European Commission
Information Society and Media

SEVENTH FRAMEWORK PROGRAMME

Large Scale Integrating Project

EXALTED

Expanding LTE for Devices

FP7 Contract Number: 258512

WP3 – LTE-M System

D3.4
LTE-M performance evaluation

Contractual Date of Delivery:	31 December 2012
Actual Date of Delivery:	31 January 2013
Responsible Beneficiary:	CEA
Contributing Beneficiaries:	VGSL, EYU, ALUD, TKS, CEA, UNIS, CTTC, TUD, UPRC
Estimated Person Months:	45
Security:	Public
Nature	Report
Version:	1.0

PROPRIETARY RIGHTS STATEMENT
This document contains information, which is proprietary to the EXALTED Consortium.

Document Information

Document ID:	EXALTED_WP3_D3.4.doc
Version Date:	30 January 2013
Total Number of Pages:	127
Abstract	This report provides the complete and final specification of LTE-M and defines the utilized evaluation methodology. All design objectives of LTE-M are discussed and the corresponding solutions are presented. The evaluated algorithms and protocols comply with the EXALTED architecture in D2.3 and are meant as input for D2.4 that will give an overall view of the EXALTED system concept and its performance. Furthermore, this report serves as guideline for complementing measurement campaigns in WP7.
Keywords	EXALTED, LTE-M, Machine-to-Machine (M2M), Machine Type Communications (MTC) evaluation methodology, spectrum efficiency, energy efficiency, cost efficiency

Authors

Name	Organisation	Email
Cédric Abgrall (Editor)	Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA)	cedric.abgrall@cea.fr
Stephan Saur	Alcatel-Lucent Deutschland AG (ALUD)	stephan.saur@alcatel-lucent.com
Dejan Drajic	Ericsson d.o.o. (EYU)	dejan.drajic@ericsson.com
Nemanja Ognjanovic	Telekom Srbija (TKS)	nemanjao@telekom.rs
Dimitri Kténas	Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA)	dimitri.ktenas@cea.fr
Bertrand Devillers	Centre Tecnològic de Telecomunicacions de Catalunya (CTTC)	bertrand.devillers@cttc.es
Giuseppe Cocco	Centre Tecnològic de Telecomunicacions de Catalunya (CTTC)	giuseppe.cocco@cttc.es
Christian Ibars	Centre Tecnològic de Telecomunicacions de Catalunya (CTTC)	christian.ibars@cttc.es
Prakash Bhat	Vodafone Group Services Limited (VGSL)	prakash.bhat@vodafone.com
Hong Chen	University of Surrey (UNIS)	hong.chen@surrey.ac.uk
Yi Ma	University of Surrey (UNIS)	y.ma@surrey.ac.uk
Hongfei Du	University of Surrey (UNIS)	hongfei.du@surrey.ac.uk
Walter Nitzold	Technische Universität Dresden (TUD)	walter.nitzold@ifn.et.tu-dresden.de
Athanasios Lioumpas	University of Piraeus Research Center (UPRC)	lioumpas@unipi.gr
Petros Bithas	University of Piraeus Research Center (UPRC)	pbithas@unipi.gr
Antonis Gotsis	University of Piraeus Research Center (UPRC)	agotsis@unipi.gr

Approvals

	Name	Organisation	Date	Visa
Internal Reviewer 1	Nenad Gligoric	EYU	15/01/2013	OK
Internal Reviewer 2	Javier Valiño/Juan Rico	TST	15/01/2013	OK
Technical Manager	Pirabakaran Navaratnam	UNIS	31/01/2013	OK
Project Manager	Djelal Raouf	SWIR	31/01/2013	OK

Executive Summary

At the beginning of the project EXALTED, one of the main research challenges was to decide whether an evolutionary or a revolutionary approach is the adequate solution for the LTE-M system design. While the first has its benefits with respect to backward compatibility to previous LTE releases, re-use of existing hardware and network components, and a relative seamless continuation of the on-going standardisation work, the latter is superior if the stringent performance requirements for energy- and spectrum efficiency must be met. A common demand for both options was that LTE-M is a system coexisting with LTE in the same spectrum. Intensive analysis at industrial and academic research organizations involved in EXALTED found out that as much evolutionary solutions as possible and as much revolutionary approaches as required is the pragmatic compromise achieving the EXALTED objectives in the best possible way. The document at hand presents the resulting specification of LTE-M and details, how much gains were observed in comparison to the baseline LTE system.

The first part of this report (section 2) is dedicated to LTE-M as radio system framework. Major components were already introduced in the previous project deliverable D3.3 [1]. They are repeated in a condensed format for the sake of completeness, partly refined based on the latest results, and complemented with some additional aspects, e.g. the definition of LTE-M communication bearers. LTE-M fully complies with the EXALTED architecture defined in the project deliverable D2.3 [2] and is one key part of the overall EXALTED system concept. User plane and control plane of the radio protocol architecture are very much based on LTE. The only substantial modification is the option that the eNodeB (Enhanced Node B) may serve as IP client for the LTE-M device and map the IP address to a short local address. However, LTE-M utilizes its own logical- , transport-, and physical channels. Within the existing LTE frame structure, so-called Multicast-Broadcast Single Frequency Network (MBSFN) subframes are available for LTE-M, and the different physical channels are mapped in form of an LTE-M super-frame to these radio resources, a solution that is fully compatible to previous LTE releases and still opens the opportunity to implement specific algorithms and protocols tailored for the needs of Machine-to-Machine (M2M) communications. Optimizations of the random access procedure and of the Hybrid Automatic Repeat Request (HARQ) functionality are the essential modifications in the MAC layer. Radio Resource Control (RRC) and Packet Data Convergence Protocol (PDCP) were adapted as well according to the specific needs of a system suitable for a multitude of short messages from devices with diverse capabilities and requirements. Key enabler is here the registration of information about the devices and their context in the network and the selection of the right mechanisms. As an example, sensors installed at a fixed position don't need any form of mobility management. The actual intelligence of LTE-M is to recognize the situation and to activate or deactivate a set of simple features.

As preparation for the actual performance analysis, two different evaluation scenarios are defined in section 3. The reason is to realize comparable conditions for the investigation with different methods and tools. It was found out that two different scenarios are required, one related to power restricted devices aiming at energy efficiency, and the other for a huge number of devices intending spectrum efficiency. One main outcome of the work package was the fact that it is rather challenging to achieve both objectives simultaneously. Thus, the differentiation of the two evaluation scenarios is a realistic and pragmatic approach. Each scenario consists of agreements concerning the used traffic-, channel-, and path loss models and additional common assumptions like the carrier frequency, cell geometry, properties of the used antennas and others. It was not the intention of EXALTED to derive new models,

but to recruit and to combine suitable existing ones that are well established in the research community. The majority of the models originate from 3GPP documents.

The remaining sections present the complete performance analysis of the LTE-M system. At first, the co-existence with LTE is discussed. Two project objectives are reflected, namely that the proposed solutions have to be supported by existing eNodeB hardware platforms (O3.1 – see Table 1-1) and that backward compatibility to LTE Release 8 [3] is maintained (O3.7). The LTE-M system design as a whole already underlines these requests, e.g. through the separation of radio resources for LTE and LTE-M. But also some individual solutions were explicitly specified to support the co-existence. Registering information about terminals is the key enabler. Thereby, it is possible for the network to distinguish between LTE and LTE-M terminals. This is necessary to exploit the performance potential of innovative scheduling techniques. Other solutions protect the network against sudden overload situations, simply re-use already existing hardware components, or maintain the performance of LTE UEs in the presence of LTE-M traffic.

One of the major objectives in EXALTED is the simultaneous support of a big number of LTE-M devices. As the amount of the overall available radio resources for LTE-M is fixed through the LTE frame structure and the LTE-M super-frame principle, the challenge is to transmit the same information by using fewer resources. This can be achieved either by minimizing the size of control and feedback messages (O3.2), or by optimizing the resource utilization by traffic aggregation or novel signal formats (O3.4). Again the diversity of device capabilities and requirements plays an important role (O3.5). All in all, it was found out that the number of supported devices can be extended by one order of magnitude.

Cost- and complexity reduction are sublime objectives. However, they cause some drawbacks. The main problem is the degradation of the link budget, particularly in the uplink. At the end, this leads to insufficient coverage of LTE-M devices because the cellular layout is dimensioned according to the LTE specification. Therefore, in section 6 two solutions are presented that ensure wide area coverage also for LTE-M. This is another important aspect of the EXALTED objective to support devices with diverse capabilities and requirements in one system (O3.5). CDMA-overlay in the uplink exploits the simple principle of a spreading gain. For the downlink a collaborative broadcast architecture is proposed that achieves a very high level of coverage.

Besides spectrum efficiency, energy efficiency is the second big challenge in EXALTED. In other words, the lifetime of the battery shall be extended from a couple of days to the duration of up to one year. In order to be able to analyse this complex problem, at first the sources of energy consumption in a conventional LTE UE were identified, and a relative breakdown was derived. Thereby it is distinguished if the device is in ACTIVE mode or in IDLE mode. In a subsequent step, several solutions were analysed with respect to the specific source of energy consumption that they try to reduce or to avoid. Apart from the obvious project objective of energy minimization in the device (O3.6), also the optimization of paging and polling of devices (O3.3) was handled in section 7. The main result is that significant energy savings are possible. They can actually lead to a battery lifetime in the order of one year.

Finally, in section 8 the possible reduction of the device cost is discussed. Similar as for the energy consumption, the assessment is based on a breakdown of the contributions from the different components. For this, EXALTED adopted the work of 3GPP, summarized in the technical report 36.888 [4], which already provides a broad set of means for cost reduction for M2M devices. The added value from EXALTED, aiming at the objective to minimize the device complexity (O3.2) is basically the proposed MIMO scheme.

Table of Contents

1. Introduction and Background

This report provides the final specification and the performance evaluation results of the LTE-M system that has been proposed and developed within the EXALTED project. In particular, it demonstrates that the proposed LTE-M system solutions achieve the WP3 objectives listed in Table 1-1, whilst they all fit into the EXALTED architecture elaborated in D2.3 [2]. Detailed descriptions of the proposed solutions are given in the report D3.3 [1] and therefore are not presented again. Instead, only the basic idea is briefly sketched.

The five most pertinent design objectives which are detailed in [1] are as follows:

- Co-existence with LTE
- Simultaneous support of a big number of machine devices and spectrum efficiency
- Provision of wide area coverage
- Energy savings to enable a long battery lifetime
- Cost efficiency

The above design objectives can be mapped to the WP3 objectives, as shown in Table 1-1.

Table 1-1: Mapping of the WP3 objectives with the design objectives.

WP3 objectives	Design objectives from D3.3 [1]	Overall KPI (see Appendix A1)	Explanation of the mapping
O3.1: Proposed solutions have to be supported by existing eNodeB hardware platforms	Co-existence with LTE (section 4)	Co-existence achieved (yes or no)	LTE-M solutions should ensure a reuse of LTE infrastructure with no or only minimal changes
O3.2: Minimization of complexity and feedback signalling	Simultaneous support of a big number of machine devices and spectrum efficiency (section 5) and cost efficiency (section 8)	• Device cost relative to LTE • Energy consumption relative to LTE or battery lifetime	By minimizing complexity, devices have to perform less computational tasks; this permits to use cost effective components. Reduction of feedback signalling permits to serve more efficiently a higher number of devices that request a channel opportunity.
O3.3: Enable paging and polling of M2M devices	Energy savings to enable a long battery lifetime (section 7)	• Paging and polling possible (yes or no) • Energy consumption relative to LTE or battery lifetime	Optimised paging and polling processes permit devices to operate in sleep periods mostly and then avoid energy depletion.
O3.4: Optimization of resource utilization including traffic aggregation at eNodeBs, relay stations and gateways	Simultaneous support of a big number of machine devices and spectrum efficiency (section 5)	Maximum number of active devices	Efficient radio resource usage and management permit to transmit the same information on fewer resources. This includes methods like cooperation and aggregation. Then devices can either enhance their spectrum efficiency or their robustness. This permits also to satisfy more devices.

O3.5: Support of devices with diverse capabilities and requirements in one system	Simultaneous support of a big number of machine devices and spectrum efficiency (section 5) and wide area coverage (section 6)	• Maximum number of active devices • Coverage	Considering the heterogeneity of devices and traffics in the scheduler permits to exploit resources more efficiently. Then more devices can be satisfied. Moreover, heterogeneous devices are characterized by different link budgets. This may be prejudicial for LTE-M devices in regard to coverage.
O3.6: Minimization of the energy consumption of devices	Energy savings to enable a long battery lifetime (section 7)	Energy consumption relative to LTE or battery lifetime	The mapping is straightforward.
O3.7: Maintenance of backward compatibility to LTE Release 8	Co-existence with LTE (section 4)	Co-existence achieved (yes or no)	LTE-M solutions should ensure backward compatibility with legacy LTE standard without affecting LTE system.

The five LTE-M system design objectives are mapped with corresponding key performance indicators (KPIs) in Table 1-1. These KPIs are used to asses if the proposed LTE-M system solutions meet spectrum-, energy- and cost-efficiency while providing wide area coverage for a multitude of machine devices transmitting sporadically short messages, ideally with no or minimal changes on the existing LTE infrastructure. To go more in depth with the evaluation, an extended set of subordinated metrics was also used that can be found in Appendix A1.

Section 2 details the specification of the LTE-M system on physical, medium access control, radio link control, packet data convergence protocol, and radio resource control (sub-) layers. LTE-M device classes as well as bearer- and QoS classes are also discussed. Section 3 presents the evaluation framework used for performance assessment. Two scenarios are defined: The first one aims to implement low complex and energy efficient M2M communications, and the second one is tailored for a deployment with a large number of LTE-M devices with heterogeneous requirements and capabilities. Channel propagation and traffic models as well as system parameters are summarized for each scenario.

Finally, sections 4 to 8 provide the performance evaluation results of the proposed solutions. The solutions are assessed according to the evaluation methodology described in section 3 and against the KPIs in Table 1-1. Section 4 demonstrates the co-existence of LTE-M solutions with LTE. Since the expected number of LTE-M devices is much higher than the number of legacy LTE terminals, section 5 evaluates the number of supported LTE-M devices for the proposed solutions on the second evaluation scenario. The last three sections evaluate the solutions on the first scenario. In section 6 two solutions, namely CDMA overlay and collaborative broadcast architecture, are assessed for provision of wide area coverage. Section 7 introduces a power consumption breakdown in order to evaluate the solutions for energy efficiency. Finally, cost efficiency is studied in section 8 with a low complex MIMO solution for M2M communications. As some solutions aim at more than just one single objective, they appear in several sections with different focus. These repetitions are intentional and shall simplify the practical use of this document for developers.

Further performance results of a HARQ solution that does not necessarily fit within the evaluation framework described in section 3, but underline its usability for specific problems are provided in Appendix A2.

2. Specification of LTE-M

The LTE-M system is designed to “co-exist” with LTE. EXALTED strives to follow the evolutional approach of LTE-M implementation, meaning that additional features are introduced to the existing LTE/LTE-A system components. The idea behind this approach is seamless transition from pure LTE/LTE-A to LTE/LTE-A/LTE-M, by software upgrade of access and core network elements. However, the feasibility of this transition will be driven by industry, for example by the readiness of existing eNodeB cabinets for this upgrade. To achieve the required performance with respect to energy- and spectrum efficiency, also some major modifications, far in excess of a pure evolution, are proposed in the LTE-M system design. An example is the introduction of Generalized Frequency Division Multiplexing (GFDM) instead of Single Carrier Frequency Division Multiple Access (SC-FDMA) as access scheme in the LTE-M uplink.

The preliminary system concept is derived from technical requirements in the project report D2.1 [5] and the conference paper [6]. It is furthermore developed by creating algorithms running on the elements of the architecture, at specific protocol layers. The complete description of algorithms is given in the project report D3.3 [1]. They are presented in this section in a consolidated way – new algorithms that resulted from the work in EXALTED are explained, and the reference to the existing LTE/LTE-A algorithms is given in case they are used as already specified in 3GPP.

Subsection 2.1 introduces the overall protocol layered architecture. Subsections 2.2 – 2.6 describe layers and sub-layers in more details, presenting the relevant algorithms. The overview of device classes, capabilities and categories is given in 2.7.

Figure 2-1 shows the overall EXALTED architecture and highlights the LTE-M radio access network, which is the subject of this report. It has an interface to the Evolved Packet Core (EPC), denoted with I-3 (S1 Interface in 3GPP). The air interface between LTE-M enabled eNodeB and stand-alone LTE-M devices or M2M gateways is indicated with I-4 and I-5, respectively.

From the perspective of the access network I-4 and I-5 are identical. Therefore we do not distinguish between both identifiers throughout this report. Moreover, the terminology *LTE-M device* used in this specification refers to both components LTE-M device and M2M gateway shown in Figure 2-1.

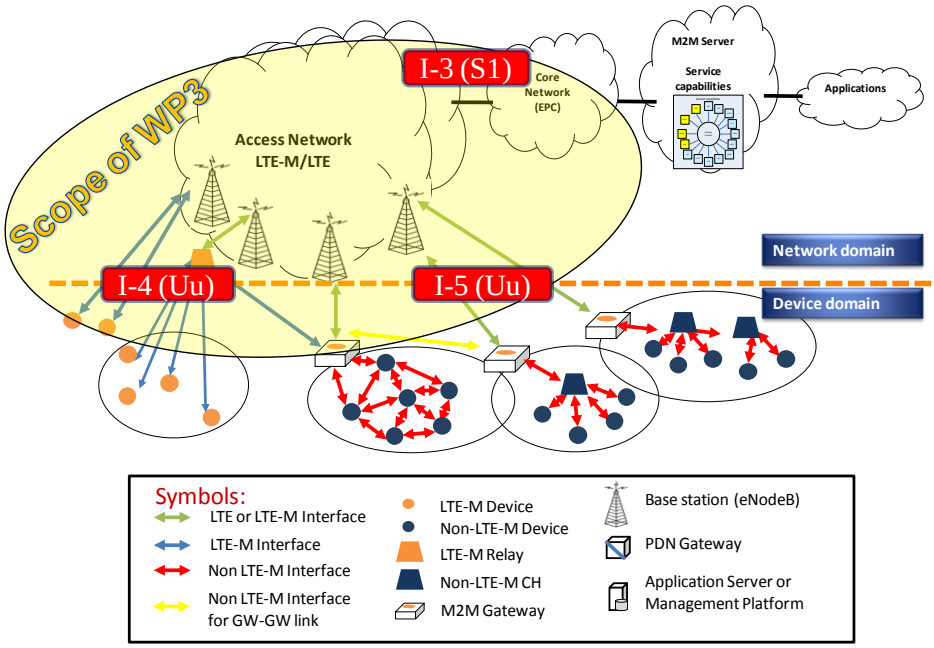


Figure 2-1: Prominence of LTE-M within the EXALTED system architecture.

