

 Open access • Journal Article • DOI:10.1109/MWC.2016.1500364WC

A Flexible 5G Wide Area Solution for TDD with Asymmetric Link Operation

— [Source link](#) 

Klaus I. Pedersen, Gilberto Berardinelli, Frank Frederiksen, Preben Mogensen





Institutions: Nokia Networks

Published on: 01 Apr 2017 - IEEE Wireless Communications (IEEE)

Topics: Telecommunications link, Duplex (telecommunications), Transmission time, Control channel and Hybrid automatic repeat request

Related papers:

- [Performance of strong interference cancellation in flexible UL/DL TDD systems using coordinated muting, scheduling and rate allocation](#)
- [Joint user scheduling and transceiver design for cross-link interference suppression in MU-MIMO dynamic TDD systems](#)
- [The Coverage-Latency-Capacity Dilemma for TDD Wide Area Operation and Related 5G Solutions](#)
- [Aligned Reverse Frame Structure for Interference Mitigation in Dynamic TDD Systems](#)
- [Enhanced HARQ Design for 5G Wide Area Technology](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/a-flexible-5g-wide-area-solution-for-tdd-with-asymmetric-25b1cq9y9q>



Aalborg Universitet

AALBORG UNIVERSITY
DENMARK

A Flexible 5G Wide Area Solution for TDD with Asymmetric Link Operation

Pedersen, Klaus I.; Berardinelli, Gilberto; Frederiksen, Frank; Mogensen, Preben Elgaard

Published in:
I E E E Wireless Communications Magazine

DOI (link to publication from Publisher):
[10.1109/MWC.2016.1500364WC](https://doi.org/10.1109/MWC.2016.1500364WC)

Publication date:
2017

Document Version
Accepted author manuscript, peer reviewed version

[Link to publication from Aalborg University](#)

Citation for published version (APA):
Pedersen, K. I., Berardinelli, G., Frederiksen, F., & Mogensen, P. E. (2017). A Flexible 5G Wide Area Solution for TDD with Asymmetric Link Operation. *I E E E Wireless Communications Magazine*, 24(2), 122-128.
<https://doi.org/10.1109/MWC.2016.1500364WC>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

A FLEXIBLE 5G WIDE AREA SOLUTION FOR TDD WITH ASYMMETRIC LINK OPERATION

KLAUS I. PEDERSEN, GILBERTO BERARDINELLI, FRANK FREDERIKSEN, PREBEN MOGENSEN

ABSTRACT

A flexible multi-service 5G wide area (WA) solution for time division duplex (TDD) operation is outlined in this article. In particular, the associated frame design is in focus. Given the fundamental tradeoffs between capacity, coverage, latency, and reliability, a flexible solution that allows optimization on a per-link basis is proposed. The solution encompasses the possibility to schedule users with different transmission time intervals to best match their service requirements and radio conditions. Due to the large downlink/uplink transmission power imbalance for each link, asymmetric link operation is proposed, where users operate with different minimum transmission times for the two link directions. This is achieved by using a highly flexible asynchronous hybrid automatic repeat request (HARQ) scheme, as well as a novel solution with in-resource control channel signaling for the scheduling grants. Performance results for the proposed 5G WA TDD solution show clear benefits over current LTE, for example, reduced latency and more scalable control overhead to better support users with different QoS requirements.

INTRODUCTION

In this article we focus on the design of a 5G multi-service air interface with wide area (WA) coverage, using time division duplex (TDD). In this context, WA coverage is achieved with high-power macrocells, deployed at relative low carrier frequencies due to the more favorable radio propagation properties at these bands. As an example, roughly half of the available bands below 6 GHz that could be made available for future 5G deployments is unpaired, that is, for TDD. 5G is set to support a wide range of highly diverse services. This includes enhanced mobile broadband (MBB) with peak data rates of 10–20 Gb/s, offering spatial uniform availability of end-user data rates of 100 Mb/s. Moreover, efficient expedition of MBB services should also support flexible scheduling of smaller payloads, and thus requires a scheduling framework that supports high dynamic range of scheduled payload sizes. Furthermore, machine type communication (MTC) with massive machine communication (MMC) and mission critical communication (MCC) are other use cases. MCC is particularly challenging as it requires both low latency and ultra reliable

communication. Among others, these service requirements translate to the need for short transmission time intervals for MCC, large bandwidth for MBB, and low bandwidth operation for MMC devices with low cost and energy consumption requirements. For more information on 5G requirements, we refer to [1, 2].

