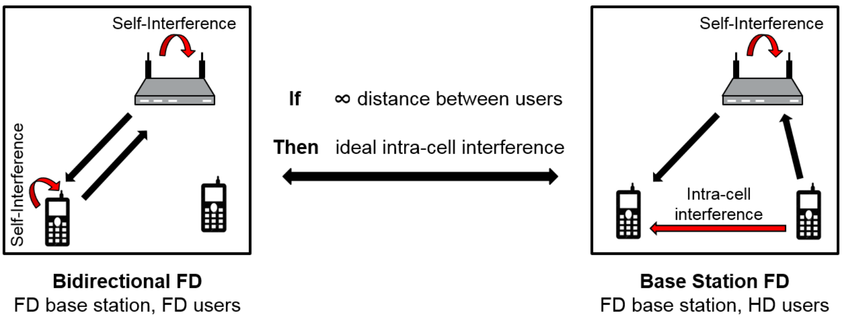


Fig. 4.14: Equivalence between bidirectional FD and base station FD.

In the multi-cell scenario, where the ICI, the intra-cell and the traffic profile affect the FD performance, results show that the DL direction would perform better with HD transmission mode. On the other hand, the gain in UL is rather minor, thus not compensating the losses in DL. This situation is observed with both symmetric and asymmetric traffic.

The main conclusion of this part of the dissertation is that FD may not be the most appropriate technology for interference limited scenarios, such as the studied 5G indoor small cell network, since simultaneous transmission and reception doubles the number of interfering streams. Nevertheless, FD shows potential in applications where the lightly loaded link needs to be enhanced. In that case, FD is able to improve both throughput and delay.

3 Included articles

A journal article composes the main body of this part of the thesis. It is a comprehensive extended and unified version of the four conference papers included in the Appendix of the dissertation.

Paper C: Analyzing the Potential of Full Duplex in 5G Ultra-Dense Small Cell Networks

The contributions of this journal article are twofold. In the first place, an evaluation of the SI cancellation capabilities of current FD radios is performed using the Nokia Solution and Networks testbed in Ulm (Germany). The evaluation is done at 2.4 GHz, with a system bandwidth of 80 MHz and a transmit power of 23 dBm. However, it is applicable to 3.5 GHz, a bandwidth of 200 MHz and a transmit power of 10 dBm. The propagation conditions are similar at these two frequencies and the transmit power reduction makes SI cancellation easier. Increasing the bandwidth leads to a higher receiver noise power and a more complex design of the transceiver.

In the second place, a detailed performance evaluation via system level simulations is provided. The analysis of the ICI and traffic profile constraints is performed individually and jointly, in order to quantify the impact of each constraint. The traffic ratios described in the previous section are used, as well as several wall penetration losses to understand the ICI effect on the FD performance. Finally, the interaction between FD and TCP is provided.

The article provides a detailed view of the FD performance in indoor small cell network, showing that in interference limited scenarios, FD provides limited gain, both in throughput and delay.

References

[1] P. Mogensen *et al.*, "Centimeter-wave concept for 5G ultra-dense small cells," in *IEEE 79th Vehicular Technology Conference (VTC Spring)*, May 2014.

[2] M. Heino *et al.*, "Recent advances in antenna design and interference cancellation algorithms for in-band full duplex relays," *IEEE Communications Magazine*, vol. 53, no. 5, pp. 91–101, 2015.

[3] K. M. Thilina *et al.*, "Medium access control design for full duplex wireless systems: challenges and approaches," *IEEE Communications Magazine*, vol. 53, no. 5, pp. 112–120, May 2015.

[4] W. W. I. N. Radio, "WINNER II channel models," February 2008.

[5] 3rd Generation Partnership Project TR 36.814 V9.0.0, "Further advancements for E-UTRA physical layer aspects (Release 9)," March 2010.

[6] D. Catania, "Performance of 5G small cells using flexible TDD," Ph.D. dissertation, Department of Electronic Systems, Aalborg University, October 2015.

[7] J. Postel, "Transmission control protocol," September 1981, updated by RFCs 1122, 3168, 6093, 6528, [Online]. Available: http://www.ietf.org/rfc/rfc793.txt.

Paper C

Analyzing the Potential of Full Duplex in 5G Ultra-Dense Small Cell Networks

Marta Gatnau Sarret, Gilberto Berardinelli, Nurul Huda Mahmood, Marko Fleischer, Preben Mogensen, Helmut Heinz

The paper has been submitted to the
EURASIP Journal on Wireless Communications and Networking - Special issue: Full-Duplex Radio: Theory, Design, and Applications, 2016.

Abstract

Full duplex technology has become an attractive solution for future 5th Generation (5G) systems for accommodating the exponentially growing mobile traffic demand. Full duplex allows a node to transmit and receive simultaneously in the same frequency band, thus, theoretically, doubling the system throughput over conventional half duplex systems. A key limitation in building a feasible full duplex node is the self-interference, i.e., the interference generated by the transmitted signal to the desired signal received on the same node. This constraint has been overcome given the recent advances in the self-interference cancellation technology. However, there are other limitations in achieving the theoretical full duplex gain: residual self-interference, traffic constraints and inter-cell and intra-cell interference. The contribution of this article is twofold. Firstly, achievable levels of self-interference cancellation are demonstrated using our own developed test bed. Secondly, a detailed evaluation of full duplex communication in 5G ultra-dense small cell networks via system level simulations is provided. The results are presented in terms of throughput and delay. Two types of full duplex are studied: when both the station and the user equipments are full duplex capable, and when only the base station is able to exploit simultaneous transmission and reception. The impact of the traffic profile and the inter-cell and intra-cell interference is addressed, individually and jointly. Results show that the increased interference that simultaneous transmission and reception causes is one of the main limiting factors in achieving the promised full duplex throughput gain, while large traffic asymmetries between downlink and uplink further compromise such gain.

1 Introduction

Wireless communication is stimulating a networked society, where data is exchanged anytime, everywhere, between everyone and everything. In 2000, only 10 gigabytes of mobile data traffic was reached per month, whereas in 2015 such amount represented 3.7 billions of gigabytes [1]. This enormous traffic increase was generated by several causes: the introduction of new services and applications, the massive use of social networks and the utilization of smart devices with mobile data connection, such as smartphones and phablets, among others. The amount of carried data will continue to grow, and it is expected to be eightfold in 2020, with reference to 2015. A new 5$^{\text{th}}$ generation (5G) radio access technology is expected to accommodate the exponentially growing demand of mobile traffic. Several strategies may be considered for boosting capacity, such as cell densification or multiple-input multiple-output (MIMO) technology with a large number of antennas. Recent advances in transceiver design have also attracted the attention of the research community on full duplex (FD) technology. FD allows a device to transmit and receive simultaneously in the same frequency band, thus, theoretically, doubling the throughput over traditional half duplex (HD) systems. Given the capabilities of this technology, it is considered as a potential candidate for future 5G systems.

A 5G concept tailored for small cells was proposed in [2], optimized for dense local area deployments. The system assumes the usage of 4×4 MIMO transceivers and receivers with interference suppression capabilities. Though originally designed as a HD time division duplex (TDD) system, the proposed concept can easily support FD communication. In order to have an operational FD node, the self-interference (SI), i.e., the interference caused by the transmit antenna to the receive antenna located in the same device should be attenuated as much as possible, ideally below the receiver noise power level. Several techniques were proposed to provide high levels of self-interference cancellation (SIC) [3–6]. Recent results show that SI can be reduced of around 100 dB [6, 7]. This may suffice for considering FD a realistic option, at least according to transmit power constraints.

The promised FD throughput gain may be compromised by several limitations. First, the residual SI may still negatively affect the reception of the desired signals. In addition, the increased interference caused by FD and the traffic profile may further compromise such theoretical FD gain. FD doubles the amount of interfering streams, leading to an increased inter-cell interference (ICI). Furthermore, exploiting FD is only possible when there is data traffic in both link directions, uplink (UL) and downlink (DL).

A number of studies analyzes the FD performance in small cell scenarios [8–12] and in a macro cell network [13] based on interference levels, disregarding the type of traffic in the network. In [8], the gain that FD provides compared to HD, assuming ideal SIC, is analyzed from a signal to interference plus noise ratio (SINR) perspective. The authors conclude that the FD gain is below the promised 100%. The authors in [9, 10] study the achievable bit rate depending on different residual SIC levels and interference conditions. Both works analyze the SINR region where FD outperforms HD, concluding that in highly interfered scenarios switching between FD and HD provides the optimal results. In [11], the FD throughput performance using different type of receivers and ideal SIC in a multi-cell scenario is studied. Results show an average throughput gain of 30-40%. In [12], results comparing MIMO HD and FD are presented, assuming full buffer traffic. The authors conclude that, without interference, FD can provide up to 31% and 36% gain in terms of throughput and delay, respectively, while in case of interfered scenarios, HD may outperform FD due to MIMO spatial multiplexing gains. A power control algorithm to maximize the sum rate of DL and UL via an efficient switching between HD and FD is proposed in [13]. The authors show that there is a SINR region where HD outperforms FD.

The impact of the traffic type is addressed in the studies [6, 7, 14–17]. The authors in [14] propose a hybrid FD/HD scheduler that selects the mode that maximizes the network throughput. The evaluation is carried out considering asymmetric traffic, showing that FD always outperforms HD. However, a strong isolation between the cells is assumed, which may downgrade the ICI impact. Malik et al. propose a power control algorithm to accommodate asymmetric traffic [15]. The proposed scheme, evaluated in a single cell scenario, shows an improvement in DL at the expenses of lowering the UL rate. The authors in [16] study the impact of symmetric and asymmetric traffic in a multi-cell scenario. Throughput results show that the FD gain reduces with the perceived ICI and the traffic ratio. The authors in [6] conclude that in dense deployment of small cells, where transmit powers are low and distances among nodes are short, 100 dB of SIC is sufficient to consider ICI as the main limiting factor for achieving the promised FD gain. Moreover, they remark that large asymmetric traffic ratios between DL and UL data may compromise the usage of FD and hence its gain. These challenges are also described in [7, 17].

The above-mentioned works study the performance of FD assuming User Data Protocol (UDP) traffic. However, most of the Internet traffic is carried over Transport Control Protocol (TCP) flows, with a small percentage of UDP flows [18]. TCP [19] is used to provide a reliable communication and reduce packet losses. Its congestion control mechanism limits the amount of data that can be pushed into the network, based on the reception of positive acknowledgments (ACKs) [20]. This procedure causes an increase in the delay and a reduction of the system throughput. FD may mitigate such drawbacks since it may allow to accelerate the TCP congestion control mechanism, given the possibility of transmitting and receiving simultaneously. It is important to notice that the previously mentioned works disregards the usage of features such as link adaptation and recovery and congestion control mechanisms.

In this paper, we perform a system level evaluation of the full duplex performance in dense small cells, where the impact of the traffic profile and the inter-cell and intra-cell interference is addressed, individually and jointly. The study is carried using a system level simulator which implements both the lower and the upper layers of the Open Systems Interconnection (OSI) model, and features mechanisms such as link adaptation and recovery and congestion control mechanisms. The contribution of this paper is twofold. Firstly, achievable levels of SIC are demonstrated using our own developed test bed. Secondly, a detailed evaluation of FD communication in 5G ultra-dense small cell networks is provided. Two types of FD communication are studied: when both the base station (BS) and the user equipment (UE) are full duplex capable, and when only the BS is able to exploit simultaneous transmission and reception. We consider the cases where the traffic is symmetric in DL and UL and when the offered load between both links is asymmetric. Furthermore, the analysis of the traffic constraints is provided with both TCP and UDP traffic. The results are presented in terms of throughput and delay and they show that the increased interference that simultaneous transmission and reception causes is one of the main limiting factors in achieving the promised full duplex throughput gain. Large traffic asymmetries between DL and UL further compromise such gain. Nevertheless, FD shows potential in asymmetric traffic applications where the lightly loaded needs to be improved, both in terms of throughput and delay.

The structure of the paper is as follows: Section 2 presents our own developed test bed and the most recent results; Section 3 describes the envisioned 5G system featuring FD communication; Section 4 introduces the simulation environment, including the simulation tool and the simulation setup; Section 5 discusses the results; Section 6 describes the future work and finally Section 7 concludes the paper.

2 Self-Interference Cancellation

A FD node generates a self-interference signal power that could easily exceed the power of the desired signal by 100 dB or more [6]. For this reason, providing a high level of SIC is a fundamental requirement to build an operational FD node. In order to identify the potential limits of SIC, we have developed a demonstrator system at Nokia Solutions and Networks in Ulm. The concept proposed in [21] and depicted in Figure C.1 has been build and studied. Such concept consists of a pre-mixer with an additional transmit chain for analogue compensation and a final digital cancellation stage. Up to 100 MHz contiguous bandwidth can be handled by the system, which is typically operating in the 2.4 GHz band. The practical antenna isolation from the transmitter (TX) to the receiver (RX) is $\sim$50 dB, and is based on physical antenna separation, as shown in Figure C.2, and the appropriate passive means. Additionally, to limit the impact of the phase noise, it is essential to provide a common clocking domain, same mixer stage for up and down conversion and radio frequency (RF) delay compensation [22].

The used hardware has the capability of canceling maximum $\sim$70 dB for a 20 MHz LTE signal (LTE20) with respect to phase noise. The achievable active cancellation is limited by the power amplifier (PA) non-linearity and the auxiliary transmitter resolution. Under these two limitations, a total active cancellation gain of 63 dB for a LTE20 signal could be demonstrated, with a joint usage of the analogue cancellation and the time domain digital cancellation stages. There are two approaches to achieve such gain. The first one is the option A depicted in Figure C.1, that uses a nonlinear intermodu-

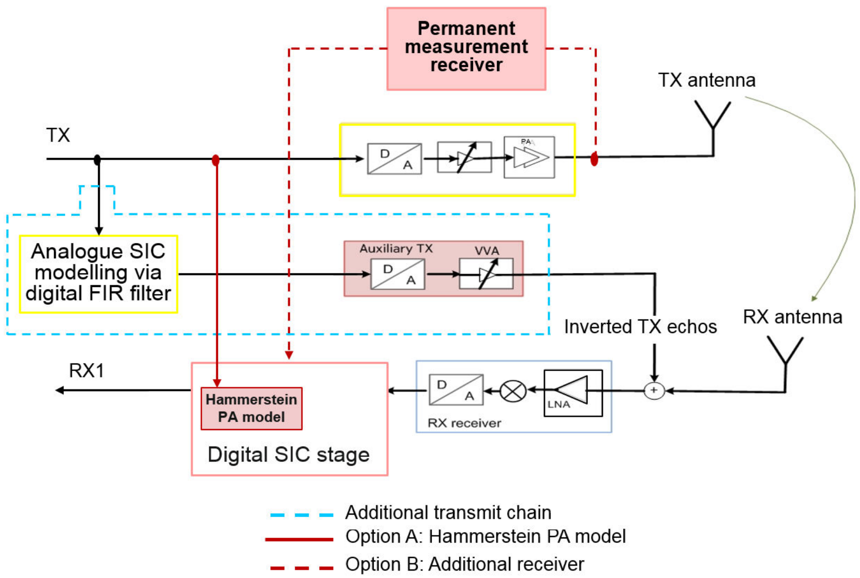


Fig. C.1: Diagram showing the self-interference cancellation procedure.

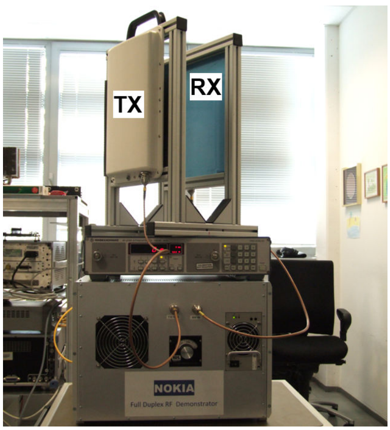


Fig. C.2: Self-interference cancellation test bed from Nokia in Ulm.

lation approach via Hammerstein PA model [23] within the digital SIC stage. This option employs the digital transmit signal as input [24]. The second approach, plotted as the option B in Figure C.1, uses the PA signal as direct input to the digital SIC stage with the need of an additional receiver, named the *permanent measurement receiver*, which contains the transmitter RF impairments and is common in a typical commercial RF design for PA linearization purposes.

The design shown in Figure C.1 also requires the usage of an additional transmit chain. Such additional transmit block has the purpose to protect the receiver against saturation, and it has the advantage that scales only with the number of transmit antennas, which is highly appropriate in MIMO systems. Furthermore, to avoid extra complexity and provide simpler hardware integration, all transmitted antenna streams are input to the same analogue and digital SIC modeling block.

A total cancellation of $\sim$100 dB for a 20 dBm 4$\times$LTE20 signal has been demonstrated, as shown in Figure C.3. The result shows the SI level close to receiver noise floor limits (-85 dBm, considering a noise figure of 10 dB), thus demonstrating the potential of the described hardware concept. Achieving a large level of SIC at higher frequencies beyond today's LTE limits, wider frequency bands of hundreds of MHz, and large number of antennas is still an open research topic.

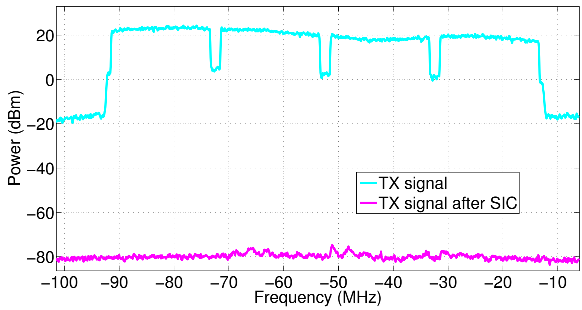


Fig. C.3: Self-interference cancellation result extracted from the test bed.

3 Full Duplex in 5G Small Cells

Featured 5G system design

Since the goal of this work is to study FD in dense small cell networks considering system level aspects, in this section we are going to describe the small cell concept which will be the reference for our evaluation.

The small cells concept presented in [2] was originally designed as a HD TDD system with orthogonal frequency division multiplexing (OFDM) as modulation scheme, but it can easily accommodate FD communication. Nodes are assumed to be synchronized in time and frequency and equipped with 4x4 MIMO transceivers with interference rejection combining (IRC) capability [25]. A novel frame structure of duration 0.25 ms is introduced, which is defined as the transmission time interval (TTI) and is shown in Figure C.4. The first two OFDM symbols are dedicated to the DL and UL control, respectively. The remaining symbols are allocated for the data, including the demodulation reference signal (DMRS) symbol, which is used for channel estimation. The IRC receiver requires information about the channel responses of the desired and the interfering signals to provide a good performance. Such channel information can be obtained by relying on orthogonal reference sequences transmitted by multiple devices in the DMRS symbol. Then, exploiting such information, it suppresses a number of the interfering streams according to the available degrees of freedom in the antenna domain [25]. Furthermore, recovery mechanisms such as hybrid automatic repeat and request (HARQ) and automatic repeat and request (ARQ) are used to deal with the residual ICI. For further details regarding the system design, the reader should please refer to [2, 26].

Using the same frame structure for both UL and DL allows for a straight-

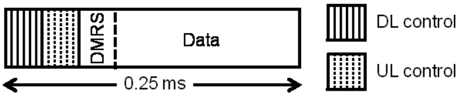


Fig. C.4: Envisioned 5G frame structure

forward extension of the envisioned 5G concept to FD transmission. Note that the control part remains as HD, in order to support different types of FD communication. The cell operations are as follows: firstly, the BS sends the scheduling grant (SG) in the DL control symbol of TTI_n. The SG includes the scheduled UE and the transmission parameters, i.e., the direction (UL or DL), the modulation and coding scheme (MCS) and the number of spatial streams used for transmission, often referred as transmission *rank*. The configuration specified in the SG is applied in TTI_{n+1} assuming a certain processing time. Consequently, there is one TTI delay between the scheduling and the corresponding data transmission. The UEs send the scheduling request (SR) in the UL symbol, including buffer information, HARQ feedback and the MCS and rank derived from their channel measurements. Notice that there is a delay between the instant when the channel is measured and the TTI when the transmission occurs, which may affect the link adaptation procedure. In addition, since the transmission direction may change at each TTI, creating sudden changes in the interference pattern, it further compromises the link adaptation procedure.

In this study, two FD techniques are investigated, which are depicted in Figure C.5. In the figure, full lines represent the intended transmissions and dashed lines refer to interfering streams. Figure C.5a shows the bidirectional FD case, where both the BS and the UEs are FD capable. In this case, the communication is performed always between the same pair BS-UE, and therefore both nodes only perceive their own SI. The second FD mode is the BS FD, shown in Figure C.5b, where only the BS is FD capable. In this case, the DL and UL scheduled UEs are different. Therefore, the intra-cell interference, i.e., the interference from the UL UE to the DL UE, also affects the system performance. Notice that in case of a multi-cell scenario, the ICI would affect the performance of the system.

When FD is exploited, the number of interfering streams compared to

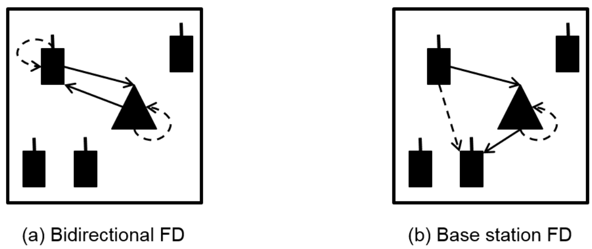


Fig. C.5: Full duplex types. The dashed lines represent interference and the full lines the desired signal.

HD is doubled. Therefore, the network interference is larger in FD than in HD, and the performance of the IRC receiver may be jeopardized since it may not have enough degrees of freedom in the antenna domain to deal with the enlarged interference. On the other hand, FD transmission will be only exploited in case there is data available at both BS and UE. Hence, the theoretical gain that FD can provide over HD may be compromised by the following limitations:

- Residual self-interference. For a FD node to be operational, a high level of isolation between the transmitting antenna and the receiving antenna located in the same device is required. Current levels of achievable SIC may not sufficient to bring the SI power below the receiver noise power level, thus leaving a residual interference that affects the SINR.
- Increased interference. The number of interfering streams with FD are doubled compared to HD, thus leading to an increased network interference (inter-cell and intra-cell interference). Then, when the interference is stronger, the data rates are lower and consequently a larger number of TTIs is needed to transmit the same amount of data.
- Simultaneous UL and DL data. The availability of simultaneous UL and DL traffic dictates the probability of exploiting FD. Hence, large asymmetries between UL and DL may jeopardize the FD gain.

Radio resource management architecture

In order to support FD communication, a design for the radio resource management (RRM) module, shown in Figure C.6, is proposed. The RRM module decides which transmission mode is going to be used at each TTI (HD or FD), the transmission direction in case of HD, and which is(are) the node(s) that is(are) going to be scheduled. The module is divided into two blocks to reduce the complexity and the computational time. As the first step, the *direction decision block* decides the optimal transmission direction per each UE. This decision is extracted based on the information received from the physical (PHY), medium access control (MAC) and radio link control (RLC) layers. Such information includes SINR measurements, HARQ feedback, buffer status reports and link quality information provided by each UE to the BS. The set of decisions for all UEs extracted from the *direction decision block* is sent to the *user decision block*. Then, as the second step, the transmission mode (HD or FD) and the UE(s) to be scheduled will be decided by the *direction decision block*.

The optimal transmission direction, determined by the *direction decision*

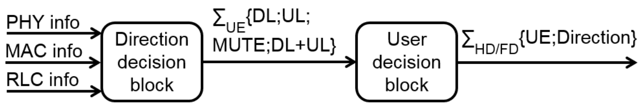


Fig. C.6: RRM module. The figure shows the design of the RRM module that supports both types of FD communication and HD.

block, can be *DL*, *UL*, *DL+UL* or *MUTE*, and it is extracted differently depending on the type of communication:

- **HD and BS FD**: for these two cases, the procedure to extract the optimal link direction is the same. In BS FD a UE cannot be scheduled in both links because it operates in HD transmission mode. The transmission direction is decided based on the offered load of each link, and thus the amount of dedicated resources is proportional to the offered load. For example, let us assume asymmetric traffic, where the highly loaded link (DL) offers k times more load than the lightly loaded link (UL). In this case, the DL would get, in average, k times more resources than the UL, and it would have higher priority. Consequently, the UL would have to wait longer to be scheduled. Furthermore, the algorithm also takes into account fairness, by granting a minimum amount of resources to a link, in order to avoid its starvation. For more details about the used scheme, the reader should refer to [27]. The possible output directions in this case are *DL*, *UL* or *MUTE*. The latter corresponds to the case where both UL and DL buffers are empty.

- **Bidirectional FD**: the transmission direction is based only on the buffer state. For each user, the *direction decision block* checks if there is data in both the DL and UL buffers. In case of bidirectional FD, simultaneous transmission and reception will only be exploited in case a UE can be scheduled in both links, which will happen only when both UL and DL buffers are filled with data. Then, if this is the case, the transmission direction for that user is *DL+UL*. Otherwise, it is *DL*(*UL*) if the UL(DL) buffer is empty and the DL(UL) is not, or *MUTE* if the UL and DL buffers are both empty.

In case of BS FD, a FD transmission is performed if two different UEs with opposite link directions can be scheduled; otherwise, the TTI is going to be HD. In case of bidirectional FD, it will be possible to exploit FD if at least one user has associated the *DL+UL* state. Note that in both cases, scheduling a FD transmission is always given priority over scheduling a HD one.

Interaction between full duplex and TCP

TCP [19] is a high layer protocol that aims at providing reliability by using a congestion control mechanism [20]. The amount of data that can be sent through the channel is limited based on the reception of positive acknowledgments (ACKs). The feature in charge of controling such limitation is the congestion window, shown in Figure C.7. Within the *Slow Start* stage, the congestion window grows exponentially according to the received TCP ACKs. When the congestion window reaches the *Slow Start Threshold*, the *Congestion Avoidance* phase starts. In this stage, the growth of the congestion window is linear, following the same principle as the *Slow Start* phase based on the reception of TCP ACKs. However, the TCP protocol has an inherent impact on the system throughput and delay because the amount of transmitted data is limited by the reception of ACKs, which will increase only if the channel conditions are favorable.

We believe that the TCP drawbacks may be mitigated by FD. Given the ability of simultaneous transmission and reception, the congestion window might grow faster and it might reach the *Congestion Avoidance* phase sooner, where a larger amount of data is transmitted within a single TTI. For clarification, an example of the congestion window growth for HD and FD in a single cell scenario with one AP and one UE is shown in Figure C.8. Both nodes have a 2 megabytes file to transmit and FD is exploited in all TTIs. In this example, shadowing and fast fading have been disabled to avoid the im-

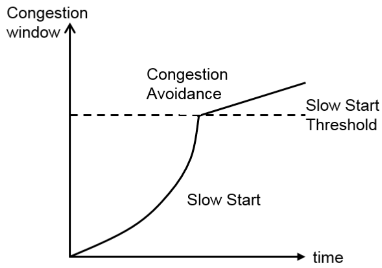


Fig. C.7: TCP congestion window

pact of the channel. The general simulation parameters are listed in Table C.1 and they will be further discussed in Section IV. From the figure we observe that FD transmits the 2 megabytes file faster than HD because the congestion window in case of FD is able to grow faster. In this example, the transmission time is reduced by nearly 45%.

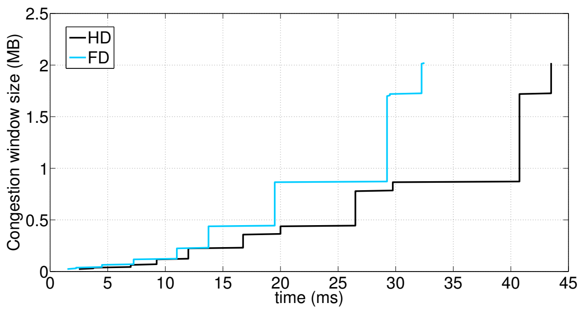


Fig. C.8: Simulation example of a congestion window growth

4 Simulation Environment

Simulation Tool

The results presented in this study are extracted from our own developed event-driven based system level simulator. It includes the design of the envisioned 5G PHY and MAC layers presented in Section II. The RLC, the TCP and the UDP mechanisms are entirely modeled, whereas the Internet protocol (IP) is only modeled as overhead. A vertical RRM layer gathering information from the PHY, MAC and RLC layers is implemented. The RRM layer includes the module described in Section II and decides the transmission parameters. The link adaptation feature extracts the most accurate MCS from the log-average of the last 5 SINR samples. The simulator supports 32 MCSs, the lowest being Quadrature Phase Shift Keying (QPSK) with a coding rate of 1/5 and the highest being 256-Quadrature Amplitude Modulation (QAM) with a coding rate of 9/10. The transmission rank can be either fixed or set dynamically according to a taxation-based rank adaptation algorithm [28]. Such algorithm runs in all nodes, and decides the rank according to the perceived interference. The goal of the algorithm is to reduce the overall network interference level by applying a higher taxation to transmissions with higher ranks. The algorithm is further detailed in [28].

At the receiver side, both desired and interfering streams arrive at the antennas and the IRC performs interference suppression based on the available degrees of freedom in the antenna domain. The SINR extracted from this procedure is input to the decoding module. Such module decides, based on a block error rate (BLER) target of 10%, whether the packet can be decoded or not. In case of failure, the HARQ mechanism will notify the RRM module that a retransmission is required. On the other hand, if the packet is successfully decoded, it is sent to the higher layers up to the statistics module where the delay and throughput are computed. A SINR soft combining model extracts the effective SINR upon retransmissions. Soft combining keeps memory of previous transmissions of the same packet to achieve SINR gain and improve the probability of correct detection [29]. The model can be expressed as follows:

$$SINR_{effective} = \sum_{i=1}^{n} SINR_i \cdot \eta^{n-1} \tag{C.1}$$

where n refers to the transmission number, $SINR_i$ is the SINR for the i^{th} transmission/retransmission of the same packet, and η is the combining efficiency, used to model the non-ideality of the combining process. In this study, it is set to 1.0 for simplicity.