2.1 LTE-M Radio Interface Protocol Architecture

This section outlines the radio protocol architecture of LTE-M. Similar to LTE, it is distinguished between user plane and control plane.

Figure 2-2 and Figure 2-3 show two different options for the user plane protocol stack. The difference between these two options is the capability of the LTE-M device to support an IP connection to the Evolved Packet Core (EPC), more precisely to the Packet Data Network (PDN) Gateway (P-GW).

The protocol architecture in Figure 2-2 assumes that the LTE-M device is able to execute IP over the S1 Interface just as LTE UEs. The EXALTED architecture denotes this interface as I-3 (see EXALTED project report D2.3 [2]). Another particularity of LTE-M is that the Radio Link Control (RLC) sub-layer operates always in transparent mode. The functionality of the RLC is not required in LTE-M (see section 2.4). Medium Access Control (MAC) protocol and Packet Data Convergence Protocol (PDCP) over the Uu Interface, in the EXALTED architecture indicated with I-4 and I-5, respectively, are optimized according to the particular needs and characteristics of LTE-M devices (see sections 2.3 and 2.5). Major changes compared to LTE are proposed on the Physical Layer (PHY). LTE-M utilizes its own Physical Channels and a special frame structure (see section 2.2).

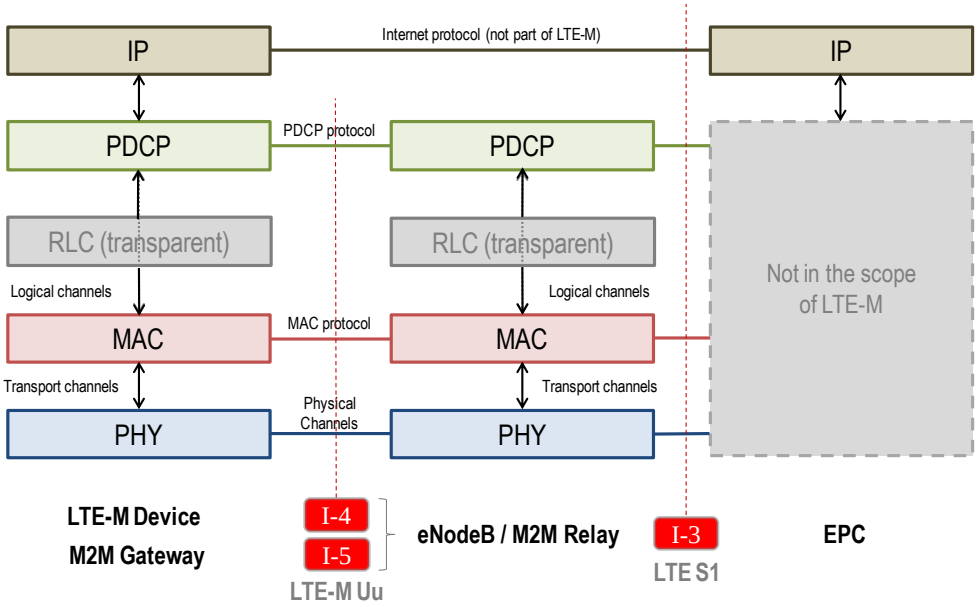


Figure 2-2: LTE-M user plane protocol stack (option 1).

Figure 2-3 illustrates the user plane protocol stack valid for LTE-M devices that are not able to support an IP connection with the EPC. From the perspective of the EPC the IP connection has to look equivalently as in the first case because its functionality shall remain completely unchanged. The solution proposed by EXALTED is to terminate the IP connection at the eNodeB and to map the IP address to a local address with much less overhead on the air interface (see section 2.5). Address translation has been also intensively studied in the project reports D4.1 [7] and D4.2 [8]. The counterpart of the address mapping function is carried out at the device. Thus, an emulation of the IP layer is achieved. The other layers in Figure 2-3 (RLC, MAC, and PHY) behave similarly as described above.

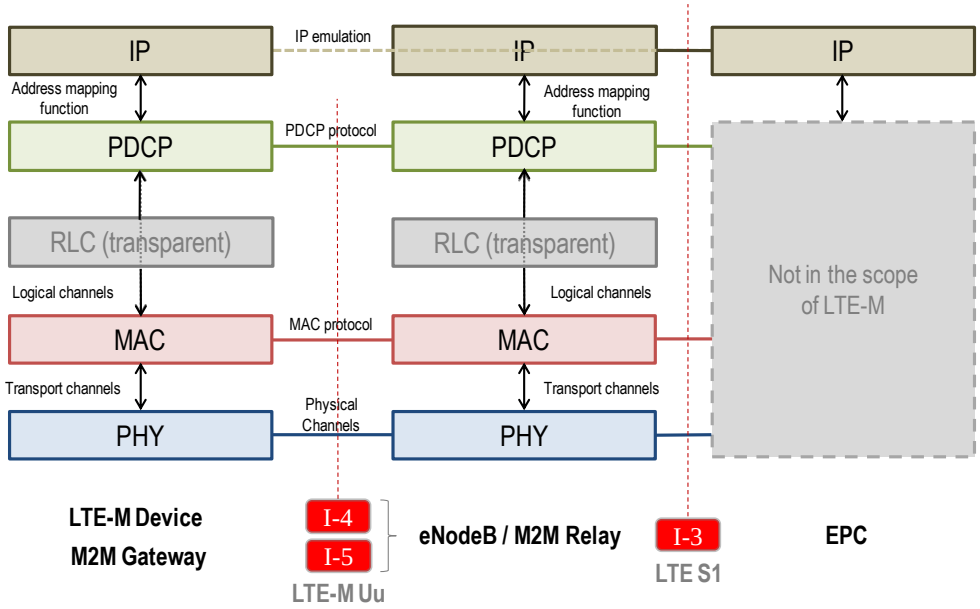


Figure 2-3: LTE-M user plane protocol stack (option 2).

Figure 2-4 shows the control plane protocol stack of LTE-M. In contrast to LTE, the S1 Interface is enriched with the feature to register LTE-M specific information about the devices in the EPC. This knowledge is used to activate only the minimal required set of functions for the respective device. As an example, mobility support is not needed for a stationary sensor.

This registering of information is an important means to optimize protocols and to save energy and spectrum. More details are given in section 2.6.

The LTE-M control plane RLC protocol operates in acknowledged mode (see section 2.4).

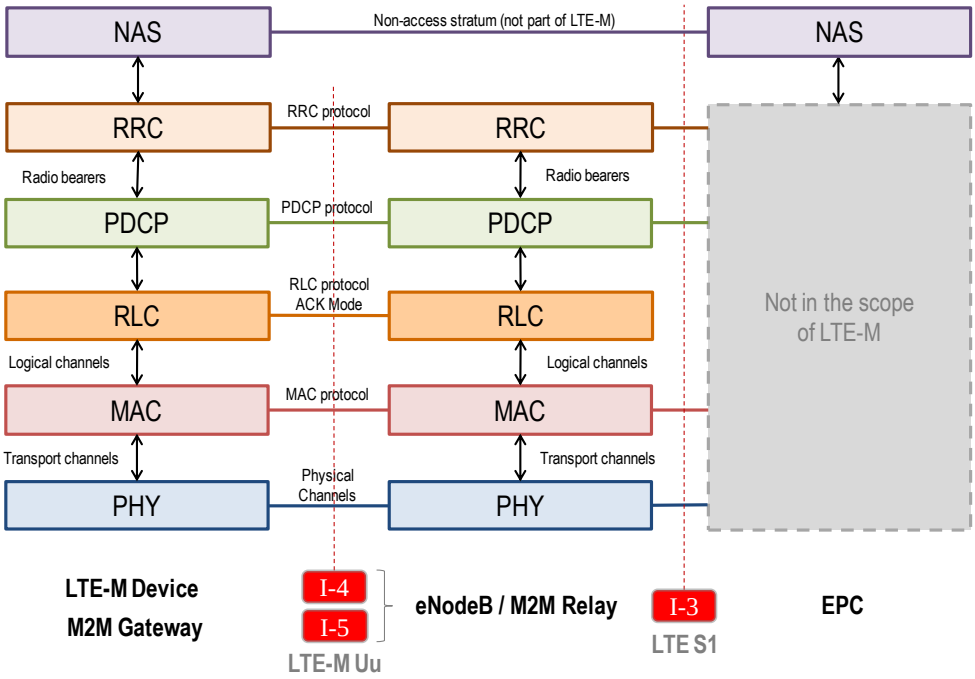


Figure 2-4: LTE-M control plane protocol stack.

2.2 LTE-M Physical (PHY) Layer

This section introduces the Physical Layer of LTE-M. The specification primarily covers aspects that differ from LTE in some respects, namely services provided to higher layers, access schemes, physical channels and signals, the basic frame structure, bit- and symbol processing together with signal generation, and, finally, required measurements.

2.2.1 Services provided to higher layers

Just as explained in the LTE specification 36.201 [9], the physical layer of LTE-M offers data transport services to higher layers. The interface between MAC sub-layer and physical layer is realized through the use of transport channels as shown in the protocol architecture in section 2.1. LTE-M uses its own transport channels, which are mapped to LTE-M specific physical channels. Apart from that, the functions listed in [9] (section 4.1.2) are valid for LTE-M as well. However, the specification of the physical channels is different from LTE and considers the particular needs of M2M communications. As an example, in the downlink convolutional coding is applied instead of Turbo coding to reduce the computational complexity in the LTE-M device. The specification of the physical channels in LTE-M is detailed in section 2.2.3.

2.2.2 Access schemes

The physical layer of LTE-M uses Orthogonal Frequency Division Multiplexing (OFDM) with Cyclic Prefix (CP) as multiple access scheme in the downlink, just as LTE as mentioned in the specification 36.201 [9]. In the uplink Generalized Frequency Division Multiple Access (GFDM) is applied. The main advantage of GFDM compared to Single Carrier Frequency Division Multiple Access (SC-FDMA), the preferred solution in LTE, is its ability to occupy small spectrum chunks with minimal distortion of the surrounding LTE signals.

The preferred, but not mandatory method to separate downlink and uplink signals in LTE-M is half-duplex operation in Frequency Division Duplex (FDD) systems, which are operated in many European countries and in North America. Downlink and uplink use different time intervals and separated frequency bands. Compatibility with full-duplex LTE is ensured. The fact that an LTE-M device does not need to transmit and receive simultaneously is an essential factor to reduce cost and complexity of the device, because expensive duplex filters are not required.

2.2.3 Physical channels and signals

LTE-M uses its own physical channels. This has two important advantages: Firstly, the existing physical channels defined for LTE remain unchanged, i.e. legacy LTE UEs are not affected by any new functionality, and the channels are not overloaded by M2M traffic. Secondly, all degrees of freedom to specify PHY layer algorithms tailored for M2M can be exploited for the new LTE-M channels if a significant performance benefit can be achieved.

In the uplink, the definition of physical signals differs from LTE. Sounding is not supported and the Demodulation Reference Signals (DM-RS) depend on the GFDM signal processing. In the downlink, physical signals defined in LTE are re-used in LTE-M and extended on LTE-M radio resources.

The physical channels and signals used in LTE-M are defined and explained in the project report D3.3 [1]. For the sake of completeness of this section, they are briefly summarized in the following.

2.2.3.1 Uplink

2.2.3.1.1 Uplink signals

In the uplink, LTE-M specific DM-RS within the uplink data sub frame are introduced. A preamble is prepended to the GFDM symbol as reference signal.

In LTE, the sounding signal is sent over the complete system bandwidth. LTE-M devices cannot support this feature because their transmission bandwidth is reduced to 1.4 MHz. Therefore, the sounding signal is not present in LTE-M. The vendor specific scheduler can partly compensate this drawback by applying a frequency hopping pattern to exploit frequency diversity.

2.2.3.1.2 Physical MTC Uplink Shared Channel (PMUSCH)

The PMUSCH is the physical uplink shared channel. This channel carries the payload information and some control information from the LTE-M devices to the eNodeB. A detailed specification of the PMUSCH bit- and symbol processing is given in sections 2.2.5 and 2.2.6, respectively.

2.2.3.1.3 Physical MTC Uplink Control Channel (PMUCCH)

The PMUCCH is the physical uplink control channel when there is no uplink data to be sent. If the LTE-M device has uplink data to send, the control information is multiplexed with data in the PMUSCH. It is envisaged to carry primarily feedback information as well as handling for HARQ procedures.

2.2.3.1.4 Physical MTC Random Access Channel (PMRACH)

The PMRACH is the physical uplink random access channel. It is used for the initial attachment to an eNodeB e.g. after switching on the LTE-M device for the first time and also

for the transmission of sporadic short messages. Efficient proposals for using the PMRACH are described in the project report D3.3 [1], sections 2.5.1.3 and 4.1.

2.2.3.2 Downlink

2.2.3.2.1 Downlink signals

LTE-M will re-use the Primary Synchronisation Signal (PSS) and Secondary Synchronisation Signal (SSS) in sub frames 0 and 5 of the LTE system for slot timing detection, radio frame timing detection, physical layer cell ID detection, and cyclic prefix length detection. These signals are defined in the 3GPP specification 36.211 [10].

LTE-M implicitly derives the number of antenna ports from the Cell specific Reference Signals (CRS) that are adopted from the LTE specification 36.211 [10].

2.2.3.2.2 Physical MTC Downlink Shared Channel (PMDSCH)

The PMDSCH is the physical downlink shared channel. It is used by the eNodeB to transmit data dedicated to certain devices. The MCS used for transmission in this channel is defined in the PMDCCH and therefore can support all envisaged Modulation and Coding Schemes (MCS) of LTE-M for efficient use of the spectrum.

2.2.3.2.3 Physical MTC Downlink Control Channel (PMDCCH)

LTE Physical Downlink Control Channel (PDCCH) capacity is insufficient to schedule large number of machine devices and the PMDCCH is required to be introduced within the configured Multicast-Broadcast Single Frequency Network (MBSFN) subframe. The PMDCCH is the physical downlink control channel and carries information about the scheduling of the resources for PMDBCH, PMDSCH and PMPCH for a specific LTE-M device or a group of LTE-M devices in up- and downlink as well as the respective MCS schemes. Information in this channel might also be encoded with the MCS for worst channel conditions but optionally the appropriate MCS for the specific cell might also be specified in the PMDBCH. A new System Information Block (SIB) identifier MTC System Information Radio Network Temporary Identity (M-SI-RNTI) is introduced for LTE-M SIB's in the PMDBCH.

2.2.3.2.4 Physical MTC Downlink Broadcast Channel (PMDBCH)

The PMDBCH is the physical downlink broadcast channel for LTE-M. It carries information about the device timers, configurations to enable access to the cell such as configuration of the PMDCCH, PMRACH and PMDMCH. This channel is commonly encoded with a fixed MCS, in a specific fixed subframe and has fixed periodicity that allows the decoding with prior knowledge of only frame timing even at cell edges. The information in the PMDBCH might also be encoded using incremental redundancy codes as described in the project report D3.3 [1], section 6.1.

2.2.3.2.5 Physical MTC Downlink Multicast Channel (PMDMCH)

The PMDMCH is the physical downlink multicast channel. It carries information that is transmitted to a multitude of nodes, specifically a multicast group. Configuration such as periodicity of occurrence and radio resources of this channel is indicated in the PMDBCH. The transmission in this channel is encoded with the incremental redundancy multicast code defined in the project report D3.3 [1], section 6.1.

2.2.3.2.6 Physical MTC Paging Channel (PMPCH)

Since the subframes with LTE paging occasions cannot be configured as MBSFN subframes (the LTE-M downlink may use MBSFN subframes as specified later), dedicated paging

channel are required for LTE-M. To facilitate deep sleep, subframes 1 and 6 are configured as MBSFN subframes for LTE-M and paging channel (PMPCH).

2.2.4 Frame structure

The sharing of spectrum with the primary LTE system is a key component of the LTE-M system design. Therefore a specific frame structure is used, that accounts for the flexible use of the primary resources of LTE. The specification of the LTE-M frame structure is shown in D3.3 [1] but for completeness of the specification, a summary of the frame structure is given in this document.

The downlink frame structure relies on the use of dedicated MBSFN subframes. Time multiplexing of LTE Multimedia Broadcast Multicast Services (MBMS) on the Multicast Channel and LTE-M is exemplary shown in the project report D3.3 [1] in Figure 2-5. The MBSFN subframes are configured by higher layers to be used for LTE-M. The pattern of MBSFN subframes is provided by the SIB2 of LTE which yields a maximum of six subframes per 10ms frame. The periodicity of MBSFN subframes is 40ms. To account for the different physical channels of LTE-M, time interleaving is done over multiple MBSFN subframe resources. A super frame structure is therefore realized with a super frame duration expressed in multiples of the MBSFN repeat period. The super frame structure is shown in Figure 2-5.

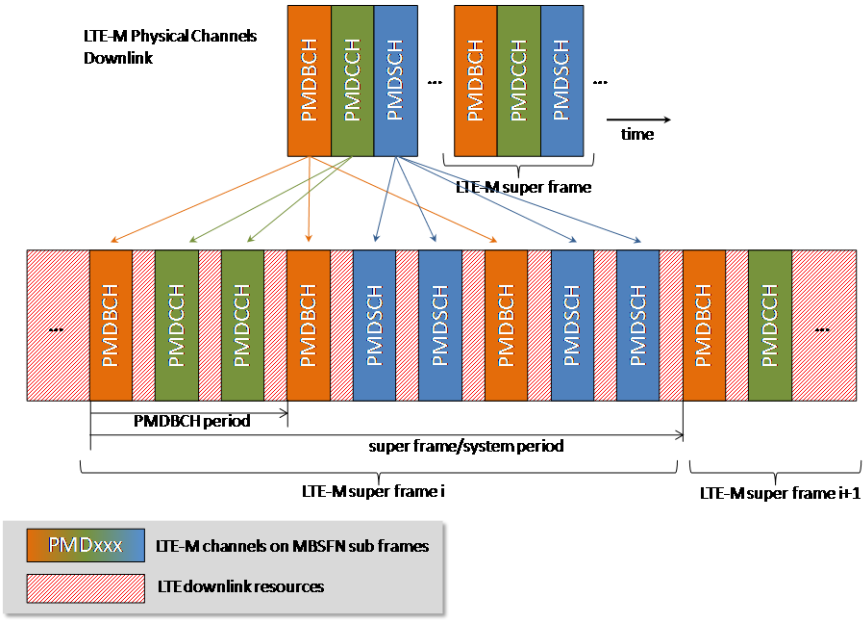


Figure 2-5: Mapping of LTE-M super frame structure to MBSFN frames in LTE downlink resource grid.