It is well known that there are fundamental tradeoffs between capacity, latency, reliability, and coverage [3]. This basically means that optimizing for one metric results in a loss for the other metrics. As an example, this can be illustrated with the effective capacity, which expresses the maximum source data arrival rate that a certain channel process can support, while fulfilling a latency constraint [4]. With no latency constraints, the effective capacity equals the Shannon capacity, while it decreases asymptotically as stricter latency constraints are enforced. From a system design point of view, this tells us that we should not optimize the air interface to, for example, always fulfill strict latency requirements, as this will incur a loss in capacity (spectral efficiency), and vice versa. Instead, the focus of this study is on a flexible system design that allows optimizing each link in coherence with its service requirements. Despite the relative short time that 5G has been researched, the open literature already includes an impressive number of 5G related studies. Examples of 5G studies include: the METIS project [5]; use of centralized network architectures has been suggested in [6]; small cell optimized TDD design for MBB has been proposed in [7]; and 5G cell densification in [8].

In our effort to design a flexible multi-service 5G WA TDD concept, we start by first identifying the fundamental objectives and constraints in the radio design. The tradeoffs between capacity, coverage, and latency are studied, and solutions aiming at making the best compromises between these metrics are suggested. In particular, a configurable 5G TDD frame structure for efficient multiplexing of users is outlined, which comprises cell-specific configurations, as well as flexibility for user scheduling within the cells. Throughout the study, we will use the Long Term Evolution (LTE) Rel-12 solution with enhanced interference mitigation and traffic adaptation (eIMTA) as the reference for today's 4G cellular [9]. In short, eIMTA is the LTE TDD solution with adaptive adjustment of downlink/uplink transmission patterns. Although being a powerful concept, LTE eIMTA

The authors are with Nokia-Bell Labs.

Digital Object Identifier: 10.1109/MWC.2016.1500364WC

does not fulfill all of the requirements of 5G, as it was designed mainly for MBB use cases with more modest data rates. As an example, the lowest possible radio hybrid automatic repeat request (HARQ) round trip time (RTT) for LTE eMTC is on the order of 9.8 ms–12.4 ms, depending on the downlink/uplink configuration [10]. Our performance analysis shows that the proposed 5G WA TDD concept can achieve significantly lower latency, and higher flexibility for scheduling of users with extreme diverse service requirements. The article is closed with concluding remarks and outlook.

FUNDAMENTAL CONSTRAINTS

UPLINK COVERAGE AND BANDWIDTH CONSTRAINTS

Providing good coverage is obviously a priority for a 5G WA design. Due to the lower user equipment (UE) transmit power (as compared to the base station, also known as eNB), the coverage is typically determined by the uplink. For a TDD system, the coverage is further challenged as UEs are allowed to transmit only at certain time-intervals. Coverage challenged UEs will need to transmit for a certain minimum time-duration to allow the receiving eNB to collect a sufficient amount of energy to successfully decode the transmission. In LTE link budget studies, it was, for example, found that the physical uplink control channel (PUCCH) transmitted during a 1 ms time-interval has a coverage range of 1.4 km and 1 km for suburban and dense urban non line of sight conditions (NLOS), respectively [9]. These results are obtained for four receive antennas at the eNB, assuming a carrier frequency of 2 GHz. Reducing the transmission time from 1 ms to 0.2 ms as assumed in recent 5G (small cell) TDD concept studies [7], is estimated to reduce the coverage range to ~300 meters. Hence, it is of paramount importance that a new 5G WA TDD concept is designed with the flexibility to allow configuration of (continuous) uplink transmit opportunities to meet the desired coverage target. For the downlink, the coverage is obviously better due to the higher eNB transmit power. It is therefore desirable to have support for asymmetric link operation, where the transmission times of data and control can be set differently for the two link directions on a per user basis, depending on its coverage.

For MMC we also consider device cost and energy constraints in our design. More specifically, we aim at designing a system that supports concurrent operation of low bandwidth MMC devices on wider bandwidth 5G carriers. We consider support of low cost and energy efficient MMC devices with a transceiver bandwidth of no more than a couple of hundred kHz to a few MHz for the downlink, and only a single antenna. For the uplink, even lower transmission bandwidth is considered. Due to the lack of frequency and space (i.e. antenna) diversity, additional time diversity is desirable for MMC with relaxed latency requirements.

MULTI-CELL COORDINATED TDD OPERATION

It is well known from numerous studies that dynamic TDD operation is feasible and attractive for small cell scenarios, allowing each cell

to autonomously decide the transmission direction depending on the needs within the cell (also known as uncoordinated TDD). Dynamic TDD operation is feasible for small cell scenarios due to the balanced output power level from eNBs and UEs, making it possible to manage cross-link interference with advanced receiver interference suppression techniques (e.g., [7, 8]). However, for a WA setting, the eNB transmit power is typically on the order of ~49 dBm, having antenna gain of at least ~14 dBi, and thus resulting in an equivalent isotropic radio power (EIRP) of ~63 dBm, while the EIRP for the UE is typically only 23 dBm, assuming a maximum transmit power of 23 dBm and 0 dBi antenna gain. The large output power imbalance (~40 dB in EIRP and ~26 dB without antenna gains) between UEs and eNBs for a WA scenario sets additional restrictions on the TDD operation, since closely coupled cells will have to coordinate the use of uplink/downlink transmission patterns to avoid severe cross-link interference problems. This essentially calls for some degree of multi-cell coordinated TDD operation for WA scenarios. Thus, each cell does not have the full freedom to determine if the cell resources are used for uplink or downlink, as some alignment and coordination with other WA cells in the vicinity is required. The use of massive MIMO and advanced interference suppression techniques can, however, help relax the requirements for tight inter-cell coordination.