The simulator includes different traffic models, such as full buffer or File Transfer Protocol (FTP) [30]. The FTP traffic model generates payloads according to a negative exponential distribution. Such payloads, defined as sessions, have an average size of 2 megabytes and arrive every $t_{inter-arrival}$ seconds. The parameter $t_{inter-arrival}$ is also generated according to a negative exponential distribution. It is composed of the period of time when the application generates the packets for a particular session (t_{on}) plus the amount of time when no packets are being generated (t_{off}). The values of t_{on} and t_{off} reflect the load in the system. So, for a fixed t_{on}, increasing t_{off} will translate into a lower load in the system, and vice versa. The carried system load dictates the network resource utilization (RU), i.e., the channel occupancy, defined as:

$$RU = \frac{\sum_{i=1}^{\text{SimTTIs}} TTI_{i=TX}}{\sum_{i=1}^{\text{SimTTIs}} TTI_{i=TX} + \sum_{i=1}^{\text{SimTTIs}} TTI_{i=MUTE}} \tag{C.2}$$

where TTI_{TX} refers to a DL HD, UL HD or FD transmission, and TTI_{MUTE} refers to the case where there is no data to be transmitted in any of the two link directions. The upper limit in the summation *SimTTIs* represents the total number of simulated TTIs. For example, a RU of 50% means that half of the time the channel is free and a RU of 100% indicates that the channel is always busy.

Several key performance indicators (KPIs) can be extracted from the simulator: SINR, statistics on the MCS and transmission rank selection, FD probability, average session throughput (TP), packet delay, etc. The session TP is defined as the amount of time required to successfully transmit a session. Then, the average session TP is the mean of all the computed session TPs. The packet delay is the time between the generation of a packet and its successful reception, including the buffering time. Finally, the probability of exploiting FD is defined as:

$$Prob\{FD\} = \frac{\sum_{i=1}^{\text{SimTTIs}} TTI_{i=FD}}{\sum_{i=1}^{\text{SimTTIs}} TTI_{i=FD} + \sum_{i=1}^{\text{SimTTIs}} TTI_{i=HD}} \tag{C.3}$$

where TTI_x refers to the type of communication performed on a TTI. Then, x can be FD or HD.

Simulation Setup

The performance of FD is evaluated in different scenarios. A single cell network is defined as a 10×10 m^2 room, containing one BS and four UEs randomly deployed. The UEs are always affiliated to the BS in the same cell (closed subscriber group). The multi-cell scenario refers to a 10×2 grid of single cell networks, as shown in Figure C.9. Ideal SIC is considered, given the current SIC capabilities [6], the short distances among nodes and the low transmit power, which is set to 10 dBm for all the nodes. The RLC mode is set to Acknowledged (AM) [31]. In our simulator, the RLC ACK is sent within the control channel to avoid generating additional overhead. The TCP implementation is New Reno [32], and it includes the recovery and congestion control mechanisms, whereas handshake procedures are not considered since they are not relevant for our studies. TCP parametrization and the remaining simulation parameters are listed in Table C.1. Finally, the selected scheme for the *user decision block* is time domain round robin, so frequency multiplexing is not considered.

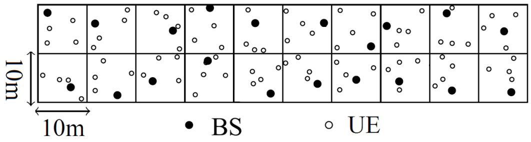


Fig. C.9: Multi-cell scenario. It corresponds to a grid of 10×2 single cell networks.

Table C.1: Simulation parameters

Parameter	Value/State/Type
System parameters	BW = 200MHz; f_c = 3.5GHz
Frequency reuse	1 (whole band)
Propagation model	WINNER II A1 w/fast fading [33]
Antenna configuration	4x4
Receiver type	IRC
Transmission power	10 dBm (BS and UE)
Link adaptation filter	Log average of 5 samples
Transmission rank scheme	Fixed or taxation-based
UL/DL decider	Metric (HD and BS FD) and traffic based (Bidirectional FD)
HARQ max retransmissions	4
HARQ combining efficiency η	1
Resource utilization	$\sim$ 25%, 50% and 75% if symmetric or asymmetric traffic 100% if full buffer traffic
RLC mode	Acknowledged
Transport protocol	UDP and TCP
Simulation time per drop	Up to 20 seconds
Number of simulation drops	50

The performance of FD is compared against that of HD. We consider two types of FTP traffic, symmetric and asymmetric. Symmetric traffic refers to the case where the offered load is the same in DL and UL (1DL:1UL). On the other hand, asymmetric traffic case corresponds to the situation in which the offered load in DL is 6 times larger than in UL (6DL:1UL). Three loads are simulated: low, medium and high, which refer to a RU of nearly 25%, 50% and 75% under ideal conditions, respectively. The results are presented in three formats: as numerical tables; as the cumulative distribution function (CDF) of the average session throughput (TP) and the packet delay; and as bar plots showing the comparison between the HD and FD performance with TCP and UDP. The latter protocol acts as a transparent layer, sending all the received data to the upper layers, without performing error checking or congestion control [34]. Finally, the gain in percentage that FD provides over HD is calculated as follows:

$$Gain_{FD}[\%] = \left(\frac{\text{FD average performance}}{\text{HD average performance}} - 1 \right) \cdot 100 \tag{C.4}$$

5 Performance Evaluation

The results provided in this section are presented in an order that aims at analyzing the impact of the increased interference caused by FD and the traffic constraints. In the first subsection, we focus on the analysis of the single cell network to avoid the impact of the inter-cell and intra-cell interference.

The multi-cell scenario will be analyzed in the second and third subsections. In first place, only the impact of ICI is quantified. For this reason, the bidirectional FD performance is analyzed by varying the penetration wall loss. Then, in the last subsection, the jointly effect of the ICI, the intra-cell interference (only for BS FD) and traffic constraints are evaluated.

Analysis of the traffic constraint limitation

In this analysis we analyze a single cell network with the transmission rank fixed to one. Bidirectional FD is considered. As a first step, the traffic generator is parametrized to generate symmetric traffic with a probability of having simultaneous traffic in UL and DL of 100%, i.e., FD can be exploited with 100% probability, and UDP is set as the transport layer. Figure C.10 shows the average cell session TP and average packet delay. From the figure we can observe that, under ideal interference conditions, the delay can be reduced by 50% and the TP can be increased by 93%, very close to the theoretical FD TP gain. This small difference in FD TP gain between the simulation results and the theoretical maximum is caused by the HD resource allocation algorithm used as a baseline, since it allocates the data optimally, as discussed in Section II. The FD gain would be 100% if the HD baseline is set to a fixed 1DL:1UL time slot allocation.

From this first result we can conclude that it is possible to achieve the

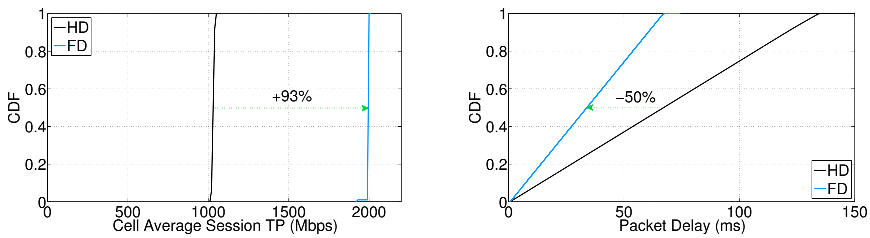


Fig. C.10: Bidirectional FD performance in a single cell network with 100% probability of exploiting FD.

promised gain from FD, but only under very specific conditions. The case of BS FD shows approximately the same performance (since the IRC receiver has sufficient degrees of freedom for suppressing the intra-stream interference given the usage of rank 1), and is not reported here. Let us evaluate the same scenario but in this case considering the low, medium and high loads introduced in Section IV. Both the symmetric (1DL:1UL) and asymmetric (6DL:1UL) traffic cases will be addressed. Figure C.11 shows the cell average session TP and the average packet delay for the symmetric traffic case. In such case, both link directions show approximately the same performance because the offered load is the same in UL and DL and interference is not present. The results show that FD always outperforms HD, and the gain that FD provides increases as the load grows. This gain increase is caused by a higher probability of exploiting FD.

Let us now consider TCP. The TCP protocol shapes the dynamics of the

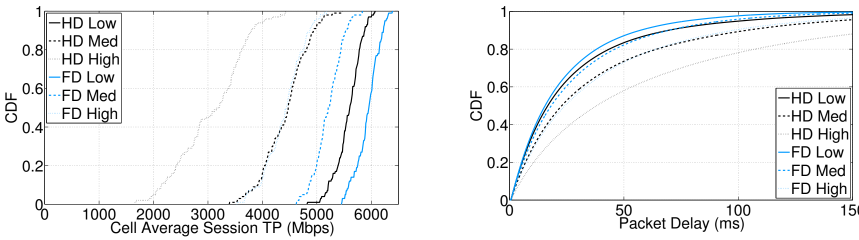
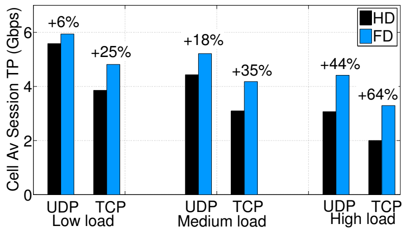


Fig. C.11: Bidirectional FD performance in a single cell network with symmetric UDP traffic.

system by limiting the amount of data that can be sent by using a congestion control mechanism. Figure C.12 shows the system performance (in terms of average cell session TP and average packet delay) with UDP and TCP, assuming symmetric traffic. The percentage numbers represent the gain that FD provides over HD. From the result we can observe that the FD gain is larger when TCP is used. The reason is twofold: firstly, FD allows the TCP congestion window to grow faster, thus being able to transmit a larger amount of data than HD, as explained in Section II; secondly, the probability of exploiting FD with TCP (from 85% up to 97%) is larger than with UDP (from 4% up to 37%). The FD probability is larger with TCP because data cannot be transmitted freely but under the constraints of the TCP congestion control mechanism, thus making the data accumulate in the buffer. In addition, since data is transmitted faster due to simultaneous transmission and reception, TCP ACKs have less chances of being piggybacked with data and hence are transmitted individually, like a normal data packet. So for example, in a single cell scenario, the number of non-piggybacked TCP ACKs with FD can be up to 2.7 times larger than with HD. Even though the TCP ACKs can be transmitted without delay with FD, they generate larger overhead if they cannot be piggybacked with data. Finally, it is important to notice that even if FD always outperforms HD in this specific scenario, the gain that FD provides is always below the theoretical one.

The asymmetric traffic case is shown in Figure C.13. Numerical results



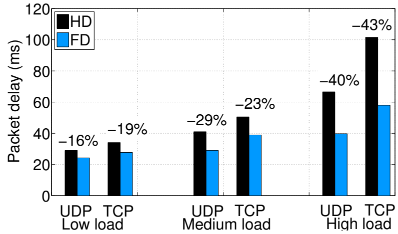


Fig. C.12: TP gain and delay reduction of bidirectional FD over HD with symmetric TCP and UDP traffic in a single cell network.

show the average session TP and packet delay in DL and UL separately, and for both UDP and TCP. First of all, we can observe that, independently of the transport layer, the gain in UL and DL is now different. This is because in HD, according to the offered load of each link, six out of seven TTIs will be allocated to DL and one to UL, in average. In FD, since UL and DL can occur at the same time, DL can obtain, in average, one extra TTI compared to HD, while the UL can get six more. The results show the same trends as the symmetric traffic case: an increase of the FD gain for a larger offered load and a higher FD gain with TCP than with UDP. It is interesting to notice that in UL at high load, FD is able to eliminate the buffering or waiting time, being able to transmit all the data from the buffer. Furthermore, the DL data can be transmitted faster since the UL TCP ACK can be transmitted immediately by exploiting FD communication. On the other hand, it generates a larger overhead due to not being able to piggyback it with data.

Analysis of the inter-cell interference limitation

To analyze how ICI affects the FD performance, we consider the multi-cell scenario shown in Figure C.9. The traffic model is now set to full buffer since we want to avoid the impact of the traffic constraints in the FD gain, the transport layer is UDP and the transmission rank is fixed to one for simplicity. The results are extracted by varying the penetration wall loss, which dictates the isolation between the cells, from 0 dB to 25 dB. In case the penetration wall loss is set to 0 dB, the simulated scenario would correspond to an open space network, while if it is set to 25 dB it would refer to an almost isolated cell. The TP gain that FD provides over HD is depicted in Figure C.14. In the figure, the 5th, 50th and 95th percentiles gain are presented. The 5th percentile represents to the outage performance, i.e., the performance of the users perceiving the worst channel conditions. The results show that, as

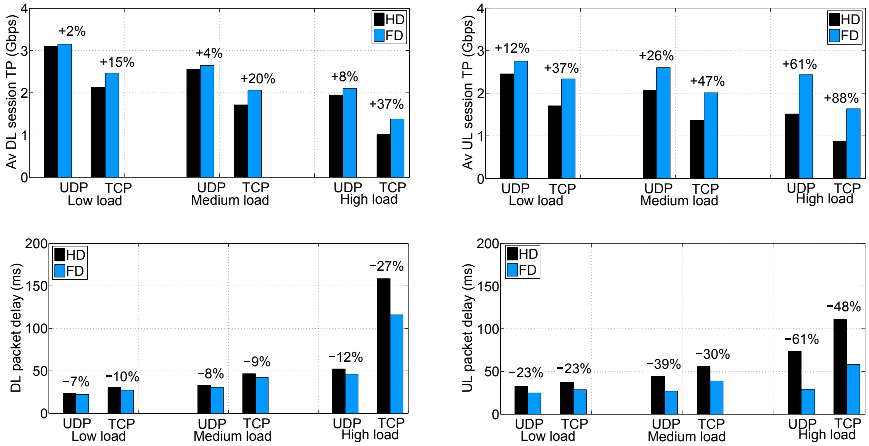


Fig. C.13: TP gain and delay reduction of bidirectional FD over HD with asymmetric TCP and UDP traffic in a single cell network.

the isolation among cells increases, the gain that FD provides over HD increases. When the isolation among cells lowers, FD perceives larger ICI than HD because FD doubles the amount of interfering streams compared to HD. Notice that, even when the penetration wall loss is set to 0 dB, corresponding to the worst case, FD shows an improvement of 9% over HD for the outage users. In addition, the 95th percentile, defining the users perceiving the best channel conditions, is improved by 56% with FD.

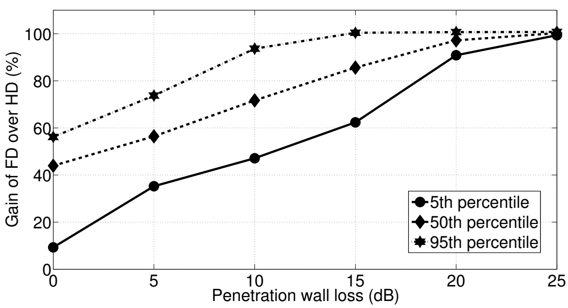


Fig. C.14: Throughput gain of FD over HD in the multi-cell scenario with full buffer traffic.

FD performance under the impact of increased interference and traffic constraints

In this last analysis, the jointly impact of the increased interference caused by FD communication and the traffic constraints is analyzed. To that purpose, the multi-cell scenario with symmetric (1DL:1UL) and asymmetric (6DL:1UL) traffic and the rank adaptation algorithm described in Section IV are used. The performance of HD and both types of FD communication with UDP and TCP for the medium load case (HD RU $\approx$ 50%) is presented.

Figure C.15 shows the CDF of the DL and UL average session TP. Starting with the UDP performance, we observe that the UL and DL results with bidirectional FD are nearly the same. This is because the traffic is symmetric and thus both links would get the same amount of resources, and the interference conditions perceived by all the nodes is in average the same. In this case, FD performs always better than HD, even showing an improvement of the outage users performance. However, for the BS FD case, the UL and DL directions show rather difference performance. The reason of such difference is the intra-cell interference. The DL user is highly interfered by the UL users. Therefore, the perceived interference conditions in the two links are different, and this affects the choice of MCS and transmission rank. Furthermore, the number of DL retransmissions is larger than in UL, creating an originally non-existing asymmetry in the traffic. This asymmetry causes the over-prioritization of the DL over the UL because the buffer size is larger, even though the offered load is the same. In this case, the DL is negatively impacted by the use of FD, since HD performs always better. The UL direction is barely optimized, showing that the outage users are negatively affected by the use of FD, while from the 50$^{\text{th}}$ percentile, FD outperforms HD. By analyzing the system behavior with TCP, we can observe that the results for the bidirectional FD communication are completely the opposite as the ones with UDP. The reason for this turnaround is the increased interference caused by a probability of exploiting FD of 81%, compared to 15% with UDP. Doubling the amount of interfering streams in almost every single TTI causes an average SINR difference of 9 dB between HD and FD, which has a repercussion on the MCS selection, the transmission rank and the link failures. HD is able to use a 12 times higher rate than FD, in average. Furthermore, the IRC receiver performance is jeopardized in case of FD given the increased interference, making the system limited to use rank 1, while HD is still able to switch to rank 2 sporadically. Finally, the HARQ retransmissions are triggered more often with FD because the SINR reaches a level below the decodable threshold. For BS FD, the TCP trends are similar to the UDP ones because the probability of exploiting FD is nearly the same (25% in UDP and 32% in TCP). We can observe that the DL direction shows the best performance with HD, while the UL in this case is even closer than in case of UDP. Notice that the RRM algorithm that decides the optimal transmission direction is different for bidirectional FD and BS FD. This is a further reason for their performance difference, besides the presence of intra-cell interference in BS FD.

The CDF of the average packet delay is shown in Figure C.16. We can

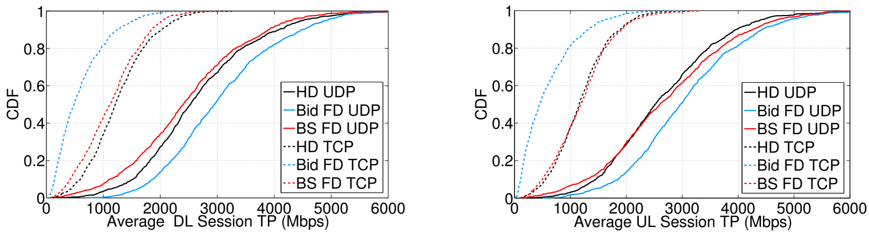


Fig. C.15: Throughput performance of HD, bidirectional FD and BS FD with symmetric TCP and UDP traffic in the multi-cell scenario.

observe that the delay shows approximately the same trends as the TP results. Bidirectional FD can reduce the delay when the transport protocol is UDP, while in case TCP is used, the delay increases dramatically. On the other hand, BS FD shows nearly the same results for UDP and TCP, but in this case, any of the two link directions can be improved by using FD. Finally, the RU is depicted in Figure C.17. The figure shows that bidirectional FD is able to reduce the channel occupancy in case UDP is used. However, with TCP, such type of FD requires a larger amount of TTIs to transmit the same amount of data than HD. In case of BS FD, the channel occupancy is slightly larger than with HD, due to the performance of the DL direction.

The numerical results when the traffic is asymmetric are presented in

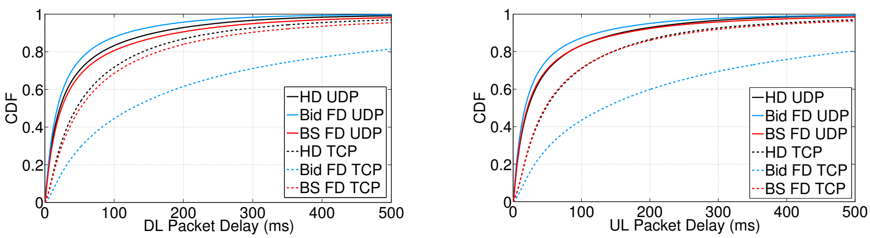


Fig. C.16: Delay performance of HD, bidirectional FD and BS FD with symmetric TCP and UDP traffic in the multi-cell scenario.

Table C.2. From previous analysis, we would expect that the UL direction can always be significantly improved by the use of FD, since with HD it gets less transmission opportunities. Starting with the bidirectional FD case, we observe that simultaneous transmission and reception can always improve

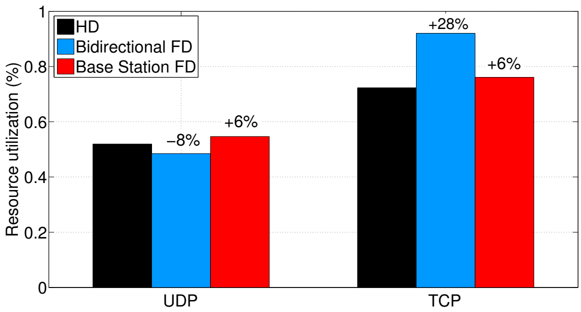


Fig. C.17: RU of HD, bidirectional FD and BS FD with symmetric UDP and TCP traffic in the multi-cell scenario.

the system TP and delay in case UDP is used, specially the UL direction. However, when TCP is enabled, the same situation as in the symmetric traffic case is repeated. An SINR difference of 9 dB in average causes the FD system to perform worse than HD. Not even the UL, which is the lightly loaded link that gets the chance of being transmitted immediately with FD can be improved. Even though FD allows the TCP congestion window to grow faster because the TCP ACKs can be transmitted immediately, the increase in the network interference has an important impact on FD. The increase of the number of HARQ retransmission and the reduction in MCS and transmission rank compared to HD compromises the performance of FD in ultra-dense small cell scenarios. Notice that such large numbers are also dictated by the fact that the absolute delay results are very low. Moving to the BS FD case, we also observe a similar behavior as in the symmetric traffic case. The main difference is that with asymmetric traffic, we can detect an improvement of the lightly loaded link. However, the gain is rather limited. This is because the DL direction, affected by the intra-cell interference, increases the HARQ retransmissions and thus enlarges the originally 6DL:1UL asymmetry. Consequently, the DL is even more over-prioritized, thus affecting indirectly the UL performance.

From this intensive analysis of the FD performance in 5G ultra-dense

Table C.2: TP gain and delay reduction of bidirectional FD and BS FD over HD with asymmetric TCP and UDP traffic in the multi-cell scenario.

Communication type	Traffic	DL TP	UL TP	DL delay	UL delay
Bidirectional FD	**UDP**	+4%	+18%	-8%	-35%
	TCP	-64%	-44%	+548%	+155%
BS FD	**UDP**	-2%	+14%	+11%	-18%
	TCP	-12%	+16%	+30%	-21%

small cell networks, we can conclude that in interference limited scenarios, the use of FD is not always beneficial. The fact that simultaneous transmission and reception doubles the amount of interfering streams has a negative impact on the system performance. However, a combination of FD and HD transmission modes may provide the optimal system performance. Finally, results indicate that FD shows potential in asymmetric traffic applications where the lightly loaded link needs to be enhanced.

6 Future Work

Future research could analyze how non-ideal self-interference cancellation and larger traffic asymmetries between the UL and DL directions impact the results presented in this work, since they provide an upper bound of the achievable FD gain. Furthermore, the use of full duplex could be studied in the context of macro-cell scenarios, where on the other side, the self-interference is much higher in macro BS and can jeopardize the performance. In this case, MAC schemes that take into account the distance among the nodes and power control can be designed to get the most benefit from the usage of full duplex communication. Finally, the potential of simultaneous transmission and reception to provide fast discovery on the context of Device-to-Device (D2D) communication can be studied.

The findings presented in this paper could be applied to design a hybrid HD/FD scheduling mechanism that obtains the maximum benefit from both types of communication.

7 Conclusions

This work analyzes the potential of full duplex technology in enhancing the throughput and delay of 5G ultra-dense small cell networks. The self-interference cancellation capabilities are investigated using our own developed test bed. The carried experiment proves that up to $\sim$100 dB of isolation between the transmitting and the receiving antennas placed in the same device are currently achievable, according to the used setup. Then, the potential of full duplex communication is studied via detailed system level simulations. Results show that achieving the theoretical double throughput gain that FD promises can only be achieved under specific assumptions, namely ideal self-interference cancellation, isolated cells and full buffer traffic model. However, the promised gain is reduced when realistic assumptions, such as traffic constraints and the inter-cell interference, are considered. Simulations prove that when the traffic profile allows the system to use full duplex communication, the increased interference caused by simultaneous transmission and reception becomes the main limiting factor in achieving the theoretical FD throughput gain. In case where only the base station is full duplex capable, the intra-cell interference has a significant impact on the system performance.

This work proves that full duplex communication is able to accelerate the dynamics of TCP and mitigate the drawbacks introduced by such protocol. Furthermore, results such technology has a compelling potential for applications with asymmetric traffic where the lightly loaded link can benefit in terms of throughput and delay.

References

[1] Cisco: Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2015-2020 (2016)

[2] Mogensen, P., *et al.*: Centimeter-wave concept for 5g ultra-dense small cells. In: 2014 IEEE 79th Vehicular Technology Conference (VTC Spring), pp. 1–6 (2014). doi:10.1109/VTCSpring.2014.7023157

[3] Choi, J.I., *et al.*: Achieving single channel, full duplex wireless communication. In: Proceedings of the 16th Annual International Conference on Mobile Computing and Networking (Mobicom), pp. 1–12. ACM, New York, USA (2010). doi:10.1145/1859995.1859997

[4] Aryafar, E., *et al.*: MIDU: Enabling MIMO full duplex. In: Proceedings of the 18th Annual International Conference on Mobile Computing and Networking (Mobicom), pp. 257–268. ACM, New York, USA (2012). doi:10.1145/2348543.2348576

[5] Hong, S., *et al.*: Applications of self-interference cancellation in 5G and beyond. IEEE Communications Magazine **52**(2), 114–121 (2014). doi:10.1109/MCOM.2014.6736751

[6] Heino, M., *et al.*: Recent advances in antenna design and interference cancellation algorithms for in-band full duplex relays. IEEE Communications Magazine **53**(5), 91–101 (2015). doi:10.1109/MCOM.2015.7105647

[7] Thilina, K.M., *et al.*: Medium access control design for full duplex wireless systems: challenges and approaches. IEEE Communications Magazine **53**(5), 112–120 (2015). doi:10.1109/MCOM.2015.7105649

[8] Xie, X., Zhang, X.: Does full-duplex double the capacity of wireless networks? In: Proceedings of IEEE INFOCOM, pp. 253–261 (2014). doi:10.1109/INFOCOM.2014.6847946

[9] Day, B.P., *et al.*: Full-duplex bidirectional MIMO: Achievable rates under limited dynamic range. In: Conference on Signals, Systems and Computers (ASILOMAR), pp. 1386–1390 (2011). doi:10.1109/ACSSC.2011.6190243

[10] Cirik, A.C., Wang, R., Hua, Y.: Weighted-sum-rate maximization for bi-directional full-duplex MIMO systems. In: Conference on Signals, Systems and Computers (ASILOMAR), pp. 1632–1636 (2013). doi:10.1109/ACSSC.2013.6810575

[11] Mahmood, N.H., *et al.*: On the potential of full duplex communication in 5g small cell networks. In: 2015 IEEE 81st Vehicular Technology Conference (VTC Spring), pp. 1–5 (2015). doi:10.1109/VTCSpring.2015.7145975

[12] Zhou, W., Srinivasan, K.: Sim+: A simulator for full duplex communications. In: 2014 International Conference on Signal Processing and Communications (SPCOM), pp. 1–6 (2014). doi:10.1109/SPCOM.2014.6983995

[13] R. Zhang, R., *et al.*: Investigation on dl and ul power control in full-duplex systems. In: 2015 IEEE International Conference on Communications (ICC), pp. 1903–1907 (2015). doi:10.1109/ICC.2015.7248603

[14] Goyal, S., *et al.*: Full duplex cellular systems: will doubling interference prevent doubling capacity? IEEE Communications Magazine **53**(5), 121–127 (2015). doi:10.1109/MCOM.2015.7105650

[15] Malik, H., Ghoraishi, M., Tafazolli, R.: Cross-layer approach for asymmetric traffic accommodation in full-duplex wireless network. In: Networks and Communications (EuCNC), 2015 European Conference On, pp. 265–269 (2015). doi:10.1109/EuCNC.2015.7194081

[16] Mahmood, N.H., *et al.*: Throughput analysis of full duplex communication with asymmetric traffic in small cell systems. In: 11th International Conference on Wireless and Mobile Communications (ICWMC), pp. 57–60 (2015)

[17] Wang, L., *et al.*: Exploiting full duplex for device-to-device communications in heterogeneous networks. IEEE Communications Magazine **53**(5), 146–152 (2015). doi:10.1109/MCOM.2015.7105653

[18] John, W., Tafvelin, S.: Heuristics to classify internet backbone traffic based on connection patterns. In: 2008 International Conference on Information Networking, pp. 1–5 (2008). doi:10.1109/ICOIN.2008.4472818

[19] Postel, J.: Transmission Control Protocol. [Online]. Available: http://www.ietf.org/rfc/rfc793.txt (1981, updated by RFCs 1122, 3168, 6093, 6528)

[20] Allman, M., Paxson, V., Stevens, W.: TCP Congestion Control to TCP's Fast Recovery Algorithm. [Online]. Available: http://www.ietf.org/rfc/rfc2581.txt (1999, obsoleted by RFC 5681, updated by RFC 3390)

[21] m. Duarte, Dick, C., Sabharwal, A.: Experiment-driven characterization of full-duplex wireless systems. IEEE Transactions on Wireless Communications **11**(12), 4296–4307 (2012). doi:10.1109/TWC.2012.102612.111278

[22] Sahai, A., *et al.*: On the impact of phase noise on active cancelation in wireless full-duplex. IEEE Transactions on Vehicular Technology **62**(9), 4494–4510 (2013). doi:10.1109/TVT.2013.2266359

[23] Korpi, D., *et al.*: Adaptive nonlinear digital self-interference cancellation for mobile inband full-duplex radio: Algorithms and RF measurements. In: IEEE Global Communications Conference (GLOBECOM), pp. 1–7 (2015). doi:10.1109/GLOCOM.2015.7417188

[24] Anttila, L., *et al.*: Modeling and efficient cancellation of nonlinear self-interference in mimo full-duplex transceivers. In: 2014 IEEE Globecom Workshops (GC Wkshps), pp. 777–783 (2014). doi:10.1109/GLOCOMW.2014.7063527

[25] 3rd Generation Partnership Project: Technical Specification Group Radio Access Network; Enhanced performance requirement for LTE User Equipment (UE) (2013)

[26] Sarret, M.G., *et al.*: Improving link robustness in 5g ultra-dense small cells by hybrid arq. In: 2014 11th International Symposium on Wireless Communications Systems (ISWCS), pp. 491–495 (2014). doi:10.1109/ISWCS.2014.6933403

[27] Mahmood, N., Catania, D., Lauridsen, M., Berardinelli, G., Mogensen, P., Tavares, F., Pajukoski, K.: A Novel Centimeter-Wave Concept for 5G Small Cells. C R C Press LLC, ??? (2016)

[28] Catania, D., *et al.*: A distributed taxation based rank adaptation scheme for 5g small cells. In: 2015 IEEE 81st Vehicular Technology Conference (VTC Spring), pp. 1–5 (2015). doi:10.1109/VTCSpring.2015.7145600

[29] Abramowitz, M., Stegun, I.A.: LTE for UMTS - OFDMA and SC-FDMA Based Radio Access. Wiley, ??? (2009)

[30] 3rd Generation Partnership Project: Further advancements for E-UTRA physical layer aspects (Release 9) (2010)

[31] 3rd Generation Partnership Project: Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA) Radio Link Control (RLC) protocol specification (2007)

[32] Floyd, S., Henderson, T., Gurtov, A.: The New Reno Modification to TCP's Fast Recovery Algorithm. [Online]. Available: http://www.ietf.org/rfc/rfc3782.txt (2004, obsoleted by RFC 6582)

[33] Radio, W.W.I.N.: WINNER II channel models (2008)

[34] Postel, J.: User Datagram Protocol. [Online]. Available: https://tools.ietf.org/html/rfc768 (1980)

Part IV

Autonomous Device-to-Device Communication

Overview

1 Problem Description and Assumptions

Part III of the thesis concluded that FD technology may not be the most suitable strategy for increasing the capacity of indoor small cell networks. This last part of the dissemination explores scenarios where FD can result in significant gains by studying its potential in providing fast discovery for D2D communication. D2D has been positioned as a strong candidate for future 5G systems, given its potential to offload the infrastructure and to provide support for the URLLC use case.