This super frame structure carries the three downlink physical channels (PMDBCH, PMDCCH, and PMDSCH) which are concatenated in time and additionally multiplexed in frequency. While applying the mapping of physical channels to MBSFN subframes, the PMDBCH is more frequently interleaved compared to other channels, as this ensures a fine granular wake-up and re-connect of M2M terminals in stand-by mode. The mapping of the other channels into the frame structure can be configured to allow for efficient usage of the primary LTE system resources. The parameters for the mapping of PMDCCH and PMDSCH are encoded in the PMDBCH. Additionally to the super-frame structure, the physical channels are also multiplexed in frequency as depicted in Figure 2-6.

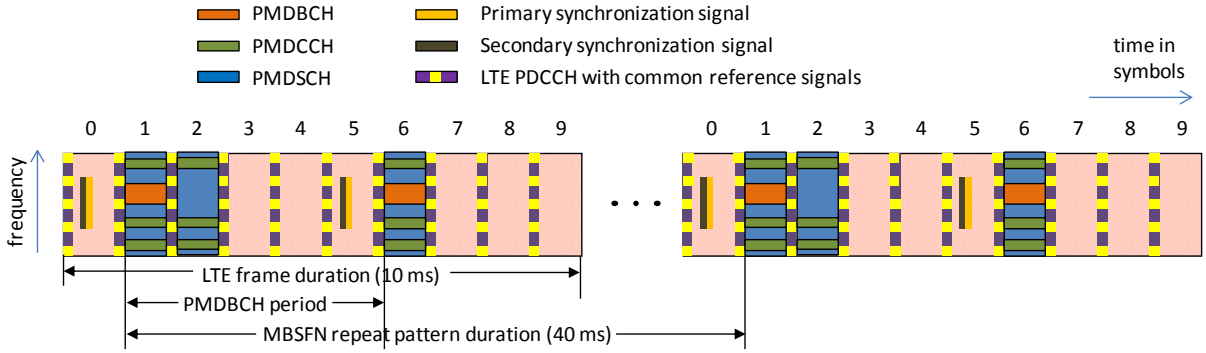


Figure 2-6: LTE-M downlink frame structure.

For the uplink no specific frame structure is specified as the uplink shared channel of LTE-M (PMUSCH) is freely scheduled within the primary LTE spectrum of PUSCH on a per M2M terminal basis. The uplink control channel of LTE-M is scheduled to the centre of the frequency band to allow maximum of orthogonality to the LTE control channels situated at the edges of the frequency band. The PMRACH is multiplexed into the PMUCCH with specified interval t_{PMRACH}. An exemplary mapping of the uplink physical channels to the resources is shown in Figure 2-7.

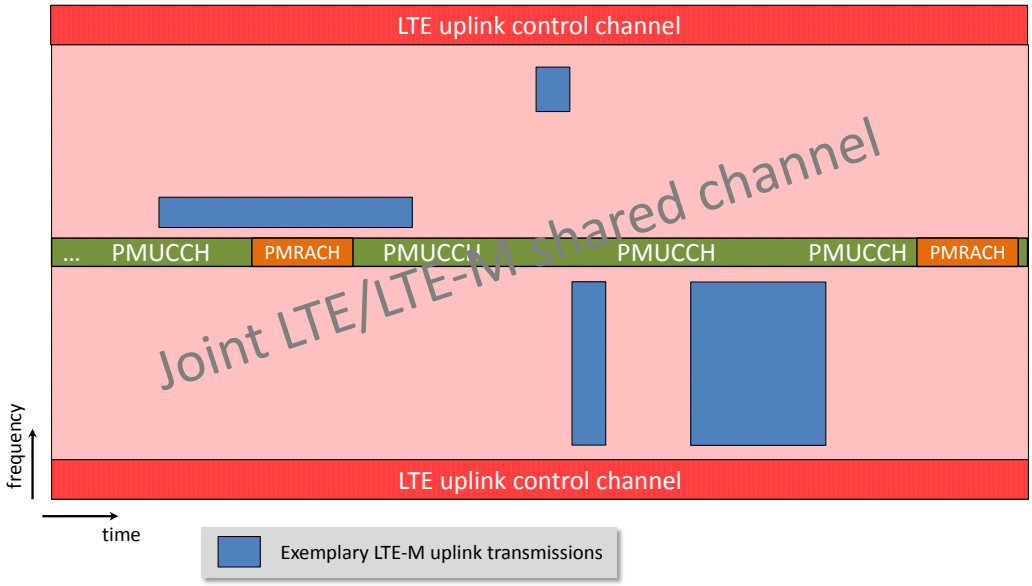


Figure 2-7: Sharing of LTE-M uplink channels with primary LTE resources.

2.2.5 LTE-M bit processing and multiplexing

The bit processing algorithms of LTE are defined in the 3GPP specification 36.212 [11]. Bit processing includes all steps performed with a transport block handed down from the MAC layer before the mapping of the bit sequence on modulation symbols is done. In the Physical Uplink Shared Channel (PUSCH), these are:

- Cyclic Redundancy Check (CRC) attachment to detect erroneous transport blocks;
- Code block segmentation and further CRC attachments to code blocks;
- Forward Error Correction (FEC) to protect the information bits with redundancy;
- Rate Matching (RM) to adapt the size of the coded transport block to the available channel bit rate, i.e. the number of available radio resources;
- Code block concatenation;

- FEC of the following types of control information: Channel Quality Indicator (CQI), Precoding Matrix Indicator (PMI), rank of the channel matrix and acknowledgements (ACK or NACK);
- Multiplexing of the data bit sequence and the coded control information;
- Channel interleaving to protect the bit sequence against deep fades.

The processing of other uplink and downlink Physical Channels does not always include the full set of the above mentioned steps. Also the realization of the algorithms differs from case to case. As an example, PUSCH applies Turbo coding as FEC, while convolutional coding is utilized for the Physical Broadcast Channel (PBCH).

In order to reuse existing eNodeB hardware, which is one of the core EXALTED objectives, the Physical Channels of LTE-M adopt many algorithms specified in 3GPP 36.212 [11], but with several simplifications and restrictions that reduce the complexity of transmitter and receiver while the performance is still sufficient. The only reason for applying a completely different algorithm is a significant performance benefit with respect to one of the system KPIs shown in sections 4-8 in this report.

Figure 2-8 shows the block diagram of the bit processing for the PMUSCH as an example. Input to the bit processing chain is one transport block of a MTC Uplink Shared Channel (M-UL-SCH) with A bits. While in LTE the size A of the transport block is more or less an arbitrary number with a lower and an upper bound, we introduce a restriction in LTE-M. As the data payload size is expected to be very small, the maximal transport block size in LTE-M is set to $A_{\max} = 6120$ bits. This number ensures that only one single code block has to be processed. In the following, the single steps are explained and differences to the LTE specification 36.212 [11] are highlighted:

- The CRC attachment consists of 24 bits. The generation of the CRC bits is equivalent as in LTE and explained in the 3GPP specification 36.212 [11], section 5.1.1. Output is a bit sequence of size $B = A + 24$ with $B_{\max} = 6144$ bits.
- Code block segmentation and code block concatenation (not shown in Figure 2-8) are not required in LTE-M because the maximal code block size $C_{\max} = 6144$ is never exceeded.
- In LTE-M the PMUSCH code block is Turbo coded exactly as defined in the 3GPP specification 36.212 [11], section 5.1.3.2. The rate of the coder is 1/3, i.e. the output bit sequence consists of the B information bits and $2 * B$ parity bits from the two constituent encoders. Moreover 12 termination bits are added. In total the output size is $D = 3 * B + 12$ bits. The adoption of the LTE FEC algorithm allows the reuse of the Turbo decoder as existing and well established functional unit in the eNodeB receiver for LTE-M.
- The purpose of rate matching is to adapt the size of the coded block to the available radio resources, also referred to as channel bit rate. It depends on the number of assigned resource blocks, and the selected modulation scheme. Moreover, it must be considered that some of the available resources are reserved for control information. The number of coded data bits that can be mapped on the modulation symbols is defined as E. This number can be smaller than, equal to, or greater than D. If $E > D$, some of the bits in the coded block are simply repeated. If $E < D$ selected bits must be removed. In LTE, the rate matching algorithm is closely related with the HARQ by writing the bits in a virtual cyclic buffer and reading a different subset of them during each (re-) transmission. As LTE-M utilizes a different HARQ algorithm, also the rate matching in LTE-M works differently. The rate matching for the LTE-M adaptive HARQ algorithm depends on a Look-Up Table (LUT), which is generated by an off-line training process and stores the coding rates for different Signal to Noise Ratios (SNRs). The turbo coding based on these coding rates in the LUT may achieve a specific Packet Loss Ratio (PLR)

under the corresponding SNRs. For example, the coding rate of 1/3 may achieve a PLR of 0.1 for the SNR of 2.1dB over an AWGN channel. In the LTE-M HARQ scheme, the system CQI message is firstly fed back to the transmitter before the transmission. By searching the LUT, a coding rate is found corresponding to this CQI message. In order to match this coding rate, puncturing operations are applied to the turbo encoded packet having an initial rate of 1/3. The resultant punctured packet will be sent out for the first transmission.

- In LTE-M the amount of control information is reduced. The number of CQI bits is significantly smaller than in LTE because only a 1.4 MHz sub-band and not the complete system bandwidth is scanned and reported. PMI and rank information are not needed at all because of the restricted multiple antenna support in LTE-M. ACK/NACK is sent as defined in the LTE specification 36.212 [11].
- The bit interleaving function of LTE-M is adopted from LTE. It is defined in the specification 36.212 [11], section 5.2.2.8.

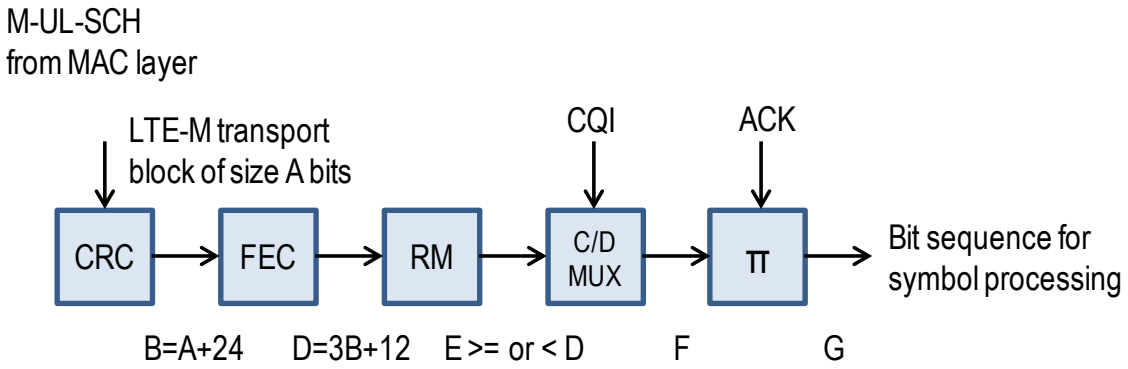


Figure 2-8: LTE bit processing and multiplexing.

2.2.6 LTE-M symbol processing and signal generation

The symbol processing of LTE is specified in the 3GPP document 36.211 [10]. Symbol processing includes all steps after channel interleaving (see previous section 2.2.5) until the generation of the physical transmit signal. In the PUSCH, these are:

- Scrambling: Firstly, the interleaved bit sequence is multiplied with a pseudo random sequence generated with shift registers with device individual initialization. The objective is to improve the robustness against inter-cell interference.
- Modulation mapping: Scrambled bit sequence is mapped to a sequence of complex modulation symbols. Depending on the channel quality, QPSK, 16-QAM and 64-QAM is possible.
- Layer mapping: Depending on the multiple antenna operation modes, modulation symbols are mapped to layers. Each layer can be considered as an independent transmission. Layers are transmitted simultaneously on the same frequency resources.
- Transform precoding: LTE applies SC-FDMA. The transform precoder is the distinguishing element between OFDMA and SC-FDMA. Actually, the transform precoder is a Discrete Fourier Transform (DFT) within the sub-band used by the UE. The objective is to achieve a more or less equal distribution of amplitudes at the output of the subsequent IFFT, which leads to a low Peak-to-Average Power Ratio (PAPR).
- Mapping to physical resources: The output values of the transform precoder are mapped to physical resources (12 out of 14 OFDM symbols available for data, and 12 subcarriers per resource block). Two full OFDM symbols, the middle symbol in each slot, are reserved for DM-RS.

- SC-FDMA signal generation: After the mapping to physical resources is done, an IFFT is performed for each OFDM symbol. Afterwards, the Cyclic Prefix (CP) is added into the sequence of time samples.
- Modulation and up-conversion: Finally, the sequence of complex-valued time samples is converted to a RF signal and filtered.

The symbol processing of other uplink and downlink physical channels partly follows different rules, which are not detailed in this report. It is referred to the LTE specification 36.211 [10].

LTE-M adopts some of the algorithms mentioned above, but taking into account the specific characteristics of LTE-M devices and the drawbacks of LTE with respect to the simultaneous support of a multitude of short messages, which are discussed in the project report D3.1 [12]; several new approaches are introduced in the LTE-M symbol processing.

Figure 2-9 shows the exemplary symbol processing for the LTE-M uplink shared channel PMUSCH. The introduction of GFDM as radio access scheme changes the specification of uplink symbol processing of LTE-M compared to LTE.

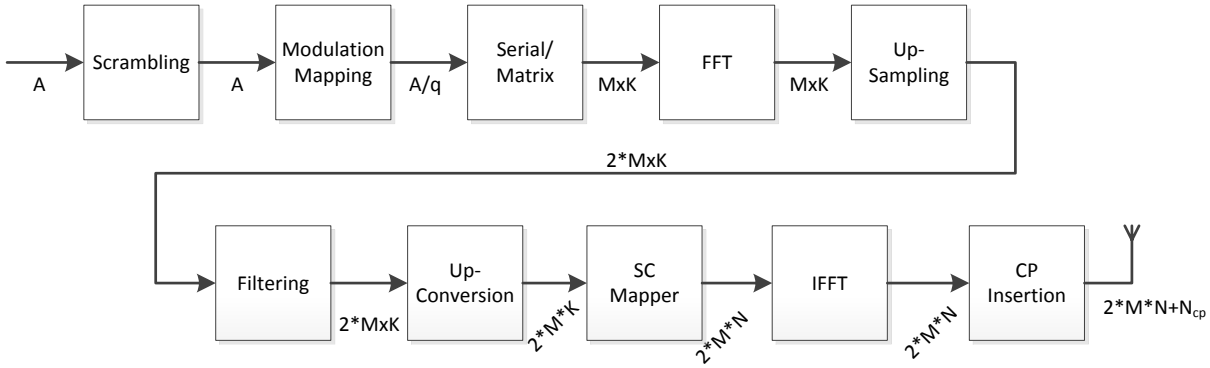


Figure 2-9: Symbol processing for LTE-M.

The input to the symbol processing chain is the transmission block generated by the preceding bit processing chain.

- Scrambling is applied in the same way as done in LTE.
- Afterwards, the bit sequence is mapped to complex modulation symbols depending on the modulation scheme used for transmission. For the PMUSCH, BPSK, QPSK, and 16-QAM are allowed in contrast to the LTE PUSCH, where the possible modulation schemes are QPSK, 16-QAM, and 64-QAM. The usage of BPSK is one important means to improve the link robustness and to preserve wide area coverage. The very high data rates with 64-QAM are not needed in LTE-M.
- The modulation mapping is followed by a conversion from a serial complex valued vector into a matrix of size M*K according to the number of subcarriers K that will be used.
- The subsequent processing is a transformation into frequency domain done by a FFT. The used FFT is an M point FFT.
- The provided blocks are the twofold upsampled and filtered with a root-raised cosine filter with a roll-off factor of 0.2.
- The succeeding processing block is the upconversion. It maps the different data onto the specific subcarriers within the GFDM symbol.
- This mapping is followed by an IFFT that transforms the signal back into time domain and complex valued time samples are obtained.
- A cyclic prefix is prepended to account for channel interference.

- Afterwards the complex valued time samples are handed over to the RF chain for further processing and transmission over the air.

The symbol processing is the same for PMUSCH and PMUCCH.
LTE-M foresees the additional option to multiply the complex modulation symbols with Code Division Multiple Access (CDMA) codes as defined in the project report D3.3 [1] in order to extend the coverage. This overlay can be applied independently from the symbol processing described above for all uplink and downlink channels.

2.2.7 Physical layer measurements

LTE-M devices will use downlink Common Reference Signals (CRS) specified in LTE for downlink channel estimation and mobility measurement albeit with low duty cycle.

Unlike in LTE, LTE-M only supports semi-static link adaptation. For power control, timing and uplink channel estimation the LTE-M eNodeB uses long term channel properties and terminal specific demodulation reference signals transmitted along with uplink data. Moreover, LTE-M devices support the following measurements specified in the 3GPP specification 3GPP TS 36.214 [13]:

- Reference Signal Received Power (RSRP) based on CRS
- Reference Signal Received Quality (RSRQ) based on Evolved UMTS Terrestrial Radio Access (E-UTRA) Received Signal Strength Indicator (RSSI)

Inter Radio Access Technology (RAT) measurements to the Global System for Mobile communications (GSM) and CDMA networks are not supported.

The eNodeB can make use of the following measurements for LTE-M specified in 3GPP TS 36.214 [13]:

- Received interference power in the uplink
- Thermal noise power
- Timing advance
- Angle of Arrival

2.3 LTE-M Medium Access Control (MAC) Layer

This section specifies the MAC layer of the LTE-M system. It specifies the MAC entities at the UE and eNodeB, as well as the MAC architecture from a functional point of view. This specification shares several aspects with 3GPP TS 36.321 [14] in order to maximize the compatibility and reusability of LTE-defined entities and protocols in LTE-M.

2.3.1 General MAC architecture

In LTE-M there are two MAC entities, one at the UE and another at the E-UTRAN, equivalently to LTE [14]. The main purpose of the MAC is to map logical channels, offered to the RLC layer, to transport channels offered by the PHY layer.

In addition, the MAC layer provides the following services to upper layers:

- Data transfer;
- Radio resource allocation.

The MAC expects the following services from the physical layer:

- Data transfer services;
- Signalling of HARQ feedback;
- Signalling of Scheduling Request;
- Measurements (e.g. Channel Quality Indication (CQI).

2.3.1.1 Uplink channels and channel mapping

The MAC layer provides data transfer services on Logical Channels to the RLC. Similarly to the LTE uplink, the following Logical Channels are defined in LTE-M uplink:

- MTC Common Control Channel (MCCCH), used to transfer control information between the UE and the Network when no RRC connection exists;
- MTC Dedicated Control Channel (MDCCH), a dedicated control channel between UE and Network;
- MTC Dedicated Traffic Channel (MDTCH), carrying user data.

LTE-M also provides data transfer through the Random Access Channel (RACH). To that end, the following Logical Channel is defined:

- MTC Random Access Traffic Channel (MRATCH), carrying small packet user data.

The uplink mapping between Logical and Transport Channels is shown in Figure 2-10.