SYSTEM CONSTRAINTS

Finally, there are system related constraints that need to be considered. Among these, there need to be regular downlink transmission resources available for sending the broadcast channel with the most essential system information, as well as physical layer discovery signals. There also needs to be resources available for uplink random access (RA) [12]. Again, for large WA cells, the required time-duration of the resources for RA should approximately equal 1 ms for UEs at the cell-edge to perform successful access. Moreover, even for a cell with heavy downlink user plane traffic, it is desirable to have frequent opportunities for uplink transmission of various physical layer related control information. The latter includes positive and negative acknowledgments for HARQ, and various channel quality information feedback that the eNB needs for link adaptation and scheduling purposes, as well as MIMO adaptation. Thus, having long time periods without uplink opportunities is undesirable.

SUMMARY OF THE MAIN CONSTRAINTS

The identified main constraints for a flexible multi-service 5G WA TDD design are summarized in Table 1. In line with the study in [11], our hypothesis is that efficient scheduling of the considered services requires the support for different transmission time intervals (TTIs). As a few examples, users with tight latency constraints (e.g. MCC) require short TTIs, while MMC users scheduled on a narrow bandwidth are most efficiently served with longer TTIs. Moreover, users with MBB traffic could also benefit from variable TTI sizes. During the initial MBB data transmission session, the end-user experienced performance is primarily determined by the RTT due to the

Providing good coverage is obviously a priority for a 5G WA design.

Due to the lower user equipment (UE) transmit power (as compared to the base station, also known as eNB), the coverage is typically determined by the uplink. For a TDD system, the coverage is further challenged as UEs are only allowed to transmit at certain time-intervals only.

There are system related constraints that need to be considered. Among these, there need to be regular downlink transmission resources available for sending the broadcast channel with the most essential system information, as well as physical layer discovery signals.

Category	Related constraints	
Service related constraints	Mobile broadband (MBB)	Flexibility to schedule MBB users with variable bandwidth and TTI sizes is of importance. Large dynamic range of the user plane data payload sizes to be scheduled, ranging from several tens of bytes (e.g., for application-layer control messages) to gigabytes for large data file transmissions.
	Massive machine communication (MMC)	Support for low cost and energy efficient MTC devices that only operate on a narrow bandwidth, but also support for wideband MTC devices. Typically moderate size payloads.
	Mission critical communication (MCC)	Support for low latency is essential, calling for short TTI sizes when needed. Ultra high reliability for some MCC use cases. Typically moderate size payloads.
Uplink coverage constraints	<ul style="list-style-type: none"> • Users need a certain minimum continuous uplink transmission time to have reliable uplink reception at the eNB. • Examples: with 0.2 ms transmission time, coverage is on the order of ~300 meters, while 1 ms transmission time offers 1.4 km coverage. 	
Inter-cell TDD coordination	<ul style="list-style-type: none"> • High-power macro cells in the same local area shall use coordinated TDD configurations of downlink/uplink transmissions to avoid cross-link interference problems. • The performance is at risk if cells in the same geographical area use different transmission directions at the same time. • Use of massive MIMO and advanced interference suppression techniques can, however, help relax the requirements for tight inter-cell coordination. 	
System constraints	<ul style="list-style-type: none"> • Desirable to have frequent downlink and uplink transmission opportunities to have fast control loops, for example, for HARQ ACK/NACK's, channel state information, radio resource control. • Known (semi-statically configured) downlink transmission opportunities for common cell control information, for example, broadcast channel with system information and cell discovery signals for mobility purposes. • Known (semi-statically configured) uplink transmission occurrences for random access (RA), and potentially also uplink contention based data transmission access for MMC [12]. • Flexibility for configuration of different downlink and uplink transmission switching patterns. 	
Link asymmetry	<ul style="list-style-type: none"> • Support link configurations where the downlink and uplink transmission time intervals can be either equal (symmetric) or take different lengths (asymmetric). • Asymmetric configurations with short transmission times for downlink, while using longer transmission times in the uplink is relevant for uplink coverage reasons. 	

Table 1. Summary of identified main constraints for a 5G WA TDD design.

TCP slow start procedure (i.e. TCP flow control). Therefore, it would be advantageous to first perform scheduling of the MBB TCP data with short TTIs, followed by longer TTI sizes when reaching steady state operation. Given the requirements in Table 1, corresponding solutions are outlined in the following section.