To initiate a D2D communication, it is previously required that devices discover their peers. This procedure is known as discovery phase, and it can be performed with the involvement of the infrastructure or autonomously by the devices. With the former option, the latency and the control overhead are increased, since devices need to exchange information with the base station in order to set up the communication. However, having the base station controlling the network discovery procedure may bring benefits in terms of interference management. When devices work autonomously, the exchange of messages is performed directly among them, thus reducing the overhead and the latency. According to the requirements for next generation access technologies, the control latency should not exceed 10 milliseconds [1]. For this reason, this work focuses on the second strategy, since the most effective solution to provide fast discovery is to avoid the involvement of the infrastructure.

To be able to set up the communication, a device should be aware whether it has been discovered by its peers. Consequently, a feedback mechanism should be introduced. Ideally, the considered procedure should not increase the control overhead, to avoid impacting the discovery time. This last part of the thesis introduces a piggybacking mechanism for the feedback. The acknowledgment is embedded in the discovery signal and therefore it does not require a dedicated transmission.

Conventional HD systems may not be sufficient to meet the 10 milliseconds latency target. The main problem of the HD transmission mode is that a device cannot listen and transmit at the same time, thus posing a trade-off between the time spent transmitting and the time dedicated to listening. To avoid the constraint that HD poses on the transmission probability, FD technology is considered. This transmission mode allows a device to be continuously listening to transmissions from neighbors while being able to transmit its own discovery message.

Commonly, the transmission of the discovery message is based on a predefined periodicity or a fixed transmission probability. However, this configuration does not allow to control neither the collisions nor the idle slots. Providing dynamism in terms of transmission probability would be then the most appropriate solution. However, there are two constraints. A high transmission probability would cause a large number of collisions, thus increasing the discovery time. On the other hand, if such probability is too low, then a large number of idle slots would be generated and the discovery time would also increase. Note that the transmission probability should be related to the number of neighbors: a large set of peer devices would lead to a low transmission probability and vice versa. Nevertheless, in a real network, the information on the number of neighbors is not available. Hence, an algorithm to estimate the number of neighboring devices is required.

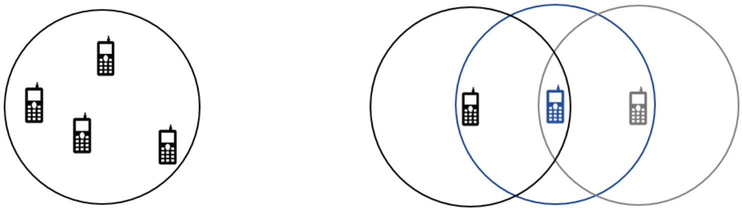
The main assumptions considered in this part of the dissertation are listed below:

- **Nodes are synchronized in time and frequency.**
  The discovery procedure is performed autonomously by the devices, i.e., the cellular infrastructure is not involved in that phase, but still provides time and frequency synchronization.

- **Dedicated spectrum for the D2D discovery.**
  No interference from the cellular network or the D2D devices exchanging data is considered. Furthermore, a pool of frequency resources is available. The maximum pool size considered in this work is 4. During reception, the devices can listen to all the resources simultaneously.

- **Devices are randomly deployed in a certain area at the same time.**
  A *cluster* is defined as the set of neighbors within the coverage range of a device, plus the own device. A simplified channel model which considers only the path loss is assumed. Then, the coverage range is determined taken into account the transmit power, the path loss and the noise power. When every device is able to reach all the other devices in the network, all the devices' clusters coincide. This case refers to the single cluster network.

The opposite case is a multi-cluster network. An example is shown in Figure 5.18.



(a) Example of a single cluster network. **(b)** Example of a multi-cluster network.

Fig. 5.18: Types of scenario analyzed in this part of the dissertation.

- **4×4 MIMO technology with interference cancellation receivers.**
  Ideal interference cancellation and transmission rank one are used in this work. The interfering streams that are not cancelled are treated as noise. For the sake of comparison, a receiver which treats all interfering streams as noise is also considered.

- **Ideal SI cancellation.**
  All devices are assumed to be FD capable with ideal SI cancellation. This assumption is in line with the previous part of the thesis.

- **The time slot duration is 0.25 milliseconds.**
  In line with the 5G concept presented in [2], since D2D communication is foreseen to be part of it.

- **The discovery message and piggybacked feedback is contained in a single time-frequency resource.**
  The quantification of the overhead and the design of the discovery message is left for future work. The assumption is that the information needed for the discovery phase, e.g. the device identifier, and the acknowledgments to one or more neighbors can be mapped in a single time-frequency resource block. The discovery message is assumed to be transmitted with a fixed MCS and transmission rank one.

2 Main Findings

The proposed system design is evaluated in the single and the multi-cluster scenarios. A basic receiver that treats interference as noise and one which is able to ideally cancel the three strongest interfering streams, according to the transmission rank one assumption, are used. In addition, the number of deployed devices and the size of the frequency resources pool is varied. The single cluster results presented in Paper D show that the discovery time follows a 'U' shape curve as a function of the transmission probability. In the left side of the curve, corresponding to low transmission probabilities, a large number of idle slots is generated, thus wasting resources. On the other hand, the right side of the curve represents the high transmission probability region where collisions occur often. The minimum of the curve corresponds to the optimal transmission probability leading to the minimum discovery time, which depends on the network conditions and the scenario.

The target of varying the number of deployed devices and the size of the frequency resources pool is to have different representations of the system congestion. This parameter is defined as the number of network devices divided by the size of the frequency resources pool. Therefore, fixing the number of frequency resources and increasing the number of network devices increases the system congestion. On the other hand, if the number of devices is fixed, increasing the frequency resources leads to a reduction of the system congestion. Based on the system congestion parameter, the optimal transmission direction and the minimum discovery time are extracted, indicating that providing dynamism in terms of transmission probability could bring benefits to the system in reducing the discovery time.

An important finding of Paper D is that, in order to fulfill the strict latency target for next generation systems of 10 milliseconds, one of the requirements is to equip devices with MIMO technology and interference cancellation receivers. Under this assumption, FD technology is able to further reduce the discovery time and meet the latency target. Otherwise, the gain that FD can provide is rather limited.

The multi-cluster scenario is more challenging than the single cluster in terms of interference management. In the single cluster scenario, the number of interfering streams is the same for all the receiving devices. On the contrary, in the multi-cluster scenario, devices could perceive different number of interfering streams. Consequently, devices need to carefully select the most appropriate transmission probability to optimize the discovery time but to avoid harming the neighboring devices. It should be highlighted that the conclusion on the advanced receivers extracted from the single cluster scenario is still valid in the multi-cluster network.

Given the potential of FD technology to reduce the discovery time, under the assumption of using interference cancellation receivers, two solutions are proposed. Firstly, an algorithm to estimate the number of neighbors and select the most appropriate transmission direction. Secondly, a signaling exchange mechanism to reduce the network interference. The results presented in Paper E demonstrate the potential of the proposed solutions in reducing the network interference at the expenses of a minor increase in the overhead, leading to a reduction in the discovery time. In addition, with the proposed scheme, a larger system congestion can be handled while meeting the 10 milliseconds latency target.

3 Included articles

The main body of this last part of the thesis is composed by 2 articles.

Paper D: Can Full Duplex reduce the discovery time in D2D Communication?

This article investigates the potential of FD in reducing the discovery time in D2D communication. A detailed analysis of how the system behaves assuming different transmission probabilities and receiver types is carried out. An important outcome of this analysis is that advanced receivers with interference cancellation capabilities are required to meet the 10 milliseconds target defined for 5G. In that case, FD technology can further reduce the discovery time, showing significant gains over HD and meeting the strict latency target. Furthermore, the study shows that providing dynamism in terms of transmission probability helps the system at adapting to network changes, e.g., the number of neighbors, and at optimizing the time it takes to complete the discovery phase. Finally, challenges in terms of interference management are described.

Paper E: Providing Fast Discovery in D2D Communication with Full Duplex Technology

This article is an extension of Paper D. The contribution of this work is twofold: an algorithm to estimate the number of neighboring devices to select the most appropriate transmission probability, and a signaling exchange to reduce the network interference. The target is to adapt to changes in the network and minimize the number of collisions and idle slots. Simulation results show the effectiveness of the proposed solution, being able to reach the strict latency target of 10 milliseconds and supporting a larger number of devices in the network.

References

[1] 3rd Generation Partnership Project TR 38.913 V0.3.0, “Study on scenarios and requirements for next generation access technologies,” March 2016.

[2] P. Mogensen *et al.*, “Centimeter-wave concept for 5G ultra-dense small cells,” in *IEEE 79th Vehicular Technology Conference (VTC Spring)*, May 2014.

Paper D

Can Full Duplex reduce the discovery time in D2D Communication?

Marta Gatnau Sarret, Gilberto Berardinelli, Nurul H. Mahmood, Beatriz Soret, Preben Mogensen

The paper has been published in the
IEEE 13th International Symposium on Wireless Communication Systems (ISWCS), 2016.

Abstract

Device-to-device (D2D) communication is considered as one of the key technologies to support new types of services, such as public safety and proximity-based applications. D2D communication requires a discovery phase, i.e., the node awareness procedure prior to the communication phase. Conventional half duplex transmission may not be sufficient to provide fast discovery and cope with the strict latency targets of future 5G services. On the other hand, in-band full duplex, by allowing simultaneous transmission and reception, may complete the discovery phase faster. In this paper, the potential of full duplex in providing fast discovery for the next 5th generation (5G) system supporting D2D communication is investigated. A design for such system is presented and evaluated via simulations, showing that full duplex can accelerate the discovery phase by supporting a higher transmission probability compared to half duplex. Simulation results show that, in order to meet the strict 5G control plane latency target, advanced receivers are required. In that case, full duplex can reduce the latency up to 80%.

1 Introduction

The massive increase of mobile devices and the explosion of new services has caused the mobile data traffic to increase very rapidly [1]. The next 5$^{\text{th}}$ generation (5G) system envisions direct device-to-device (D2D) communication as one of the key technologies to accommodate such services. In D2D communication, devices exchange data among them without the involvement of the infrastructure. By allowing devices to work autonomously, benefits such as network offloading, potential coverage extension and reduction of control overhead and latency can be achieved. However, to allow nodes to communicate directly with each other, a discovery phase is required beforehand. The discovery procedure has to be completed in a short time given the strict 5G control plane latency target of 10 milliseconds [2].

The D2D discovery procedure can be controlled by the infrastructure at the expense of extra control overhead, or the devices can operate in an ad hoc manner, independently of the network. The most effective solution to provide fast discovery is to avoid the involvement of the infrastructure, allowing the devices to operate autonomously. Previous research works [3–5] have evaluated the discovery procedure in this context. The authors in [3] propose a device discovery mechanism able to discover few thousands of devices within 8 seconds. Even though the system shows significant improvements over WiFi, it cannot meet the 5G targets. In [4], a discovery mechanism for contention-based networks is proposed, defining an initiator or group owner, which sends discovery messages in a dedicated channel, and the joiners, that send requests in the channel specified by the group owner. Results show that having a dedicated discovery channel improves the discovery time. A design for the discovery message is proposed in [5]. It contains the identifier of the latest k nodes that have successfully transmitted the discovery message to compensate for link failures and reduce the overhead.

Although the literature on device discovery is large, the proposed solutions use mainly half duplex (HD) transmissions. With HD, a transmitting device cannot listen to the messages from neighboring devices, thus increasing the discovery time. Such limitation can be overcome with full duplex (FD) technology. FD allows a device to transmit and receive simultaneously in the same frequency band, thus potentially accelerating the discovery procedure. This technology was historically considered unfeasible due to the dominant self-interference (SI), but given the recent advances in the SI cancellation technology, FD has become practically feasible [6]. Few works in the literature study FD in the context of fast discovery for D2D communication, e.g., [7–9]. The authors in [7] present a strategy to reduce idle slots and collisions, where FD is used to detect the activity of other devices. The work assumes that a device stops transmitting when it is discovered, but this assumption may not be valid in networks with dynamic (de)activation of the nodes where transmitting the discovery message is always required. In [8], FD is combined with compressed sensing to overcome the drawbacks from HD and single packet reception, completing the discovery in one time slot. However, a very limited number of neighbors is considered and the feedback procedure is not addressed. The authors in [9] evaluate FD with directional antennas, where each device selects a transmission direction randomly at each time slot. Nevertheless, the feedback procedure is not addressed.

This paper studies how to provide fast discovery in 5G systems supporting D2D communication. The envisioned system design is presented, considering FD as a key technology component to meet the strict 5G latency requirements. The feedback mechanism and the discovery message design are addressed for facilitating fast discovery and decreasing the control overhead. The discovery time is evaluated in a single cluster and in a multi-cluster scenario, given the different interference conditions on both scenarios. Simulation results show that, depending on several factors, the transmission probability leading to the minimum discovery time varies, which may be different for HD and FD. The analysis reflects that FD reduces the discovery time compared to HD and that interference cancellation receivers are required to obtain further benefits from FD. Finally, the challenges to be addressed in the multi-cluster scenario are also identified.

The paper is organized as follows. Section II describes the envisioned D2D fast discovery. Section III presents the system model. The simulation results are discussed in Section IV and Section V concludes the paper and states the future work.

2 D2D Fast Discovery

We consider an autonomous ad hoc network where the cellular infrastructure is not involved in the discovery procedure among the devices, but still provides time and frequency synchronization.

We foresee that a dedicated part of the spectrum will be allocated for the D2D discovery, since sharing it with the cellular system would cause an increase in the interference, calling for solutions to manage the cellular-to-D2D interference. The system design is time slotted with a certain number of orthogonal frequency resources. On reception mode, devices can listen to all the frequency resources. Hereafter, a *resource* will refer to a single time-frequency block.

An important part of the system design is the discovery message, which is transmitted in a broadcast manner and contains relevant information needed for the discovery procedure. The transmission of this message should be robust to minimize link failures. Consequently, it should be always transmitted with a fixed modulation and coding scheme (MCS) and using one spatial stream, i.e., transmission *rank* one. As such, link adaptation mechanism is not required for the discovery phase and the control overhead is reduced since the devices do not have to exchange information regarding the transmission parameters to be able to successfully decode the messages. A device will be successfully discovered only when a positive acknowledgment from a neighbor is received. Therefore, the feedback mechanism should be carefully designed to reduce the control overhead. We propose a discovery message that includes a field with the identifier of the transmitting node and a feedback field containing the identifiers of the devices that the transmitting node has discovered. Figure D.1 shows the format of such a message, which may also include other relevant discovery information. Since the discovery message is broadcast, a receiving device would check if its identifier appears in the feedback field of the received message. If the identifier appears, such device will know that it has been discovered by the transmitting node. Further optimization of the discovery message design is under study, e.g., the number of bits occupied by each field and the feedback field mapping to provide scalability.

The proposed discovery protocol assuming ideal decoding is shown in Figure D.2. The actual decoding mechanism is described in Section III. In the figure, the diagram is shown from the device d_i side, which has three neighbors n_1, n_2 and n_3. Device d_i transmits according to a certain transmission probability ρ_i, whereas n_1, n_2 and n_3 transmit with ρ_1, ρ_2 and ρ_3, respectively.

Device identifier	Other relevant discovery info	Feedback field
x bits	y bits	z bits

Fig. D.1: Proposed discovery message format

The transmission probability can be different for each device. Assuming d_i is the first node to transmit, it broadcasts the discovery message with the feedback field empty, since it has not received any discovery message yet. Then, n_1 broadcasts the discovery message, including in the feedback field the identifier of d_i, so when d_i receives the message, it will be aware that it has been discovered by n_1. The next device to transmit is d_i, which piggybacks the identifier of n_1 in the discovery message so when n_1 receives the message it will know that it has been discovered by d_i. The same principle applies to the remaining steps of the figure. Note that the feedback field can contain more than one identifier, as in the last step of Figure D.2, where d_i piggybacks the identifier of n_2 and n_3.

Finally, the discovery time is defined as the time needed for a node to be discovered by all its neighbors, based on the feedback reception time. It is extracted as the maximum time among all the neighbor feedback reception times and it depends on the transmission probability ρ. A collision will occur if the signal-to-interference-plus-noise ratio (SINR) is below a certain decoding threshold, as described in Section III. If a too high ρ is used, a large number of collisions will occur and the discovery process will take longer. On the other hand, if ρ is too small, the inactivity of nodes will generate idle slots, thus increasing the time it takes to finish the discovery phase. In addition, there is another trade-off in the choice of ρ for HD transmission mode. The necessity of transmitting the discovery message with a certain probability leads inevitably to a reduction of the opportunities for listening and therefore to an increase of the discovery time. For this reason, we believe that FD technology is a potential candidate to satisfy the strict latency requirements given the possibility of simultaneously transmit and receive, which may reduce the processing time. An example of a three devices scenario comparing HD and FD is depicted in Figure D.3, where four orthogonal frequency resources and ideal decoding are assumed. Following the same principle as Figure D.2, HD is able to complete the discovery time in six time slots, while FD can finish it in four time slots because of the continuous listening state.

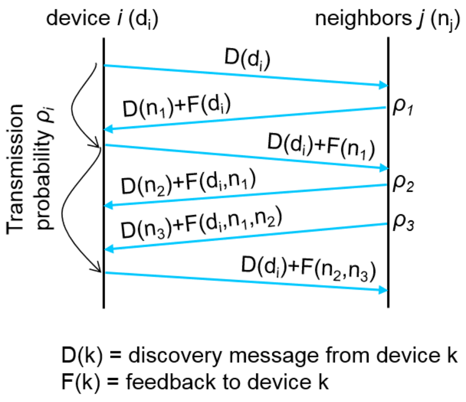


Fig. D.2: Proposed discovery protocol for D2D communication

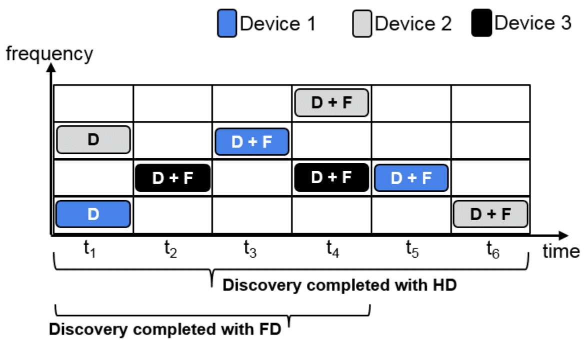


Fig. D.3: Example of the discovery procedure with 3 nodes and assuming 4 frequency resources

3 System model and simulation setup

The proposed system has been simulated with our own developed Matlab simulator. Some of the concepts used in this work are extracted from our envisioned 5G system [10], since D2D communication is expected to be part of it.

Orthogonal frequency division multiplexing (OFDM) is used as modulation scheme. The format of a resource block is shown in Figure D.4, composed of 14 OFDM symbols and occupying 10 MHz in frequency and 0.25 milliseconds in time. The first two OFDM symbols are used for synchronization with the base station. The remaining, including the demodulation reference signal (DMRS) symbol used for channel estimation, are dedicated for broadcasting the discovery message. Transceivers with 4×4 multiple-input multiple-output (MIMO) are considered.

We assume a simple path loss model $\gamma = d^{-\alpha}$ and no fading. In the simulation results presented in the next section, α is set to 4. The transmission rank is set to one and the MCS corresponds to Quadrature Phase Shift Keying (QPSK) with a code rate of 1/3. Then, assuming a subcarrier spacing of 60 kHz and 11 OFDM symbols in time, 1210 useful bits can be embedded in the discovery message. The decoding is based on a SINR threshold of 0 dB, which corresponds to a block error rate (BLER) target of 0.01% according to the used MCS. A maximum of 4 frequency resources are used and the noise power in each resource is -95 dBm. Finally, ideal SI cancellation in FD nodes is assumed [6].

Two different receivers are considered in this study:

- *Interference as noise (IAN) receiver*: this type of receiver considers the interfering streams as noise. Since a device may receive more than one desired signal, a computation of the SINR per each received signal is performed, assuming one stream as the desired one and the rest of them as noise. The computation of the SINR is shown in Equation (D.1):

$$SINR_{IAN} = \frac{\gamma_d \cdot P_T}{\sum_{i=1}^{rxSignals} I_i + N} \tag{D.1}$$

  where γ_d is the path loss between the desired transmitter and the receiver, P_T is the transmit power, $rxSignals$ are the number of received signal streams, $I_i = \gamma_i \cdot P_T$ is the power of the interfering stream i and N is the noise power.

- *Interference cancellation (IC) receiver*: this receiver is able to cancel the $N - K$ strongest interfering streams, where N is the number of MIMO receive antennas and K is the transmission rank. Then, for the remaining interfering streams, the working principle of the IAN receiver is applied. For example, according to our simulation setup, the three strongest interfering streams will be canceled since the transmission rank is one and each node is equipped with 4 receiving antennas. Equation (D.2) describes how the SINR is calculated with the IC receiver:

$$SINR_{IC} = \frac{\gamma_d \cdot P_T}{\sum_{i=k}^{rxSignals} I_i + N} \tag{D.2}$$

  where $k = N - K + 1$ and the interfering streams I_i are sorted based on the signal strength from the strongest to the weakest one.

The devices are randomly deployed, forming *clusters*. A cluster is defined

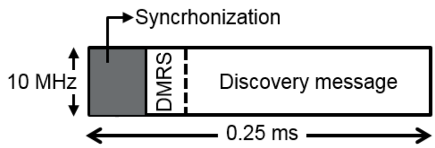


Fig. D.4: Resource block format

as the set of neighbors that a device can reach based on the transmit power, the path loss and the noise power, plus the own device. Therefore, a cluster size is extracted for each individual node. Two scenarios are analyzed, a single and a multi-cluster scenario. The former refers to the case in which every node can reach all the other nodes in the network, and suffers from intra-cluster interference. The latter is affected by intra and inter-cluster interference, since clusters may be partially overlapped or totally isolated. The discovery message is randomly mapped in frequency at every transmission opportunity.

Results are presented in terms of the average discovery time among all the deployed nodes in the scenario. As explained in Section II, the device discovery time is extracted as the maximum time needed to discover all the device neighbors. The variable ρ refers to the transmission probability and θ is used as an indication of the system congestion, calculated as shown in Equation (D.3):

$$\theta = \frac{\text{number of network devices}}{\text{number of frequency resources}} \tag{D.3}$$

4 Performance Evaluation

This section presents numerical results evaluating the discovery latency. The deployed area and the transmit power are set accordingly to the simulated scenario. To ensure statistical reliability, 1000 redeployments are considered.

4.1 Single cluster scenario

In this section we study the case where all the devices perceive the same cluster, i.e., all of them have to discover each other. The number of devices ranges from 10 to 50 in steps of 10, and they are randomly dropped with an uniform distribution over a 100×100 m^2 area, transmitting with 23 dBm power. As

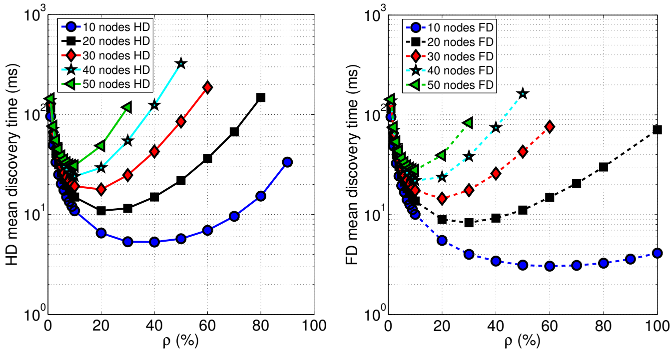


Fig. D.5: Single cluster mean discovery time with HD and FD, assuming the IAN receiver and 4 frequency resources

explained in Section II, there is a trade-off in selecting the transmission probability ρ. A high ρ increases the number of collisions and thus the discovery time. On the other hand, a low ρ generates many idle slots and also leads to a higher discovery time. In addition, HD has the limitation of not being able to receive messages from neighboring devices when a device is transmitting. This HD limitation can be overcome with FD, since it allows simultaneous transmission and reception. However, the ρ choice effect on the number of collisions and idle slots still needs to be taken into consideration.

As a first step, we have analyzed the discovery time varying ρ from 1% to 90% in case of HD, and from 1% to 100% for FD. This exercise has been repeated varying the number of available frequency resources (1, 2 and 4) and using the IAN and the IC receiver. Figure D.5 shows the HD (left) and FD (right) discovery time assuming 4 frequency resources and the IAN receiver. From the figure we can observe that the discovery time describes a 'U' shape curve, independently of the transmission mode and the number of devices. The minimum of this curve provides the optimal ρ that leads to the minimum discovery time. As the number of devices increases, the 'U' shape curve is shifted to the left to avoid increasing the link failures, and thus the optimal ρ decreases. We observe that for the 10 and 20 devices cases, FD ρ is higher than HD ρ. Except at very low ρ where the performance for FD and HD is the same, FD provides lower discovery time than HD. For example, assuming 20 devices and $\rho = 50\%$, FD shows a gain of 49% over HD.

Such 'U' shape curve is observed independently of the number of frequency resources or the receiver type. The impact of varying the amount of frequency resources is equivalent to varying the system capacity. If the frequency resources are less than 4, the 'U' shape curve from Figure D.5 shifts to the left because ρ has to be decreased to avoid generating too many collisions and therefore link failures. Using the IC receiver would allow a device

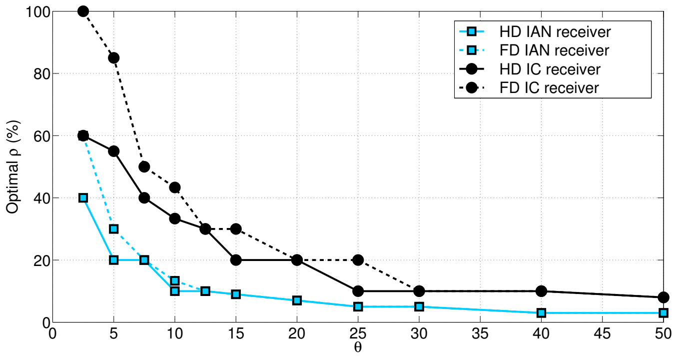


Fig. D.6: Optimal ρ for HD and FD assuming the IAN and the IC receiver

to use a higher ρ compared with using the IAN receiver because canceling interference improves the SINR and therefore larger number of messages can be received without increasing the number of link failures.

From the first result we can already identify that the optimal transmission probability changes depending on the cluster conditions. Therefore, the second step is to obtain the curve that describes the optimal ρ. Figure D.6 presents such a curve, depending on the system congestion (see Equation (D.3)), for 4 cases: HD and FD with the IAN receiver, and HD and FD with the IC receiver. For all the cases, the optimal ρ decreases as the system congestion increases. This is due to the trade-off between collisions and discovery time. We can observe that the transmission probability is higher when receivers with interference suppression capabilities are used. This is the case when the largest difference between HD and FD ρ is seen. On the other hand, when using IAN receivers, the HD and FD ρ is rather close or the same.

As pointed out previously, the optimal ρ leads to the minimum discovery time, which is shown in Figure D.7. The green line indicates the latency target, which is 10 milliseconds [2]. The figure shows that, in order to meet the latency requirement, advanced receivers with interference cancellation capabilities are required. Using IAN receivers, the latency target can be only reached if few devices are deployed. In any case, in the region where the discovery time does not exceed 10 milliseconds, we can observe that FD allows for a higher system congestion θ than HD. Figure D.8 shows the gain in discovery time of FD over HD for the considered receivers, as a function of the system congestion. We can observe that the gain with the IC receiver ranges from 20% to 80% in the region of θ where the latency target can be met. On the other hand, using the IAN receiver and restricting the gain where the latency requirements according to θ are achieved, FD gain goes from 25% to 42%.