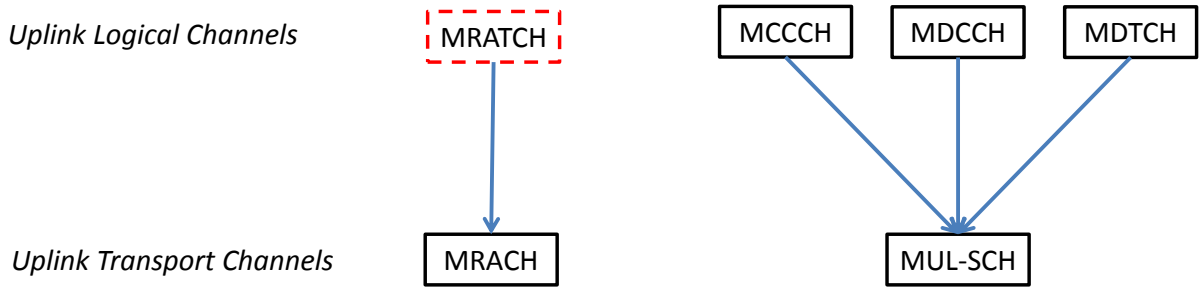


Figure 2-10: Uplink mapping of Logical to Transport Channels in LTE-M.

2.3.1.2 Downlink channels and channel mapping

In the downlink, similarly to LTE, the following Logical Channels are defined both in LTE-M:

- MCCCH, MDCCH, MDTCH, which have the same functionality as their uplink counterparts;
- MTC Broadcast Control Channel (MBCCH), carrying broadcast information for access to the system;
- MTC Multicast Control Channel (MMCCH), carrying control information for point-to-multipoint connections;
- MTC Multicast Traffic Channel (MMTCH), carrying data for point-to-multipoint connections;
- MTC Paging Control Channel (MPCCH), carrying paging information which allows the network to page a device.

The downlink mapping between Logical and Transport Channels is shown in Figure 2-11.

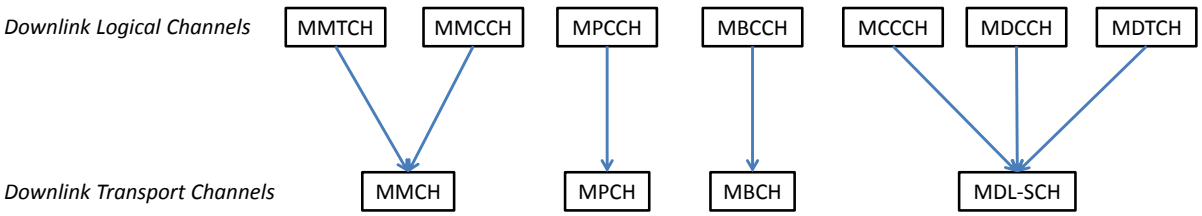


Figure 2-11: Downlink mapping of Logical to Transport Channels in LTE-M.

2.3.1.3 MAC functions

The following functions are supported by MAC:

- Mapping between logical channels and transport channels;
- Multiplexing of MAC SDUs from one or different logical channels onto transport blocks (TB) to be delivered to the physical layer on transport channels;
- Demultiplexing of MAC SDUs from one or different logical channels from transport blocks (TB) delivered from the physical layer on transport channels;
- Scheduling information reporting;
- Error correction through HARQ;
- Priority handling between UEs by means of dynamic scheduling;
- Priority handling between logical channels of one UE;
- Logical Channel prioritisation;
- Transport format selection.

2.3.2 Random access procedure

The role of Random Access (RA) in LTE-M is different from LTE. In LTE random access is mainly used to place requests for transport channel allocations. In LTE-M it is also used to transmit small data packets. This random access transmission, corresponding to the uplink, is particularly effective when either the UE or the network may not request PUSCH allocations for a batch of packets. In this case it becomes inefficient to request resources for a single packet, and LTE-M provides the capability to simply append it to a RA procedure.

The RA procedure described here may be initiated either by MAC itself or upon reception of a PDCCH order. This covers device-initiated RA (e.g. alert message) and network-initiated RA (e.g. spontaneous polling). It is assumed that the device knows the available set of PRACH resources as well as other relevant parameters described in 3GPP TS 36.321 [14], Section 5.1.1. The UE follows the RA resource selection procedure and transmits a RA preamble, optionally followed by a data payload. In case a data payload is appended, and no additional data transfer capability is needed, the RA request will simply request no bandwidth.

The RA preamble is CDMA-encoded in order to enable preamble detection in the event of a collision. Furthermore, both preamble and data payload are multiplied by a random coefficient in a finite field GF^q. CDMA encoding employs a short code which is determined through a fixed mapping using the Random Access Radio Network Temporary Identifier (RA-RNTI). The processing gain of the CDMA code will be determined by the network. The goal of adding CDMA encoding to the RA request is to enable collision recovery at the receiver, as described in EXALTED Deliverable D3.3, Section 4.2.2 [1].

Once the RA preamble is transmitted, the UE monitors the PDCCH for a Random Access Response.

The MAC carries out contention resolution similarly to [14]. Based on a timeout and feedback from the PDCCH, a contention phase will be considered successful and unsuccessful. If unsuccessful, a new RA procedure will be initiated after a random back-off period between 0 and the back-off Parameter Value. If contention resolution is successful, the UE flushes the HARQ buffer and discards all RA signalling.

2.3.3 HARQ operation

HARQ schemes combine Forward Error Correction (FEC) and Automatic Repeat reQuest (ARQ) to achieve more reliable transmissions in a communication system. In the LTE-M, an

adaptive HARQ scheme [15] is designed to overcome the two disadvantages of long-delay and throughput loss in the fixed-rate HARQ schemes.

The LTE-M adaptive HARQ scheme works on the basis of the original turbo channel code in the LTE, which has an initial coding rate of 1/3. With the aid of puncturing operations, this 1/3-rate turbo encoded rate may match any coding rate higher than this. Rate matching in section 2.2.5 is employed before the first transmission.

Using an appropriate coding rate selected from the LUT as discussed in section 2.2.5, the first transmission may fail with low probability. The turbo decoder same as the one in the LTE performs iterative decoding for a 1/3-rate packet at the receiver. Therefore, those parity bits punctured during the rate matching at the transmitter may be padded by zeros before the decoding. If the turbo decoding succeeds, an ACK message is sent back to the transmitter. Otherwise, if the transmitter times out, it will send the whole 1/3-turbo encoded packet for the second transmission. At the receiver, chase combining will be applied to those repeated bits during both transmissions. Furthermore, incremental redundancy is also applicable for those parity bits transmitted during the second transmissions but not in the first transmission. Then, the turbo decoding is activated again. If the CRC is satisfied after the decoding, an ACK message is fed back to stop the transmission of this packet. If it fails again, the packet will be discarded. The flow chart of the LTE-M HARQ may be illustrated in Figure 2-12.

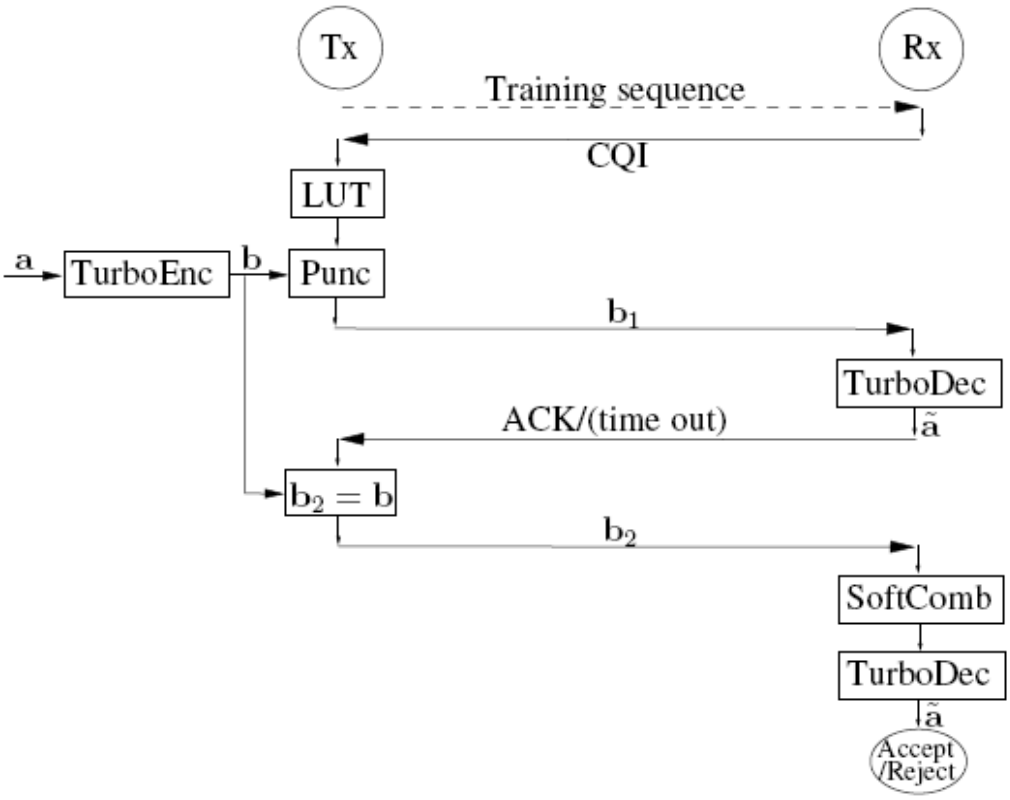


Figure 2-12: The flow chart of LTE-M optimal HARQ process.

Only two transmissions are allowed in the LTE-M HARQ. However, a large number of packets may be successfully delivered during the first transmission. Therefore, the average number of retransmission is between one and two. Meanwhile, simulation results in the project report D3.3 [1] have shown that the throughput of the LTE-M adaptive HARQ is higher than those of classical turbo coded HARQ schemes, while maintaining similar packet loss ratio.

Additionally, fountain codes become a best candidate for an adaptive HARQ scheme due to their rateless property. In the Appendix A2, conventional Raptor codes are improved to assist HARQ schemes for transmissions of short block lengths. Simulation results in A2 demonstrate that the proposed HARQ scheme based on improved Raptor codes may achieve a significantly low PLR, despite a low decoding complexity.

2.3.4 Exchange of scheduling information between LTE-M device and eNodeB

Generally, the assignment of PMDSCH and PMUSCH radio resources is carried out in the LTE-M enabled eNodeB, and the way how this is done is vendor specific, i.e. not part of the specification. Different scheduling algorithms have been investigated and presented in the EXALTED report D3.3 [1]:

- QoS class based scheduling (see [1] section 4.1.2);
- Access Grant Time Interval (AGTI) Scheduling Framework (see [1] section 4.1.3);
- Scheduling algorithm for heterogeneous traffics (see [1] section 4.1.4).

However the protocol defining how scheduled resources are requested and granted is subject to standardisation. LTE-M foresees two mechanisms, namely *dynamic resource assignment* and *semi-static resource assignment*. The difference is shown in the schematic view in Figure 2-13.

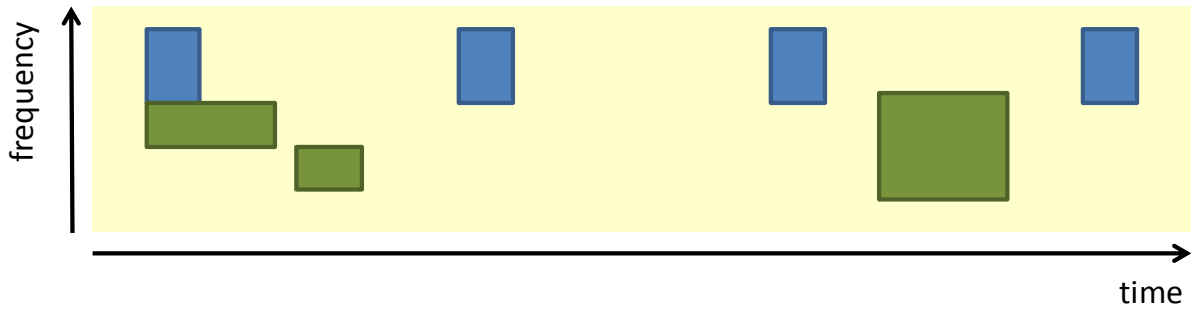


Figure 2-13: Semi-statically (blue) and dynamically (green) scheduled radio resources.

Dynamically scheduled resources shown in green colour in Figure 2-13 are valid for one single message only and are the appropriate and required solution for inhomogeneous and event-driven traffic flows and for fast changing radio propagation conditions, which demand a continuous adaptation of link parameters like modulation scheme and code rate. This can happen if the LTE-M device is mobile. LTE-M adopts the mechanism in the 3GPP specification 36.321 [14], but due to the drawback that the signalling overhead is relatively big compared with the small payload of one single message, in particular if messages of a huge number of devices must be scheduled, dynamic resource assignment is not the preferred operation mode.

Persistent or semi-static resource assignment shown in blue colour in Figure 2-13 is valid for more than just one single message. Radio resources, i.e. resource blocks, are granted on a regular time grid. During the validity period of this assignment no additional scheduling information needs to be exchanged between eNodeB and LTE-M device. It is the preferred operation mode for time-driven traffic flows and static propagation conditions that typically occur if the LTE-M device is installed at a fixed location. In contrast to the LTE specification [14], in LTE-M semi-static resource assignment is always enabled. It can be applied in both uplink and downlink.

After the LTE-M device has requested semi-statically assigned resources, the eNodeB must decide the following parameters and exchange them with the device:

- Sequential numbers of the resource blocks (size of allocation and its position within the radio frame);
- Number of empty subframes between two transmissions (frequency of transmissions);
- Overall number of transmissions for the assignment (validity period).

2.3.5 Discontinuous reception (DRX)

Discontinuous reception (DRX) is introduced to allow the UE to periodically switch off the receiver circuitry and save battery power. LTE supports long DRX of 2560 subframes (2.56 seconds) for IDLE mode UEs. To allow further power saving LTE-M longer DRX should be supported. UEs whilst in DRX mode are not expected to send Precoding Matrix Indicator (PMI), Channel Quality Indicator (CQI) and Rank Indicator (RI) and hence will also result in uplink power saving (see section 2.5.3. in the EXALTED report D3.3 [1]).

An efficient paging procedure allows the terminal to sleep with no receiver processing most of the time and to briefly wake up at predefined time intervals to monitor paging information from the network. DRX cycle for paging is shown on Figure 2-14.

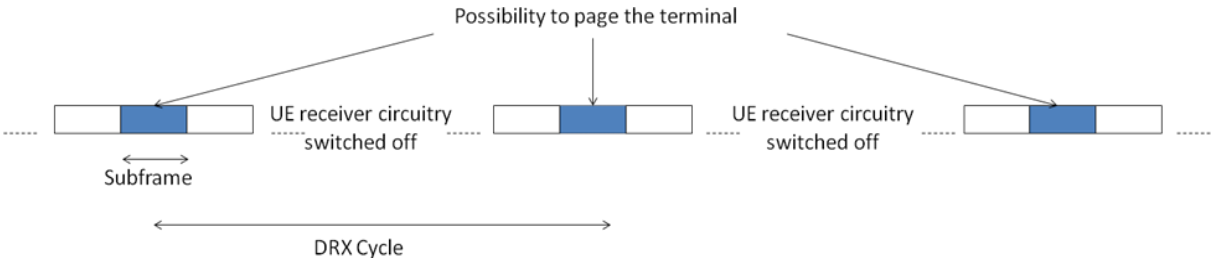


Figure 2-14: DRX for paging.

Monitoring of the paging channel uses significant part of the energy in the IDLE mode. To cope with this issue, for LTE-M devices it should be possible to increase the paging cycle (DRX cycle). The LTE-M device determines the paging DRX cycle and other relevant paging parameters by processing the serving cell paging parameters broadcast by the serving cell, and LTE-M device will periodically wake up on every paging occasion to check for paging messages. In order to increase the paging cycle for specific LTE-M device (DRX cycle) the following algorithm is proposed (for more details see section 5.2.2. in D3.3 [1]):

- Introduce Paging flag (factor), and multiply DRX cycle by this factor;
- Paging factor should be defined in HSS per device;
- During attach procedure factor will be downloaded on MME and delivered to LTE-M device.

New algorithm allows extension of DRX cycle based on multiplication of DRX cycle with predefined factor which is saved on HSS as a part of device subscription, and is delivered to device during registration to the network.

The same idea is used for optimization of the number of necessary radio measurements, and the possibility to allow the LTE-M terminals to control the timing and to perform the required measurements only when needed. Frequency of radio measurements depends on DRX cycle, so it is proposed to introduce mobility flag (factor), and multiple DRX cycle by this factor. Mobility factor should be defined in HSS per device as the subscription parameter, and should be delivered to LTE-M device (for more details see section 5.2.2. in [1]). This factor could be different for:

- RRC_IDLE;
- RRC_CONNECTED.

2.3.6 LTE-M MAC functions and services that are equivalent as in LTE

A number of functions defined in the 3GPP specification 36.321 [14] for LTE MAC have been adopted without further change for the proposed LTE-M system. These are listed in the following:

- Maintenance of Uplink time alignment;
- DL-SCH data transfer;
- UL-SCH data transfer;
- MAC reconfiguration;
- MAC Reset;
- Semi-Persistent Scheduling;
- Handling of unknown, unforeseen and erroneous protocol data.

2.4 LTE-M Radio Link Control (RLC)

The protocol stacks of LTE-M including the RLC sub-layer are shown in Figure 2-2 (user plane option 1), Figure 2-3 (user plane option 2), and Figure 2-4 (control plane).

The user plane of the LTE-M RLC sub-layer is always operating in transparent mode as defined in the 3GPP document 36.222 [16]. Thus, it is completely bypassed. This means that it doesn't have any functionality and doesn't cause any overhead.

In the control plane the LTE-M RLC sub-layer is essential in order to ensure that Non-Access Stratum (NAS) messages are protected in the acknowledged mode as defined in 3GPP document 36.222 [16]. LTE-M adopts the existing LTE specification.

2.5 LTE-M Packet Data Convergence Protocol (PDCP)

LTE-M re-uses the integrity protection of LTE to detect unintended packet replacement or insertion for data on signalling radio bearer. Ciphering protects message confidentiality and is applied to both data and signalling bearer.

2.5.1 Services provided to other layers

Generally, the PDCP provides its services to the RRC and user plane upper layers at the terminal (see 3GPP specification 36.323 [17]). In LTE-M the PDCP supports following functions.

- Transfer of user plane data
- Transfer of control plane data
- Header compression in LTE-M configuration when the Address Mapping Function (AMF, see section 2.5.3) is not used
- Ciphering;
- Integrity protection.

The maximum supported size of a PDCP Service Data Unit (SDU) for LTE-M is significantly less (e.g. 2047 octets) compared to 8188 octets in LTE because LTE-M is not expected to support RLC Acknowledged Mode (AM) and RLC Unacknowledged Mode (UM) user plane bearers.