TDD FRAME STRUCTURE DESIGN

SUBFRAME CONSTRUCTS

The proposed solution is based on a series of bi-directional TDD blocks that consists of an integer number of subframes. Given the WA coverage requirements discussed in the previous section, we consider a minimum block size of 1 ms, but also options of using longer blocks of, for example, 4 ms duration. Each block is having downlink transmission in the start, followed by a short guard period, and uplink transmission. The time resolution for the switching point between downlink and uplink in a block is on subframe resolution, assuming a 0.2 ms subframe duration as our default setting. Figure 1 shows a possible bi-directional TDD block configuration, where the first three subframes are configured for downlink, and the last two subframes are for uplink transmission (minus the fraction that is punctured for guard). Within each block, users are flexibly time-frequency multiplexed on subframe and physical resource block (PRB) resolution in coherence with their service requirements and radio conditions. The

per-user resource allocation is facilitated by adopting the principle of in-resource control channel (CCH) signaling for physical layer scheduling grants [11], as illustrated in Fig. 1. Among other things, this allows scheduling users with variable effective length transmission time intervals (TTI). Referring to Fig. 1, User #3 is scheduled in the downlink with an effective TTI size corresponding to three subframes. User #1 is scheduled with an effective TTI size of one subframe in the start of the downlink part, while User #2 is scheduled on the last subframe of the downlink part. Notice that in the spirit of the in-resource CCH method, the scheduling grant for users 1–3 appears in the start of their downlink transmission. Those scheduling grants contain information on the physical resources for the data transmission, as well as the corresponding modulation and coding scheme, HARQ, and MIMO transmission information (e.g. if the user is scheduled with multiple streams). The in-resource CCH is transmitted with quadrature phase shift keying (QPSK), allowing a modest set of different effective coding rates as also assumed for the LTE physical dedicated control channel (PDCCH) [10]. In the interest of UE complexity to monitor for in-resource CCH transmissions, the network can configure UEs to only search for such scheduling grants with a certain time-frequency resolution (see more details in [11]).

Notice from Fig. 1 that User #1 is scheduled in both the downlink and uplink in the same bi-directional TDD block. The in-resource CCH

that appears in the start of the block for User #1 contains a joint scheduling grant for the downlink and uplink allocation to that user. The uplink transmission for User #1 contains both data (a.k.a. payload) as well as the corresponding HARQ acknowledgment (ACK), or negative ACK (NACK). The ACK/NACK for User #1 is multiplexed with uplink data. Similarly, the ACK/NACK for the downlink data transmission to User #2 is also sent in the uplink part of the block. In this particular example, it is transmitted during the last subframe of the block. Following the same spirit as used for LTE, several ACK/NACK's can be transmitted on the same set of uplink resources by using different code signatures. From the perspective of users #1 and #2, the bi-directional TDD block is said to be self-contained, as the downlink data transmission and the corresponding uplink ACK/NACK's appear within the same block.

However, depending on the block configuration (duration and switching point between downlink and uplink) and the users' coverage requirement, it is not always possible to transmit the users' ACK/NACK in the same block as the downlink data transmission. First of all, the UE needs time to process the received downlink data before it can send the corresponding ACK/NACK in the uplink. This is referred to as the UE processing time. For 5G, some terminal modem vendors have reported that the UE processing time can be as short as the guard period (i.e. less than 0.1 ms), while more conservative assumptions for UE processing time is on the order of 0.3 ms (e.g., [7, 11]). Second, as discussed earlier, a certain minimum uplink transmission time is required for sending the ACK/NACK, depending on the users' uplink link budget conditions. Hence, users in the close vicinity of their serving cell will need only one subframe (or less) for the uplink ACK/NACK transmission to have it reliably decoded at the eNB, while users at the macro cell-edge need longer transmissions. Referring to Fig. 1, only the downlink data transmission for User #3 appears in the block, while the corresponding uplink ACK/NACK is postponed until a subsequent block where there is sufficient uplink transmission time available (assuming this user is coverage limited).

TDD RADIO FRAME CONFIGURATION

An integer number of bi-directional TDD blocks (and hence also an integer number of subframes) forms the TDD radio frame, as illustrated in Fig. 2. The bi-directional TDD blocks within each TDD radio frame can have different configurations, for example, different switching points between downlink and uplink transmissions. Within each cell, the TDD radio frame structure is periodically repeated. As discussed earlier, downlink transmission resources for system information broadcast appears at least once per TDD radio frame at a predefined location (time-frequency) that is known by the users in the cell. Similarly, the locations for the cell to transmit cell discovery signals for mobility purposes are fixed within each TDD radio frame. The same applies for uplink resources for random access [12]. The exact location and design of common downlink signals for system information broadcast, cell discovery, as well as uplink random access, are outside the scope of this article.