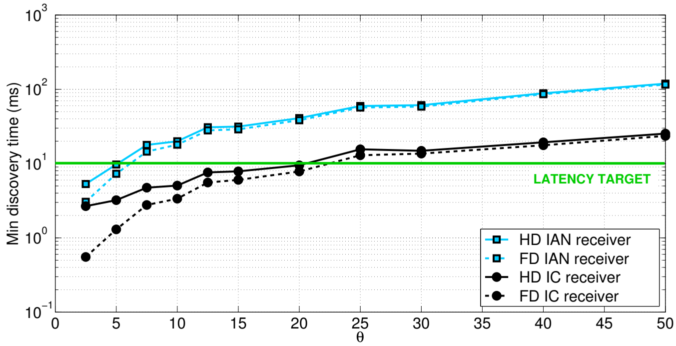


Fig. D.7: Minimum discovery time for HD and FD assuming the IAN and the IC receiver

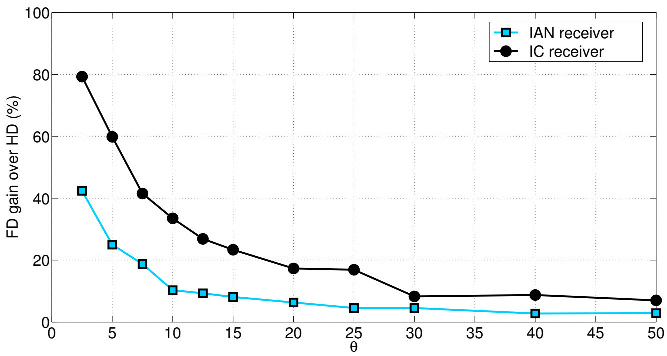


Fig. D.8: FD gain over HD with the IAN and the IC receiver

4.2 Multi-cluster scenario

To analyze the multi-cluster scenario, 50 devices transmitting with 0 dBm power are deployed in an area of 1000×1000 m^2, resulting in a discovery range of 237 meters and a cluster size that ranges from 3 to 16, being 8.28 on average. Such distribution is highly related to the considered simulation parameters.

Figure D.9 shows an example of this scenario, where each color represents a different cluster size and the number after each device is its own cluster size. Let us focus the analysis on the three devices inside the red square: the green device that has a cluster of size 3 (i.e., 2 neighbors), the yellow device having 8 neighbors and the white device with 13 neighbors. In principle, the green device should be able transmit with higher ρ because it has only two neighbors, thus completing the discovery procedure very fast. However, if the green device increases its own ρ, it will generate high interference to the yellow and white device because these two devices have a larger cluster size,

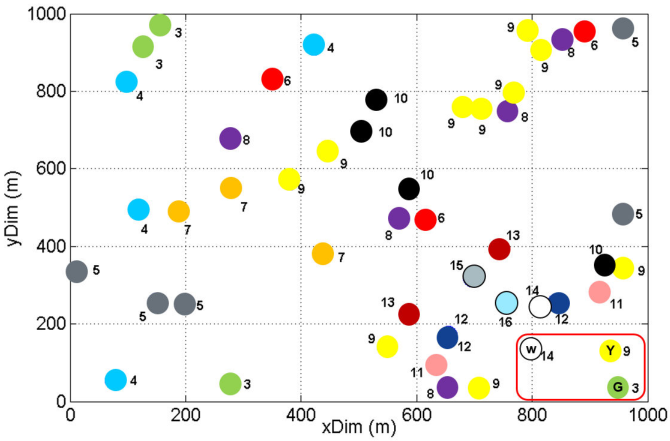


Fig. D.9: Example of a multi-cluster deployment

thus increasing the number of collisions occurring in these two devices. This observation raises a question regarding the mechanisms that are required to manage interference among neighboring devices. Addressing this challenge is left for future work, as we intend to investigate a dynamic algorithm to distributively determine the best ρ for each node.

Following the same analysis like in the single cluster scenario, the first step is to evaluate the mean discovery time depending on the transmission probability ρ. However, in this scenario the cluster size is variable, meaning that the mean discovery time is not a valid key performance indicator anymore. Instead, we should focus our analysis on the 5^{th}, 50^{th} and 95^{th} percentiles of the discovery time. Figure D.10 shows such result, assuming 4 frequency resources and the IC receiver. The most relevant result is the 95^{th} percentile because it corresponds to the users perceiving worse interference conditions, i.e., the outage users. From the figure we can observe that the same 'U' shape curve can be extracted, meaning that there is an optimal ρ that leads to the minimum discovery time. In addition, this results also leads to the conclusion that advanced receivers are required to meet the latency target of 10 milliseconds. From the figure we can see that the optimal ρ for HD is 40% and for FD is 60%, providing a FD gain of 42%. It is important to notice that in the multi-cluster scenario, in contrast with the single cluster scenario, not using advanced receivers does not allow to meet the latency target for the 95^{th} percentile users.

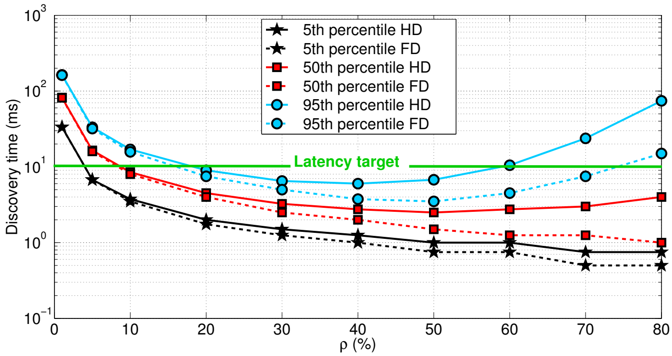


Fig. D.10: Discovery time for the multi-cluster scenario assuming 4 frequency resources and the IC receiver

5 Conclusion and future work

In this paper we have investigated the role of full duplex technology and ideal interference cancellation receivers in aiding fast discovery for D2D communication. We have shown that there is a compromise in selecting the transmission probability, that dictates the amount of collisions and the generation of idle slots. Furthermore, half duplex shows another trade-off on such selection because spending too much time transmitting does not allow a node to receive messages from the neighboring nodes. To overcome this limitation, full duplex is considered since it allows for continuous reception. Results show that devices are discovered faster with full duplex nodes, and that there is a optimal transmission probability that leads to the minimum discovery time. However, in order to meet the strict latency requirements, the use of advanced receivers with interference cancellation capabilities is required. In that case, further benefits can be obtained from full duplex.

Future work will be focused on developing an algorithm that dynamically adjusts the transmission probability to achieve the minimum discovery time, and on addressing the challenges identified in the multi-cluster scenario.

Acknowledgement

This work has been performed within the framework of the Horizon 2020 project FANTASTIC-5G (ICT-671660) receiving funds from the European Union. The authors would like to acknowledge the contributions of their colleagues in the project, although the views expressed in this contribution are those of the authors and do not

necessarily represent the project.

References

[1] Cisco, "Cisco visual networking index: Global mobile data traffic forecast update, 2015-2020," Feb. 2016.

[2] 3rd Generation Partnership Project, "Study on scenarios and requirements for next generation access technologies," Mar. 2016.

[3] F. Baccelli *et al.*, "On the design of device-to-device autonomous discovery," in *Fourth International Conference on Communication Systems and Networks (COMSNETS)*, Jan 2012.

[4] H. B. Li and R. Miura, "Discovery protocol for peer aware communication networks," in *IEEE 82nd Vehicular Technology Conference (VTC Fall)*, Sept 2015.

[5] Y. Kang and X. Xu, "A fast ad-hoc neighbor discovery algorithm based on a friend trust mechanism," in *IEEE 12th International Conference on Networking, Sensing and Control (ICNSC)*, April 2015.

[6] M. Heino *et al.*, "Recent advances in antenna design and interference cancellation algorithms for in-band full duplex relays," *IEEE Communications Magazine*, vol. 53, no. 5, pp. 91–101, 2015.

[7] G. Sun *et al.*, "PHED: Pre-handshaking neighbor discovery protocols in full duplex wireless ad hoc networks," in *IEEE Global Communications Conference (GLOBECOM)*, Dec 2012.

[8] X. Yang *et al.*, "Full-duplex and compressed sensing based neighbor discovery for wireless ad-hoc network," in *IEEE Wireless Communications and Networking Conference (WCNC)*, March 2015.

[9] R. Xu *et al.*, "A neighbor discovery algorithm for full duplex ad hoc networks with directional antennas," in *The 27th Chinese Control and Decision Conference (CCDC)*, May 2015.

[10] P. Mogensen *et al.*, "Centimeter-wave concept for 5G ultra-dense small cells," in *Proc. of IEEE 79th Vehicular Technology Conference*, May 2014.

Paper E

Providing Fast Discovery in D2D Communication with Full Duplex Technology

Marta Gatnau Sarret, Gilberto Berardinelli, Nurul H. Mahmood, Beatriz Soret, Preben Mogensen

The paper has been submitted to the
Springer 9th International Workshop on Multiple Access Communications (MACOM), 2016.

Abstract

In Direct Device-to-Device (D2D), the device awareness procedure known as the discovery phase is required prior to the exchange of data. This work considers autonomous devices where the infrastructure is not involved in the discovery procedure. Commonly, the transmission of the discovery message is done according to a fixed probability. However, this configuration may be not appropriate to meet the 10 milliseconds control plane latency target defined for the next 5^{th} generation (5G) system. In this work, we propose a distributed radio resource management framework supporting full duplex technology to provide D2D fast discovery. Such framework provides an algorithm to estimate the number of neighbor devices and to dynamically decide the transmission probability, for adapting to network changes and meeting the 10 milliseconds target. Finally, a signaling scheme is proposed to reduce the network interference. Results show that our framework performs better than a static approach, reducing the time it takes to complete the discovery phase. In addition, supporting full duplex allows to further reduce the discovery time compared to half duplex transmission mode.

1 Introduction

Device-to-Device (D2D) communication has drawn significant attention for the design of 5^{th} generation (5G) systems to offload the infrastructure and to cope with the continuous growth of wireless applications and services. In D2D communication, devices are allowed to communicate directly, without the involvement of the infrastructure. However, prior to the establishment of such communication, devices must discover their peers. This device awareness procedure is known as discovery phase. According to the latest specifications for next generation systems [1], the control plane latency cannot exceed 10 milliseconds. Such requirement poses challenges on how to facilitate fast device discovery. Full duplex (FD) technology, which allows for simultaneous transmission and reception in the same frequency band, may speed-up the discovery process.

The execution of the D2D discovery procedure can be controlled by the infrastructure or performed autonomously by the devices. The first option requires the exchange of messages with the base station, generating additional control overhead and increasing the latency. The latter option, where devices send the discovery message periodically, has potential to diminish the control overhead and provide lower latency [2].

Autonomous device discovery using conventional half duplex (HD) transmission mode has been studied by the research community [4, 5, 7], whereas few works consider FD technology [8–10]. A synchronous distributed ad-hoc network is studied in [5], focusing on optimizing the discovery latency and the number of discovered devices. The authors propose a resource structure as well as a resource selection. However, the feedback mechanism is not considered and their system operates in a larger time scale than that specified by [1]. A discovery message design to minimize collisions is proposed in [4]. The work analyzes an autonomous D2D system where devices transmit the discovery message with a fixed probability, showing an improvement in the number of discovered devices. The authors in [7] propose using a small portion of the resources for new devices appearing in the network, such that their discovery message can be transmitted with a shorter delay. In [8], a strategy to reduce idle slots and collisions is presented. FD is used to detect the activity of other devices. The work assumes that a device stops transmitting when it is discovered. Nevertheless, this assumption may not be valid in networks with dynamic (de)activation of the nodes where transmitting the discovery message is always required. In [10], FD is combined with compressed sensing to overcome the drawbacks from HD and single packet reception. The authors claim that the discovery phase is completed in a single time slot. However, a very limited number of neighbors is considered and the feedback procedure for discovery acknowledgment is not addressed. The authors in [9] evaluate FD with directional antennas, where each device selects a transmission direction randomly at each time slot. It is important to notice that the mentioned works assume the transmission of the discovery message with a fixed probability. This principle does not allow to control the generated idle slots and collisions, thus posing critical challenges in meeting the latency requirements.

In our previous work [2], we showed that adapting the rate of discovery message transmission to the number of active devices may be beneficial, and we identified challenges in terms of interference management in a large network. In this paper, we propose a radio resource management (RRM) framework for autonomous D2D communication supporting FD technology. It provides a mechanism to estimate the number of neighbors as this information is not available in realistic ad-hoc networks, and an adaptive scheme to select the most appropriate transmission probability for the discovery messages. The interference can be better coordinated by allowing the devices to exchange their transmission probability, which captures the number of neighbors in our proposal. Thus, each terminal can dynamically set the most appropriate transmission probability using not only the current value and own information but also information from the neighbors. Results show that our solution achieves lower latency than a static approach. Moreover, supporting FD allows to further reduce the discovery time compared to HD transmission mode.

The paper is organized as follows. Section II describes the proposed RRM framework. Section III presents the system model and discusses the simulation results. Finally, Section IV concludes the paper and states the future work.

2 D2D Fast Discovery

2.1 General system overview

We focus on autonomous ad hoc networks with a dedicated band of the spectrum for the discovery procedure. Devices communicate directly with each other and the infrastructure is not involved in the discovery phase, but still provides time and frequency synchronization. This design allows to avoid interference between cellular and D2D users.

A time slotted system is considered. At each transmission opportunity, there is possibility of exploiting a pool of orthogonal frequency resources, where the resource to be used is randomly chosen. It is assumed that, on reception, devices can simultaneously listen to all frequency resources. The discovery message is transmitted in a broadcast manner according to a certain transmission probability ρ and it contains the information required to perform the discovery phase, e.g., the device identifier and its position. Since the discovery procedure needs to be completed in a short time to meet the strict control plane latency requirements [1], the number of link failures should be minimized. This can be achieved by transmitting the discovery message with a robust modulation and coding scheme (MCS) at the expense of a larger message, and by using one spatial stream, often referred as transmission *rank* one, assuming that devices are equipped with 4×4 multiple-input multiple-output (MIMO) transceivers. The dimensioning of the discovery message is left for future work. In this paper, we assume that the discovery message can be mapped over a single time/frequency resource.

The discovery phase is required to set a unicast/multicast communication. Therefore, the devices involved in such communication should be acknowledged of the fact that their peers are aware of their presence. We propose a design for the discovery message that includes a *feedback* field, containing the identifiers of the devices that have been discovered by the transmitting device. Since the discovery message is broadcast, a device that receives and decodes the message will check if its identifier is piggybacked. If so, the receiving device will know that it has been discovered by the transmitting device. The discovery time is then based on the feedback reception time, and it depends on the transmission probability ρ. Using a high ρ causes a large number of collisions which increases the discovery time. On the other hand, using a small ρ creates a large number of idle slots due to the inactivity of the devices, which also increases the time needed to complete the discovery procedure. Furthermore, in case of HD transmission, the necessity of transmitting the discovery message leads to a reduction of the opportunities for listening to neighbors' transmissions. We investigate the potential of FD technology in reducing the discovery time, since it eliminates the HD constraint by allowing simultaneous transmission and reception on the same frequency band.

2.2 RRM design

In our previous work [2], we showed that the transmission probability that leads to the minimum discovery time depends on the scenario, e.g., on the number of neighbors. Such result indicates that a dynamic choice of ρ can be beneficial for the system. In addition, we identified challenges in terms of interference management in large networks. Let us define *cluster* as the set of neighbors within the coverage range of a device, plus the own device. Therefore, the cluster and its size is a device-specific parameter. In case every device is able to reach all the other devices in the network, all the devices' clusters coincide. We refer to this case as *single cluster* network. The opposite case is a *multi-cluster* network. Figure E.4 shows an example of a portion of a multi-cluster network, where the clusters from two devices, C and G, are highlighted. In particular, the number beside each device refers to their cluster size. In this specific example, C only reaches G, while the latter reaches C, Y and W. Let us focus on G, which has two neighbors perceiving a larger cluster size (W and Y) and C, which only reaches G. Since G is not aware of the overall interference perceived by W and Y and it has only three neighbors, it would benefit from using a high ρ. However, using a high ρ may increase the number of collisions to W and Y, who have a larger number of neighbors, consequently increasing their discovery time. On the other hand, C will benefit from such a high ρ because it has a cluster of size 2.

From the previous example we can extract that an exchange of information among devices can be beneficial to reduce the overall network interference and to avoid increasing the discovery time. Furthermore, in a realistic network, the information related to the number of neighbors is not available. To solve the mentioned problems, we propose a RRM framework to dynamically adjust the transmission probability allowing devices to adapt to network changes. It consists of two parts: the instantaneous estimation of the number of neighbors, and the dynamic adjustment of ρ based on network information exchange. The proposed solution is distributed, so it does not require a centralized controller that collects information from the network.

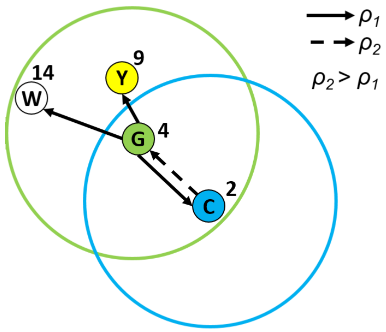


Fig. E.1: Example of a multi-cluster network with 2 highlighted clusters

Algorithm 2 Algorithm for estimating the number of neighbors

$\rho \leftarrow$ Current transmission probability. Initial value $\rho = 0.5$
$\tilde{M} \leftarrow$ Estimated number of neighbors. Initial value $\tilde{M} = 0$
repeat At each time slot
  Extract ρ according to the selected information exchange approach (Table E.1)
  if Transmission time, based on ρ, **then**
    Transmit the discovery message
  else
    Receive discovery messages from neighbors, and estimate $\tilde{M}$ as:

$$\tilde{M} = \frac{\#\text{decoded signals}}{\rho} \tag{E.1}$$

  end if
until The device turns off

Estimating the number of neighbors

The estimation of the number of neighbors is done based on the available information at each device: the own ρ and the number of signals successfully decoded in each receiving time slot. The estimation of the number of neighbors with HD is done as described in Pseudo-code 2. In case of FD, the *if* statement encapsulates only the transmission part, since a FD device is continuously receiving. In Equation E.1, the *#decoded signals* refers to the number of instantaneous messages that a device on reception can successfully decode. Then, the equation is equivalent to the number of active neighbor devices that a node can detect.

Signaling scheme

To reduce the network interference, we propose that devices send ρ within the discovery message, since it is related to the number of estimated neighbors: a low number of estimated devices leads to a high transmission probability, and vice versa. The value of ρ can be represented with different number of bits, depending on the desired resolution and the allowed control overhead. Table E.1 lists the proposed approaches to extract ρ, according to the signaled information. The difference among these approaches is the amount of extra information sent within the discovery message and how this information is utilized to decide the ρ to be used. With the *selfish* approach, devices do not signal any information about their ρ, and they behave in a selfish manner. Hence, the control overhead is not increased but the network interference is uncontrollable. With the *cooperative minimum* option, devices signal the used ρ, which is extracted as indicated in Table E.1. In this case, the control overhead is slightly higher but the interference is reduced. Finally, with the *cooperative maximum* approach, devices transmit two values of ρ: the one used for transmission, extracted by applying the *cooperative maximum* approach, and the one extracted from the estimation of the number of neighbors, i.e., the *selfish* ρ_{sf}. The difference between the second and the third approaches is that the latter avoids for the minimum ρ to spread across the network, at the cost of limited extra overhead.

Let us focus again in the example depicted in Figure E.1. Using the *selfish approach*, G and C will use a high transmission probability, hence G will be highly interfering W and Y, causing unsuccessful transmissions. If the *cooperative minimum approach* is used, the ρ extracted by W will be the one that the other devices in the network will use. However, C should be able to transmit with a higher ρ because it is not interfering W. This situation can be solved with the *cooperative maximum approach*.

3 Performance Evaluation

The RRM framework techniques proposed in Section II are evaluated using our own developed Matlab simulator. Firstly, we want to prove that a dynamic ρ selection has advantages over using a fixed one. Secondly, the performance of the single and the multi-cluster scenarios is discussed.

For our D2D network, we assume a simple path loss model $\gamma = d^{-\alpha}$, where α is set to 4, and no fading. The packet decoding is threshold-based: a collision occurs if the signal to noise-plus-interference ratio (SINR) is below such threshold, which is set to 0 dB, assuming Quadrature Phase Shift

Table E.1: Information exchange approaches

Approach	Signaling	Principle
Selfish (ρ_{sf})	None	Use ρ extracted from the estimated number of devices without considering information from the neighbors. Section III.A describes the function that, depending on the estimated number of devices, provides the most appropriate ρ.
Cooperative minimum	ρ_u	Set ρ as the minimum between: a. the ρ extracted from the estimated number of neighbors (ρ_{sf}) b. the minimum ρ received from the neighbors
Cooperative maximum	ρ_u and ρ_{sf}	Set ρ taking into account: a. the ρ extracted from the estimated number of neighbors (ρ_{sf}) b. the used ρ received from the neighbors ($\rho_{u,nb}$) c. the estimated ρ received from neighbors ($\rho_{sf,nb}$) The decision is taken as follows: 1. Extract the minimum of $\rho_{u,nb}$ and $\rho_{sf,nb}$ of all the received messages from my neighbors 2. If $\rho_{u,nb} = \rho_{sf,nb} = \rho_{nb}$: select the minimum between ρ_{sf} and ρ_{nb} 3. If $\rho_{u,nb} < \rho_{sf,nb}$: select the minimum between $\rho_{sf,nb}$ and ρ_{sf}

Keying (QPSK) with a coding rate of 1/3 and a block error rate of 0.01%. The representation of ρ is ideal in this study, i.e., it is represented with maximum resolution. Finally, ideal self-interference cancellation in FD devices is assumed [3].

According to the findings in [2], one of the requirements to meet the strict 5G control plane latency target is to use interference cancellation (IC) receivers. These receivers are able to suppress the $N - K$ strongest interfering streams, where N is the number of MIMO receive antennas and K is the transmission rank. For example, according to our simulation setup (transmission rank one and devices equipped with 4 receiving antennas), the three strongest interfering streams can be suppressed. We assume ideal IC in this work. The SINR reads then:

$$SINR_{IC} = \frac{\gamma_d \cdot P_T}{\sum_{i=k}^{rxSignals} I_i + N_0} \tag{E.2}$$

where γ_d is the pathloss between the receiver and the desired transmitter, P_T is the transmit power, N_0 is the noise power, $k = N - K + 1$ and the interfering streams I_i are sorted based on the signal strength from the strongest to the weakest one. In this work, the transmit power is 0 dBm and a noise power in each time/frequency resource is -95 dBm.

The devices are randomly deployed in a certain area. Two scenarios are analyzed, a single and a multi-cluster scenario. The former suffers only from intra-cluster interference, while the latter is affected by intra and inter-cluster interference, since clusters may be partially overlapped or totally isolated. The single cluster scenario refers to a 100×100 m^2 area with the number of deployed devices ranging from 10 to 50. An area of 1000×1000 m^2 is considered for the multi-cluster scenario, where the number of deployed devices goes from 10 to 300. In particular, the average cluster size ranges from 3 to 44.

We assume that a discovery message opportunity occurs every 0.25 milliseconds. This is consistent, for example, with the assumption in terms of frame duration of our 5G small cell concept presented in [6]. The discovery time is defined as the time needed for a device to be discovered by all its neighbors, based on the feedback reception time. It is extracted, individually for each node, as the maximum time among all the neighbor feedback reception times. For the single cluster scenario, results are presented in terms of average discovery time, since the interference conditions for all the devices are, in average, the same. In case of the multi-cluster scenario, results are presented in terms of the 95$^{\text{th}}$ percentile of the discovery time. For both scenarios, the performance of FD is compared against the HD performance. Finally, the variable θ provides an indication of the system congestion, and it is defined as:

$$\theta = \frac{\text{number of network devices}}{\text{number of frequency resources}} \tag{E.3}$$

3.1 Dynamic transmission probability

Let us consider the single cluster scenario with its corresponding parametrization, and setting the size of the frequency resource pool to 1, 2 and 4, to have different system level congestion representations. In our previous work we proved that the optimal ρ which minimizes the discovery time is scenario-dependent. Figure E.2 shows such optimal ρ as a function of θ, for the ideal case where the devices know the exact number of neighbors. From the figure we can observe that, as the system congestion increases, the optimal ρ diminishes due to the constraint on the number of collisions. We can also see that FD allows for a higher ρ in some cases, specially at low system congestion. This is because, as explained in Section II, FD solves the constraint that HD poses on the ρ selection. Therefore, since FD operates in a larger range of ρ, we expect to have larger gains from using a dynamic approach with FD com-

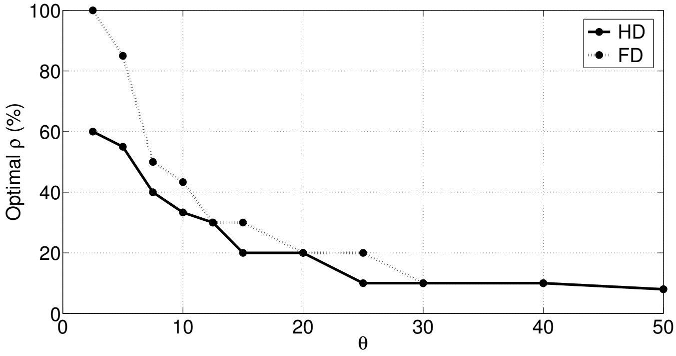


Fig. E.2: Optimal ρ as a function of the system congestion θ

pared to HD. This two curves can be easily approximated by a simple multiplicative inverse function $\rho = f(\tilde{M}, \text{frequency resources pool size}) = f(\theta) \approx \frac{a}{\theta^b}$, where a and b are fitting parameters. Such approximations allow us as to have a representation of the optimal ρ based on the system congestion.

3.2 Single cluster performance

As the next step, we evaluate the proposed RRM framework in the single cluster scenario, assuming a frequency resource pool of size 4. In this case, the considered approach is the selfish one, since the interference conditions of all the devices are, in average, the same. Consequently, the three approaches described in Table E.1 show the same performance. The evaluation is done by comparing: the optimal discovery time, extracted under the assumption of ideal information at the devices; the performance of the proposed RRM framework, extracting ρ_{SF} from the inverse approximation of the curves shown in Figure E.2; and a fixed ρ of 40%, since it provides a good trade-off on the HD performance given its limitation of not being able to receive messages while transmitting. Figure E.3 shows the performance comparison between the mentioned cases. We can observe that the HD performance is barely affected from the usage of a dynamic ρ, except when the number of network devices is large. This is caused by the small operational range of ρ , given the HD constraint. However, in case of FD, we can observe that a dynamic ρ selection allows to get very close to the optimal system performance. The maximum difference between the optimal and the algorithm performance is 0.83 milliseconds, at high congestion. At low congestion, the maximum difference is 0.13 milliseconds.

We want to emphasize that the robustness of the approximated ρ curves have also been evaluated, by varying the path loss model or the deployed

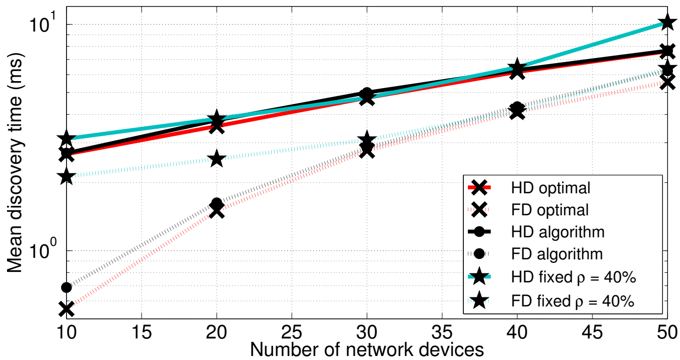


Fig. E.3: Single cluster. Performance comparison between the proposed algorithm, the optimal performance and a fixed ρ

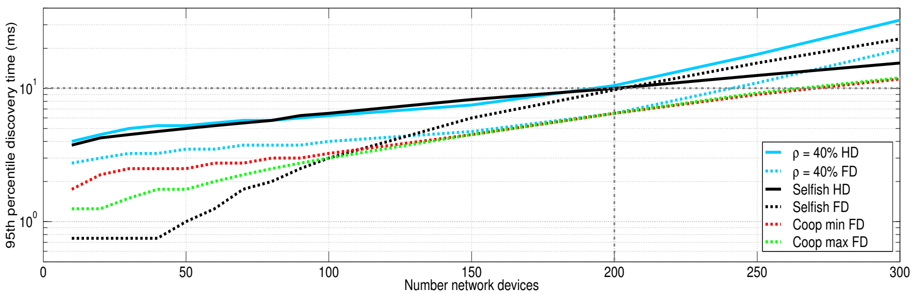


Fig. E.4: Multi-cluster. Performance comparison between the proposed algorithm, the optimal performance and a fixed ρ

area. Results show that the maximum difference between the ideal discovery time and the one extracted by using our proposed framework and the approximations is ∼0.5 milliseconds.

3.3 Multi-cluster performance

We focus here on the multi-cluster scenario, assuming a frequency resource pool of size 4, and we will analyze the performance of the three approaches presented in Table E.1 and the fixed 40% transmission probability. For HD, only two performance curves are presented, the selfish and the fixed 40% one, since the performance with the other two proposed algorithms is nearly the same as the selfish one. The reason for that, as explained before, is the smaller range of ρ with HD since a device cannot listen while transmitting, leading to an increase of the discovery time if a too high ρ is used. Figure E.4 shows the 95$^{\text{th}}$ percentile of the discovery time. The results show that,

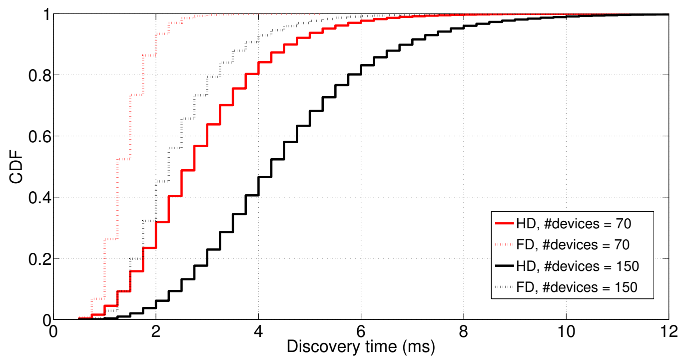


Fig. E.5: Multi-cluster. Cooperative maximum algorithm CDF

with HD, the latency requirement of 10 milliseconds can be achieved with up to 200 network devices, but the discovery time is always larger than the one achieved with FD. Focusing on the FD performance, we can observe that providing a dynamic solution for the ρ selection brings benefits in terms of reduction of the discovery time, independently of the network density. In case of few devices in the network, the selfish approach shows the best performance in terms of discovery time, since it is the most aggressive scheme and using interference cancellation receivers allows the system to use a high ρ. The cooperative minimum approach is too conservative at low network density, since the ρ used by the device in the worst interference conditions spreads across the network. Note that the drawback of such approach can be solved by using the cooperative maximum approach, since it allows to increase the transmission probability in case the neighboring nodes are not affected by this ρ increase. As the network density increases, we observe that the cooperative minimum and maximum approaches reach the optimal system performance. In this case, both approaches perform equally since the optimal ρ is already small, to avoid generating a large interference and hence a large number of collisions. Notice that with the proposed cooperative algorithms, the system can support higher network density while still meeting the 10 milliseconds latency requirement.