2.5.2 PDCP procedures

Unlike LTE, LTE-M is a system optimised for small packets, user plane data are not expected to be segmented to several IP packets, and lossless handover is not expected to be supported. Hence unlike PDCP in LTE, for PDCP in LTE-M sequence numbers are not required. Other LTE-M PDCP procedure for NAS messages remain the same as LTE.

2.5.3 Address mapping function (AMF)

PDCP in LTE-M includes an address mapping function that maps in co-ordination with the MAC the IP address (128 bits) to the Cell Radio Network Temporary Identifier (C-RNTI) with 16 bits, which is an identity uniquely identifying the terminal and the RRC connection. AMF allows for communication without source and destination IP addresses, thus avoiding this overhead. The associated header checksum is explicitly transmitted over the radio interface. The principle of AMF is illustrated in Figure 2-15.

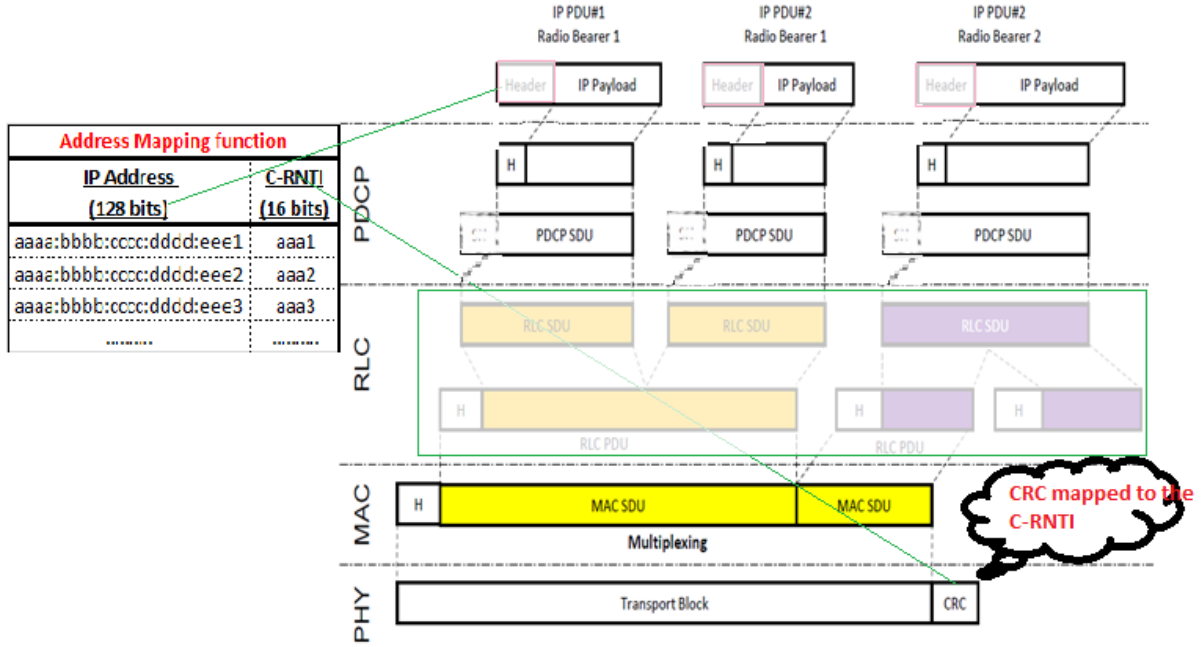


Figure 2-15: Principle of the Address Mapping Function (AMF).

2.6 LTE-M Radio Resource Control (RRC) Protocol

The protocol stack of LTE-M including the RRC sub-layer is shown in Figure 2-4 (LTE-M) control plane protocol stack. In the access stratum protocol stack, the RRC protocol is also known as 'Layer 3'. It is the main controlling function in the access stratum, being responsible for establishing the radio bearers and configuring all the lower layers using RRC signalling between the eNodeB and the UE. The RRC protocol supports the transfer of common NAS information (i.e. NAS information which is applicable to all UEs) as well as dedicated NAS information (which is applicable only to a specific UE).

2.6.1 General RRC architecture

Regarding UE states in LTE, which is also adopted in LTE-M, a UE could be in RRC_CONNECTED when a RRC connection has been established. If no RRC connection is established, the UE is in RRC_IDLE state. The RRC states can further be characterised as follows 3GPP TS 36.331 [18]:

- **RRC_IDLE**:
  - A UE specific DRX may be configured by upper layers;
  - UE controlled mobility;
  - The UE:
    - Monitors a Paging channel to detect incoming calls, system information change, for ETWS capable UEs, ETWS notification, and for CMAS capable UEs, CMAS notification;
    - Performs neighbouring cell measurements and cell (re-)selection;
    - Acquires system information;
    - Performs logging of available measurements together with location and time for logged measurement configured UEs.

- **RRC_CONNECTED**:
  - Transfer of unicast data to/from UE;
  - At lower layers, the UE may be configured with a UE specific DRX;
  - For UEs supporting CA, use of one or more SCells, aggregated with the PCell, for increased bandwidth;
  - The UE:
    - Monitors a Paging channel and/ or System Information Block Type 1 contents to detect system information change, for ETWS capable UEs, ETWS notification, and for CMAS capable UEs, CMAS notification;
    - Monitors control channels associated with the shared data channel to determine if data is scheduled for it;
    - Provides channel quality and feedback information;
    - Performs neighbouring cell measurements and measurement reporting;
    - Acquires system information.

2.6.2 Services provided to other layers

The RRC protocol offers the following services to upper layers 3GPP TS 36.331 [18]:
- Broadcast of common control information;
- Notification of UEs in RRC_IDLE, e.g. about a terminating call, for ETWS, for CMAS;
- Transfer of dedicated control information, i.e. information for one specific UE.

In brief, the following are the main services that RRC expects from lower layers:
- PDCP: integrity protection and ciphering;
- RLC: reliable and in-sequence transfer of information, without introducing duplicates and with support for segmentation and concatenation.

2.6.3 RRC procedures and data

RRC performs the function of LTE-M specific broadcast, paging of LTE-M devices, RRC connection management, radio bearer control, and device measurement reporting and control. Proposals for improvements of RRC layer protocols for LTE-M are presented in section 5 in the EXALTED report D3.3 [1]. The basic idea to register information about the terminals, which is accessible by the eNodeB, is presented in detail (section 5.1). To ensure a proper segmentation between LTE and LTE-M devices specific registering procedures are analysed. This is important for efficiently serving the specific needs of the high-data rate LTE world as well as the low data rate LTE-M world.

A lot of EXALTED solutions benefit from the basic idea to register information about the LTE-M devices and M2M gateways and their respective capabilities at the HSS. Examples are the differentiation whether or not a device supports IP, whether a device is mobile or installed at a fixed location, which type of traffic it causes and to which delay class it belongs. Such information is primarily exploited to simplify the RRC protocols. Knowledge of these extra

information in comparison to the LTE, allows optimization of the paging procedure (monitoring of the paging channel) in LTE-M and optimization of the mobility support (section 5 [1]). Also these information can be used in improvements of scheduling techniques (section 4.1 [1]). MAC scheduling of the devices can be optimized based on their priority, and three different proposals are presented showing significant gains with respect to the overall QoS, in particular if the number of devices is big.

2.7 LTE-M Device Classes and Capabilities

The Device Domain includes various types of devices, and therefore the treatment by the access and core network depends on device's characteristics, required access network conditions and type of the application they run. Some additions to the existing specifications are introduced, but at the same time a high level of backward compatibility with LTE is achieved.

2.7.1 Summary of differentiators for LTE-M device classes

The concept of LTE-M device classes has already been introduced in deliverable D3.3 [1]. Capabilities additional to the ones already existing in 3GPP document TS 36.306 [19] are presented, and the mechanism for conveying them to the network is the same as described in 3GPP document TS 36.331 [18]. For the sake of completeness of this Deliverable, details already presented in [1] are elaborated again, with introduction of an additional device category to the ones described in [1] and [19].

Among the components of the EXALTED system, device classes and capabilities are applicable to those devices that use LTE-M interface, namely LTE-M Devices and M2M Gateways. Devices in capillary networks are "hidden" behind Gateways from the network's perspective, and it is a Gateway's role to report the capabilities in a way that reflects the features of a capillary network behind it.

The main purpose of the usage of device capabilities is to notify the core network of the features and functionalities that a particular device (UE) has, so that a proper treatment for that device on the access network is granted.

Taking into account types of traffic, applications, protocol stack running on the EXALTED components etc., some features of devices are recognized as the main candidate capabilities. These are summarized in the Table 2-1.

Table 2-1: Device capabilities in EXALTED.

Capability	Type	Comment
IP support	Boolean (true/false)	This capability should reflect the existence of simple non-IP LTE-M devices. The eNodeB to which they are attached performs IP (PDP) connection termination on their behalf.
Mobility support	Boolean (true/false)	Mobility management is still under discussion in EXALTED, since it introduces several negative impacts on signalling and complexity of the system, but in case it is needed, a simple capability denoting whether the IP session continuity should be retained in case of a handover should reflect it.
Ability to perform channel measurements and reporting	Boolean (true/false)	Devices having this capability can be polled for the purpose of monitoring etc.

Power restricted	Boolean (true/false)	Power restricted devices should generate and receive less traffic, and often be in a sleep mode.
Type of traffic	e.g. Integer 1 – time driven 2 – event driven	The event driven traffic is not expected on regular time basis by the network.

2.7.2 Definition of LTE-M device classes

According to TS 36.331 [18], UE informs the Evolved Universal Terrestrial Radio Access Network (E-UTRAN) of its capabilities through the *UE capability transfer* procedure. This procedure includes two basic messages: *UECapabilityEnquiry*, sent from the network to a UE in order to request the transfer of UE radio access capabilities, and *UECapabilityInformation*, containing the reply with capabilities, sent from UE to the network. The procedure is depicted on Figure 2-16 (source: [18]).

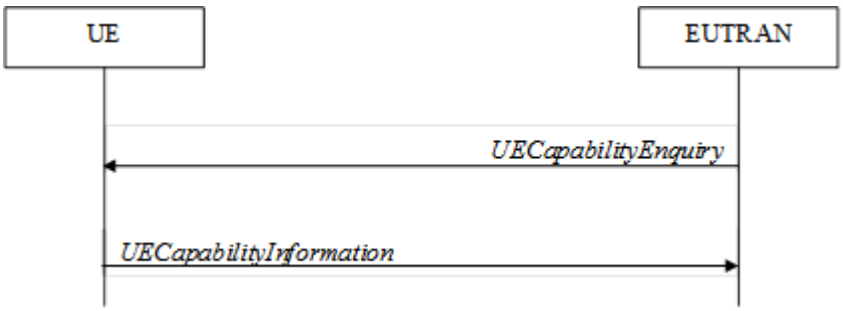


Figure 2-16: UE capability transfer.

The messages are exchanged through the Dedicated Control Channel (DCCH). If the UE has changed its radio access capabilities, necessary Non-Access Stratum (NAS) procedures must be initiated by higher layers, and that would result in the update of UE radio access capabilities using a new RRC connection [18].

After the procedure is initiated by E-UTRAN, the UE must set the content of *UECapabilityInformation* message, and submit it to lower layers for transmission. The enquiry message specifies the radio access technology, for which the information is required (GPRS, LTE, CDMA2000, etc.), and the reply message contains specific *fields* (i.e. variable values) depending on the technology required.

The relevant fields in case of E-UTRAN release 8 (LTE) are presented in Figure 2-17 in Abstract Syntax Notation (ASN.1). Further extensions to release 9 and 10 (LTE-A) are given in the same specification [18].

Among the fields presented, there is a field which can be of particular interest for EXALTED, namely *ue-Category*, and also *ue-Category-v1020* (for LTE-A categories), which specifies the category of UE in respect to its supporting physical uplink and downlink capability. Release 8 assumes 5 categories, while release 10 introduces 3 more. Depending on the device and a particular use case, this category should be properly set. The list of relevant parameters describing each of the categories is given in TS 36.306 [19].

An additional category for LTE-M, that reflects the devices with very low requirements, is introduced. In order to fit to the current numerical assignment, this category is given the number “0”, and it is known as *Category 0*. Table 2-2 summarizes the parameters for this category.

```
UE-EUTRA-Capability ::=          SEQUENCE {
    accessStratumRelease                AccessStratumRelease,
    ue-Category                         INTEGER (1..5),
    pdcp-Parameters                     PDCP-Parameters,
    phyLayerParameters                  PhyLayerParameters,
    rf-Parameters                       RF-Parameters,
    measParameters                      MeasParameters,
    featureGroupIndicators              BIT STRING (SIZE (32))                  OPTIONAL,
    interRAT-Parameters             SEQUENCE {
        utraFDD                             IRAT-ParametersUTRA-FDD             OPTIONAL,
        utraTDD128                          IRAT-ParametersUTRA-TDD128          OPTIONAL,
        utraTDD384                          IRAT-ParametersUTRA-TDD384          OPTIONAL,
        utraTDD768                          IRAT-ParametersUTRA-TDD768          OPTIONAL,
        geran                               IRAT-ParametersGERAN                OPTIONAL,
        cdma2000-HRPD                       IRAT-ParametersCDMA2000-HRPD        OPTIONAL,
        cdma2000-1xRTT                      IRAT-ParametersCDMA2000-1XRTT       OPTIONAL
    },
    nonCriticalExtension                UE-EUTRA-Capability-v920-IEs            OPTIONAL
}
```

Figure 2-17: UE-EUTRA-Capability fields.

Total buffer size is from TS 36.213 [20] Table 7.1.7.2.1-1, assuming 3 MHz (15 resource blocks). Remaining values assume Peak data rate of 1 Mbps in uplink and downlink. Total number of Soft channel bits is an approximation.

Table 2-2: Category 0 parameters.

Parameter	Value
Maximum number of DL-SCH transport block bits received within a TTI	1064
Total number of soft channel bits	~25k
Maximum number of supported layers for spatial multiplexing in DL	1
Maximum number of bits of an UL-SCH transport block transmitted within a TTI	1064
Support for 64QAM in UL	No
Total layer 2 buffer size [bytes]	30 000

Furthermore, in LTE-M it is envisaged to introduce additional fields in *UE-EUTRA-Capability* or *UE-EUTRA-Capability-v1020-IEs*, or even a completely new set of fields (e.g. *UE-EUTRA-M-Capability*), which will address capabilities needed for the LTE-M system and EXALTED use cases.

One possibility to map capabilities listed in Table 2-1 to fields, following the ASN.1 notation and specification [18], is presented in Table 2-3.

Table 2-3: A proposal for LTE-M capabilities.

Capability	Type	ASN.1 name
IP support	Boolean	ipSupport
Mobility support	Boolean	mobilitySupport
Ability to perform channel measurements and reporting	Boolean	channelMonitoring
Power restricted	Boolean	powerRestricted
Type of traffic	Integer	typeOfTraffic

In case of the EXALTED architecture, besides the basic differentiation between an LTE-M Device and an M2M Gateway, one of the most relevant capabilities that can be used for

further differentiation is the IP support, since it implies major impacts to device connectivity and the ways it is reachable. Therefore, the following device classes are suggested:

- IP LTE-M device;
- Non-IP LTE-M device;
- M2M Gateway.

Table 2-4 summarizes the possible values of LTE-M capability fields for those device classes.

Table 2-4: Proposed LTE-M device classes and capabilities.

Capability	IP LTE-M Device	Non-IP LTE-M Device	M2M Gateway
IP support	true	false	true
Mobility support	true	true	true
Ability to perform channel measurements and reporting	false	false	true
Power restricted	true	true	false
Type of traffic	*application specific*	*application specific*	*application specific*

In this way the LTE/LTE-M core network is able to treat IP devices, non-IP devices and M2M Gateways differently, while maintaining the strong backward compatibility with LTE specification.

2.8 LTE-M Bearer- and QoS Classes

LTE-M follows the same philosophy as in LTE and distinguishes between two bearer classes for Guaranteed Bit Rate (GBR) services and non-GBR services. Each bearer class consists of several QoS classes that are characterized by technical requirements such as message delay and message loss rate. In LTE, nine different classes are available. In order to re-use the existing addressing scheme the number of classes is adopted, i.e. there are also nine different classes in LTE-M. If a certain terminal is associated with, e.g., QoS class 3, and it is attached as an LTE UE to the network, the LTE parameters of class 3 are valid. If it is attached as LTE-M device, the corresponding LTE-M parameters are decisive. However, the addressing and the order of priority of the QoS class are identical in both cases.

Important parameters are, e.g., the tolerable message delay, message size, message loss rate and eventually the guaranteed bit rate. Specific values for these requirements are not part of this specification, but under the responsibility of the operator or service provider. The entries in Table 2-5 are *examples* and reflect the requirements for the EXALTED use cases Smart Metering and Monitoring (SMM) and eHealth introduced in the project report D2.1 [5]. A characteristic of many M2M applications is the need to maintain a certain delay-throughput product, which means that maximum throughput and maximum delay cannot be applied at the same time. The impact of the definition of QoS classes on the number of LTE-M devices that can be supported has been intensively studied in EXALTED, and the results are presented in section 5.

Table 2-5: Examples for QoS classes in LTE-M.

QoS Class Indicator (QCI)	Bearer Class	Priority	Max. Message delay	Guaranteed bit rate	Services
1	GBR	2	1 s	25 Kbit/s	**High demanding SMM**: Use cases which are both sensitive to delays and throughput (for instance actuators for monitoring stock)
2	GBR	4	1 s	10 Kbit/s	**High demanding SMM**
3	GBR	3	10 s	50 Kbit/s	**Bandwidth demanding SMM**: Applications monitoring smooth signals but demanding higher bandwidth (for example bridge sensors monitoring several samples at the same time)
4	GBR	5	10 s	25 Kbit/s	**Bandwidth demanding SMM, eHealth**: Vital signs sensing / monitoring
5	Non-GBR	1	100 ms	-	**Time demanding SMM**: (Nodes monitoring critical variables that need to be sent in precise moments (for instance presence sensors) but without sending high amounts of data)
6	Non-GBR	6	1 minute	-	**Low demanding SMM**: Monitoring simple variables varying slowly in time (for example temperature)
7	Non-GBR	7	10 minutes	-	**Low demanding SMM**
8	Non-GBR	8	1 hour	-	**Low demanding SMM**
9	Non-GBR	9	1 day	-	**Low demanding SMM**

3. Evaluation Methodology

The evaluation methodology described in this section aims to define a common framework towards assessing the EXALTED innovations regarding the LTE-M access network. Common evaluation assumptions, parameters, models and key performance indicators (KPIs) are identified to ensure the comparability of the technical innovations and serve as a guideline for assessing M2M solutions over cellular network, outside the EXALTED project as well. Two evaluation scenarios are defined, namely:

- **Scenario 1: Low complexity and energy efficient M2M communications**: This scenario is linked to the energy efficiency objective and tailored for LTE-M devices that are extremely power limited and only have a restricted function volume.
- **Scenario 2: Supporting a large number of LTE-M devices with heterogeneous requirements and capabilities**: This second scenario covers the spectrum efficiency objective and assumes a huge number of devices and also a big variety of different types of devices.