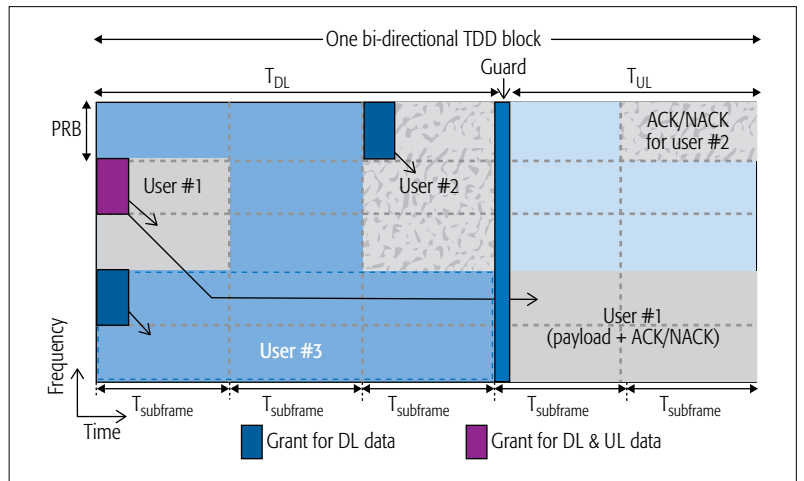


FIGURE 1. Basic bi-directional TDD block construct illustrating time-frequency multiplexing of users.

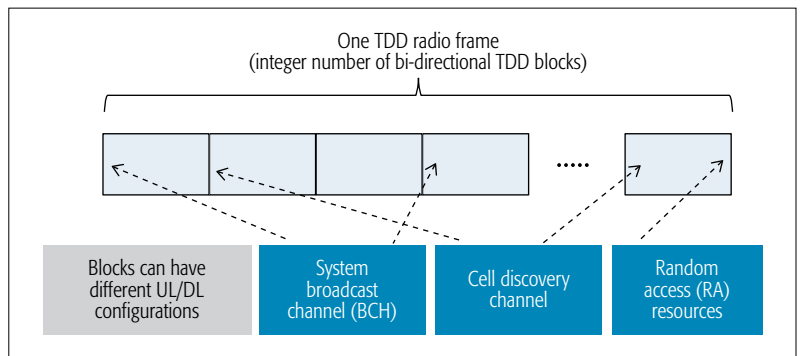


FIGURE 2. TDD radio frame constructed of an integer number of bi-directional TDD blocks.

The TDD radio frame is assumed to be coordinated between neighboring cells, such that the same synchronized downlink and uplink switching pattern is used by the cells. Depending on how fast cells in the same geographic area are able to coordinate, the TDD radio frame configuration can be semi-dynamically adjusted to best match the time-variant offered traffic for the two link directions. Hence, for a traditional distributed macrocellular network structure where the base stations are inter-connected via a backhaul, configuration of the TDD radio frame structure is only adjusted on a slow time-scale, using self organizing network (SON) type of solutions [13]. For centralized network architectures with virtually zero-latency fronthaul connections [6], adaptation of the TDD radio frame structure configuration can be faster, exploiting more efficient adaptation in coherence with the time-variant traffic conditions. Notice furthermore that the TDD radio frame structure can consist of bi-directional TDD blocks of different lengths, for example, a mixture of 1 ms and 4 ms blocks. The longer 4 ms block offers lower relative overhead from the guard period, and options for using longer TTI durations (e.g. improved time-diversity). The 4 ms block with relative long uplink transmission time is also attractive for allocation of the random access resources, and for serving uplink coverage challenged users that require longer transmission times for maintaining their uplink connectivity.

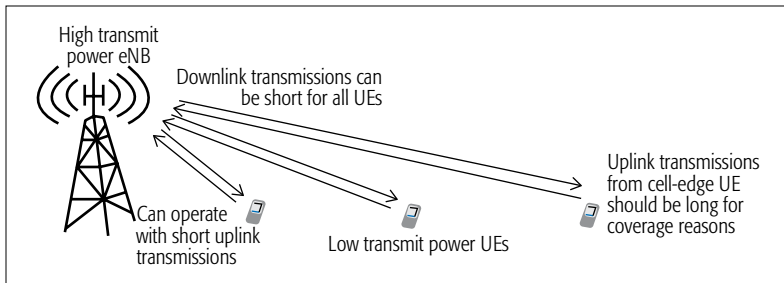


FIGURE 3. The basic principle of asymmetric link operation.