Finally, Figure E.5 shows the cumulative distribution function (CDF) of the discovery time with HD and FD for the cooperative maximum approach, for the cases of 70 and 150 number of network devices. The curves show that, beyond the mean gain of FD over HD, HD suffers from larger time variances, with some devices perceiving large discovery time and others finishing their process much faster.

4 Conclusion and future work

In this paper we have proposed a RRM framework supporting FD technology to provide fast discovery in autonomous D2D communication. The framework provides a mechanism to estimate the number of neighbor devices and a scheme to dynamically adjust the transmission probability, in order to reduce the number of collisions and idle slots and decrease the latency for discovery. The proposed signaling exchange mechanism allows to reduce the network interference and improve the system performance. System level results show that our proposed dynamic solution allows to reduce the discovery time compared to traditional static approaches, especially in case full duplex technology is used. Future work will focus on the control message design and on applying the framework in scenarios with high mobility such as vehicle-to-anything (V2X) applications.

Acknowledgment

This work has been performed within the framework of the Horizon 2020 project FANTASTIC-5G (ICT-671660) receiving funds from the European Union. The authors would like to acknowledge the contributions of their colleagues in the project, although the views expressed in this contribution are those of the authors and do not necessarily represent the project.

References

[1] 3GPP TR 38.913 V0.3.0: Study on scenarios and requirements for next generation access technologies (Mar 2016)

[2] Gatnau, M., Berardinelli, G., Mahmood, N.H., Soret, B., Mogensen, P.: Can full duplex reduce the discovery time in D2D communication? In: 13th International Symposium on Wireless Communication Systems (ISWCS) (June 2016), accepted

[3] Heino, M., et al.: Recent advances in antenna design and interference cancellation algorithms for in-band full duplex relays. IEEE Communications Magazine 53(5), 91–101 (2015)

[4] Hong, J., Park, S., Choi, S.: Neighbor device-assisted beacon collision detection scheme for D2D discovery. In: 2014 International Conference

on Information and Communication Technology Convergence (ICTC) (Oct 2014)

[5] Jung, S., Chang, S.: A discovery scheme for device-to-device communications in synchronous distributed networks. In: 16th International Conference on Advanced Communication Technology (Feb 2014)

[6] Mogensen, P., et al.: Centimeter-wave concept for 5G ultra-dense small cells. In: IEEE 79th Vehicular Technology Conference (VTC) (2014)

[7] Park, S., Choi, S.: Expediting D2D discovery by using temporary discovery resource. In: IEEE Global Communications Conference (GLOBECOM) (Dec 2014)

[8] Sun, G., Wu, F., Gao, X., Chen, G.: PHED: Pre-handshaking neighbor discovery protocols in full duplex wireless ad hoc networks. In: IEEE Global Communications Conference (GLOBECOM) (Dec 2012)

[9] Xu, R., Li, J., Peng, L., Ye, Y.: A neighbor discovery algorithm for full duplex ad hoc networks with directional antennas. In: The 27th Chinese Control and Decision Conference (CCDC) (May 2015)

[10] Yang, X., Wang, X., Yang, R., Zhang, J.: Full-duplex and compressed sensing based neighbor discovery for wireless ad-hoc network. In: IEEE Wireless Communications and Networking Conference (WCNC) (2015)

Part V

Conclusion

Conclusion

The 5G umbrella is very broad and includes several use cases to accommodate the future traffic demands and the new arising applications. The focus of this dissertation is on the enhanced mobile broadband and the ultra-reliable low latency communication uses cases. In particular, on improving the capacity of the proposed 5G indoor small cell concept [1] by means of dynamic TDD and full duplex technology, and on providing fast discovery for D2D communication.

This chapter summarizes the main findings of the work and describes future research paths based on the presented conclusions.

1 Main Findings

Providing total flexibility in terms of time slot direction assignment would allow to optimally accommodate the traffic in a small cell. This strategy, known as dynamic TDD, suffers from an increased interference variability [2]. Even though advanced receivers such as IRC mitigate the impact of the interference [3], it is not sufficient to reduce the interference to the desired level. To mitigate the impact of such residual interference and improve the system capacity, two strategies are studied via system level simulations. Firstly, an analysis of the HARQ fast recovery mechanism is carried out. The preferred operational mode for HARQ is set as asynchronous and adaptive, to provide flexibility in time and frequency, respectively. The option of sending the feedback at the desired time slot and frequency resources is more appropriate given the challenging 5G targets. Results show that HARQ has the potential to increase the average system throughput, while improving the outage performance. In addition, HARQ can help at reducing the latency since it is the fastest recovery mechanism. Secondly, the performance of outer-loop link adaptation technique is studied. OLLA adjusts dynamically the MCS used for transmission based on the HARQ feedback. It was originally intended to compensate for systematic errors, such as delay and quantization of measurements, and it was designed as a static technique. i.e., based on a predefined block error rate. However, given the unpredictable interference of the 5G indoor small cell scenario, providing adaptability in terms of the block error rate, based on the interference conditions, can bring benefits to the system. Even though the gain observed with OLLA is rather limited, it is important to highlight that such gain does not come at the expenses of increased system complexity.

Recent advances in the transceiver design has enabled implementation of full duplex technology with viable cost. FD, by allowing simultaneous transmission and reception in the same frequency band, promises double throughput compared to conventional half duplex transmission mode. As a result, FD is being considered as a technology component for future 5G systems. Given the potential of FD in improving the throughput, this new technology is studied via system level simulations for improving the 5G indoor small cell concept [1]. The study also analyses the enhancement that FD can provide in terms of delay, which is not provided by theoretical studies. The constraints that limit the gain that FD can provide over HD are identified: the residual self-interference, the increased interference and the traffic profile. In the analysis, the SI is considered ideal [4]. Thus, the presented results provide an upper bound for the FD gain.

Two FD setups are studied. Firstly, when both AP and UE are FD capable, named bidirectional FD. In that case, results show that when the traffic profile allows for exploiting simultaneous transmission and reception in a non-isolated cell scenario, the system performance is highly impacted by the increased inter-cell interference. On the other hand, in an isolated cell, the probability of exploiting FD based on the traffic profile dictates the gain that FD can provide over HD. When asymmetric traffic is considered, the probability of exploiting FD is reduced. However, the lightly loaded link highly benefits from such asymmetry compared to HD transmission mode, both in terms of throughput and delay.

The second FD type is when only the BS or AP is able to exploit FD. In this case, apart from the increased inter-cell interference, the DL user is highly interfered by the UL user. Such UE-to-UE or intra-cell interference creates differences between the two link directions. Results show that the DL is negatively impacted by the use of FD, whereas the UL can be slightly improved. In case of asymmetric traffic, the lightly loaded link shows a gain below the theoretical one, even though the potential observed for the bidirectional FD case.

Finally, the analysis shows that FD can accelerate the TCP protocol, by increasing faster the congestion window given the rapid transmission of TCP acknowledgments. However, this advantage of FD over HD is hidden in case of a strong inter-cell interference. In that case, the amount of data which is possible to be transmitted depends on the selected MCS and the transmission rank, showing that HD could perform better than FD.

Given the conclusions of the limited benefits provided by FD in indoor small cell networks, other applications where this technology can provide higher benefits is studied. In particular, the potential of FD in reducing the discovery time in D2D communication. This type of communication allows devices to communicate directly among them, thus offloading the infrastructure. The discovery mechanism is required to detect peer devices in the surroundings, and should be completed within 10 milliseconds. To achieve such strict control latency requirement, FD technology is considered, since it overcomes the HD constraint of not being able to listen to neighboring transmissions while transmitting the discovery message.

The analysis shows that FD has potential in reducing the discovery time if interference cancellation receivers are assumed. Results indicate that the optimal transmission probability leading to the minimum discovery time is different depending on the scenario, e.g., number of neighbors. In addition, interference management techniques are required to avoid increasing the latency. The proposed solution includes the dynamic selection of the transmission probability based on an estimation of the number of neighboring devices and a signaling scheme to reduce the network interference. Simulations show that the proposed strategy reduces the discovery time and can meet the latency target defined for 5G at the expenses of a minor increase in control overhead.

2 Future Work

The work presented in this dissertation provides many possibilities for future research on small cell networks and applications where full duplex technology may bring large benefits.

When dynamic TDD is considered to improve the capacity of 5G small cells, mechanisms to stabilize the interference are required. Part II of the thesis describes interference management techniques as a solution to deal with the drawback introduced by dynamic TDD. The approach of using coordination by means of a centralized controller could be evaluated, for example following the coordinated multi-point principle, where users receive the data from multiple base stations to improve the signal quality. The design should be carefully addressed, since the generated control overhead and the increased delay may bring more harm than benefits.

The study of FD in 5G indoor small cells led to the conclusion that under strong interference, HD may outperform FD. Consequently, the design of a hybrid HD/FD scheduler that chooses the most appropriate transmission mode depending on the interference conditions could benefit the system in terms of throughput and delay. An analysis of the SINR range where each transmission mode brings higher benefits to the system can be carried out.

The studies involving FD were performed assuming ideal SI cancellation. Therefore, an interesting work could be to quantify the impact of non-ideal SI cancellation. A model which depends on the transmit power and the bandwidth could be derived and evaluated. In addition, 5G is targeted to be cost-effective. Therefore, the model could also contemplate simpler cancellation techniques and study its impact on the FD performance.

A third FD application could be studied: the BS acting as a FD relay, i.e., the BS forwarding data from one user to another one. This case faces similar conditions as the BS FD case, since intra-cell interference is also present. However, from a traffic perspective, it is required that two users need to exchange data. There are two options on implementing such a relay. The first one is to allow the BS to decode the data and store it until the receiving node acknowledges it. This approach would improve the throughput but not the delay. The second options is to forward the data as it arrives at the BS, without performing error check. The main drawback of this approach is that in case of link failure, both UL and DL transmissions need to be rescheduled. On the other hand, the delay can be improved if the decoding is performed successfully at the first transmission attempt.

The indoor small cell network targeted in this work is an interference limited scenario. Therefore, FD could be evaluated in other scenarios, e.g., a macro cell. In this case, since the transmit power of the base station is higher than in the indoor scenario, a model to take into account the impact of non-ideal SI cancellation should be included. Interference management and power control techniques would also be required to limit the interference among base stations. Furthermore, a smart scheduler that selects the pair of users which reduces the harm from the UL to the DL user would also be needed.

In terms of D2D communication, future work could focus on the design of the discovery message and how feedback is piggybacked: which is the maximum number of acknowledgments that could be embedded in the discovery message, which is the information required to be transmitted or how to identify each device. In this work, an assumption for the discovery message transmission was to use a fixed MCS so no link adaptation was required and the control overhead could be reduced. However, it could be interesting to consider link adaptation techniques and compare both approaches, in terms of control overhead, robustness and discovery time.

The proposed FD solution of adapting dynamically the transmission probability and signaling such parameter could be evaluated in scenarios with mobility such as V2X. This scenario is challenging since the number of neighbors varies rapidly and often.

References

[1] P. Mogensen *et al.*, "Centimeter-wave concept for 5G ultra-dense small cells," in *IEEE 79th Vehicular Technology Conference (VTC Spring)*, May 2014.

[2] D. Catania, "Performance of 5G small cells using flexible TDD," Ph.D. dissertation, Department of Electronic Systems, Aalborg University, October 2015.

[3] F. M. L. Tavares, G. Berardinelli, N. H. Mahmood, D. Catania, T. B. Sørensen, and P. E. Mogensen, "Interference-robust air interface for 5G ultra-dense small cells," *Journal of Signal Processing Systems*, pp. 1–14, 5 2016.

[4] M. Heino *et al.*, "Recent advances in antenna design and interference cancellation algorithms for in-band full duplex relays," *IEEE Communications Magazine*, vol. 53, no. 5, pp. 91–101, 2015.

Part VI

Appendix

Paper F

Full Duplex Communication Under Traffic Constraints for 5G Small Cells

Marta Gatnau Sarret, Davide Catania, Gilberto Berardinelli,
Nurul H. Mahmood, Preben Mogensen

The paper has been published in the
IEEE 82nd Vehicular Technology Conference (VTC Fall), 2015.

Abstract

Full duplex (FD) communication promise of doubling the throughput of half duplex (HD) communication makes such type of system an attractive solution to cope with the expected mobile data traffic increase. Nevertheless, simultaneous transmission and reception in dense deployment scenarios increases the inter-cell interference compared to a traditional HD communication, due to a larger number of nodes simultaneously transmitting. Moreover, FD communication can only be exploited when there is traffic in both uplink and downlink directions simultaneously. In this paper, FD communication is studied within the framework of 5th generation (5G) small cell systems in order to address its effective gain in such specific scenarios. The factors that affect FD performance are analysed, and its performance is evaluated against a traditional HD communication. System level simulations show that the gain of FD over HD in the considered scenarios is lower than the expected 100% gain, with a strong dependency on the traffic and the interference conditions.

1 Introduction

Wireless mobile data traffic is increasing exponentially, and according to the latest forecast, it is expected to reach a factor of $\sim$x1000 by 2020 (with reference to 2010) [1]. It is widely established that such traffic requirement can be met by exploiting larger frequency bands, multiple-input multiple-output (MIMO) technology with a large number of antennas, and cell densification. We believe that a low complexity system which combines all the mentioned approaches would cope with future traffic requirements. In that sense, we have proposed a new radio access technology (RAT) for a 5$^{\text{th}}$ generation (5G) small cells system optimized for dense local area deployments [2]. Time division duplex (TDD) has been chosen as the operational mode, and a minimum 4x4 MIMO transceiver configuration with advanced receivers is assumed.

Recently, full duplex (FD) communication has also been considered as a potential option to accommodate future mobile data traffic growth [3]. FD allows each node to transmit and receive in the same frequency band at the same time. Therefore, the throughput can ideally be doubled, compared to a traditional half duplex (HD) TDD system. However, to enable an efficient FD communication, the self-interference (SI), i.e, the interference caused from the transmitter antennas to the receiver antennas placed on the same device, should be cancelled. Recent results show that designing a FD system is feasible, providing a system performance with only 8% loss with respect to the performance of an ideal FD system where SI is fully suppressed. Several techniques were proposed to achieve such close-to-ideal implementation [3–5].

The FD performance in wireless networks has been previously studied, considering ideal self-interference cancellation and residual SI cancellation error. In [6], Xie et al. analysed the gain that FD provides compared to HD in WLAN systems, from a signal to interference plus noise ratio (SINR) perspective, considering ideal SI cancellation. Their conclusion is that the FD gain is below the expected 100% in the majority of the cases. A detailed study of the achievable bit rate depending on different levels of non-ideal SI cancellation and SINR conditions is presented in [7] and [8]. Both works address the SINR conditions which allow FD to outperform HD. They conclude that, in highly interfered scenarios, combining FD and HD is the most appropriate solution. Nonetheless, all studies mainly focus their analysis on interference levels, disregarding the type of traffic in the network. In [9], full buffer traffic is considered, and the FD throughput performance using different type of receivers is studied, providing an average throughput gain of 30-40%.

In this paper, an evaluation of the FD performance in the context of our proposed 5G small cell system is presented. An unpredictable bursty traffic with large fluctuations is expected, due to the low number of users per cell. Such traffic burstiness is expected to impact the expected throughput gain of a FD system. Realistic system level mechanisms such as finite buffer traffic models, non-ideal link adaptation, retransmission mechanisms and dynamic selection of the number of transmission streams are taken into consideration. Several study cases have been considered, to clearly understand the factors that impact FD performance and the scenarios where FD can outperform HD systems. The main difference between our study and [9] lies on the fact that the latter uses ideal link adaptation and full buffer traffic model. Moreover, retransmission mechanisms are not considered, and the throughput is calculated from the SINR in the physical layer (Shannon throughput). In the presented work, we have moved a step forward and analysed FD performance considering a complete system, from application to physical layers, and with more realistic assumptions.

The paper is structured as follows. Section II describes FD operation in the envisioned 5G system. Section III presents the simulation setup, and system level results are discussed in Section IV. Finally, Section V concludes the paper and describes the future work.

2 Full Duplex in 5G Small Cells

The envisioned 5G system, described in [2], is designed as a new RAT with the aim of coping with the future wireless traffic demand. It has been originally proposed as a TDD system because of its flexibility and the possibility of exploiting unpaired frequency bands. In that sense, a new frame structure,

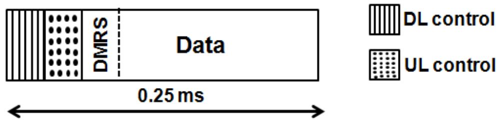


Fig. F.1: Envisioned 5G frame structure

shown in Figure I.1, is introduced, with a duration of 0.25ms, defined as the Transmission Time Interval (TTI). It consists of the downlink (DL) and uplink (UL) control symbols, followed by the data part. The first symbol of the data part is dedicated to the transmission of the demodulation reference signals (DMRS), used to enable channel estimation at the receiver. Although it was designed for TDD, the 5G frame also adapts perfectly to FD communication. Having separate symbols for the UL and DL control channels allows the system to support networks where all the nodes, access points (APs) and user equipments (UEs), are FD capable (bidirectional FD) or networks where only the AP is FD capable (relay FD). The data part would carry UL and DL traffic simultaneously in both cases, and the actual transmission(s) would be one TTI delayed from the corresponding scheduling grant.

Interference rejection combining (IRC) receivers, which can potentially suppress a number of interfering streams at the expense of reducing the number of desired transmission streams [10], and retransmission mechanisms such as hybrid automatic repeat and request (HARQ) and automatic repeat and request (ARQ), are used to deal with the inter-cell interference. When FD transmissions occur, the inter-cell interference may become a major problem, due to the higher number of simultaneously active nodes (APs or UEs). In this case, the IRC receivers may not have sufficient degrees of freedom for suppressing the increased number of interfering streams, since they may be doubled compared to the traditional HD system.

This may jeopardize the achievement of the double throughput with respect of HD systems. In order to quantify the gain of FD communication, three factors need then to be taken into account:

- Self-interference cancellation. A high level of isolation between transmitting and receiving ports in the same device is required.
- SINR conditions. As the inter-cell interference increases, the SINR that a node perceives decreases, thus leading to a reduction of the sustainable rates, since a higher number of TTIs is required to transmit the same amount of data.
- Simultaneous UL and DL data. The lower the load offered by each link, the smaller is the probability that simultaneous UL and a DL transmis-

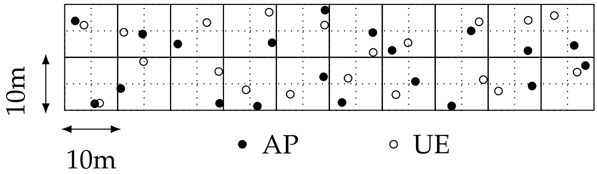


Fig. F.2: Simulated scenario

sion can happen at the same time, thereby exploiting FD capabilities.

Next sections will present the simulation setup that has been used for this study, providing a performance analysis of FD and HD systems in different type of scenarios.

3 Simulation setup

To study the impact of the factors described in Section II in isolation, the presented 5G system has been analysed with HD and with bidirectional FD. The results are extracted from our event driven based system level simulator, described with more detail in [11], which implements the envisioned 5G small cells system design [2].

The scenario, shown in Figure I.5, consists of 20 square cells of 10m x 10m arranged in a 10 by 2 grid. The cell densification, also referred as the deployment ratio (DR), is defined in each simulation drop. Then, if a cell is active, one AP and one UE are randomly deployed, with the UE always being affiliated to the AP in the same cell (closed subscriber group). This means that, for example, a DR of 0.6 implies having, in average, a pair of AP and UE deployed in 12 cells out of 20.

Each node is FD capable and is equipped with a 4x4 MIMO transceiver and an IRC receiver. For simplicity, in this work we consider ideal SI cancellation. The number of used spatial streams, often referred as the transmission *rank*, is decided according to the rank adaptation algorithm presented in [12], which chooses the most suitable rank depending on the perceived SINR and a set of weights. These weights apply a higher taxation to the high ranks, discouraging their selection, in order to reduce the overall inter-cell interference in the network. Finally, all nodes use the entire bandwidth to transmit (200 MHz), and the transmission power is fixed and it is the same for the AP and the UE.

A vertical radio resource management (RRM) layer, which gathers information from the physical (PHY), medium access control (MAC) and radio link control (RLC) layers, provides the scheduling decisions and the transmission parameters. The transmission direction in case of HD is obtained using a Head-of-Line (HOL) delay based UL/DL decider [2], which considers the volume of data in the transmission buffers and the amount of time that the first packet in the buffer has been waiting to be transmitted. For the bidirectional FD case, the decision is being done from a traffic perspective. This means that whenever there is data in both UL and DL buffers, both links are going to be scheduled for the next TTI. On the contrary, if there is only data in one buffer, such link is going to be scheduled, i.e. the system will behave as a HD system, without exploiting the FD capabilities. The frame structure and scheduling delays described in Section II are modelled in the simulator. Therefore, a causal non-ideal link adaptation which takes a decision regarding the modulation and coding scheme (MCS) and the transmission rank, based on previous measurements, is used. On the receiver side, if the packet is not successfully decoded, an HARQ retransmission will be triggered, allowing up to four retransmissions. If after four HARQ retransmissions the packet is not still received correctly, an ARQ RLC retransmission will then be triggered.

The traffic in UL and DL is symmetric, meaning that the offered load in both directions is the same. For this study, a finite buffer traffic model has been used; this model generates payloads according to a negative exponential distribution, with an average size of 2 megabytes (defined as sessions), arriving every $t_{inter-arrival}$ seconds [13]. $t_{inter-arrival}$ is also generated according to a negative exponential distribution, and it is composed of the period of time when the application generates the packets for a particular session (t_{on}) plus the amount of time when no packets are being generated (t_{off}). The values of t_{on} and t_{off} reflect the load in the system. So, for a fixed t_{on}, increasing t_{off} will translate into a lower load in the system, and vice versa. The carried load in the system is characterized by the resource utilization (RU) of the network, which indicates the amount of TTIs that have been used for data transmission, i.e.

$$RU = \frac{\sum TTI_{DL} + \sum TTI_{UL}}{\sum TTI_{DL} + \sum TTI_{UL} + \sum TTI_{MUTE}} \tag{F.1}$$

where TTI_x refers to the type of data carried on a TTI, and x can be UL, DL or MUTE (in case there is no data in any of the two link directions). For example, a RU of 50% means that half of the time the channel is free, and a RU of 100% refers to a full buffer traffic model, where the channel is always occupied. The RU shown in the results is extracted from the HD case, since it is our baseline to extract the FD gain. The remaining simulation parameters are described in Table I.1.

Table F.1: Used parameters to run the simulations

Parameter	Value/State/Type
System parameters	BW = 200MHz; f_c = 3.5GHz
Frequency reuse	1 (whole band)
Propagation model	WINNER II A1 w/fast fading [14]
Antenna configuration	4x4
Receiver type	IRC
Transmission power	10 dBm (AP and UE)
Link adaptation filter	Log average of 5 samples
UL/DL decider	HOL (HD) and traffic based (FD)
HARQ max retransmissions	4
Resource utilization	25%, 50% and 75%
RLC mode	Acknowledged
Transport protocol	UDP
Simulation time per drop	4 seconds
Number of drops	50

In terms of key performance indicators (KPIs), we define the post-IRC SINR and the session throughput. A session is defined as the time needed to receive a generated payload, according to [13]; then, the session throughput is defined as the average of each individual session from all the nodes in the network, i.e, gathering the samples from both link directions.

4 Performance Evaluation

In order to provide a clear understanding of the FD capabilities in 5G systems, several cases have been analysed. The first two studies aim at analysing separately the impact of the traffic and the interference, respectively, in the FD performance. The third study case provides the analysis of FD in a more realistic scenario, considering the joint impact of the two factors mentioned previously.

4.1 Impact of Finite Buffer Traffic Model

The goal of this first study case is to verify the impact of the traffic on the expected performance of FD. For this analysis, a single cell scenario with finite buffer is considered, thus avoiding inter-cell interference effects. AP and UE use fixed rank 1 transmissions, and the results are extracted by varying the

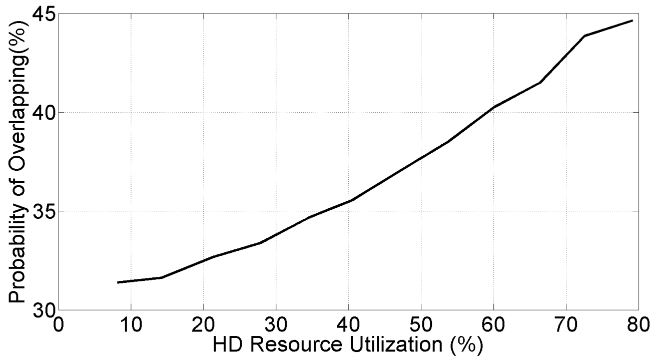


Fig. F.3: Probability of simultaneous UL and DL traffic

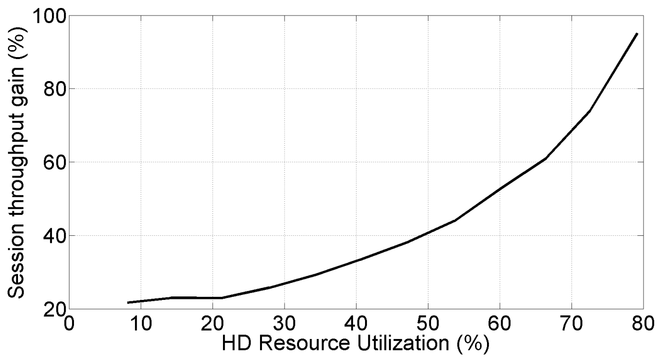


Fig. F.4: Session throughput gain of FD over HD

RU in the network.

The probability of simultaneous UL and DL traffic, depicted in Figure F.3, shows that considering the realistic traffic model specified in [13] has a negative impact on the FD gain, since the probability of exploiting FD communication is up to 45% for a RU of 79%. Notice that the RU in case of FD is smaller than the HD RU (59% instead of 79%), since the nodes can get rid of the data faster, thus occupying less time the channel. The session throughput gain of FD over HD as a function of the RU is shown in Figure F.4. As the RU increases, the gain from FD is higher because the probability of having simultaneous UL and DL traffic is also higher, reaching 95% throughput gain for a RU of 79%. Notice that, even if the channel is not always occupied, the throughput gain approaches 100% because of the buffering effect, which has a lower impact for FD because data can be transmitted faster. Furthermore, note that the HD case, which is used as a reference to extract the FD gain, adopts a scheduler (the HOL decider mentioned above) which exploits traffic

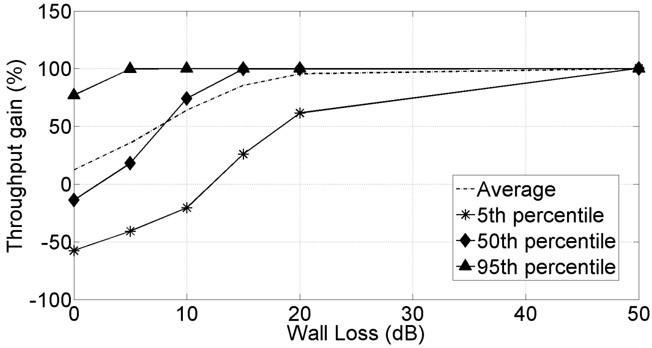


Fig. F.5: Throughput gain of FD over HD

information for optimizing the HD performance. It is important to remark that such gains are obtained from an isolated cell, i.e., inter-cell interference does not appear. Furthermore, even if not addressed in this paper, having asymmetric traffic is expected to have a further negative impact on the FD performance.

4.2 Impact of Inter-Cell Interference

In this second study, the increase of inter-cell interference given by FD communication is addressed. To isolate the impact of the inter-cell interference, the multi-cell scenario with DR equal to 1 and a full buffer traffic model is considered. In this way, all nodes perform simultaneous transmission and reception in each TTI. In addition, like in the previous study, fixed rank 1 transmissions are considered. The analysis is done by varying the wall loss, i.e., the level of isolation among cells.