They are employed to examine the performance of the LTE-M system with respect to the predetermined objectives and system requirements. In Figure 3-1, the evaluation methodology is illustrated, which includes three main parts:

- The evaluation models (e.g. traffic and channel models);
- The evaluation tools (e.g. Link and system level simulations) and the KPIs;
- The performance analysis and the evaluation of the results.

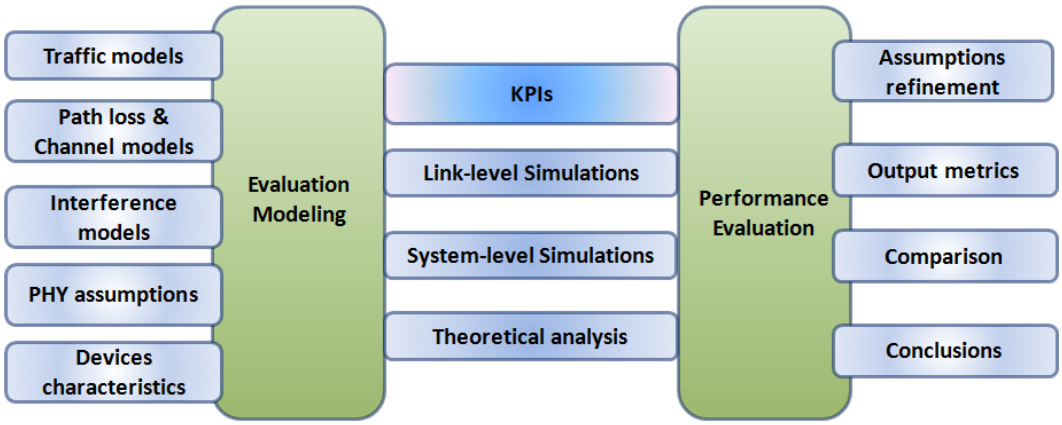


Figure 3-1: The LTE-M evaluation methodology.

At first, the commonalities of both scenarios are presented, afterward the respective particularities.

3.1 Topology and channel models

Although each evaluation scenario focuses on different objectives, for both of them the same cellular layout is assumed as depicted in Figure 3-2. The yellow circles indicate the seven positions of three eNodeBs serving three 120° cells illustrated in green, red and blue. System simulations use a wrap-around, i.e. propagation paths leaving the layout, re-enter it at the point 180° on opposite side. The distance between the eNodeB sites is 500m.

LTE-M devices and LTE UEs are dropped at random positions within the layout. It is assumed that each LTE-M device or UE is connected with the eNodeB that exhibits the best channel quality. This evaluation scenario only consists of stationary devices. The frequency re-use factor is 1, i.e. the same frequency band is used in all cells.

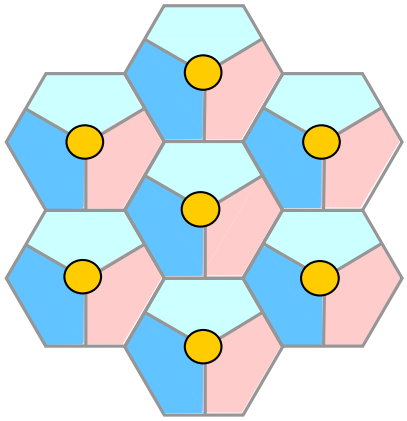


Figure 3-2: Cellular layout.

Moreover, the propagation and channel modelling of both scenarios consists of the following components

- AWGN channel model: It is required to perform link level simulations to be used as link abstraction in system level simulations. Density of thermal noise is set to -174dBm/Hz.
- Additionally, AWGN can be combined with a Rayleigh fading channel model. As a special case, the channel gain from the eNodeB to each antenna of the LTE-M device is Rayleigh distributed, with each path having the same average power.
- Fast fading is modelled as a 6-tap delay line based on the Typical Urban scenario (3GPP TS 45.005 [21], section C.3.3) with a mobile speed equal to 3km/h. Additionally, channel parameters tailored for a typical M2M propagation scenario have been proposed in the 3GPP contribution R1-130463 [22]. However, for the evaluation it was not yet available.
- For system level simulation, two path loss models have been adopted. They aim at covering at best all use cases envisaged in EXALTED, i.e. Smart Metering and Monitoring (SMM), Intelligent Transport System (ITS) and e-health (D2.1 [5]). Environmental monitoring would require a macrocell channel model, while ITS is covered by microcell and macrocell models.
  - The microcell channel model is based on the 3GPP Spatial Channel Model (SCM) described in 3GPP TR 25.996 [23] and fits an urban microcell environment. We assume Line-of-Sight (LOS) conditions. The SCM includes path loss and shadowing parameters as well. According to this, the path loss PL is $PL(d) = -35.4 + 26\log_{10}(d) + 20\log_{10}(f_c)$, where d [m] is the distance between LTE-M device and eNodeB and $f_c = 800\text{MHz}$ is the carrier frequency. The log-normal shadowing standard deviation is 4dB.
  - The macrocell channel model is the one proposed in 3GPP TR 36.814 [24], section A.2.1.1 as well as in 3GPP TR 36.888 [4]. The agreed model is $PL(d) = 15.3 + 37.6 * \log_{10}(d)$, where d [m] is the distance between LTE-M device and eNodeB and $f_c = 2\text{GHz}$ is the carrier frequency. A penetration loss of 20dB can also be added if relevant. The generic model for any carrier frequency is $PL(d) = -53.97 + 37.6 * \log_{10}(d) + 21 * \log_{10}(f_c) + WL$, d [m], f_c [MHz] and $WL = 20\text{dB}$ for penetration loss of external walls if needed.
- Interference model: white interference coming from surrounding cells, which transmit with maximum power.

3.2 Scenario 1: Low Complexity and Energy Efficient M2M Communications

The purpose of this scenario is the performance evaluation of LTE-M in the important case where the LTE-M devices, e.g. sensors or actuators, are extremely power limited and have only a restricted function volume. The considered devices are very cheap, and it is not foreseen to replace their batteries. Therefore, the battery lifetime is the most important metric related with this scenario. Once deployed in the field, they act autonomously according to a certain function defined in advance as long as their battery permits. It is assumed that the devices are located at a fixed position. All considered devices are widely spread in a joint LTE/LTE-M radio cell. It may happen that some devices have to cope permanently with bad radio propagation conditions if they are located in a shadowed area, e.g. behind a wall. Among the applications considered in EXALTED and defined in the public report D2.1 [5], this scenario is primarily linked with monitoring use cases. Examples are sensors measuring and reporting the air pollution or fine dust concentration in a city, stability of buildings, temperature and humidity in a wide area, ground motions in a region endangered by earthquake, and volcanism activities. The data rates are very small and the reporting is based on a regular time grid, e.g. each sensor transmits one message per hour.

The evaluation of the proposed solutions with respect to the overall KPIs and further subordinated metrics (see Table 1-1) can be found in the following sections:

- Co-existence with LTE in section 4;
- Provision of wide area coverage in section 6.2;
- Energy savings to enable a long battery lifetime in section 7.2;
- Cost efficiency in section 8.

Table 3-1 summarizes the simulation parameters and assumptions that are commonly used in this evaluation scenario as long as not indicated differently. Other parameters like data rates or number of supported devices are irrelevant in this scenario.

Table 3-1: Common simulation parameters and assumptions.

Parameter	Value
Distance between eNodeBs	500 m
eNodeB antenna parameters	According to 3GPP TR 36.814 [24], Table A2.1.1-2
eNodeB transmit power	43 dBm
Maximal LTE UE transmit power	23 dBm
Maximal LTE-M device transmit power	17 dBm
LTE-M device function volume	1 isotropic Rx and Tx antenna, no IP support, able to perform CSI reporting
CSI estimation accuracy	Perfect CSI available
Synchronization accuracy	Perfect synchronization
Carrier frequency	800 MHz

3.2.1 Traffic models

The evaluation of Scenario 1 is based on one of the following optional traffic models.

- Greedy source model: This model is commonly used for link level simulations. In order to minimize the simulation time and to get statistically stable results, all available radio resources are permanently assigned to the one single device under investigation. For system level simulations that typically consider multiple LTE-M devices, the frequency resources are split in equal parts and each sub-band is assigned to one device per cell.

- Time-driven short messages: In this model it is assumed that each device transmits a short message consisting of maximal 1064 bits, which is the biggest possible transport block size in LTE-M. This transmission occupies one single subframe. The transmissions happen in fixed time intervals. The default value of this time interval is 10s without loss of generality, unless it is not indicated differently. In a scenario with multiple devices, the transmission of the first message from each device happens at a random point of time within the time interval. As an example, the regular transmissions over time for three LTE-M devices are shown in
- Figure **3-3**. The model is required for system level simulations.

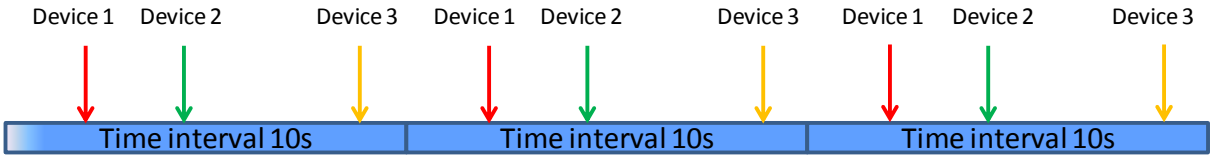


Figure 3-3: Regular time-driven transmission of short messages.

3.2.2 Performance metrics

The performance metrics for Scenario 1 are defined as follows:

- Co-existence of LTE-M in a LTE system is achieved if the minimum performance requirements valid for LTE UEs, which are defined in the specification 36.101 [25], are fulfilled, and if the performance of LTE-M is reasonable in the sense of the EXALTED project objectives and the technical requirements in the report D2.1 [5]. The reasoning for this may be given in a quantitative or qualitative style in section 4 e.g. by using the metrics listed in Appendix A1.

- In EXALTED, coverage is defined as the probability that a randomly located LTE-M device experiences a mean SINR leading to a Transport Block Error Ratio (BLER) below 10%. Alternatively, for broadcast signals, coverage is defined as the probability that a receiver is able to decode an entire broadcast message.

- Energy savings are given in percentage. The metric is defined as $E_{\text{energy}} = 1 - \frac{\text{Energy consumption of a LTE-M device}}{\text{Energy consumption of a LTE UE}}$. For a fair comparison we assume the same function volume in both devices, i.e. the display and other features that are irrelevant for LTE-M applications are not considered. The breakdown of energy consumption is explained in section 7.

- Cost savings are given in percentage. The metric is defined as $E_{\text{cost}} = 1 - \frac{\text{Cost of a LTE-M device}}{\text{Cost of a LTE UE}}$. For a fair comparison we assume the same function volume in both devices, i.e. the display and other features that are irrelevant for LTE-M applications are not considered. The breakdown of costs is explained in section 8.

Apart from these overall KPIs, several subordinated metrics have been evaluated. They are summarized in Appendix A1.

3.3 Scenario 2: Supporting a Large Number of LTE-M Devices with Heterogeneous Requirements and Capabilities

The purpose of this evaluation scenario, which is mainly related to the PHY and MAC layers of the LTE-M system, is to provide a guideline for candidate solutions to be evaluated with respect to two special characteristics of the M2M ecosystem:

- Large number of active devices within a cell
- Wide diversity of the machine's performance requirements and capabilities, compared to the conventional cellular communications.

The evaluation of the proposed solutions with respect to the overall KPIs and further subordinated metrics (see Table 1-1) can be found in the following sections:

- Co-existence with LTE in section 4;
- Support of a big number of devices / Spectrum efficiency in section 5;

3.3.1 Traffic models

The following traffic models are taken into account in the evaluation scenario 2.

- Event-Driven traffic modelling: Memory-less data packet arrivals per M2M device are assumed, in order to accurately model the traffic behaviour of several M2M applications such as intelligent transportation and collision reporting/avoidance systems, where data are triggered by random events. Event-driven traffic arrivals are modelled by the Poisson distribution with a single statistical parameter expressing the average packet arrival rate λ. Based on an assigned average arrival rate, random traffic patterns obeying the Poisson distribution are easily generated for each M2M device utilizing random number generators (for an extensive treatment of Poisson traffic modelling & analysis refer to paper [26]).
- With respect to QoS requirements, real-time classes are assumed based on the 3GPP 23.203 [27] QCIs with 1) Priority level, 2) Delay Threshold and 3) Threshold Violation probability or Packet Dropped Rate.
- QoS requirements of the LTE-M devices with respect to the delay tolerance are uniformly distributed between a minimum and a maximum delay tolerance.
- For representing some of the main traffic data flows that occur over LTE networks, following real-time (RT) and non-real-time (NRT) traffic types are considered:
  - RT traffics: Voice over IP (VoIP) and Near Real-Time Video (NRTV) are modelled respectively according to 3GPP documents [28] and [29].
  - NRT traffics: HyperText Transfer Protocol (HTTP) and File Transfer Protocol (FTP) are modelled according to [29].
- M2M traffics are considered and modelled according to 3GPP TR 37.868 [30]. Two types of traffics are considered at the point of view of the core network:
  - Standalone LTE-M devices that have a traffic flow as described in [30];
  - Gateways (GW) that need to aggregate the traffic flow for the whole capillary network they represent. Hence this traffic is modelled as follows. A first parameter sets the number of devices that form the capillary network. Traffic flows are generated for all these devices behind the GW as described in [30] and stored in a buffer. A second parameter sets the period with which a ‘super-packet’ is generated based on all packets that have been stored in the buffer during the current period.

3.3.2 Other evaluation assumptions and parameters

The scope of this section is to outline other common simulation assumptions for independent assessment of the candidate innovations (Table 3-2).

Table 3-2: Other simulation parameters and assumptions.

Parameter	Value
Carrier frequency	800 MHz or 2GHz
Operating bandwidth	Up to 20 MHz
CSI estimations	Perfect
Synchronization	Perfect
Number of transmit antennas of eNodeB	1-4
Number of eNodeBs in Tracking Area (TA)	10-100
Number of transmit/receive antennas of M2M devices	1 RF chain, more than one antennas available
Number of transmit /receive antennas of LTE UEs	1 Tx, 2 Rx
eNodeB Tx power per sector	46dBm for BW 10MHz, 20MHz and 43dBm for BW 1.25MHz, 5MHz
LTE-UE Tx power	26 dBm
LTE-M device Tx power	17 dBm
Antenna parameters (eNodeB)	Antenna gain: 13 dB, Angle for 3 dB antenna attenuation: 70°, max. antenna attenuation: 20dB, See 36.814 [24]
Other assumptions (eNodeB)	Data power ratio: 0.857

3.3.3 Performance metrics

- Co-existence achieved (yes or no): Co-existence of LTE-M in a LTE system is achieved if the minimum performance requirements valid for LTE UEs, which are defined in the specification 23.203 [27], are fulfilled, and if the performance of LTE-M is reasonable in the sense of the EXALTED project objectives and the technical requirements in the report D2.1 [5]. The reasoning for this may be given in a quantitative or qualitative style in section 4 e.g. by using the metrics listed in Appendix A1;
- The maximum number of active devices that are supported by the LTE-M system is compared to the corresponding number when employing the conventional LTE system.

Apart from these overall KPIs, several subordinated metrics have been evaluated. They are summarized in Appendix A1.

4. Objective 1: Co-Existence with LTE

Co-existence of LTE and LTE-M systems integrated within the same eNodeB and sharing the same carrier is desirable for an operator to minimise the OPEX (operational expenditure) and CAPEX (capital expenditure). In a first subsection it is proven that the specification of LTE-M achieves this essential objective of EXALTED. An important aspect is the assessment if and how existing infrastructure used by both systems can handle the additional traffic. Based on this general framework, the second subsection presents a selection of algorithms that facilitate or enable co-existence in a particular way.

4.1 Summary of Co-Existence Aspects in the LTE-M Specification

Co-existence of both systems is ensured with the time and frequency multiplexing of resources for LTE and LTE-M that was introduced in section 2.2. Moreover, the co-existence of LTE UEs and LTE-M devices including gateways is achieved with the introduction of a new UE category and registering information about LTE-M terminals to facilitate the network to differentiate LTE-M terminals from normal LTE terminals.

The optimization of the system for low data rate and high delay tolerant M2M applications requires the introduction of new physical channels for some of the functionalities. EXALTED has also introduced LTE-M bearer classes and several new logical channels for addressing the M2M requirements.

As shown in Figure 2-6, several MBMS Traffic Channels (MTCHs) are time multiplexed with normal LTE channels and decoded by UEs configured to receive MBMS. One or more of MTCHs are identified by the network for M2M services and allocated for the new physical channel (PMDSCH), where the optimized logical channels identified by EXALTED are multiplexed.

A LTE-M device receives the MBMS configuration in SIB 13. The scheduling information is provisioned on the MCCH and indicates an MBSFN subframe in which a MTCH is transmitted. The proposed physical channels for LTE-M are introduced in section 2.2.3 and listed for the sake of completeness in Table 4-1.

Table 4-1: Physical channels in LTE-M.

Name	UL/DL	Comment
PMDBCH	DL	Downlink Broadcast channel
PMDCCH	DL	Downlink control channel
PMDSCH	DL	Downlink shared channel
PMDMCH	DL	Multicast channel
PMRACH	UL	Random access channel
PMUCCH	UL	Uplink Control Channel
PMUSCH	UL	Uplink Shared Channel

The channels and procedures of LTE listed in Table 4-2 are re-used with little or no modification.

Table 4-2: Physical channels and signals in LTE that are re-used in LTE-M.

Physical Channels and Signals	Description
SS (Synchronisation Signal)	Used for cell search. LTE-M devices are allowed to combine energy over multiple samples to improve coverage at very low signal levels. Physical cell ID and SS sequence has a one-to-one mapping.
DLRS (Downlink Reference Signal)	Used for timing advance estimation, symbol timing synchronisation, received quality measurement.
PBCH (Physical Broadcast Channel)	Provides initial information subsequent to cell search (system bandwidth, system frame number, number of transmit antennas). Additional broadcast information specific to LTE-M is provisioned in PMDBCH.
PCFICH (Physical Control Format Indicator Channel)	Indicates number of symbols for control channel.
PDCCH (Physical Downlink Control Channel)	Provides scheduling information for system information and paging messages.
PMCH (Physical Multicast Channel)	Used for MBSFN operation and for new physical channels of LTE-M system.
UL DM RS (Demodulation Reference Signal)	Multiplexed with PUSCH and PUCCH, used for UL timing advance estimation, symbol timing synchronisation and received quality measurement.