ASYMMETRICAL OPERATION

As illustrated in Fig. 3, the proposed TDD radio frame structure allows asymmetric link operation, where users can be scheduled in the downlink with short TTI sizes (if desirable), while the minimum TTI size for the uplink is adjusted in coherence with the users' coverage conditions. Hence, the eNB packet scheduler is responsible for scheduling cell-edge coverage challenged users with longer TTIs in the uplink. On a similar note, cell-edge users will also be configured to send uplink control information (e.g. ACK/NACK) using longer transmission times. The transmission time length for uplink control is assumed to be configured with higher-layer signaling per UE. Asynchronous HARQ is assumed for both the downlink and uplink in order to have full support for asymmetric link operation. The use of asynchronous HARQ offers more flexibility for the timing of ACK/NACK and retransmissions as compared to synchronous HARQ. Finally, in line with the findings in [11], it is worth mentioning that the use of the in-resource CCH solution also unleashes more flexible use of time-frequency domain inter-cell interference coordination, as well as it allows the same beamforming for physical-layer control and data downlink transmissions.

PERFORMANCE ANALYSIS

In the following, we present example performance results to further illustrate the merits of the proposed solution. We start by addressing latency performance, followed by presenting examples of how the overhead varies depending on the use of different TTI sizes.

LATENCY RESULTS

Table 2 summarizes the one-way downlink user plane latency for different TDD radio frame configurations. UEs at different coverage ranges, enumerated as short, medium, large, and extreme coverage ranges, are considered. The definition of latency follows that used in 3GPP for the latest eIMTA studies [9]. The user plane latency is the sum of five different components, namely the eNB and UE processing times, the frame alignment time, the TTI duration, and the average HARQ RTT. The frame alignment time is the average waiting time from when data arrives at the eNB until it can start to transmit the corresponding payload in the downlink. Hence, the value of this depends on the TDD radio frame configuration. The downlink HARQ RRT is defined as time from the eNB starting to transmit a payload, until it can start to send the corresponding retransmission

on the same stop-and-wait (SAW) channel. It is assumed that the average block error rate (BLER) for a first transmission equals 10 percent, so in calculating the average HARQ RTT, weighting by a factor 0.1 is applied. The HARQ latency depends both on the TDD radio frame configuration, as well as on the UEs' coverage conditions, since users at larger coverage ranges can only transmit their ACK/NACK in a bi-directional TDD block with a sufficiently long time for uplink resources. Results are presented for two different TDD radio frame configurations. The TDD radio frame configuration with 4x1 ms blocks is composed of a downlink heavy (D), balanced (B), downlink heavy (D), and uplink heavy blocks (U). As the name indicates, the D block has a majority of the resources for downlink, the B block has similar downlink/uplink resources, while the U block has the most resources for uplink transmission. Users at the short range can transmit their ACK/NACK in a single subframe (i.e. requires only few uplink resources), while medium range users can only transmit their ACK/NACK in B and U blocks, and large range users can only transmit their ACK/NACK in U blocks. Given these assumptions, the values for the frame alignment and the average HARQ retransmission RTT are obtained from simple Monte-Carlo simulations. The results in Table 2 clearly show the benefits of having a scheme that allows asymmetric link operation. The users at short range can benefit from operation with low latency, while the users at medium and large coverage ranges tend to experience longer latencies as they are subject to more constraints on when they have opportunities for ACK/NACK transmissions in the uplink. But still, the one-way user plane latencies for the considered TDD radio frame of 4x1ms bi-directional TDD blocks is only on the order of 1.20–1.26 ms, which is significantly lower than the best case 5–6 ms radio latency for LTE eIMTA.

Table 2 also includes results for a TDD radio frame configuration consisting of 4x1ms D blocks, followed by a 1x4 ms D block that contains approximately 1.2 ms (i.e. corresponding to six 0.2 ms subframes) of time for uplink transmission. The longer time for uplink transmission is required for extreme coverage UEs to be able to transmit their ACK/NACK. The results for this TDD radio frame configuration therefore further show the difference in experienced latency for the short range and extreme range user. It should furthermore be noticed that the results in Table 2 assume a conservative setting (0.3ms) for the eNB and UE processing times. If the eNB and UE processing times are reduced to ~0.1 ms, the total one-way downlink user plane latency is reduced to less than 1 ms (e.g., as discussed for tactile Internet use cases).

SCHEDULING OVERHEAD VS LATENCY TRADEOFFS

The low latency results are primarily achieved by using the 1 ms bi-directional TDD blocks and the even shorter scheduling allocations of 0.2 ms TTIs (i.e. corresponding to one subframe), using conservative link adaptation setting to ensure that first transmissions have a relatively high success rate (i.e. low BLER). Scheduling with short TTIs, does, however, come at a cost of increased CCH overhead. Fig. 4 shows a bar chart that summa-