Figure F.5 shows the throughput gain, in average and in the 5^{th}, 50^{th} and 5^{th} percentiles, of FD over HD as a function of the wall loss. The former throughput is defined as the average of the individual throughputs of all nodes, extracted as the total amount of received bits over the whole simulation time. We can see that, as the wall loss increases, each cell starts behaving as an isolated cell (no inter-cell interference). When the wall loss is 20 dB, the average throughput gain is 96%, and in case of full isolation (50dB wall loss), FD achieved the theoretical double throughput compared to HD, since SINR conditions are the same for both types of communication. On the other hand, the throughput gain in an open space scenario (0 dB wall loss) is 12% in average, way far from the expected factor of $\sim$x2. Analysing the detailed percentile throughputs, the figure shows that the nodes perceiving the poorest SINR conditions (5^{th} percentile) will worsen their performance compared

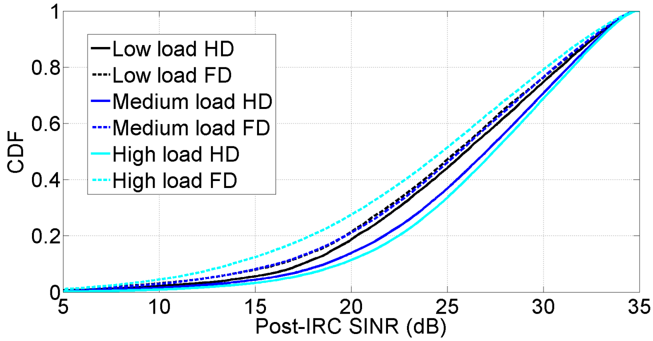


Fig. F.6: Post-IRC SINR for a DR of 0.6

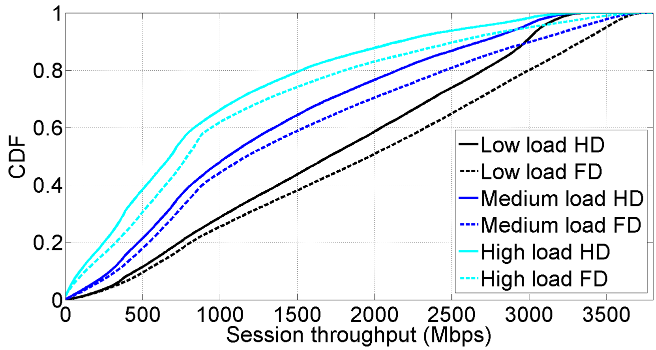


Fig. F.7: Session throughput for a DR of 0.6

to HD unless the isolation between cells is at least 20 dB. On the other hand, the users perceiving good SINR conditions (95th percentile) will be always getting the theoretical 100% gain, except in case of an open space scenario, where the gain is limited to 76%. For a wall loss of 5 dB (as defined in [14]), we observe that the users in outage get 41% less throughput when operation with FD, while the users in the 50th percentile get a limited gain of 18% and the 95th percentile users achieve the expected double throughput gain.

From this analysis we can conclude that, without traffic constraints, the limiting factor in achieving the expected double throughput gain of FD in dense scenarios is the inter-cell interference. The presented results show that in case of no inter-cell interference (ideal scenario), FD communications achieves, in fact, the theoretical ∼x2 throughput.

4.3 Realistic Multi-Cell Analysis

In this section, the performance of FD is evaluated under more realistic conditions, considering the multi-cell scenario with different DRs, the finite buffer traffic model, and the rank adaptation described in Section III. The wall loss is set to 5 dB, according to the used channel model [14], and three loads are used: low, medium and high, defined as 25%, 50% and 75% RU, respectively. In this case, the results include the impact of both traffic and inter-cell interference, and they are presented in terms of cumulative distribution function (CDF).

Figure F.6 shows the CDF of the post-IRC SINR for DR equal to 0.6 (i.e. in average, a pair of AP and UE will be deployed in 12 cells out of 20). It can be observed that the difference between the post-IRC SINR in FD and HD increases with the load. This is because a high load increases the probability of nodes exploiting FD, thus increasing the amount of interference. In average, the FD SINR is 2.2%, 5.3% and 10.9% lower than the HD SINR with low, medium and high load, respectively. The most affected users will be the ones perceiving poor SINR conditions, usually referred as the 5^{th} percentile or outage users. In this case, the SINR reduction when comparing FD against HD is 13.3%, 19.9% and 37.0% with low, medium and high load, respectively.

Figure F.7 shows the CDF of the session throughput for the same DR of 0.6. The FD gain is up to 30.3% when considering high load, but in general, the gain is ∼15%, thus being far from the expected ∼x2 factor. On the other hand, it is important to notice that FD always outperforms HD, and when considering high load, i.e., when there is a higher probability of exploiting FD communication, the gain is 29.4%, 15.8% and 13.8% for the 5^{th}, 50^{th} 95^{th} percentile respectively. Such result shows that FD can improve the performance of the users perceiving the poorest SINR conditions percentile while also improving the performance of rest of the users when the network is not over-densified.

Table F.2 summarizes the gains for different DRs at low, medium and high load. For low DRs, FD outperforms HD in all regions, providing average gains between 15% and 22%. Nevertheless, for high DRs, FD impacts negatively the system performance, specially the users in outage. As we can see from Table F.2, there is a loss in the 5^{th} percentile in case of high load. This is due to the increased inter-cell interference. In the case of a low DR, the occurrences of significant interfering streams at each cell diminish, thus leveraging the gains of the IRC receiver. But on the other hand, for high DRs, IRC receivers cannot cope with such inter-cell interference because they do not have enough degrees of freedom to suppress the interfering streams.

From the full set of presented results, several conclusions can be extracted. First of all, the gain that FD can provide under realistic conditions is not the

Table F.2: Session throughput gain of FD over HD depending on the deployment ratio

	0.2	**0.4**	**0.6**	**0.8**	**1.0**
5th	18.1%	14.1%	8.1%	4.9%	0.9%
50th	10.8%	14.4%	15.3%	10.4%	8.1%
95th	13.2%	13.5%	13.6%	12.7%	12.0%

(a) Low load

	0.2	**0.4**	**0.6**	**0.8**	**1.0**
5th	20.2%	18.8%	12.4%	-0.4%	1.7%
50th	15.1%	20.5%	11.5%	7.4%	4.6%
95th	13.7%	13.6%	12.4%	11.7%	11.4%

(b) Medium load

	0.2	**0.4**	**0.6**	**0.8**	**1.0**
5th	28.6%	30.3%	29.4%	-16.4%	-
50th	20.8%	14.0%	15.8%	21.5%	14.5%
95th	13.9%	12.4%	13.8%	18.0%	-1.8%

(c) High load

expected ~x2. Secondly, it has been proved that there is a high dependency on the traffic load in the FD gain. Finally, when the traffic conditions allow benefiting from FD, the inter-cell interference becomes the barrier to achieve the FD theoretical gain.

5 Conclusions and future work

In this paper, the use of full duplex communication under traffic constraints for 5G small cell systems has been studied. An analysis of the performance of such type of communication for different scenarios and under realistic conditions is presented, showing a strong dependency between the full duplex system performance and the type of traffic in the network. Furthermore, the study shows that when the traffic conditions allow exploiting simultaneous transmission and reception, the increased inter-cell interference becomes the limiting factor in leveraging the gain of full duplex communication. System level results show that full duplex provides a significant gain (though inferior to the expected doubling) over half duplex systems for scenarios which are not severely interference limited. Nevertheless, the full duplex gain is rather limited in the case of very dense deployments due to the higher inter-cell interference levels.

Future work will focus on the study of FD with asymmetric traffic and multi-user cells. In particular, the latter allows us to investigate situations where only access points may be full duplex capable, while user equipments adopt traditional half duplex transmission. In addition, power control in both uplink and downlink is currently under investigation.

References

[1] "Cisco visual networking index: Global mobile data traffic forecast update, 2014-2019," Feb 2015.

[2] P. Mogensen *et al.*, "Centimeter-wave concept for 5G ultra-dense small cells," in *IEEE 79th Vehicular Technology Conference (VTC Spring)*, May 2014, pp. 1–6.

[3] S. Hong *et al.*, "Applications of self-interference cancellation in 5G and beyond," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 114–121, February 2014.

[4] J. I. Choi *et al.*, "Achieving single channel, full duplex wireless communication," in *Proceedings of the 16th Annual International Conference on Mobile Computing and Networking (Mobicom)*. New York, USA: ACM, 2010, pp. 1–12.

[5] E. Aryafar *et al.*, "MIDU: Enabling MIMO full duplex," in *Proceedings of the 18th Annual International Conference on Mobile Computing and Networking (Mobicom)*. New York, USA: ACM, 2012, pp. 257–268.

[6] X. Xie and X. Zhang, "Does full-duplex double the capacity of wireless networks?" in *Proceedings of IEEE INFOCOM*, April 2014, pp. 253–261.

[7] B. Day *et al.*, "Full-duplex bidirectional MIMO: Achievable rates under limited dynamic range," in *Conference on Signals, Systems and Computers (ASILOMAR)*, November 2011, pp. 1386–1390.

[8] A. Cirik, R. Wang, and Y. Hua, "Weighted-sum-rate maximization for bidirectional full-duplex MIMO systems," in *Conference on Signals, Systems and Computers (ASILOMAR)*, November 2013, pp. 1632–1636.

[9] N. Mahmood *et al.*, "On the potential of full duplex communication in 5G small cell networks," in *IEEE 81st Vehicular Technology Conference (VTC Spring)*, May 2015, pp. 1–5, accepted.

[10] "Technical specification group radio access network; enhanced performance requirement for LTE user equipment (UE)," Jan 2013.

[11] M. Gatnau *et al.*, "Improving link robustness in 5g ultra-dense small cells by hybrid arq," in *IEEE 11th International Symposium on Wireless Communications Systems (ISWCS)*, Aug 2014, pp. 491–495.

[12] D. Catania *et al.*, "A distributed taxation based rank adaptation scheme for 5G small cells," in *IEEE 81st Vehicular Technology Conference (VTC Spring)*, May 2015, pp. 1–5, accepted.

[13] "Further advancements for E-UTRA physical layer aspects (Release 9)," Mar 2010.

[14] "WINNER II channel models," Feb 2008.

Paper G

Can Full Duplex Boost Throughput and Delay of 5G Ultra-Dense Small Cell Networks?

Marta Gatnau Sarret, Gilberto Berardinelli, Nurul H. Mahmood, Preben Mogensen

The paper has been published in the
IEEE 83rd Vehicular Technology Conference (VTC Spring), 2016.

Abstract

Given the recent advances in system and antenna design, practical implementation of full duplex (FD) communication is becoming increasingly feasible. In this paper, the potential of FD in enhancing the performance of 5^{th} generation (5G) ultra-dense small cell networks is investigated. The goal is to understand whether FD is able to boost the system performance from a throughput and delay perspective. The impact of having symmetric and asymmetric finite buffer traffic is studied for two types of FD: when only the base station is FD capable, and when both the user equipment and base station are FD nodes. System level results indicate that there is a trade-off between multiple-input multiple-output (MIMO) spatial multiplexing and FD in achieving the optimal system performance. Moreover, results show that FD may be useful for asymmetric traffic applications where the lightly loaded link requires high level performance. In such cases, FD can provide an average improvement of up to 116% in session throughput and 77% in packet delay compared to conventional half duplex transmissions.

1 Introduction

Recent advances in system and antenna design have been attracting the attention of the research community to make full duplex (FD) systems a reality. The promise of FD to double the throughput of traditional half duplex (HD) systems may help the future 5^{th} generation (5G) radio access technology (RAT) to accommodate the expected mobile traffic demands [1]. Industry and academia are involved in the design of such a new RAT. In [2], we have proposed a new 5G concept optimized for ultra-dense small cells using time division duplex (TDD) as operational mode. The usage of a 4x4 multiple-input multiple-output (MIMO) antenna configuration equipped with interference suppression receivers is assumed for all the nodes in the network.

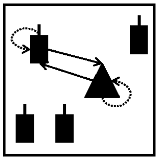
FD allows a device to simultaneously transmit and receive in the same frequency band. Nevertheless, FD requires high levels of self-interference cancellation (SIC), i.e., a significant attenuation of the interference caused at the receive antenna from its own transmitted signal. In [3] and [4], the authors show that a $\sim$100 dB reduction of the residual self-interference can be obtained by combining passive and active suppression techniques. Then, assuming such level of SIC and considering a small cell scenario with low power nodes located at short distances, inter-cell interference (ICI) may dominate over the residual self-interference. Since simultaneous transmission and reception doubles the number of interfering streams, we believe that the theoretical double gain that FD promises might not be always possible to achieve, as also discussed in [5]. In order to reduce such ICI, current work at system level focuses mainly on scheduling algorithms and power control mechanisms. In [6], the design of a hybrid FD/HD scheduler is presented. The algorithm selects the mode that maximizes the network throughput, obtaining promising gains in a Long Term Evolution (LTE) TDD system. The authors assume full buffer traffic and FD capable base stations (BSs). Nevertheless, such result might be biased since the impact of the increase in ICI due to FD communication is downplayed by assuming relatively isolated cells. Furthermore, traffic constraints are not considered. A power control algorithm is proposed in [7], which leads to an efficient switching between HD and FD in order to maximize the sum rate of downlink (DL) and uplink (UL). Nevertheless, the mentioned works target a single cell scenario with larger cell dimensions than ours. In our ultra-dense small cell network, such protocols may not be useful given the short distances and the low transmit powers.

Few works in the literature focus on the FD performance considering a complete system which includes all the layers active. In [8], the authors present a 802.11 system with FD capable nodes. Results comparing MIMO HD and FD are presented, considering both an isolated single cell and an interfered multi-cell scenario. They claim that, without interference, FD can provide up to 31% and 36% gain in terms of throughput and delay, respectively, while in case of interfered scenarios, HD may outperform FD due to MIMO spatial multiplexing gains. Our previous study [9] analyzed the performance of FD versus HD in a multi-cell scenario. Each cell was deployed with a BS and a single user equipment (UE), both FD capable, assuming ideal SIC and symmetric finite buffer traffic. Our conclusions are similar to [8], showing that FD may not be able to achieve the theoretical throughput gain in certain scenarios.

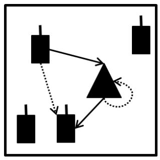
In this work, we take a step forward and investigate the performance of two different types of FD in the envisioned 5G ultra-dense small cell networks with multiple users. The first case refers to all devices being FD capable (bidirectional FD) and the second one to only the base station operating in FD (BS FD). Furthermore, the work is expanded by considering both symmetric and asymmetric traffic. The latter refers to the case where one of the links is offered more load than the other. As discussed in [4], such type of traffic might have a negative impact on the FD performance.

The paper is structured as follows. Section II presents the design of FD in the envisioned 5G system. Section III describes the simulator and the simulation parameters. System level results are discussed in Section IV. Finally, Section V concludes the paper and states the future work.

Fig. G.1: Envisioned 5G frame structure



(a) Bidirectional FD



(b) Base station FD

Fig. G.2: Full duplex types

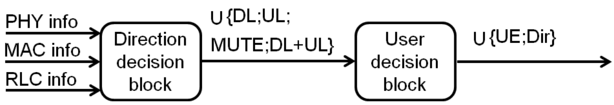


Fig. G.3: FD module structure

2 Full Duplex in 5G Small Cells

Our envisioned 5G system, described in detail in [2], uses TDD as operational mode and orthogonal frequency division multiplexing (OFDM) as modulation scheme. Nodes are synchronized in time and frequency and are equipped with 4x4 MIMO transceivers and interference rejection combining (IRC) receivers [10]. A novel frame structure of duration 0.25 ms, shown in Figure I.1, is defined as the transmission time interval (TTI). The first two OFDM symbols are dedicated to the DL and UL control, respectively. The remaining symbols are dedicated to the data, including the demodulation reference signal (DMRS) symbol, used for channel estimation. Notice that only time domain scheduling is considered.

The cell operations are here described. In the first place, the base station sends the scheduling grant (SG) in the DL control symbol of TTI_n, including the intended UE, the link direction and the transmission parameters, i.e., the modulation and coding scheme (MCS) and the number of spatial streams, often referred as transmission *rank*. Such decision is applied in TTI_{n+1} due to the processing time, thus leaving one TTI delay between the scheduling and the corresponding transmission. Then, in the UL symbol, UEs send the scheduling request (SR). The SR includes buffer information, hybrid automatic repeat and request (HARQ) feedback and the MCS and rank derived from their measurements. It is important to notice that the delay between the time when the channel is measured and the time when the transmission occurs may affect the link adaptation performance, since the transmission direction (DL HD, UL HD or FD) may change at each TTI.

Since all the nodes in the network are time aligned, they will be transmitting their reference sequences in the DMRS symbol, allowing every receiver to estimate both desired and interfering channel. Then, the IRC receiver will exploit such information to suppress part of the interfering streams according to the transmission rank and the number of antennas [10].

The usage of the same frame structure for both UL and DL allows for a straightforward extension of the envisioned 5G concept, originally designed for TDD, to FD transmission. The FD techniques investigated in this study are depicted in Figure G.2. In this figure, full lines represent the intended transmissions and dotted lines symbolize interference. Note that the ICI is not depicted in the figure and it would be added on top of the own-cell interference. The first type is the bidirectional FD (Figure G.2a), where both BS and UEs are FD capable. In this case, since the communication is always between the same pair BS-UE, both nodes only perceive their own self-interference. Figure G.2b shows BS FD mode, where only the BS is FD capable. In this case, the DL and UL links are different. Then, besides the self-interference, there is the intra-cell interference, i.e., the interference from the UL UE to the DL UE.

The FD module shown in Figure H.2 decides whether FD can be exploited and which are the nodes to be scheduled. It is divided into two blocks, in order to split functionalities and thus reduce complexity. The first block, the *direction decision block*, is in charge of deciding the optimal link direction per UE. This decision is taken according to the information received from the physical (PHY), MAC and radio link control (RLC) layers, such as signal-to-interference-plus-noise ratio (SINR) measurements, HARQ feedback, buffer status reports and link information provided by the UE to the BS. The decisions made by the *direction decision block* are then transferred to the *user decision block*. This second block is in charge of determining whether FD can be exploited or not and to decide which of the UE(s) to schedule.

The output value of the *direction decision block* is a vector of pairs with length equal to the number of UEs in the cell. Each pair contains the UE identifier and its optimal transmission direction. This direction can be *DL*, *UL*, *DL+UL* or *MUTE* and it is decided differently depending on the system:

- **HD and BS FD**: for these two cases, the procedure to extract the optimal link direction is the same. In BS FD a UE cannot be scheduled in both links, i.e., it is a HD node. Therefore, since the decision is taken from a user perspective, the algorithm deciding the optimal link should be the same for both HD and FD. The transmission direction is decided based on

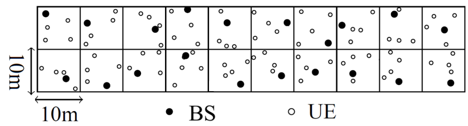


Fig. G.4: Simulated scenario

the offered load of each link, and thus the amount of dedicated resources is proportional to the offered load. For example, let us assume asymmetric traffic, where the highly loaded link (DL) offers 6 times more load than the lightly loaded link (UL). In this case, the DL would get approximately six times more resources than the UL, and it would have higher priority. Consequently, the UL would have to wait longer to be scheduled. Furthermore, the algorithm also takes into account fairness, by granting a minimum amount of resources to a link, in order to avoid its starvation. The possible output directions in this case are *DL*, *UL* or *MUTE*. The latter corresponds to the case where both UL and DL buffers are empty.

- **Bidirectional FD**: the transmission direction is based on the buffer state. For each user, the *direction decision block* checks if there is data in both the DL and UL buffers. In case of bidirectional FD, simultaneous transmission and reception will only be exploited in case a UE can be scheduled in both links, i.e., when both UL and DL buffers are filled with data. Then, if this is the case, the transmission direction for that user is *DL+UL*. Otherwise, it is *DL*(*UL*) if the UL(DL) buffer is empty and the DL(UL) is not, or *MUTE* if the UL and DL buffers are both empty.

The output value of the *user decision block* is taken according to the input provided by the *direction decision block*. It is a set containing the UE identifier(s) and its corresponding link direction(s) for HD(FD). In case of BS FD, a FD transmission is performed if two different UEs with opposite link directions can be scheduled; otherwise, the TTI is going to be HD. In case of bidirectional FD, it will be possible to exploit FD if at least one user has associated the *DL+UL* state. Note that in both cases, scheduling a FD transmission is always given priority over scheduling a HD one.

3 Simulation Environment

The results generated in this study are extracted from our event-driven based system level simulator. It includes the implementation of the PHY and MAC layers according to the 5G envisioned design presented in Section 2. The In-

Table G.1: Used parameters to run the simulations

Parameter	Value/State/Type
System parameters	BW = 200MHz; f_c = 3.5GHz
Frequency reuse	1 (whole band)
Propagation model	WINNER II A1 w/fast fading [11]
Wall penetration loss	5 dB [11]
Antenna configuration	4x4
Receiver type	IRC
Transmission power	10 dBm (BS and UE)
Self-interference cancellation	Ideal
Link adaptation filter	Log average of 5 samples
Rank adaptation	Taxation-based [12]
HARQ max retransmissions	4
RLC mode	Unacknowledged
RLC reordering timer	60 ms
Transport protocol	UDP
Traffic type	Symmetric and asymmetric finite buffer
Simulated time per drop	2-12 seconds
Number of drops	50

ternet protocol (IP) is only modeled as overhead and the RLC, transmission control protocol (TCP) and user datagram protocol (UDP) layers are completely modeled. In addition, a vertical Radio Resource Management (RRM) layer is considered. The RRM holds the FD module and decides the transmission parameters. The MCS is derived from the log-average of the last 5 SINR samples and the rank is obtained using a taxation-based rank adaptation algorithm [12]. Such algorithm decides the rank according to the perceived interference, applying higher taxation to the higher ranks to reduce the overall network interference level. Finally, the *user decision block* is implemented following a round robin fashion. Further details of the simulator are described in [13].

The simulation parameters are presented in Table I.1. In terms of deployment, we simulate a grid of 10×2 cells, each of them of dimension 10×10 m^2, as shown in Figure I.5. In each simulation drop, a cell is deployed with a BS and 4 UEs randomly placed in a closed subscriber group manner, i.e., all the UEs within a cell are connected to the BS in that cell. Ideal SIC is assumed since according to [3] [4] and considering a scenario where the distances are short and the transmission power for all nodes is low, the impact of SIC may be negligible. Finally, we consider RLC unacknowledged mode (UM) [14] and UDP to avoid the impact of the control overhead, thus HARQ being the only active recovery mechanism.

The performance of the two FD cases described in Section 2 are compared against HD performance. We consider two types of finite buffer traffic, symmetric and asymmetric. In case of symmetric traffic, the offered load is the same in DL and UL (1DL:1UL), while for the asymmetric traffic case, the offered load in DL is 6 times larger than in UL (6DL:1UL). In addition, three different loads are simulated: low, medium and high, which are defined to obtain a channel occupancy of 25%, 50% and 75% under ideal conditions, respectively. The used traffic model generates payloads of average 2 megabytes (i.e., sessions), arriving every $t_{arrival}$ seconds, and both parameters have a negative exponential distribution [9] [15]. The results are presented in terms cumulative distribution function (CDF) of the average session throughput (TP) and the packet delay. The former is calculated as the sum of the node average session throughput per cell, extracting DL and UL statistics separately. The latter is the time interval between the generation of a packet and its correct reception. Notice that the transmitted data unit is composed by a set of aggregated packets. The number of aggregated packets is determined by the available channel bandwidth and the selected MCS and transmission rank. Finally, numerical tables specify the average throughput session gains of FD over HD and the average packet delay reduction, both in percentage. When a minus symbol ($-$) appears on the throughput results (i.e., reduction) or a plus symbol ($+$) appears on the delay results (i.e., increase), it indicates the instances where HD outperforms FD.

4 Performance Evaluation

System level results are discussed in this section, which is divided into two subsections. In the first one, the performance of both types of FD considering symmetric traffic is discussed. In the second subsection, the same exercise is repeated assuming asymmetric traffic.

4.1 Symmetric traffic

This first part of the study focuses on the analysis of the HD and FD system performances with symmetric traffic. In case of the bidirectional FD, the only interference source comes from the neighboring cells, meaning that all nodes will be affected by approximately the same level of interference. In addition, both DL and UL get the same amount of resources since the offered load is the same in both links. Consequently, the performance in the two directions, DL and UL, is expected to be the same. Figure G.5 and Figure G.6

show the DL average session throughput and the DL packet delay, respectively. From the figures we can observe that the gain that FD provides over HD is higher as the load increases, since the probability of exploiting FD also increments because there is a higher chance that UL and DL data coincide in the buffers. Nevertheless, the throughput improvement is below the theoretical 100% gain. This is because the HD system is able to exploit higher transmission ranks, since it perceives less interference compared to FD. This means that the gain that FD provides over using MIMO spatial multiplexing is limited. Nevertheless, at high load the channel occupancy is high, meaning that HD will not be able to exploit MIMO spatial multiplexing, thus using the same rank as FD. In this case, even if FD perceives larger number of retransmissions, simultaneous transmission and reception provides a significant gain (36% in throughput and 26% in delay) because it allows reducing the buffering time and the piling up of packets. Notice that the slope change in Figure G.6 is caused by the RLC reordering timer [14]. Table G.2(a) summarizes the average results for all the loads.

In the BS FD case, the channel conditions are no longer the same in both links. The intra-cell interference caused by the UL UE to the DL UE within the same cell is very strong, causing a negative impact on the DL performance. Consequently, it performs worse than the UL. In addition, since DL perceives higher number of retransmissions compared to UL, it creates an asymmetry between the two links, which does not exist in the offered load. Therefore, the *direction decision block* will start allocating more resources to DL, thus reducing the number of UL TTIs and consequently worsening the UL performance. Such behavior is observed from the average numbers in Table G.2(b). We can see that, regardless of the considered load, the DL performance (throughput and delay) is always better with HD. On the other hand, the UL delay is slightly improved with FD, while only the throughput for the medium load case can be hardly increased.

Up to this point, the performance of bidirectional FD and BS FD has been analyzed. Comparing the average gains provided in Table G.2, it is obvious that their performance is different. On one hand, this is due to the different interference conditions that both systems perceive. BS FD has to deal with the intra-cell interference on top of the ICI, while bidirectional FD has to handle only the latter. On the other hand, the algorithm that dictates the optimal transmission direction (i.e., *direction decision block*) is different for both types of FD. Looking into the probability of exploiting FD, we observe that it goes from 11% to 32% for the bidirectional FD, while for the BS FD it goes from 15% up to 44%. An increase in the number of retransmissions in the DL direction of the BS FD case increases the probability of FD, but it also means that a higher level of interference is kept for longer time due to simultaneous transmission and reception. Since the targeted scenario is interference limited, a larger increase in the interference level makes HD perform better than FD.

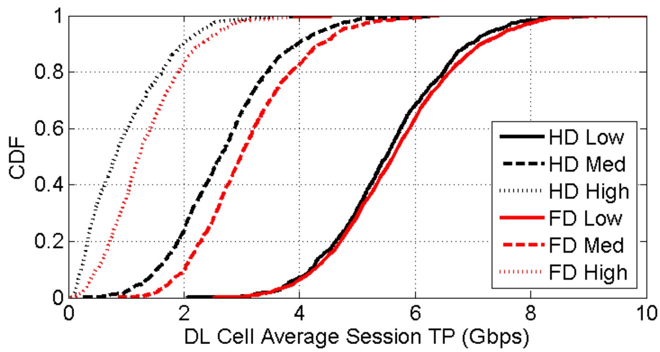


Fig. G.5: Session TP with symmetric traffic (bid. FD)

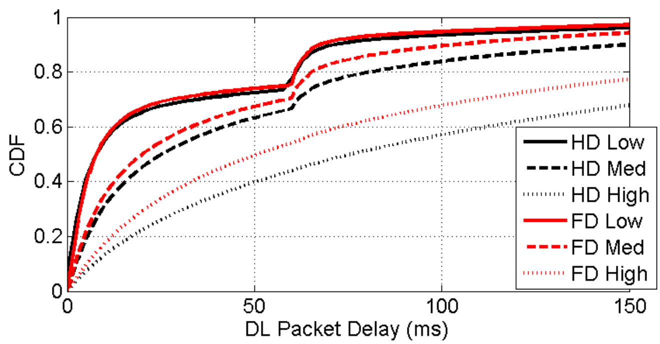


Fig. G.6: Packet delay with symmetric traffic (bid. FD)

Hence, we can conclude that a trade-off between MIMO spatial multiplexing and FD dictates the optimal system performance, which is fully related to the interference conditions.

4.2 Asymmetric traffic

The traffic asymmetry between DL and UL is 6:1, i.e., the offered load is 6 times bigger in DL compared to UL. Consequently, the expected gain from FD is different in each link. In HD, the DL gets more resources than the UL since it is the highly loaded link, thus delaying the UL transmissions and reducing its throughput. In FD, the UL direction does not need to wait, since it may be transmitted at the same time the DL occurs, thus improving the delay and throughput. The DL improvement is expected not to be significant since such link is already getting high channel occupancy.

The described behavior is observed for bidirectional FD in Table G.3(a),

Table G.2: FD vs HD performance with symmetric traffic

Load	TP	Delay
Low	3%	10%
Med	15%	22%
High	36%	26%

(a) TP gain and delay reduction (bid. FD)

Load	DL TP	UL TP	DL delay	UL delay
Low	-1%	3%	+3%	5%
Med	-4%	5%	+13%	4%
High	-14%	0	+11%	5%

(b) TP gain and delay reduction (BS FD)

where the gain is always higher in UL, and it increases with the load because the probability of FD is higher. For low and medium load, the gain is limited due to reduced FD probability (from 6 to 10%) and the already mentioned trade-off between MIMO spatial multiplexing and FD. At high load, the gain increases significantly in the UL due to the buffering. When the system perceives congestion, MIMO spatial multiplexing is not exploited, since the interference levels are too high. Furthermore, the UL in HD gets very few resources because it is the lightly loaded link, so data is accumulated in the buffer. This behavior is mitigated with bidirectional FD.

It has already been stated that the DL performance in case of asymmetric traffic might show limited improvement with FD. Nevertheless, consider the case of BS FD, where the intra-cell interference causes a negative impact on the DL direction. Table G.3(b) show that, independently of the load, both throughput and delay in DL provide better results with the HD system. Such behavior is justified by the trade-off between MIMO spatial multiplexing and FD, the extra intra-cell interference and the consequent increase in retransmissions. On the other hand, it is interesting to notice that the UL direction can be improved regardless of the load and both in terms of throughput and delay, as shown in Figure G.7 and Figure G.8, respectively. As in the symmetric traffic case, due to an increase of the DL retransmissions, the probability of exploiting FD in BS FD (from 8 to 15%) is higher than in bidirectional FD (from 6 to 10%). Such increase in retransmissions also enlarges the already existing asymmetry between DL and UL, thus reducing the amount of resources dedicated to UL and consequently worsening its performance. The slope change in Figure G.8, as previously observed in the symmetric traffic case, is caused by the RLC reordering timer.