4.2 Co-Existence with LTE Evolved Packet Core (EPC)

The LTE EPC will require scaling and densification to support a huge number of additional LTE-M devices similar to the requirement of an increased subscriber number as LTE rollout progresses. Impact on the Mobility Management Entity (MME) and the Serving Gateway (SGW) is for handling of significantly large signalling load along with large number of IDLE to/from ACTIVE transitions. EPC should be also able to support non-3GPP access for nodes behind gateways. Additional impacts to aspects of session management, policy and charging enforcement, roaming, backhaul security, multimedia support, etc. are outside the scope of EXALTED.

4.3 Further Solutions Facilitating Co-Existence

As co-existence is a general objective independent from a certain deployment scenario, the validity of the following study is not limited to one of the evaluation scenarios presented in section 3.

The presented solutions can be broadly categorised as solutions targeting optimisation of physical channels and those that reuse existing physical channels but optimise specific network and UE behaviour for efficient M2M system operation.

4.3.1 Registering information about the terminals

In order to enable the networks to differentiate between LTE and LTE-M terminals as well as between different types of LTE-M terminals and their respective requirements, a new procedure is proposed that enables terminals to describe themselves with M2M specific UE category and their application specific capabilities, traffic generation patterns and service requirements. Information that is already stored in the network about the terminals should be extended with new information. The proposed algorithm introduces extension of messages used by the terminals when registering to a network and extension of information that is supposed to be saved in the Home Subscriber Servers (HSS) will be investigated and

proposed. Since registering information about terminals actually simply represents an extension of existing LTE messages, co-existence between LTE and LTE-M is achieved. The additional parameters and their meaning are summarized in section 7.2.7. The extension leads to a small additional overhead. However, this will not dramatically affect provisioning time and HSS memory consumption.

Using this information, the network will know the type of terminals attached to a radio cell and will be in a position to use that information to configure the terminals and the cell in an appropriate manner. Proposed procedure actually is applied as an enabler for the other procedures that can benefit from this information.

4.3.2 Slotted access

Slotted access is applied solely on the PMRACH. The basic idea is that the LTE-M devices must not access the PMRACH at any possible point in time, but only if an additional condition is fulfilled. Such a condition can be that the LTE-M device may transmit on the PMRACH only in subframes, when it receives a paging signal from the eNodeB. The aim is to force the arrival distributions of PMRACH transmissions from a multitude of LTE-M devices being flat over time in order to minimize the collision probability on the PMRACH.

It does not change the PMRACH resource structure and does not impact any of the LTE channels.

Slotted access supports co-existence with LTE for the following reasons:

- It does not change the PMRACH resource structure and does not impact the LTE RACH or any other of the LTE channels.
- Slotted access is also an overload protection that prevents the case of a too big number of devices trying to enter the LTE/LTE-M access and core networks at the same time. Without such a protection mechanism a sudden overload caused by LTE-M devices may affect the performance of legacy LTE UEs.

4.3.3 *Hybrid automatic repeat request (HARQ)*

HARQ in the LTE and LTE-M systems has the same target of increasing the reliability of transmissions. LTE standard adopts a classical turbo coded HARQ having a fixed-rate for each (re)transmission, while the proposed HARQ [15] for LTE-M utilizes the Channel Quality Indicator (CQI) back from the receiver to generate the encoded bits with an appropriate coding rate for the first transmission. This proposed HARQ procedure is different with, but has no impact to the original one.

As described in the project report D3.3 [1], EXALTED's proposed HARQ in LTE-M may reuse the LTE turbo encoder. However, it will perform puncturing operations based on its output 1/3-rate systematic encoded bits. Moreover, the CQI indicator message designed in the proposed HARQ will be required by the transmitter before the first transmission. Soft combining and iterative decoding are the same for both HARQ schemes. There is no timing change for our proposed one.

To sum up, EXALTED's proposed HARQ for the LTE-M may be co-existed with the HARQ in LTE. They may not interact with each other, but they won't interfere with each other.

4.3.4 Scheduling for heterogeneous traffics

Scheduling algorithms are not part of the LTE and LTE-M specifications. Network operators can then implement any resource scheduling method that permits to satisfy the devices with respect to their specific QoS requirements. These requirements depend on traffic types

which are characterized by different constraints, either on devices (i.e. number of buffered packets) or on packets (time to live (TTL) of the packet). Two classes of traffic are considered: the Guaranteed Bit Rate (GBR) class when a minimum bit rate is requested by an application and the Non-GBR or 'best effort' class when bearers do not guarantee any particular bit rate. GBR aims at describing real-time services.

In most LTE networks, several real-time (RT) and non-real time (NRT) traffic types have been standardized (VoIP and NRTV for RT traffics – HTTP and FTP for NRT traffics) and their characteristics are used by the schedulers. Nevertheless, M2M networks are characterized by other traffic types, whose constraints plainly differ from the ones of LTE traffic types (see sections 2.8 and 3.3.1).

The scheduling algorithm for heterogeneous traffics has been detailed in the project report D3.3 [1]; it proposes to deal jointly with the legacy LTE traffics generated by conventional LTE UEs and with the M2M traffic generated by newly deployed LTE-M devices as well as the aggregated traffic issued from a M2M gateway that interfaces a capillary network beyond. Results in subsection 5.2.8 show how LTE and LTE-M can coexist in terms of traffic generation and satisfaction of UEs and M2M devices over the legacy LTE resources and infrastructure. Indeed the algorithm permits to schedule communications with respect to the specific QoS constraints of all traffic types.

Surely the legacy LTE UEs will be affected by presence of the new LTE-M traffics over the same resources but this study shows how the LTE-M devices are satisfied while the LTE UEs just suffer from a light reduction in services satisfaction. It is undeniable that the use of disjoint resources for LTE and LTE-M would permit to not affect the LTE network while satisfying more M2M devices, but it was not the goal of this study.

4.3.5 *Access grant time interval (AGTI) scheduling for M2M traffic*

AGTI is a semi-static scheduling algorithm which allocates a portion of the LTE frame (in other words part of the available LTE Physical Resource Blocks) to M2M devices. The allocation pattern is periodic, according to a scheduling rate calculated at once when the M2M devices are registered in the network and their QoS requirements become known to the eNodeB. Care has been taken in order to closely follow the LTE scheduling principles. In particular:

- LTE and M2M devices are multiplexed in the same time-frequency resources grid;
- LTE transmissions are prioritized over M2M transmission requests, and if after all LTE-UEs have been served a number of PRBs remains available, then M2M devices may be allocated these resources. Thus, inter LTE-M2M resources orthogonally is preserved;
- AGTI guarantees specific statistical delay-based QoS constraints, that are the delay-threshold violation probability, which is part of the LTE-QCI list, and thus the existing LTE QoS features may be reused.

4.3.6 *Generalized frequency division multiplexing (GFDM)*

Considering a co-existence scenario (Figure 4-1), where a LTE primary system based on OFDM/SC-FDMA is overlaid with a LTE-M secondary system either of the same kind or based on GFDM [31], the latter shows better properties when looking at the resulting bit error rates (Figure 4-2) under the assumption that both systems are not synchronized towards each other.

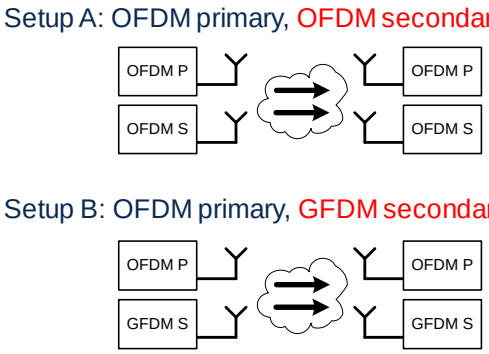


Figure 4-1: Co-existence Setup for primary and secondary systems of different access schemes.

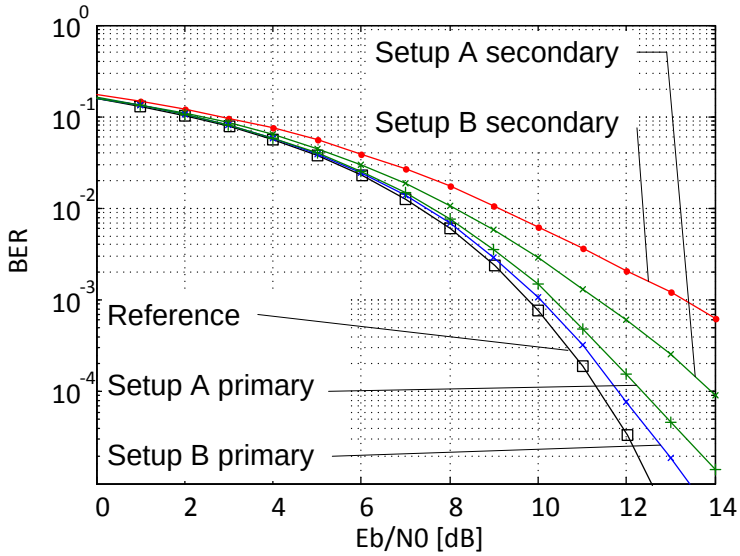


Figure 4-2: BER performance of co-existence scenarios A and B.

The reason for this can be found in the fact that, while the LTE spectrum can be shaped towards neighbouring bands with extra filters, by leaving some resources empty somewhere in the middle, a very bad interference situation occurs for the secondary system (Figure 4-3).

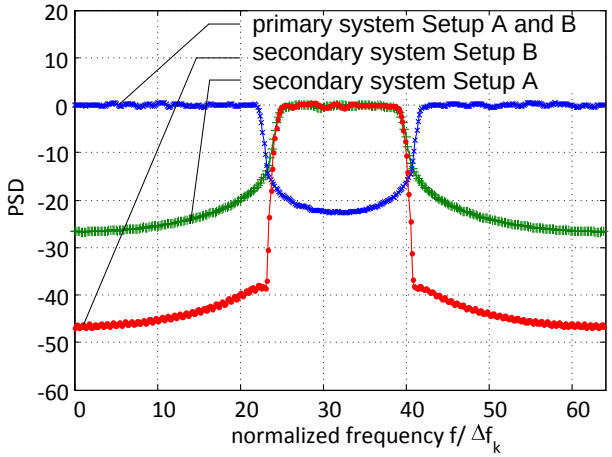


Figure 4-3: Spectral properties of secondary system within primary system.

GFDM not only provides a smaller disturbance of the primary system when compared to a LTE-like secondary signal of the same bandwidth, but if GFDM was to be used as a primary scheme, it would also allow to create better spectral holes in the signal. This effect would improve also the performance of the secondary system.

Looking at a BER of 10^{-4}, the primary system in setup B outperforms the primary system of setup A by roughly 0.5 dB which corresponds to 12.2% less energy per bit that is necessary to achieve the same error rate. Similarly, for an Eb/N0 value of 12 dB, the achieved error rate of the primary system in setup A is $1.5*10^{-4}$ for setup A and $7.5*10^{-5}$ for setup B, which is a reduction of 50%.

4.4 Conclusion

Two work package objectives are reflected, namely that the proposed solutions have to be supported by existing eNodeB hardware platforms (O.3.1) and that backward compatibility to LTE is maintained (O.3.7). LTE-M as a whole already underlines these requests through the separation of radio resources for LTE and LTE-M. But also some individual solutions were explicitly specified to support the co-existence. Registering information about terminals is the key enabler. Thereby, it is possible for the network to distinguish between LTE and LTE-M terminals and between different types of LTE-M terminals. Slotted access on the LTE-M random access channel supports the co-existence of both systems mainly by avoiding a sudden overload situation of the EPC in the case of an event, where a huge number of sensors try to connect to the network. The proposed HARQ scheme in LTE-M is different from that in LTE, but existing functional blocks may be re-used. Innovative *scheduling techniques* jointly consider the heterogeneous traffic flows of LTE and LTE-M users and their individual constraints and QoS requirements. This also includes the case of aggregated traffic from non LTE-M devices behind a M2M gateway. Finally, GFDM underlines co-existence through its spectral properties that allow assigning small spectrum chunks to LTE-M devices without affecting the performance of adjacent LTE UEs.

5. Objective 2: Support of a Big Number of Devices / Spectrum Efficiency

The support of a large number of M2M devices is one key objective to be addressed within the LTE-M system design. While typical LTE cells are designed to serve a multitude of several hundreds of LTE users, a LTE-M enabled cell has to cope with several thousand or even higher numbers of LTE-M devices. As LTE is designed to serve users with high data rate it is not optimized for servicing many nodes that only want to deliver single very short messages.

As radio resources are scarce, the LTE-M system design aims at the support of a large number of devices through two different approaches which are the reduction of signalling and the efficient radio resource usage. The mechanisms that enable LTE-M to achieve the goal of increased number of supported devices will be summarized shortly in section 5.1. Detailed information can be found in D3.3 [1]. Section 5.2 will give a comprehensive evaluation of these mechanisms with respect to the given objective while in section 5.3, conclusions for the overall performance of LTE-M on the support of a large number of devices is given.

5.1 Proposed Solutions

In order to achieve a very efficient communication a primary goal is the reduction of signalling that is used in the communication protocols. There is quite a mismatch between the payload, that has to be delivered and the additional signalling information that must be added in LTE. The LTE-M system exploits the nature of reoccurring transmissions in M2M applications such as temperature sensors. A device that does e.g. report in a periodic way does not need to get scheduling information every transmission. Therefore, LTE-M supports a semi-static scheduling rule that reduces the signalling information. The principle is shown in Figure 2-13 in section 2.3.4 where green resources are freely scheduled and need specific scheduling information. Blue resources are scheduled periodically and therefore, scheduling information has only to be provided once. In order not to jeopardize the system performance of the primary LTE system, these periodicities have to be determined in an optimized way. To take this into account, M2M traffic specific scheduling solutions are provided. The implementation is a vendor specific task and therefore not part of the LTE-M specification but to give insight in the gains, that the LTE-M system could give, these scheduling options are evaluated as well. Such a periodic scheduling approach is supported by the AGTI based scheduling. The LTE-M devices will be organized in clusters and the scheduling periodicity is calculated analytically. Even without using specific CQI information, the scheduling algorithm can gain performance in means of providing real-time communication for specified devices. The solution additionally reduces the signalling used for scheduling messages. To go one step further, another QoS based scheduling solution takes the channel quality additionally to the delay constraints into account. No classes of devices need to be formed which models the typical configuration of an M2M environment in a more realistic way. The solution is capable of meeting the exact delay constraints of different devices. This enables efficient real time communication as well as delay relaxed energy efficient communication.

Additionally, due to the possibility that exact delay constraints are met, devices can go to sleep mode faster: an advantage that positively influences the energy efficiency. To further combine different traffic characteristics of a joint LTE/LTE-M system a scheduling algorithm is provided that shows the efficient use of LTE resources for the joint LTE/LTE-M downlink.

A typical bottleneck in uplink communication is the random access channel (RACH). The LTE RACH is designed to serve a specific number of LTE devices. As the number of devices

in the M2M setting is many times higher, such a channel needs to be redesigned in a way to support even a very high number of devices with reasonable collision probabilities. This fact is discussed in the 3GPP submission [32]. Therefore, the LTE-M system introduces the use of slotted random access to the PMRACH as depicted in Figure 5-1 and proposed in the 3GPP submission [33]. This slotted approach supports different priorities in a hierarchical manner.

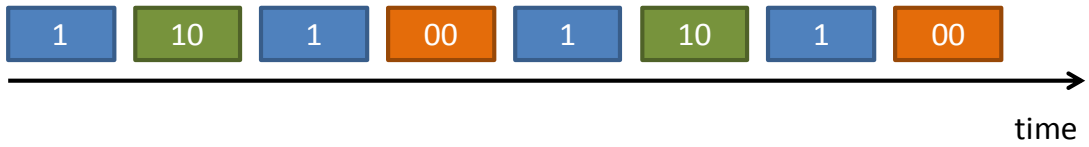


Figure 5-1: Hierarchical slotted access with two priorities.

With this solution the load of the PMRACH at peak transmission situations is distributed over different slots and significant reductions in collision probabilities can be met. Therefore, retransmissions have to occur very sporadically and the resources are used in a way more efficiently for a large number of M2M devices. Additionally, to even further increase the resource efficiency, LTE-M enables the transmission of very short messages directly within the PMRACH. With this approach, signalling can be reduced as the whole transmission has already taken place when the device attempts to ask for transmission resources.

Finally the signalling reduction can be tackled on a higher layer by the introduction of new device descriptions. These descriptions do characterise different applications of LTE-M devices and therefore limit the functionality that a specific device can support. In turn, the device can benefit from a reduced communication overhead as specific procedures typically done in every communication scenario can be omitted for this specific device within its description. Therefore, optimized functionalities can be applied to the newly defined device descriptions and an overall reduction in signalling can be achieved.

The second general approach that is used within the scope of the LTE-M system design is the optimization of radio resource usage. The efficient usage of radio resources can be combined very easily as it supports the previously mentioned techniques of LTE-M on a lower layer, namely the PHY layer.

Regardless of the reduced collision probability with slotted random access in PMRACH, collisions may still occur. Therefore a further option can be added to the PMRACH to even resolve PMRACH accesses that faced a collision. This is done via the introduction of physical layer network coding [34], [35], [36]. Using a combined decoding of the already collided packets, the spectral efficiency of the system can be improved even in situations where the load of the cell is high. To further improve the spectral efficiency especially in the uplink, a new generalized modulation scheme called Generalized Frequency Division Multiplexing (GFDM) is introduced [37]. Within the system specification the GFDM access has the advantage of having similar and even sharper spectral properties as SC-FDMA with a reduction of cyclic prefix that is usually used to account for the influence of the channel. When using GFDM as uplink radio access, more users can be served due to less CP.

The efficiency of the communication system is strongly coupled with the need for reliable communication. To enable reliable communication, LTE used HARQ schemes that can run several HARQ procedures, each with the option of several retransmissions. As resources are not that scarce in the LTE setting, the system is optimized for reliability even with many retransmissions. This is not the case in LTE-M and therefore the system design of LTE-M uses an optimized HARQ procedure. That only supports one retransmission but in turn optimizes the first transmission in that way that an almost optimal code rate is chosen. In

simulations it is shown that with this procedure radio resource usage is done in a more efficient way as with equal performance, less retransmissions are needed.

The following section evaluates the different gains introduced by different parts of the LTE-M system with respect to the objective of the support of a large number of devices.

In the adaptive paging algorithm (evaluated in section 7) minimization of network signalling for monitoring paging channel of the LTE-M is achieved. Focus of the algorithm is on the minimization of the number of paging messages sent towards LTE-M device during the paging procedure. Reducing the number of paging messages, contribute to a reduction in the network load, thus available resources can be used for handling more devices. Less paging also reduces the signalling in the radio access network. This part of the LTE-M system design is referred to as an enabling technology that will not be evaluated within the section 5.2

Also one more algorithm is proposed in section 7, namely registering information about terminals. To enable LTE-M network to differentiate between different terminals and their requirements, new parameters are introduced and stored in networks' Home Subscriber Servers (HSS) as a part of user subscription (paging factor, mobility management factor and Traffic generation pattern). Actually this algorithm is used as an enabler for power savings in the other procedures. For example by using paging factor and mobility management factor, power consumption in IDLE mode can be reduced (during attach procedure these parameters are downloaded on MME as a part of subscription and delivered to LTE-M device), while from Traffic generation pattern parameter, access schemes and resource allocations for LTE-M could use additional terminal information and provide improvements.