izes the experienced relative CCH scheduling overhead for different scheduling options, and different user experienced downlink SINR values. The results in Fig. 4 are obtained by using the same assumptions as in [11], where the required resources for the in-resource CCH depends on the user experienced SINR. It is observed that a coverage challenged UE with a post-detection SINR of -6 dB will experience a CCH overhead of ~ 75 percent if scheduled with a TTI size of 0.2 ms on a 2.5 MHz bandwidth. Scheduling such a user with a TTI size of, for example, 2 ms (as possible within the longer 4 ms TDD subframe), reduces the CCH overhead by a factor of 10 . The reduction of the CCH overhead translates to higher spectral efficiency. Thus, users with delay tolerant services (e.g. MBB and MMC) are most efficiently served with longer TTIs to achieve higher spectral efficiency. The users with a more favorable downlink experienced SINR of 2 dB generally experience lower CCH overhead, as less resources are required for the in-resource CCH (i.e. less strong coding). However, also for such users, there is a clear reduction of the CCH overhead from using longer TTIs. In fact, by using a TTI size of 2 ms for such users, the relative CCH overhead is reduced to less than 1 percent. In comparison, the physical layer CCH overhead for LTE is on the order of 7 to 21 percent, depending on whether 1 to 3 symbols are configured for control per subframe [10]. The proposed 5G solution is therefore superior to LTE, offering a more flexible and scalable solution to efficiently adjust the tradeoffs between latency and CCH overhead.

In summary, the tradeoffs between latency, capacity, and coverage are visible from the results in Table 2 and Fig. 4. A cell-edge user with -6 dB SINR loses 66 percent in throughput from reducing the TTI size from 2 ms to 0.2 ms, but only gains a factor of four in reduced user plane latency.¹ A user nearer the serving cell (with $+2$ dB SINR), loses only 8 percent in throughput from reducing the TTI size from 2 ms to 0.2 ms.

CONCLUSION

In this article we have proposed a highly flexible TDD solution for efficient multiplexing of users with highly diverse service requirements in a wide area (macro) type of environment. The proposed TDD radio frame structure is composed of a series of subframes that form bi-directional TDD blocks with self-decodable physical layer control and data channel elements. The TDD radio frame configuration is coordinated between neighboring macro sites to avoid undesirable cross-link interference. The solution allows operating each link in coherence with its service constraints. This includes scheduling of users with different TTI sizes to efficiently control latency-capacity tradeoffs, including asymmetric operation with different uplink and downlink minimum transmission times to meet coverage constraints. The presented performance results show that short latency can be achieved for users with good coverage, while the coverage challenged users tend to experience slightly higher latency. The proposed scheme is superior to LTE eIMTA, both in terms of the offered flexibility for multiplexing users with diverse services requirements, and in terms of achieved user plane latency. The most

TDD radio frame configuration	DBDU (4x1 ms)			DDDD (4X1ms) + D (1X4ms)	
	Short	Medium	Large	Short	Extreme
UE coverage position	Short	Medium	Large	Short	Extreme
eNB processing time	0.30	0.30	0.30	0.30	0.30
Frame alignment	0.27	0.27	0.27	0.17	0.17
TTI duration	0.20	0.20	0.20	0.20	0.20
UE processing	0.30	0.30	0.30	0.30	0.30
HARQ retrans.	0.13 (0.1x1.3)	0.16 (0.1x1.6)	0.19 (0.1x1.9)	0.19 (0.1x1.9)	0.50 (0.1x5)
Total one way delay	1.20	1.23	1.26	1.16	1.47

Table 2. Average downlink one-way user plane latency budget. All latency values are in ms.

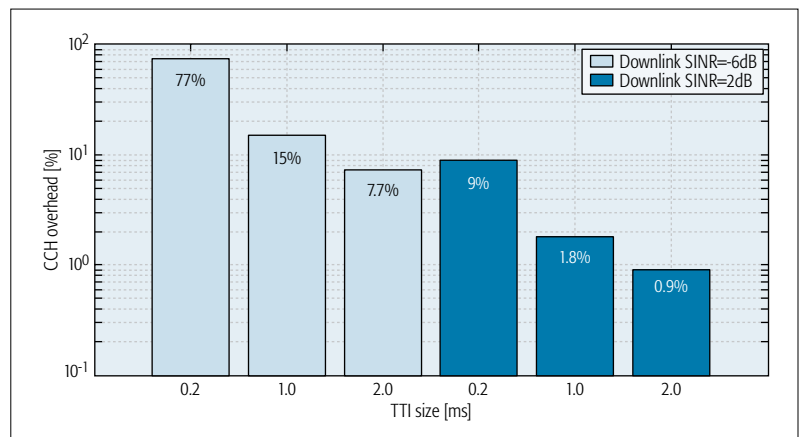


FIGURE 4. Relative CCH scheduling overhead for different scheduling options and user experienced quality, assuming a user scheduling bandwidth of 2.5 MHz.

important lesson learned in this study is, therefore, the importance and benefit of designing 5G with a highly flexible frame structure, offering efficient tradeoffs between different optimization targets to support users with highly diverse QoS requirements.

ACKNOWLEDGMENT

Part of this work has been performed in 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] ITU, "IMT Vision – Framework and Overall Objectives of the Future Development of IMT for 2020 and Beyond," Document, Radiocommunication Study Groups, Feb. 2015.
- [2] E. Dahlman *et al.*, "5G Wireless Access: Requirements and Realization," *IEEE Commun. Mag. – Communications Standards Supplement*, vol. 52, no. 12, Dec. 2014, pp. 42–47.