To summarize this second study case, we conclude that FD is a good so-

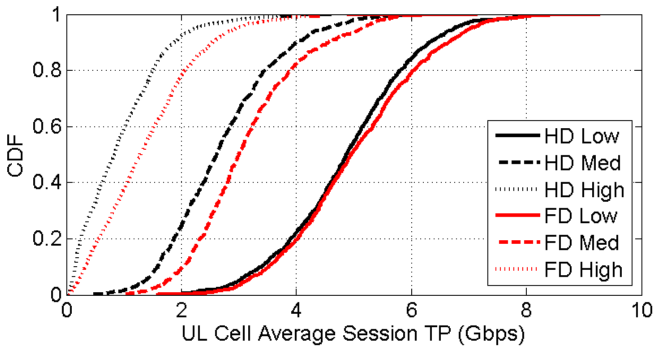


Fig. G.7: UL session TP with asymmetric traffic (BS FD)

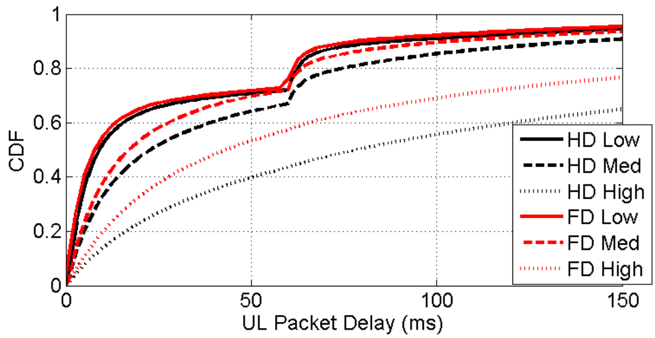


Fig. G.8: UL packet delay with asymmetric traffic (BS FD)

lution for applications with asymmetric traffic, where the lightly loaded link requires a boost in performance and the highly loaded link is not the target. Furthermore, in case the system is close to congestion, FD is an attractive solution to deal with the buffering, since simultaneous transmission and reception mitigates this effect.

After the analysis of both types of FD under traffic constraints, it is noticeable that bidirectional FD provides better results than BS FD. This is because ultra-dense small cell networks are already interference limited scenarios and BS FD has to deal with intra-cell interference on top of the ICI. Furthermore, having BS FD is more practical from a cost perspective since having FD capable UEs is still a challenge.

Table G.3: FD vs HD performance with asymmetric traffic

Load	DL TP	UL TP	DL delay	UL delay
Low	0	2%	2%	12%
Med	4%	18%	8%	29%
High	16%	116%	28%	77%

(a) TP gain and delay reduction (bid. FD)

Load	DL TP	UL TP	DL delay	UL delay
Low	-1%	3%	+3%	6%
Med	-2%	16%	+4%	19%
High	-13%	46%	+20%	29%

(b) TP gain and delay reduction (BS FD)

5 Conclusions and future work

In this paper we have investigated the potential of bidirectional FD and base station FD in ultra-dense small cell networks, under different type of traffic constraints. System level results prove that, in interference limited scenarios, there is a trade-off between the use of MIMO spatial multiplexing and simultaneous transmission and reception to obtain the optimal performance, since the latter may double the interference level. Our results show that in the targeted scenario, base station FD provides limited gain since it has to deal with intra-cell interference. Bidirectional FD shows better performance. In this case, the system is always improved in terms of throughput and delay, regardless of the offered load. Furthermore, in case of asymmetric traffic, both types of FD communication show an interesting potential for applications where the lightly loaded link requires a boost in performance. Finally, future work will focus on the impact of recovery mechanisms, such as RLC acknowledged mode and TCP.

References

[1] "Cisco visual networking index: Global mobile data traffic forecast update, 2014-2019," Feb 2015.

[2] P. Mogensen *et al.*, "Centimeter-wave concept for 5G ultra-dense small cells," in *IEEE 79th Vehicular Technology Conference (VTC Spring)*, May

2014, pp. 1–6.

[3] M. Heino *et al.*, “Recent advances in antenna design and interference cancellation algorithms for in-band full duplex relays,” *IEEE Communications Magazine*, vol. 53, no. 5, pp. 91–101, May 2015.

[4] K. Thilina *et al.*, “Medium access control design for full duplex wireless systems: challenges and approaches,” *IEEE Communications Magazine*, vol. 53, no. 5, pp. 112–120, May 2015.

[5] L. Wang *et al.*, “Exploiting full duplex for device-to-device communications in heterogeneous networks,” *IEEE Communications Magazine*, vol. 53, no. 5, pp. 146–152, May 2015.

[6] S. Goyal *et al.*, “Improving small cell capacity with common-carrier full duplex radios,” in *IEEE International Conference on Communications (ICC)*, Jun 2014.

[7] R. Zhang *et al.*, “Investigation on DL and UL power control in full-duplex systems,” in *IEEE International Conference on Communications Workshops (ICC)*, June 2015, pp. 1–6.

[8] W. Zhou and K. Srinivasan, “Sim+: A simulator for full duplex communications,” in *International Conference on Signal Processing and Communications (SPCOM)*, Jul 2014, pp. 1–6.

[9] M. Gatnau *et al.*, “Full duplex communication under traffic constraints for 5G small cells,” in *IEEE 82nd Vehicular Technology Conference (VTC Fall)*, Sep 2015, pp. 1–5.

[10] “Technical specification group radio access network; enhanced performance requirement for LTE user equipment (UE),” Jan 2013.

[11] “WINNER II channel models,” Feb 2008.

[12] D. Catania *et al.*, “A distributed taxation based rank adaptation scheme for 5G small cells,” in *IEEE 81st Vehicular Technology Conference (VTC Spring)*, May 2015, pp. 1–5.

[13] M. Gatnau *et al.*, “Improving link robustness in 5G ultra-dense small cells by hybrid ARQ,” in *IEEE 11th International Symposium on Wireless Communications Systems (ISWCS)*, Aug 2014, pp. 491–495.

[14] “Technical specification group radio access network; evolved universal terrestrial radio access (E-UTRA) radio link control (RLC) protocol specification,” Dec 2007.

[15] “Further advancements for E-UTRA physical layer aspects (Release 9),” Mar 2010.

Paper H

Impact of Transport Control Protocol on Full Duplex Performance in 5G Networks

Marta Gatnau Sarret, Gilberto Berardinelli, Nurul H. Mahmood, Preben Mogensen

The paper has been published in the
IEEE 83rd Vehicular Technology Conference (VTC Spring), 2016.

Abstract

Full duplex (FD) communication has attracted the attention of the industry and the academia as an important feature in the design of the future 5th generation (5G) wireless communication system. Such technology allows a device to simultaneously transmit and receive in the same frequency band, with the potential of providing higher throughput and lower latency compared to traditional half duplex (HD) systems. In this paper, the interaction between Transport Control Protocol (TCP) and FD in 5G ultra-dense small cell networks is studied. TCP is a well-known transport layer protocol for providing reliability, which comes at the price of increased delay and reduced system throughput. FD is expected to accelerate the TCP congestion control mechanism and hence mitigate such consequences. System level results show that FD can outperform HD and alleviate the TCP drawbacks when the inter-cell interference is not the main limiting factor. On the other hand, under strong inter-cell interference, results show that the capabilities of the system to cope with such interference dictates the gain that FD may provide over HD.

1 Introduction

Full duplex (FD) technology allows a device to transmit and receive simultaneously in the same frequency band, ideally doubling the throughput over conventional half duplex (HD) systems. However, building an operational FD node requires a high level of self-interference cancellation (SIC), i.e., a high attenuation of the transmitted signal at the own receive antenna. Current achievable levels of SIC are in the order of 100 dB [1], thus making feasible the implementation of a FD device. For this reason, FD is considered as a potential candidate for a future 5$^{\text{th}}$ generation (5G) radio access technology (RAT). Besides residual self-interference (SI), other limitations such as inter-cell interference (ICI) and traffic constraints [2] may also reduce the theoretical 100% throughput gain.

The design of the 5G RAT is still under discussion by the industry and the academia. We presented our vision in [3]. The system was originally designed as a HD time division duplexing (TDD) system but it can easily accommodate FD technology. 5G is targeting a massive and uncoordinated deployment of small cells, where all nodes are equipped with multiple-input multiple-output (MIMO) antenna technology and receivers with interference rejection capabilities.

A detailed study of the techniques for SIC is presented in [1]. The authors evaluated SI suppression using a testbed, showing $\sim$100 dB of cancellation. They conclude that in dense deployment of small cells, where transmit powers are low and distances among nodes are short, such level of SIC is enough to consider that ICI becomes a major limitation to achieve the promised FD gain. Moreover, they remark that large asymmetric traffic ratios between downlink (DL) and uplink (UL) data may compromise the usage of FD and hence its gain. These challenges are also described in [4]. Authors in [5] evaluate a FD network considering asymmetric traffic, showing that FD always outperforms HD. However, the authors assume a strong isolation between the cells, which may mitigate the ICI impact. Malik et al. propose a solution based on power control to accommodate asymmetric traffic [6]. The proposed scheme shows an improvement in DL at the expenses of lowering the UL rate. However, the analysis is carried on a single cell scenario. Finally, the authors in [2] study the impact of symmetric and asymmetric traffic in a multi-cell scenario. Throughput results show that the FD gain reduces with the perceived ICI and the traffic ratio. It is important to notice that the mentioned work disregards the usage of features such as link adaptation or recovery and congestion control mechanisms.

As previously stated, traffic constraints have an impact on the FD performance. According to [7], most of the Internet traffic is carried over Transport Control Protocol (TCP) flows, with a small percentage of User Data Protocol (UDP) flows. TCP [8] is used to provide a reliable communication and reduce packet losses as much as possible. The congestion control mechanism provided by TCP limits the amount data that can be pushed into the network, based on the reception of positive acknowledgments (ACKs) [9]. This procedure causes an increase in the delay and a reduction of the system throughput. Such drawbacks may be mitigated by FD since it may allow to accelerate the TCP congestion control mechanism, given the possibility of transmitting and receiving simultaneously.

This paper focus on the analysis of FD performance in 5G ultra-dense small cell networks with TCP traffic, considering the congestion control and recovery mechanisms defined by this protocol. To the best of the authors knowledge, this is the first work investigating the interaction between the TCP mechanism and the FD technology. This work extends our previous contribution [10] by considering multi-user cells and the TCP protocol. We study the case of bidirectional FD, where both access points (APs) and user equipments (UEs) are FD capable. Two types of traffic are considered: symmetric, when the ratio between the DL and the UL load is the same, and asymmetric, when the load in DL is larger than in UL. Results are extracted via system level simulations.

The paper is structured as follows. Section II presents the envisioned 5G system. Section III describes the interaction between TCP and FD. Section IV defines the simulation environment. System level results are discussed in Section V. Section VI concludes the paper and states the future work.

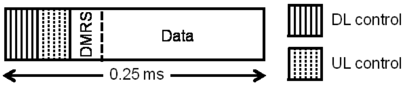


Fig. H.1: Envisioned 5G frame structure

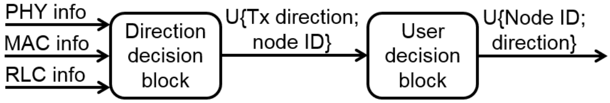


Fig. H.2: Module that decides the link direction, the transmission mode and the scheduled node(s)

2 Full Duplex in 5G Small Cells

In [3], we presented the design of our envisioned 5G system. It was originally designed as a HD TDD system, targeting a massive and uncoordinated deployment of small cells. Nodes are assumed to be synchronized in time and frequency. The system uses a novel frame structure of duration 0.25 ms, defined as the Transmission Time Interval (TTI) and shown in Figure I.1. A scheduling grant containing transmission parameters such as the link direction, the modulation and coding scheme (MCS) or the number of transmission streams (i.e., transmission *rank*) is sent within the DL control symbol. UE specific information is sent within the UL control symbol, including channel and buffer state information and Hybrid Automatic Repeat Request (HARQ) feedback. The data part carries UL or DL data in case of HD, and both UL and DL data in FD. Note that the transmission direction may change every 0.25ms. Thus, a TTI may be DL HD, UL HD or FD, independently of the decisions from previous TTIs. All nodes are equipped with 4×4 MIMO antenna configuration and advanced receivers, such as Interference Rejection Combining (IRC) [11]. These receivers use the degrees of freedom from the antenna domain to suppress incoming interference.

This work focuses on the performance of bidirectional FD, which refers to the case where both APs and UEs can simultaneously transmit and receive. In this case, a node may perceive SI and ICI, since FD is always exploited between the same pair AP-UE, thus avoiding intra-cell interference. Figure H.2 shows the structure of the module that decides the link direction, the transmission mode (HD or FD) and the scheduled node(s). This module is located in the Radio Resource Management (RRM) layer and it is divided into two blocks, *direction decision block* and *user decision block*, in order to separate functionalities and thus reduce complexity. In the first step, the optimal transmission direction per node is extracted, based on information from the

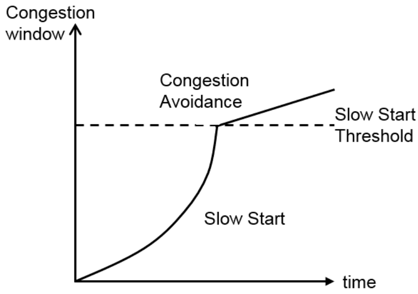


Fig. H.3: TCP congestion window

physical (PHY), medium access control (MAC) and radio link control (RLC) layers. The output from the *direction decision block*, which corresponds to the union of pairs {*optimal transmission direction, node identifier*}, is transferred to the *user decision block*, where the transmission mode and the scheduled node(s) are then decided. A FD transmission has always priority over a HD one. In case there are more than one pair of nodes that are able to use FD, different time scheduling algorithms may be applied to decide which node to schedule. The output of the FD module is then the pair {*scheduled node, direction*}.

The procedure to extract the optimal transmission direction is different for HD and FD:

- **HD**: the optimal transmission direction is decided based on the amount of data which is currently in the buffers (UL and DL), and previous decisions. For example, in case of asymmetric traffic, where the DL traffic load is six times higher than in UL, a node will decide *DL* six times more than *UL* in average. The information regarding previous slot allocations is used to avoid the starvation of the lightly loaded link, i.e., at least one slot should be allocated to such link with a certain periodicity. The optimal direction can be *DL* or *UL* if there is data in at least one of the buffers, or *MUTE* if there is not data in either of them.

- **FD**: since we want to exploit FD as much as possible, the link decision is based only on the buffer size. This means that the optimal transmission direction will be *DL+UL* if there is data in both buffers, *DL* (*UL*) if the *UL* (*DL*) buffer is empty, or *MUTE* if both buffers are empty.

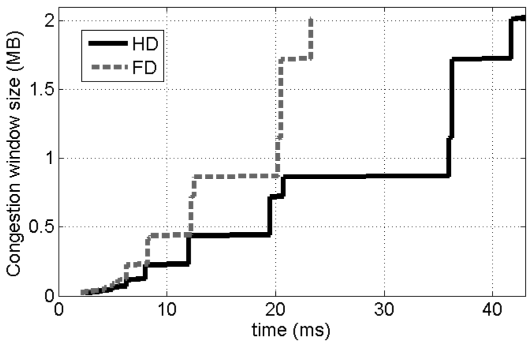


Fig. H.4: Congestion window growth

3 Interaction between Full Duplex and TCP

TCP [8] is a protocol that provides reliability by using a congestion control mechanism [9]. TCP limits the amount of data that can be sent through the channel based on the reception of positive ACKs. The congestion window, shown in Figure H.3, controls such limitation. During the *Slow Start* stage, the congestion window grows exponentially according to the received TCP ACKs. When the *Congestion Avoidance* phase is reached, the growth of the congestion window is linear, following the same principle on the TCP ACKs as the *Slow Start* phase. Nevertheless, TCP has an inherent impact on the system throughput and delay, since the amount of transmitted data is limited by the reception of positive feedback and consequently it will increase only if the channel conditions are favorable.

We believe that FD may help at mitigating the TCP drawbacks since simultaneous transmission and reception might increase the congestion window faster and help at reaching the *Congestion Avoidance* phase sooner, where a larger amount of data is transmitted within a single TTI. For clarification, a simple example is shown in Figure H.4. This figure shows the growth of the congestion window for HD and FD in a single cell scenario with one AP and one UE, where both nodes have a 2 megabytes file to transmit and the probability of exploiting FD is 100%. Shadowing and fast fading have been disabled to provide a fair comparison between both cases, and the general simulation parameters are listed in Table I.1 and they will be further discussed in the next section. We can observe that FD is able to transmit the file faster than HD since its congestion window grows faster. The transmission time is reduced by approximately 45% in this case. It is important to remark that in dense networks, ICI may slow down the growth of the congestion window. The performance of FD with TCP traffic in such networks

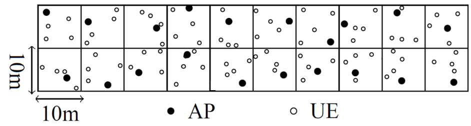


Fig. H.5: Simulated scenario

Table H.1: Used parameters to run the simulations

Parameter	Value/State/Type
System parameters	BW = 200MHz; f_c = 3.5GHz
Frequency reuse	1 (whole band)
Propagation model	WINNER II A1 w/fast fading [12]
Antenna configuration	4x4
Receiver type	IRC
Transmission power	10 dBm (BS and UE)
Self-interference cancellation	Ideal
Link adaptation filter	Log average of 5 samples
Rank adaptation	Taxation-based [13]
HARQ max retransmissions	4
RLC mode	Acknowledged
Transport protocol	UDP and TCP
TCP timer for ACK	100 ms
TCP initial retransmission timeout	1 ms
Segment size threshold	10 MB
Traffic type	Symmetric and asymmetric (6:1) finite buffer
Simulation time per drop	15-40 seconds
Number of drops	50

is discussed in section V.

4 Simulation Environment

Results are extracted from an event-based system level simulator. It implements the 5G MAC and PHY design described in Section II. Furthermore, it includes a detailed modeling of the RLC and TCP layers, and a vertical RRM layer that collects information from the PHY, MAC and RLC layers to provide the scheduling parameters. Finally, the application layer generates File Transfer Protocol (FTP) traffic [14] and the Internet Protocol (IP) layer is modeled as overhead.

In addition to the procedures described in Section II, the PHY and MAC layers also implement the HARQ retransmission mechanism and link and rank adaptation schemes. The link adaptation (LA) algorithm keeps track of the last five channel measurements to extract an accurate MCS. The rank adaptation (RA) algorithm is taxation-based [13]. It decides, according to the incoming interference, which is the most appropriate rank to reduce the overall network interference, i.e., how many MIMO antennas will be used for transmission and how many of them for IRC interference suppression [11]. The reader can refer to [13] for further details on the rank adaptation algorithm. Finally, the selected scheme for the *user decision block* is time domain round robin, i.e. frequency multiplexing is not considered.

In terms of deployment, a small cell is located in a 10×10 m^2 room, containing one AP and four UEs randomly deployed, with the UEs affiliated to the AP in the same cell (closed subscriber group). The multi-cell scenario refers to a 10×2 grid of small cells, as shown in Figure I.5. SIC is considered ideal, according to [1], given the current SIC capabilities, the short distances among nodes and the low transmit power, which is set to 10 dBm for all the nodes. The RLC mode is set to Acknowledged (AM) [15]. We assume that the RLC ACK is sent within the control channel, i.e., it does not generate additional overhead. The TCP implementation in the simulator is New Reno [16], which includes the recovery and congestion control mechanisms, whereas handshake procedures are not considered since they are not relevant for our studies. TCP parametrization and the remaining simulation parameters are listed in Table I.1.

The generated results compare HD and FD performance with TCP, whereas UDP [17] is considered for the sake of comparison. Notice that UDP acts as a transparent layer, sending everything that it receives to the upper layers, without performing error checking or congestion control. For both cases, RLC AM and HARQ are enabled. Two FTP traffic cases are studied: symmetric, where the offered load in DL and UL is the same (1DL:1UL ratio), and asymmetric, where the amount of DL data is six times higher than in UL (6DL:1UL ratio). For each case, three levels of load are considered: low, medium and high, corresponding approximately to 25%, 50% and 75% resource utilization (RU). The RU is defined as the percentage of time the medium is used. Results are presented in terms of average session throughput (TP), defined as the average of the individual session throughputs per link (UL or DL) or per cell (UL+DL), where the session throughput corresponds to the amount of time required to successfully transmit a session. Such a session is characterized by the packet size and the $t_{arrival}$ parameters, which are negatively exponential distributed [14]. The average packet size is 2 megabytes, and the average $t_{arrival}$ is set according to obtain the loads described above. The second key performance indicator (KPI) is the packet delay, defined as the time between the generation of a packet and its success-

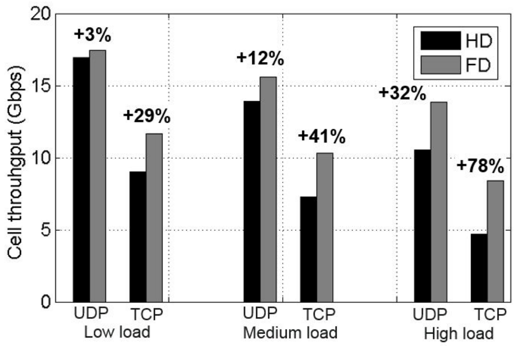


Fig. H.6: Single cell throughput with symmetric traffic

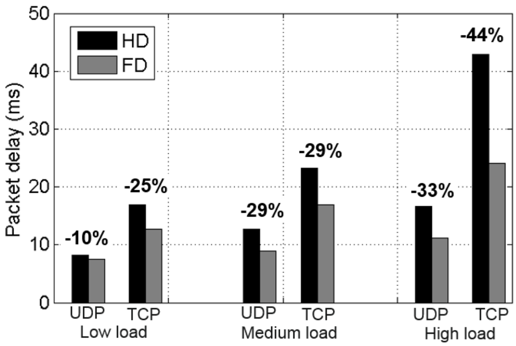


Fig. H.7: Single cell delay with symmetric traffic

ful reception, including the buffering time. Finally, percentages indicate the gain of FD over HD. In throughput gain, a plus (+) indicates that FD outperforms HD, and a minus (-) the opposite case. In delay reduction, (-) denotes better performance of FD and vice versa for (+).

5 Performance Evaluation

Results are divided according to the deployment, in order to isolate the ICI impact, since it may be the limiting factor in the achievable FD gain [1].

Table H.2: TP gain and delay reduction of FD over HD with asymmetric traffic in single cell scenario

Load	Traffic	DL TP	UL TP	DL delay	UL delay
Low	**UDP**	+1%	+10%	0	-16%
	TCP	+17%	+39%	-13%	-29%
Medium	**UDP**	+6%	+36%	-5%	-44%
	TCP	+37%	+77%	-21%	-43%
High	**UDP**	+5%	+69%	-6%	-61%
	TCP	+61%	+128%	-36%	-55%

5.1 Single cell scenario: avoiding inter-cell interference

Since this scenario is not affected by ICI, retransmissions rarely occur and the FD gain is only affected by the traffic constraints. Figure H.6 depicts the average cell throughput with symmetric traffic. This result shows that FD always outperforms HD, independent of the transport protocol, and the FD gain is higher when the load increases. However, notice that the FD gain obtained with TCP is higher than with UDP. This difference is a consequence of the TCP congestion control mechanism. Packets accumulate in the buffer because data transmission is controlled by the TCP congestion window, thus increasing the probability of having simultaneously UL and DL data. In this case, the FD probability ranges from 67% to 83% for TCP, while it goes from 4% to 11% for UDP because this protocol acts as a transparent layer. Figure H.7 shows the packet delay with symmetric traffic, which has a similar behavior as the throughput. We observe that FD also outperforms HD for all cases and the FD gain increases with the load.

For asymmetric traffic (see Table H.2), FD always outperforms HD in terms of throughput and delay, for both UDP and TCP. However, the UL gain is higher because in HD the UL gets less transmission opportunities since it is the lightly loaded link. Furthermore, in TCP, DL (UL) data needs to be acknowledged from the UL (DL) in order to increase the TCP congestion window and continue transmitting. In HD, DL suffers since UL has less transmission opportunities, hence delaying the UL TCP ACK transmission. In UL the impact is less significant because DL gets more transmission opportunities (highly loaded link) and the DL TCP ACK is transmitted with lower delay. This HD problem is solved with FD since the TCP ACK can be transmitted immediately in both UL and DL, allowing the congestion window to grow faster.

From this first analysis we can conclude that FD is able to improve TCP performance, providing better results than UDP. Such gains come from a

Table H.3: TP gain and delay reduction of FD over HD with symmetric traffic in multi-cell scenario

Load	Traffic	Cell TP	Average delay
Low	UDP	+2%	-8%
	TCP	+1%	+22%
Medium	UDP	+16%	-27%
	TCP	-50%	+284%
High	UDP	+41%	-26%
	TCP	-32%	+52%

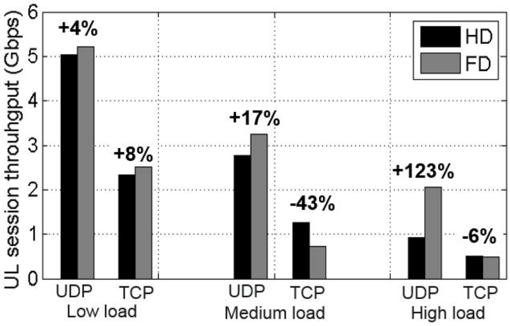


Fig. H.8: Multi-cell UL throughput with asymmetric traffic

faster growing of the TCP congestion window and the immediate transmission of the TCP ACK with FD. Moreover, without ICI, FD can always outperform HD, specially the lightly loaded link in case of asymmetric traffic.

5.2 Multi-cell scenario: impact of inter-cell interference

In the multi-cell scenario (Figure I.5), a node may perceive significant interference from its neighbors, meaning that FD will be affected by increased ICI compared to HD. Table H.3 shows the FD gain with symmetric traffic. We can observe that, with UDP, FD always outperforms HD, both in terms of throughput and delay, and the FD gain increases with the load. Nevertheless, with TCP, the situation is the opposite and FD leads to worse throughput and delay performance than HD in all cases. Such performance is caused by the interference conditions as a consequence of the FD probability. Such probability ranges from 11% to 34% with UDP and from 67% to 89% with TCP. This indicates that, the higher is the FD probability, the larger is the ICI, since

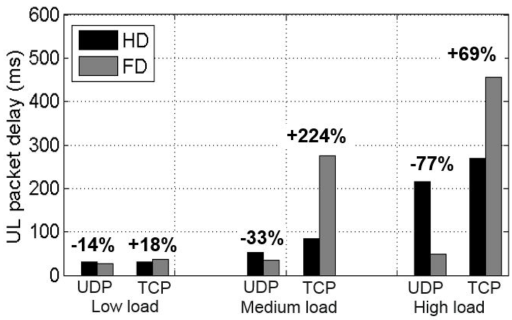


Fig. H.9: Multi-cell UL delay with asymmetric traffic

FD doubles the number of interfering streams. Notice that, according to the LA and RA algorithms, HD would then use higher rank and MCS than FD. Consequently, HD is able to transmit a larger amount of data, increase faster the TCP congestion window and get rid of the data sooner than FD. Then, HD occupies the medium for less time and reduces the interference generated to neighboring cells. Results shows that the worst case is at medium load, where the RU is 67% for HD and 92% for FD.

Figures H.8 and H.9 show the UL throughput and delay, respectively, corresponding to the performance of the lightly loaded link in the asymmetric traffic case. Results show the same trends as the symmetric traffic case. With UDP, FD shows the best performance, specially at high load (123% throughput gain an 77% delay reduction) because FD mitigates the buffering effect, since UL get more transmission opportunities than in HD. Table H.4 shows that also the DL is always improved with FD. However, with TCP, not even the lightly loaded link, which may perceive six times more resources in FD than in HD, can be improved. The reasoning is the same as for the symmetric traffic case. The FD probability, which has an effect on the ICI and hence on the transmission rank and MCS, goes from 4% to 11% with UDP and from 85% to 91% with TCP. Figure H.10 shows the UL transmission rank, where *R1* refers to rank one, *R2* to rank two, etc. From the figure, we observe that FD is already limited to rank one at medium and high load, while HD exploits MIMO spatial multiplexing at all loads (with very low probability at high load). Choosing a higher transmission rank allows the TCP congestion window to grow faster and reduce the ICI since the medium is freed before.

According to the presented results, we conclude that TCP leads to a higher FD probability, thus increasing the ICI and provoking a slower growth of the TCP congestion window. We showed that there is a trade-off between the MIMO antennas used for data transmission and the ones used for interference suppression to achieve the optimal system performance.

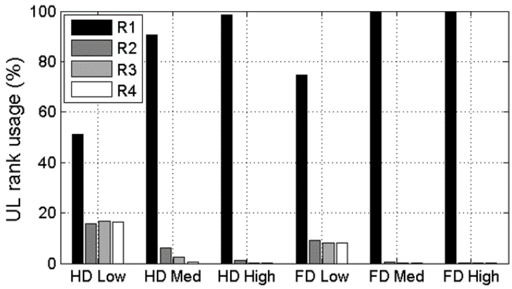


Fig. H.10: Multi-cell UL rank with TCP asymmetric traffic

Table H.4: TP gain and delay reduction of FD over HD with asymmetric traffic in multi-cell scenario

Load	Traffic	DL TP	UL TP	DL delay	UL delay
Low	UDP	+1%	+4%	-2%	-14%
	TCP	-9%	+8%	+45%	+18%
Medium	UDP	+3%	+17%	-5%	-33%
	TCP	-60%	-43%	+429%	+224%
High	UDP	+17%	+123%	-22%	-77%
	TCP	-51%	-6%	+123%	+69%

6 Conclusions and future work

In this paper we have investigated the performance of bidirectional FD in 5G ultra-dense small cell networks considering the impact of strong inter-cell interference, traffic constraints and TCP congestion control and recovery mechanisms. System level results show that, under ideal interference conditions, FD outperforms HD in terms of throughput and delay, and helps at reducing the increased latency inherent in the TCP mechanism. However, in case of significant inter-cell interference, HD may provide better system performance than FD due to the slower behavior of the TCP congestion control mechanism in FD. Therefore, we conclude that there is a trade-off between the MIMO degrees of freedom used for data transmission and the ones used for interference suppression to obtain the optimal system performance. Such trade-off is strongly linked to the probability of exploiting FD and hence to the level of inter-cell interference. Future work will focus on studying FD in device-to-device proximity discovery.