¹ The reason for not gaining a factor of ten when decreasing the downlink TTI size from 2 ms to 0.2 ms is because a constant eNB and UE processing time of 0.3 ms is assumed, as well as the need for a relative long uplink ACK/NACK transmission time for a cell-edge UE.

- [3] B. Soret *et al.*, "Fundamental Tradeoffs among Reliability, Latency and Throughput in Cellular Networks," *IEEE Proc. GLOBECOM*, Dec. 2014, pp. 1391–96.
- [4] D. Wu and R. Negi, "Effective Capacity: A Wireless Link Model for Support of Quality of Service," *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, July 2003, pp. 630–43.
- [5] A. Osseiran *et al.*, "Scenarios for 5G Mobile and Wireless Communications: The Vision of the METIS Project," *IEEE Commun. Mag.*, vol. 52, no. 5, May 2014, pp. 26–35.
- [6] C. Lin *et al.*, "Towards Green and Soft: A 5G Perspective," *IEEE Commun. Mag.*, vol. 52, no. 2, Feb. 2014, pp. 66–73.
- [7] P. Mogensen *et al.*, "5G Small Cells Optimized Radio Design," *Proc. Globecom, Int'l. Wksp. Emerging Techniques for LTE-Advanced and Beyond 4G*, Dec. 2013, pp. 111–16.
- [8] N. Bhushan *et al.*, "Network Densification: The Dominant Theme for Wireless Evolution into 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, Feb. 2014, pp. 82–89.
- [9] 3GPP Technical Report 36.912; Annex C.2 with Link Budget; Details in Document RP-090746, Mar. 2011.
- [10] H. Holma and A. Toskala, Eds., *LTE-Advanced, 3GPP Solutions for IMT-Advanced*, Wiley, 2012.
- [11] K. I. Pedersen *et al.*, "A Flexible Frame Structure for 5G Wide Area," *IEEE Proc. VTC*, Sept. 2015.
- [12] E. Paolini *et al.*, "Coded Random Access: Applying Codes on Graphs to Design Random Access Protocols," *IEEE Commun. Mag.*, vol. 53, no. 6, June 2015, pp. 144–50.
- [13] S. Hämmäläinen, H. Sanneck, and C. Sartori, Eds., *LTE Self-Organising Networks (SON)*, 1st ed., Wiley, 2012.

BIOGRAPHIES

KLAUS I. PEDERSEN (klaus.pedersen@nokia.com) received the M.Sc. degree in electrical engineering and the Ph.D. degree from Aalborg University, Aalborg, Denmark, in 1996 and 2000, respectively. He is currently a research team leader at Nokia-Bell Labs, Aalborg, and a part-time professor at Aalborg University in the Wireless Communications Network (WCN) section. He is the author/coauthor of more than 100 peer-reviewed publications on a wide range of topics, as well as an inventor on several patents. His current work is related to 5G air interface design,

including radio resource management aspects, and the continued Long Term Evolution (LTE) and its future development. He is currently part of the EU funded research project FANTASTIC-5G that focuses on research toward a new multi-service capable 5G air interface for below 6 GHz operation.

GILBERTO BERARDINELLI (gb@es.aau.dk) received his M.Sc. degree in telecommunication engineering, cum laude, from the University of L'Aquila, Italy, in 2005, and his Ph.D. degree from Aalborg University, Denmark, in 2010. He is currently an associate professor in the Wireless Communication Networks section at Aalborg University, Denmark, also working in close cooperation with Nokia-Bell Labs. His research interests are mostly focused on physical layer and radio resource management design for 5G systems. He is the author or co-author of more than 60 international publications, including conference proceedings, journal contributions, and book chapters. He is currently part of the EU funded research project FANTASTIC-5G.

FRANK FREDERIKSEN (frank.frederiksen@nokia.com) received the M.Sc. degree in electrical engineering from Aalborg University in 1994. He is currently working at Nokia-Bell Labs, Denmark, as a wireless networks specialist. During his employment, he has been working on concept creation and research within various radio access technologies, including HSPA, LTE, LTE-A, and 5G. His main focus has been on the physical layer and lower medium access control operation. He is the author or co-author of several patents.

PREBEN MOGENSEN (preben.mogensen@nokia.com) received his M.Sc. and Ph.D. degrees from Aalborg University (AAU) in 1988 and 1996, respectively. He became full professor at AAU in 1999, and he is currently heading the Wireless Communication Networks Section in the Department of Electronic Systems. Since 1995, he has also been working part time at Nokia-Bell Labs, and was nominated an NSN fellow in 2009. He has supervised more than 35 successfully finalized Ph.D. candidates, and he has co-authored more than 300 scientific publications. His current research is focused on 5G and IoT.