References

[1] M. Heino *et al.*, "Recent advances in antenna design and interference cancellation algorithms for in-band full duplex relays," *IEEE Communications Magazine*, vol. 53, no. 5, pp. 91–101, 2015.

[2] N. H. Mahmood *et al.*, "Throughput analysis of full duplex communication with asymmetric traffic in small cell systems," in *11th International Conference on Wireless and Mobile Communications*, 2015.

[3] P. Mogensen *et al.*, "Centimeter-wave concept for 5G ultra-dense small cells," in *IEEE 79th Vehicular Technology Conference*, 2014.

[4] K. Thilina *et al.*, "Medium access control design for full duplex wireless systems: challenges and approaches," *IEEE Communications Magazine*, vol. 53, no. 5, pp. 112–120, 2015.

[5] S. Goyal *et al.*, "Full duplex cellular systems: will doubling interference prevent doubling capacity?" *IEEE Communications Magazine*, vol. 53, no. 5, pp. 121–127, 2015.

[6] H. Malik, M. Ghoraishi, and R. Tafazolli, "Cross-layer approach for asymmetric traffic accommodation in full-duplex wireless network," in *European Conference on Networks and Communications*, 2015.

[7] W. John and S. Tafvelin, "Heuristics to classify internet backbone traffic based on connection patterns," in *International Conference on Information Networking*, 2008.

[8] J. Postel, "Transmission control protocol," Sep 1981, updated by RFCs 1122, 3168, 6093, 6528, [Online]. Available: http://www.ietf.org/rfc/rfc793.txt.

[9] V. P. M. Allman and W. Stevens, "Tcp congestion control to tcp's fast recovery algorithm," Apr 1999, obsoleted by RFC 5681, updated by RFC 3390, [Online]. Available: http://www.ietf.org/rfc/rfc2581.txt.

[10] M. Gatnau *et al.*, "Full duplex communication under traffic constraints for 5G small cells," in *IEEE 82nd Vehicular Technology Conference*, 2015.

[11] "Technical specification group radio access network; enhanced performance requirement for LTE user equipment (UE)," Jan 2013.

[12] "WINNER II channel models," Feb 2008.

[13] D. Catania *et al.*, "A distributed taxation based rank adaptation scheme for 5G small cells," in *IEEE 81st Vehicular Technology Conference*, 2015.

[14] “Further advancements for E-UTRA physical layer aspects,” Mar 2010.

[15] “Technical specification group radio access network; evolved universal terrestrial radio access (E-UTRA) radio link control (RLC) protocol specification,” Dec 2007.

[16] T. H. S. Floyd and A. Gurtov, “The newreno modification to tcp’s fast recovery algorithm,” Apr 2004, obsoleted by RFC 6582, [Online]. Available: http://www.ietf.org/rfc/rfc3782.txt.

[17] J. Postel, “User datagram protocol,” Aug 1980, [Online]. Available: https://tools.ietf.org/html/rfc768.

Paper I

On the Potential of Full Duplex Performance in 5G Ultra-Dense Small Cell Networks

Marta Gatnau Sarret, Marko Fleischer, Gilberto Berardinelli, Nurul H. Mahmood, Preben Mogensen, Helmut Heinz

The paper has been published in the
IEEE European Signal Processing Conference (EUSIPCO), 2016.

Abstract

Full duplex allows a device to transmit and receive simultaneously in the same frequency band, theoretically doubling the throughput compared to traditional half duplex systems. However, several limitations restrict the promised full duplex gain: non-ideal self-interference cancellation, increased inter-cell interference and traffic constraints. In this paper, we first study the self-interference cancellation capabilities by using a real demonstrator. Results show that achieving ∼110 dB of cancellation is already possible with the current available technology, thus providing the required level of isolation to build an operational full duplex node. Secondly, we investigate the inter-cell interference and traffic constraints impact on the full duplex performance in 5th generation systems. System level results show that both the traffic and the inter-cell interference can significantly reduce the potential gain of full duplex with respect to half duplex. However, for large traffic asymmetry, full duplex can boost the performance of the lightly loaded link.

1 Introduction

The number of devices requiring wireless connection is increasing day by day. Latest forecasts show that the global mobile data traffic will increase approximately eightfold between 2015 and 2020 [1]. An attractive solution to accommodate such traffic demand is full duplex (FD) technology. FD allows a device to simultaneously transmit and receive in the same frequency channel, thus, theoretically, doubling the throughput of traditional half duplex (HD) systems. Given its potential, FD is considered as a technology component of future 5th generation (5G) systems.

A FD device generates interference from the transmitter chain to the receiver chain located in the same device, named as self-interference (SI). Such type of interference must be suppressed in order to build an operational FD node. Recent studies [2, 3] show that current levels of SI cancellation (SIC) are in the order of 110 dB, which is already sufficient to build an operational FD node. However, the gain that FD can provide over HD may be affected by the residual SI among other limitations [3, 4]. Since FD doubles the amount of interfering streams, it leads to an increased inter-cell interference (ICI). Furthermore, exploiting FD is only possible when there is data traffic in both link directions, uplink (UL) and downlink (DL).

The authors in [2] evaluated SIC performance using a 20 MHz signal with a maximum transmit power of 24 dBm, showing that an isolation of ∼100 dB is sufficient to consider ideal SIC. The physical FD performance in WLAN systems is studied in [5], considering ideal SIC and bidirectional FD, i.e., the

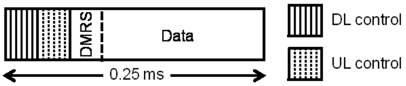


Fig. I.1: Envisioned 5G frame structure

case when user equipments (UEs) and access points (APs) are both FD capable. The authors have concluded that the gain that FD may provide over HD is below the theoretical 100% in the majority of cases. In [6], the FD performance in Long Term Evolution (LTE) time division duplex (TDD) with full buffer traffic and File Transfer Protocol (FTP) traffic is analyzed. The authors have evaluated the case where only the AP is FD capable, showing that FD always outperforms HD. Nevertheless, such results may be biased because the assumed isolation among cells may mitigate the ICI effect. In [7], the impact of symmetric and asymmetric traffic in a multi-cell scenario is presented. The authors show that the FD throughput gain reduces with the perceived ICI and the traffic asymmetry. Nevertheless, it is important to notice that the mentioned work does not consider a complete system with all layers active. Furthermore, features such as link adaptation and recovery mechanisms are not used.

The goal of the paper is twofold. First, an experimental study is carried out using a test bed to show current levels of achievable SIC. Second, an overview of the bidirectional FD performance in our envisioned 5G ultra-dense small cell network is presented. In [8], we have described such system, optimized for dense local area deployments. TDD has been chosen as the operational mode, with all the nodes in the network synchronized in time and frequency and equipped with multiple-input multiple-output (MIMO) antennas and interference suppression receivers. The results presented in this paper are extracted from a system level simulator, which includes all the system layers active and features such as link adaptation and recovery mechanisms. The gain that FD may provide over HD in different scenarios is studied, in order to evaluate the impact of the ICI and the traffic asymmetry.

The paper is structured as follows. Section II presents the envisioned 5G system and the limitations that FD brings in achieving the promised double throughput gain. Section III describes the test bed experiment and the SIC results. Section IV introduces the simulation environment and the system level results are shown in Section V. Finally, Section VI concludes the paper and states the future work.

2 Full Duplex in 5G Small Cells

The proposed 5G radio access technology (RAT) described in [8] was originally designed as a HD TDD system, targeting a massive and uncoordinated deployment of small cells, where all the nodes are synchronized in time and frequency. A novel frame structure of duration 0.25 ms is introduced. Such frame is defined as the Transmission Time Interval (TTI) and is divided in a control part followed by the data part, as shown in Figure I.1. A scheduling grant (SG) is transmitted within the DL control symbol. The SG contains transmission parameters such as the link direction, the Modulation and Coding Scheme (MCS) and the number of transmissions streams (often referred as the transmission *rank*). Within the UL control symbol, UEs transmit the scheduling request, with specific information regarding the channel and the buffer state and the Hybrid Automatic Repeat and Request (HARQ) feedback. The data part, which includes the demodulation reference signal (DMRS) used for channel estimation, carries UL or DL data in case of HD, and both DL and UL in case of FD. It is assumed that the transmission direction can change at each TTI, independently of previous decisions. Consequently, a TTI can be DL HD, UL HD or FD.

All nodes are equipped with 4×4 MIMO antenna configuration and Interference Rejection Combining (IRC) receivers [9]. Such receivers aim at suppressing incoming interference by using the degrees of freedom from the antenna domain.

In this paper we study the performance of bidirectional FD, which corresponds to the case where both the AP and the UEs are FD capable. To quantify the practical gains that FD may bring over traditional HD systems, three limitations must be considered:

- Self-interference cancellation. For a FD node to be effective, a high level of isolation between the transmitter antenna and the receiver antenna located in the same device is required.
- Inter-cell interference. FD doubles the amount of interfering streams compared to HD, meaning that FD will perceive stronger ICI than HD. The stronger the ICI is, the lower are the data rates and hence a larger number of TTIs are required to transmit the same amount of data.
- Simultaneous UL and DL data. The occurrence of UL and DL traffic at the same time dictates the probability of exploiting FD. Consequently, the FD gain will be impacted by large asymmetries between UL and DL.

In the next section, a description of the experiment carried out with our test bed will be presented, to show the current levels of achievable SIC.

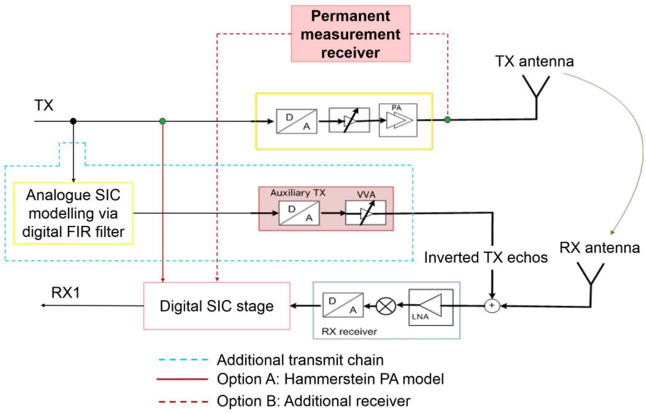


Fig. I.2: Auxiliary transmitter concept for receiver protection with two options for RF impairments modeling

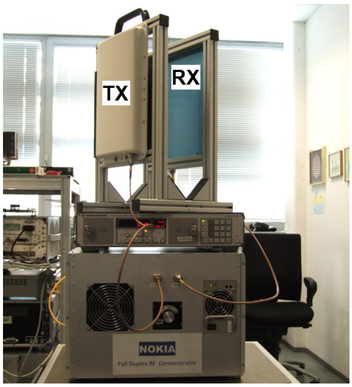


Fig. I.3: SIC hardware platform

3 Self-interference cancellation

The self-interference signal power in a FD scenario could easily exceed the receive signal power level by 100 dB or more [2]. Therefore, managing SIC is a fundamental requirement for the success of FD. The use of higher frequencies beyond today's LTE limits and intended broadband LTE channels of 80 to 100 MHz, together with massive MIMO, add additional obstacles in achieving a high degree of SIC. A test system was developed at Nokia Ulm, for demonstration purposes, to identify the potential limits of SIC. The concept of a pre-mixer approach with an additional transmit chain for analogue compensation and a final digital cancellation stage has been built and

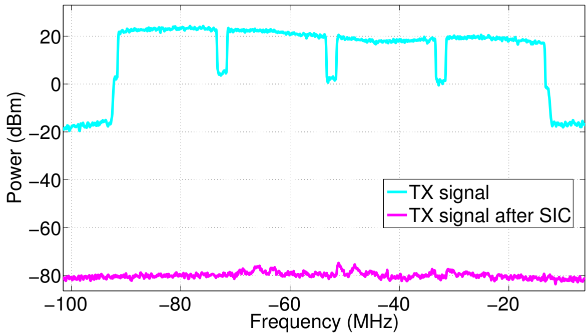


Fig. I.4: SIC for a 20 dBm 4×20 MHz LTE signal

studied. Such concept was proposed in [10] and is depicted in Figure I.2. The system can handle up to 100 MHz contiguous bandwidth and is typically operating in the 2.4 GHz band. The practical antenna isolation from the transmitter (TX) to the receiver (RX) is ∼50 dB. Such isolation is based on physical antenna separation, as shown in Figure I.3, and appropriate passive means. Furthermore, a common clocking domain, same mixer stage for up and down conversion and radio frequency (RF) delay compensation is essential to push the phase noise limits [11].

The hardware in use is upper bounded by ∼70 dB active cancellation gain for a 20 MHz LTE signal (LTE20) with respect to phase noise. The practical active cancellation limit is given by the power amplifier (PA) non-linearity and auxiliary transmitter dynamic. An active cancellation gain of 63 dB for LTE20 could be demonstrated, which is split between the analogue cancellation stage and the time domain digital cancellation. This gain demands the use of nonlinear intermodulation modeling via Hammerstein PA model by using the digital signal as input to the digital SIC stage [12] (option A in Figure I.2) or the PA signal as direct input with the need of an additional receiver (option B in Figure I.2). In the latter approach, the measurement receiver contains the transmitter RF impairments and is common in a typical commercial RF design for PA linearization purposes.

The approach of using an additional transmit chain is intended to protect the receiver against saturation. This approach has the advantage that it scales only with the number of transmit antennas, which is appropriate in a MIMO context. In addition, all transmitted antenna streams are input to the same analogue and digital SIC modeling block, thus avoiding extra complexity and providing simpler hardware integration.

Figure I.4 depicts a total cancellation of ∼100 dB for a 20 dBm 4×LTE20 signal, showing the SI level close to receiver noise floor limits and hence demonstrating the potential of this hardware concept. It justifies the ap-

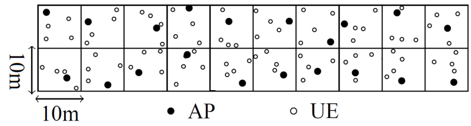


Fig. I.5: Simulated multi-cell scenario

Table I.1: Used parameters to run the simulations

Parameter	Value/State/Type
System parameters	BW = 200MHz; f_c = 3.5GHz
Frequency reuse	1 (whole band)
Propagation model	WINNER II A1 w/fast fading [13]
Antenna configuration	4x4
Receiver type	IRC
Transmission power	10 dBm (AP and UE)
Self-interference cancellation	Ideal
Link adaptation filter	Log average of 5 samples
Rank adaptation	Taxation-based [14]
HARQ max retransmissions	4
RLC mode	Acknowledged [15]
Transport protocol	UDP
Traffic type	Symmetric and asymmetric (6:1) finite buffer
Simulation time per drop	1-12 seconds
Number of drops	50

proach to treat SIC as ideal in FD networks, as long as the transmit output power does not exceed home or local area AP limits (up to 24 dBm).

4 Simulation Environment

The simulated results are extracted from our event-driven system level simulator. It includes the envisioned 5G physical (PHY) and medium access control (MAC) design described in Section II. Moreover, it includes a detailed implementation of the radio link control (RLC), the user datagram protocol (UDP) and the transport control protocol (TCP) layers, and a vertical radio resource management (RRM) layer. The role of the RRM is to collect information from the PHY, MAC and RLC layers to extract the most appropriate scheduling decision in each TTI, such as node identifier, link direction, MCS and rank. The Internet protocol (IP) layer is modeled as overhead and the application layer generates different types of traffic.

Features such as link adaptation, rank adaptation and the HARQ recovery mechanism are implemented on top of the PHY and MAC design described in Section II. The link adaptation algorithm keeps track of the latest five channel measurements to extract an accurate MCS. The rank adaptation algorithm [14] determines, based on the incoming interference, how many streams will be used for data transmission and consequently, the MIMO degrees of freedom in the receiver that will be used for interference suppression. For further details, please refer to [14].

The extraction of the transmission direction in HD is based on the amount of data which is currently in the buffers (UL and DL) and previous decisions. Thus, the optimal direction can be *DL* or *UL* if there is data in at least one of the buffers, or *MUTE* if there is no data in either of them. In case of FD, the transmission direction is only based on the buffer size. This means that the optimal transmission direction will be *DL+UL* if there is data in both buffers, *DL* (*UL*) if the *UL* (*DL*) buffer is empty, or *MUTE* if both buffers are empty. The reader can refer to [16] for further details on the algorithm to extract the optimal transmission direction. Finally, the user scheduler in time domain is round robin, and no user frequency multiplexing is considered in this work.

Two scenarios are studied, a single small cell and a multi-cell network. The latter corresponds to $10{\times}2$ grid of small cells (Figure I.5). Each small cell refers to a $10{\times}10$ m^2 room containing one AP and four UEs randomly deployed. The UEs are always affiliated to the AP in the same cell (closed subscriber group). Ideal SIC is considered, according to Section III and [2], given the short distances among nodes and the low transmit power, set to 10 dBm for both APs and UEs. The RLC mode is set to Acknowledged [15] and the transport protocol to UDP. The RLC ACK is transmitted within the control channel and its overhead is therefore not included in the throughput calculation. The remaining simulation parameters are detailed in Table I.1.

This work compares the performance of HD and FD under different conditions. Two types of traffic are considered, full buffer traffic and FTP traffic [17]. For the latter, two cases are studied: symmetric, where the offered load is the same in UL and DL (1DL:1UL), and asymmetric, where the offered load in DL is six times bigger than in UL (6DL:1UL). Furthermore, three load levels are simulated for each case: low, medium and high, which approximately correspond to 25%, 50% and 75% channel occupancy under ideal interference conditions, respectively. Results are presented numerically in a table format and/or in terms of cumulative distribution function (CDF). Tables show the gain that FD provides over HD, in percentage. The studied key performance indicators are the average session throughput (TP) and the packet delay. The former corresponds to the average of the individual session TPs per link (UL or DL) or per cell (UL+DL). The session TP is defined as the amount of time required to successfully transmit a session, and a session

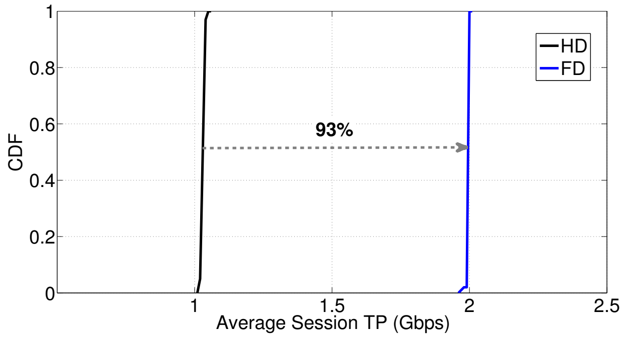


Fig. I.6: Single cell TP with 100% FD probability

is characterized by the packet size and the $t_{arrival}$ parameters, which are negatively exponential distributed [17]. The average packet size is 2 megabytes, and the average $t_{arrival}$ is set to obtain the loads described above. The packet delay is the time between the generation of a packet and its successful reception, including the buffering time.

5 Performance Evaluation

The results presented in this section show the impact of traffic and ICI. In the first three subsections, the impact of these limitations are studied in an isolated manner. Then, in the last subsection, the performance of HD and FD is evaluated considering both effects.

5.1 Single cell with 100% FD probability

In this first case we focus on the performance of FD and HD in a single cell scenario, to avoid ICI. The traffic generator is configured with fixed packet size and $t_{arrival}$ time, the same in both UL and DL, so the probability of having simultaneous traffic in both directions is 100%. Figure I.6 and Figure I.7 show the cell TP and average delay, respectively. We observe that, if neither ICI nor SI are present and the FD probability is 100%, the delay can be reduced by 50% with FD, while the theoretical FD TP gain can be almost achieved (93%), due to the fact that the HD baseline is optimized, as explained in Section IV. If the HD baseline is a fixed 1DL:1UL time slot allocation, then the achieved FD gain is 100%. Therefore, it is important to notice that the theoretical FD gain is possible to be achieved, but only under specific conditions.

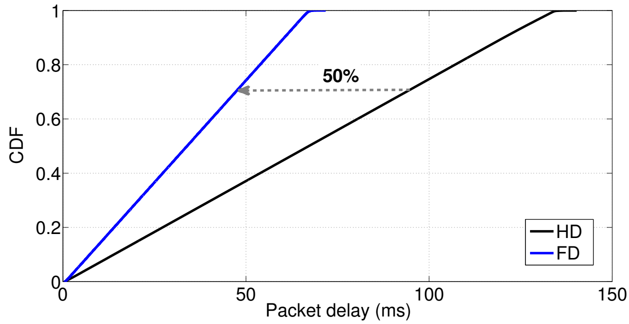


Fig. I.7: Single cell delay with 100% FD probability

Table I.2: Gain of FD over HD in a single cell scenario

	Symmetric		Asymmetric			
Load	**TP**	**Delay**	**DL TP**	**UL TP**	**DL delay**	**UL delay**
Low	6%	11%	2%	13%	6%	23%
Med	13%	19%	4%	28%	9%	35%
High	38%	41%	9%	62%	13%	57%

5.2 Single cell with less than 100% FD probability

This subsection moves a step forward from the previous one by considering the negatively exponential distributed traffic model described in Section IV. Note that from an interference perspective, ICI and SI are still not present. Table I.2 shows the TP and delay gains of FD over HD in percentage, for both symmetric and asymmetric traffic. Numerical results show that FD always outperforms HD under ideal interference conditions, and such gain increases with the offered load of the system. This is because the probability of having simultaneous UL and DL is higher when the offered load increases, and therefore FD can be exploited more often. In the asymmetric traffic case, we observe that the FD gain in UL is higher than in DL. With FD, the lightly loaded link gets, on average, six times more resources than with HD. On the contrary, the highly loaded link gets only, on average, one extra resource with FD.

From this analysis we conclude that the FD gain is negatively impacted by the traffic profile, being limited to 38% in case of symmetric traffic. This is because the probability of having simultaneous UL and DL data is below 100% and hence the chances to exploit FD decrease.

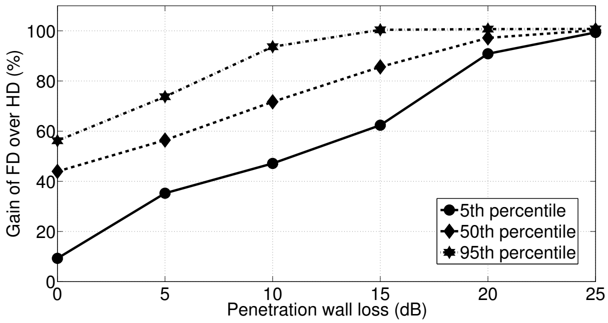


Fig. I.8: Multi-cell TP with full buffer traffic

5.3 Multi-cell with 100% FD probability

The next step after studying the impact of traffic is to analyze how ICI affects the FD performance, by considering the multi-cell scenario. To avoid the traffic impact, the full buffer model is considered and the transmission rank is fixed to one. The analysis is done by varying the penetration wall loss, which defines the isolation between the cells. Such wall loss ranges from 0 dB, which would correspond to an open space scenario, to 25 dB, which refers to an almost isolated cell. Figure I.8 shows the gain of FD over HD according to the wall penetration loss. The figure shows the gain in the 5^{th}, 50^{th} and 95^{th} percentiles, where the former refers to the outage performance, i.e., the performance of the users perceiving worse channel conditions. As expected, we observe that, as the isolation among cells is higher, the gain that FD can provide over HD increases. This is because FD doubles the amount of interfering streams, thus showing a higher ICI than HD. Finally, it is interesting to notice that, even in the worst case (open space scenario), the outage users can improve their performance with FD by 9%, while the users perceiving the best channel conditions can improve their performance by 56% by using FD.

5.4 Multi-cell without 100% FD probability

The last step is to consider both traffic constraints and ICI. Hence, the negatively exponential distributed traffic model and the multi-cell scenario with a penetration wall loss of 5 dB [13] are used. Numerical results are shown in Table I.3. As in the previous analysis, we observe that the FD gain increases with the offered load. However, in case of symmetric traffic, the maximum TP gain is 34%, thus below the theoretical FD gain, and the maximum delay

Table I.3: Gain of FD over HD in a multi-cell scenario

	Symmetric		Asymmetric			
Load	**TP**	**Delay**	**DL TP**	**UL TP**	**DL delay**	**UL delay**
Low	1%	8%	1%	3%	4%	12%
Med	14%	23%	4%	17%	6%	35%
High	34%	29%	17%	130%	28%	84%

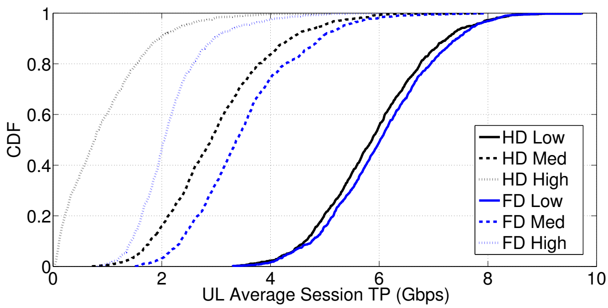


Fig. I.9: UL TP with asymmetric traffic

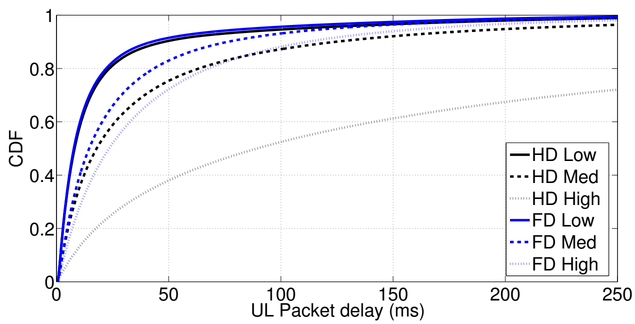


Fig. I.10: UL delay with asymmetric traffic

reduction is 29%. In case of asymmetric traffic, while the DL performance gain is limited to 23% and 16% in terms of TP and delay respectively (see Table I.3), the UL performance is boosted with FD, since it mitigates the impact of buffering, thus transmitting more data with lower delay. Figure I.9 and Figure I.10 show the CDF of the UL TP and delay, respectively, of the asymmetric traffic case. Results show that both starvation and buffering problems can be mitigated with FD, specially at high load, since in HD the UL data starves while with FD it is transmitted immediately. Therefore, FD is an attractive solution for applications where the performance of the lightly loaded link shall be improved.

6 Conclusions and future work

In this paper, we first investigated self-interference cancellation capabilities by using our own developed test bed. The carried experiment shows that up to $\sim$100 dB of isolation are currently achievable, thus validating the assumption of ideal self-interference cancellation in a dense small cell scenario. Secondly, we investigated the potential of full duplex in 5G ultra-dense small cell networks. We show that the theoretical 100% FD gain is achievable only under specific assumptions, namely ideal SIC, isolated cells and full buffer traffic model. Under realistic assumptions, the promised gains of FD are reduced by the traffic constraints and the inter-cell interference. System level results show that FD can always outperform HD in the considered scenarios, though the gains are limited. In case of symmetric traffic, such gains go up to 38% and 41% in throughput and delay, respectively. Furthermore, in case of asymmetric traffic, FD has the potential of boosting the lightly loaded link, specially in terms of delay, since it mitigates the buffering effect. Future work will focus on the usage of full duplex in providing fast discovery in device-to-device type of communication.

References

[1] "Cisco visual networking index: Global mobile data traffic forecast update, 2015-2020," Feb 2016.

[2] M. Heino *et al.*, "Recent advances in antenna design and interference cancellation algorithms for in-band full duplex relays," *IEEE Communications Magazine*, vol. 53, no. 5, pp. 91–101, 2015.

[3] K. Thilina *et al.*, "Medium access control design for full duplex wireless systems: challenges and approaches," *IEEE Communications Magazine*, vol. 53, no. 5, pp. 112–120, 2015.

[4] L. Wang *et al.*, "Exploiting full duplex for device-to-device communications in heterogeneous networks," *IEEE Communications Magazine*, vol. 53, no. 5, pp. 146–152, 2015.

[5] X. Xie and X. Zhang, "Does full-duplex double the capacity of wireless networks?" in *Proc. of IEEE INFOCOM*, Apr 2014, pp. 253–261.

[6] S. Goyal *et al.*, "Full duplex cellular systems: will doubling interference prevent doubling capacity?" *IEEE Communications Magazine*, vol. 53, no. 5, pp. 121–127, 2015.

[7] N. H. Mahmood *et al.*, "Throughput analysis of full duplex communication with asymmetric traffic in small cell systems," in *11th International Conference on Wireless and Mobile Communications*, Oct 2015.

[8] P. Mogensen *et al.*, "Centimeter-wave concept for 5G ultra-dense small cells," in *Proc. of IEEE 79th Vehicular Technology Conference*, May 2014.

[9] "Technical specification group radio access network; enhanced performance requirement for LTE user equipment (UE)," Jan 2013.

[10] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Transactions on Wireless Communications*, vol. 11, no. 12, pp. 4296–4307, 2012.

[11] A. Sahai, G. Patel, C. Dick, and A. Sabharwal, "On the impact of phase noise on active cancelation in wireless full-duplex," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 9, pp. 4494–4510, 2013.

[12] L. Anttila *et al.*, "Modeling and efficient cancellation of nonlinear self-interference in MIMO full-duplex transceivers," in *Globecom Workshops*, Dec 2014, pp. 777–783.

[13] "WINNER II channel models," Feb 2008.

[14] D. Catania *et al.*, "A distributed taxation based rank adaptation scheme for 5G small cells," in *Proc. of IEEE 81st Vehicular Technology Conference*, May 2015.

[15] "Technical specification group radio access network; evolved universal terrestrial radio access (E-UTRA) radio link control (RLC) protocol specification," Dec 2007.

[16] M. Gatnau *et al.*, "Can full duplex boost throughput and delay of 5G ultra-dense small cell networks?" in *Proc. of IEEE 83rd Vehicular Technology Conference (VTC Spring)*, May 2016.

[17] "Further advancements for E-UTRA physical layer aspects," Mar 2010.

ISSN (online): 2246-1248
ISBN (online): 978-87-7112-779-9

AALBORG **UNIVERSITY PRESS**