5.2 Performance in Scenario 2

This section will comprehensively evaluate the performance of specific mechanisms of LTE-M with respect to the objective of supporting a large number of devices. The evaluation is carried out for Scenario 2 as the parameters characteristics of Scenario 2 are specifically designed for the evaluation of this objective (see section 3.3). The goal for each LTE-M mechanism in this section is to provide a characteristic gain concerning the objective given by a relative benefit over existing solutions.

5.2.1 Random access with collision recovery

The performance of the proposed LTE-M Random Access scheme with Collision Recovery, called NCDP protocol, was evaluated using the parameters defined in section 3. In order to assess the advantage of NCDP regardless of the performance of the PHY layer, we assumed in all cases correct decoding of the packets, achieved through power control. This is a common assumption in our proposed NCDP protocol and also in the benchmark systems. A traffic model consisting of Poisson Arrivals was chosen, and the metric for performance evaluation is *Normalized Throughput*, Φ as is customary in the analysis of MAC systems. Spectral efficiency can be easily derived from Φ as $S_{MAC} = \Phi S_{PHY}$, where S_{PHY} is the physical layer spectral efficiency.

Two benchmarks for our evaluation are considered. A basic benchmark is provided by Slotted Aloha. As it is well known, the throughput of Slotted Aloha saturates at $\Phi=1/e=0.368$. A second benchmark is provided by the Contention Resolution Diversity Slotted Aloha (CRDSA) Protocol. Such a scheme, detailed in paper [38] and described also in EXALTED Deliverable D3.3 [1], consists in sending multiple instances of a packet through several transmission slots. The receiver performs successive interference cancellation, starting from a "clean burst", i.e. a burst where only one node transmitted a packet.

The proposed Network Coded Diversity Protocol, fully described in EXALTED Deliverable D3.3 [1], detects the linear combination of packets when a collision occurs, and then uses standard Linear Network Code decoding algorithms to recover all the packets. In the selected version, each packet is transmitted in a fixed number (d) of slots, chosen randomly, and each packet is multiplied by a random coefficient out of a finite field of size 2^8. In both NCDP and CRDSA a frame is defined, containing S slots, from where slots for transmission are chosen. Figure 5-2 shows the normalized throughput as a function of the traffic load. As it can be seen, NCDP exceeds all benchmarks and reaches a throughput above 80%, which is quite remarkable for a random access scheme. The best performance is obtained when a packet is transmitted with constant probability over every time slot. In this case, no retransmissions were considered, and a packet is lost if it is not successfully decoded within one frame.

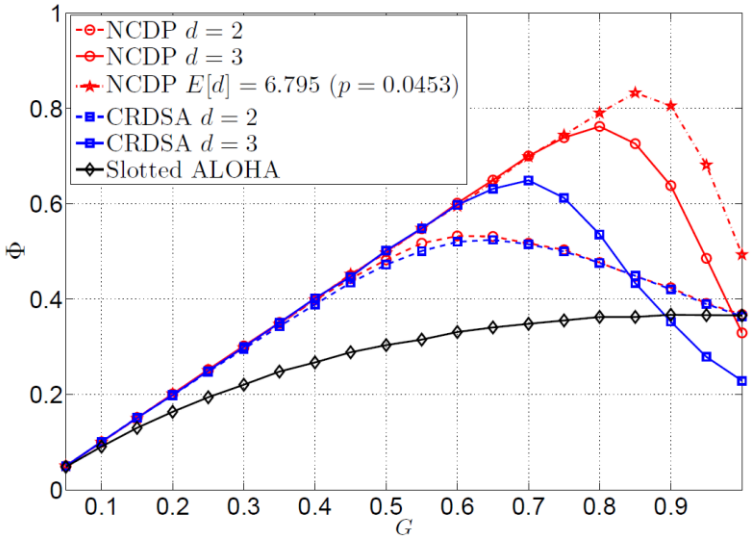


Figure 5-2: Normalized throughput Φ vs. normalized traffic load G of the proposed collision recovery NCDP protocol without retransmissions. In this simulation, each packet is transmitted either twice (d=2), three times (d=3), or with fixed probability p=0.0453 on every slot of the frame. The frame size is S=150 slots.

The NCDP protocol including retransmissions was also evaluated, and results are shown in Figure 5-3. In this simulation, unsuccessful transmissions result in a NACK sent back to the transmitter. The transmitter then enters a backlog state and randomly chooses a frame within the following B frames to retransmit its packet. As it can be seen in the picture, the throughput falls abruptly when a significant number of packets is lost, as backlogged nodes tend to saturate the channel.

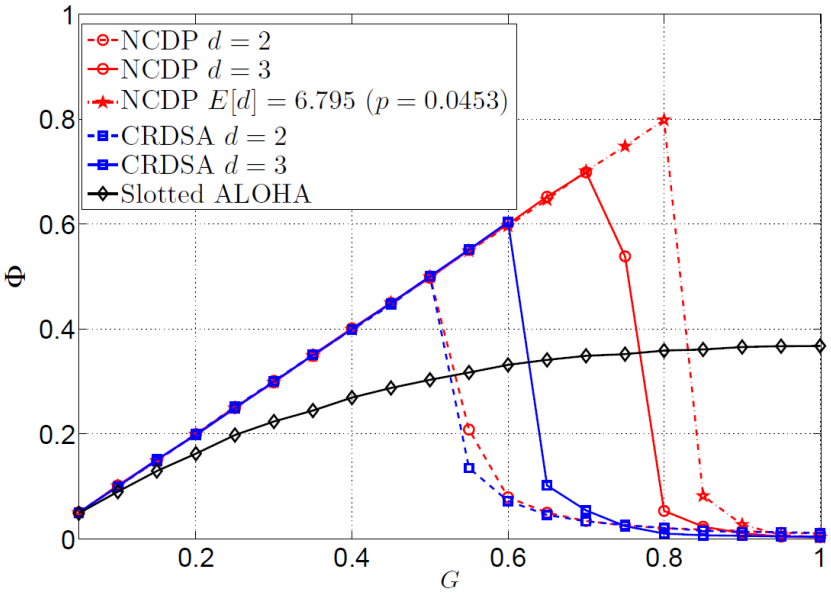


Figure 5-3: Normalized throughput Φ vs. normalized traffic load G of the proposed collision recovery NCDP protocol with retransmissions. Each packet is transmitted either twice (d=2), three times (d=3), or with fixed probability p=0.0453 on every slot of the frame. The frame size is S=150 slots and the maximum backlog is B=50 frames.

5.2.2 HARQ

HARQ is relatively independent to the assumptions and parameters in the PHY layer. Only the perfect CQI is the necessary one that EXALTED's proposed adaptive HARQ requires. Using the parameters defined in section 3, the average number of retransmissions between our proposed HARQ [15] and two classical HARQ schemes is compared. The first benchmarker is a Type-II HARQ scheme with Chase combining, which transmits the same 1/2-rate turbo encoded bit sequence for each transmission. The other one is referred to as Type-III HARQ, which transmits a 1/2-rate encoded packet for the first transmission, while the rest of parity bits will be transmitted during the second transmission. The chase combining is also performed for the Type-III HARQ, since there are a part of bits repeated during the retransmissions. Figure 5-4 illustrates the average number of transmission for each packet for the three HARQ schemes.

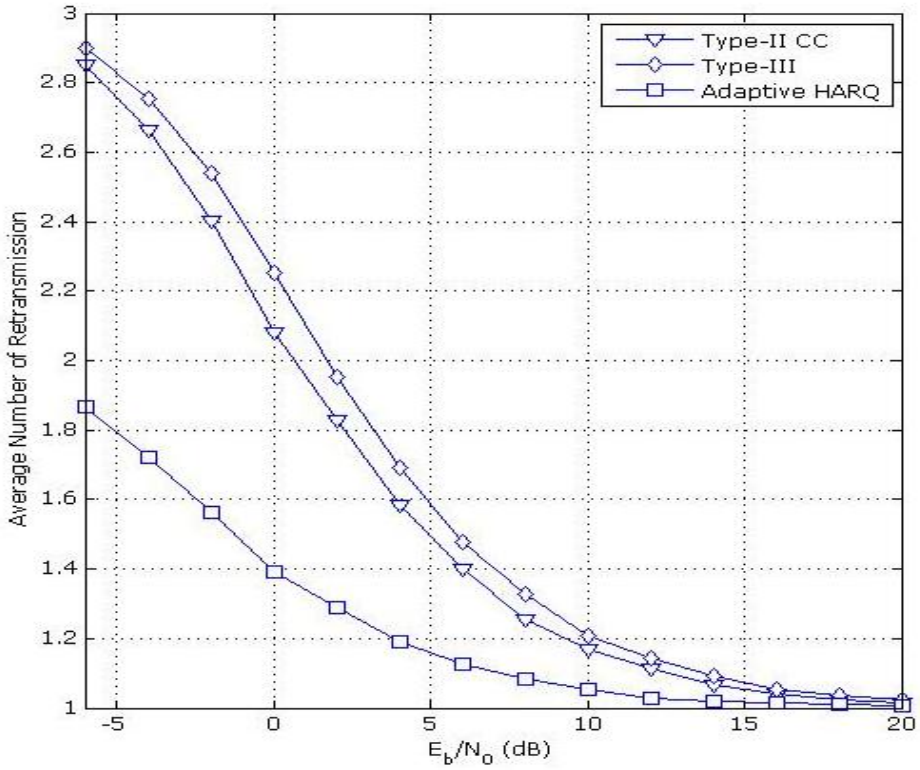


Figure 5-4: The average number of transmissions for each packet versus Eb/N0.

It may be observed in Figure 5-4 that our proposed HARQ has a significantly lower number of transmissions than those of other two benchmarks. More specifically, the number of transmissions was less than 1.9 for our proposed HARQ scheme, while the other two reached about 2.9 transmissions. Even at the high SNR of 10dB, our proposed HARQ scheme reduces the average number of retransmissions by one third, compared to the classical HARQ schemes. Since the average number of transmissions is deemed as the delay that a packet may be transmitted, its reverse is reasonably used as a metric of normalized number of users that a HARQ scheme may support. According to Figure 5-4, our proposed HARQ increases the number of users by about 30%, compared to the other benchmarks.

5.2.3 Semi-persistent scheduling

The performance evaluation is based on the scenario in section 3.2. The solution aims at a reduction of radio resources per LTE-M device on the PMDCCH required for scheduling information. As the overall PMDCCH capacity is fixed, it is obvious that a bigger number of devices can be supported if the amount of resources per device is reduced. In the following, the relative increase of supported devices applying semi-persistent scheduling over the conventional solution (dynamic scheduling) is derived.
At first the relevant parameters of the analysis are defined:

- Transmission number β: It is the average number of transmissions of one transport block until it is received successfully. This parameter is important because in contrast to a complete persistent scheduling, semi-persistent scheduling allocates radio resources for retransmissions always dynamically. It has found out that β increases significantly if retransmissions are scheduled on persistent resources as well, and that the system capacity drops off in this case.

- Time interval T_I between two transmissions on persistent resources. In LTE, semi-persistent scheduling is used for VoIP, and T_I is fixed to 20 ms for active users. In LTE-M, T_I is an arbitrary parameter that can be chosen according to the application. The minimum value of T_I is the length of an LTE-M super-frame.
- Relative portion α of first transmissions (not retransmissions) on persistent resources. It is clear that the initial transmission in an activity period of the device must be scheduled dynamically because the persistent resources are not yet known. Moreover, if the channel properties change over time, the link parameters like modulation scheme and coding rate must be adapted, and the allocation of persistent resources eventually needs to be updated. Generally, α depends on T_I. If T_I is small in relation to the validity interval of a modulation and coding scheme, α converges to the ideal value $\alpha = 1$.

A relative indication for the amount of control signalling needed for dynamic scheduling is $N_C^D = \beta / T_I$. For semi-persistent, where only retransmissions are scheduled dynamically, this value decreases to $N_C^{SP} = (\beta - 1)/T_I + (1-\alpha)/T_I$. The first part includes all retransmissions, and the second part considers first transmissions on dynamic resources. This leads to a relative gain of $G = \beta / (\beta - \alpha)$. Figure 5-5 shows the achievable gain factor G depending on the parameters α and β. Realistic values considering the assumptions in evaluation scenario 1 are $\beta = 1.5$ (one third of resources is required for retransmissions in average), and $\alpha = 80\%$, the given PMDCCH capacity is sufficient to support $G = 2.14$ times more LTE-M devices if semi-persistent scheduling is applied compared to the case with dynamic scheduling. For $\beta = 1.1$ and $\alpha = 95\%$, the number of supported devices increases even by the factor $G = 7.33$. The overall result is that semi-persistent scheduling can increase the number of LTE-M devices if the transmission number is close to the ideal value $\beta = 1$, and if α is sufficiently large. The latter condition is fulfilled for stable channel properties and small transmission time intervals T_I.

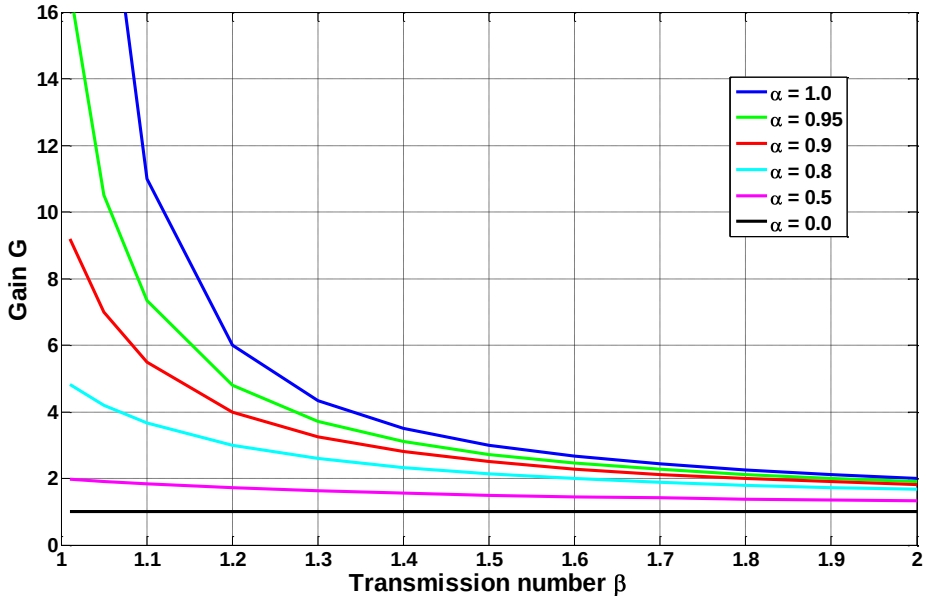


Figure 5-5: LTE-M device gain factor depending on the parameters α and β.

5.2.4 Slotted access

Slotted access addresses the reduction of collision probability if a multitude of LTE-M devices transmit on PMRACH resources after an event. The aim of this section is to show that the number of LTE-M devices with successful random access procedure increases significantly by using slotted access.
According to the evaluation scenario described in section 3.2, in case of an event, the arrival rate of short messages is Poisson distributed. Figure 5-6 shows a random snapshot of the arrival rate of PMRACH usages over time measured in multiples of the LTE-M super-frame length. The event occurs during the frame with index 0, and as a result $N = 500$ LTE-M devices try sending a short message.

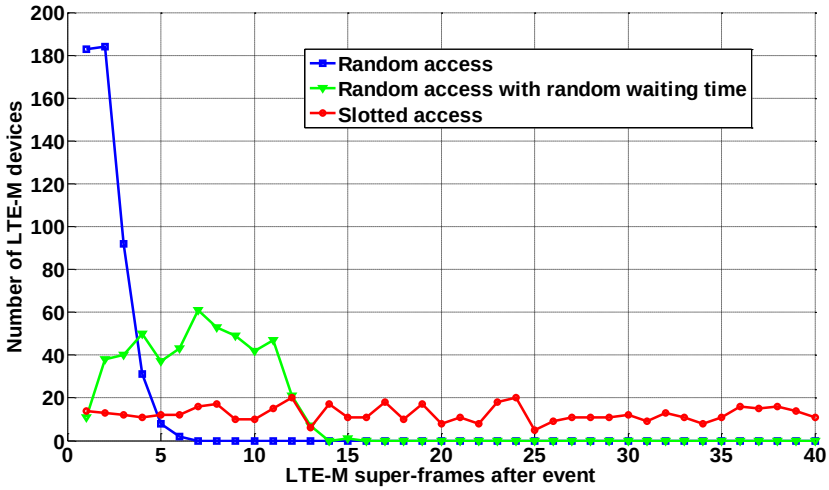


Figure 5-6: Distribution of PMRACH channel usage after an event.

The blue curve in Figure 5-6 shows the arrival rate based on the Poisson distribution with parameter $\lambda = 1$. Almost 75% of the devices start the random access procedure during the first two super-frames and the probability that two or more devices use the same PMRACH preamble is very high. A simple possibility to avoid a sudden overload of PMRACH resources is to introduce a random waiting time between the recognition of the event and the random access procedure. A waiting time equally distributed between zero and ten super-frames (green curve) mitigates the strong competition shortly after the event to some extent. The maximum number of devices transmitting their preamble simultaneously reduces from over 180 to 60.

Slotted access (red curve) forces the arrival rate being flat over time and therefore minimizes the collision probability. In the example shown, the devices are equally distributed within an interval of 40 super-frames, i.e. the expected value of devices initiating random access per super-frame is $\mathrm{E}[n] = 500/40 = 12.5$ in the considered set-up.

However, the drawback of slotted access is the additional delay of the reporting. If the application would restrict the maximal tolerable delay to e.g. 10 super-frames, the expected value devices would increase to $\mathrm{E}[n] = 500/10 = 50$ devices, and hence, the collision probability becomes considerably higher.

In the following, the collision probability of the three alternatives is analysed in more detail. Figure 5-7 shows this metric depending on the number of users varying between $N = 100$ and 1000. Moreover, the number of available PMRACH preambles is an important performance parameter. In the diagram, the cases $S = 64$ (solid lines), which is the number of preambles in the LTE RACH, and $S = 32$ (dotted lines) are depicted. The latter case reflects a possible

alternative that doesn't provide an additional random access channel for LTE-M in the specification, but the reservation of a subset of LTE RACH preambles for LTE-M devices instead. As expected, the collision probabilities for $S = 32$ are considerably higher for all three alternative solutions. This result underlines the need for a separate random access channel PMRACH for LTE-M devices.

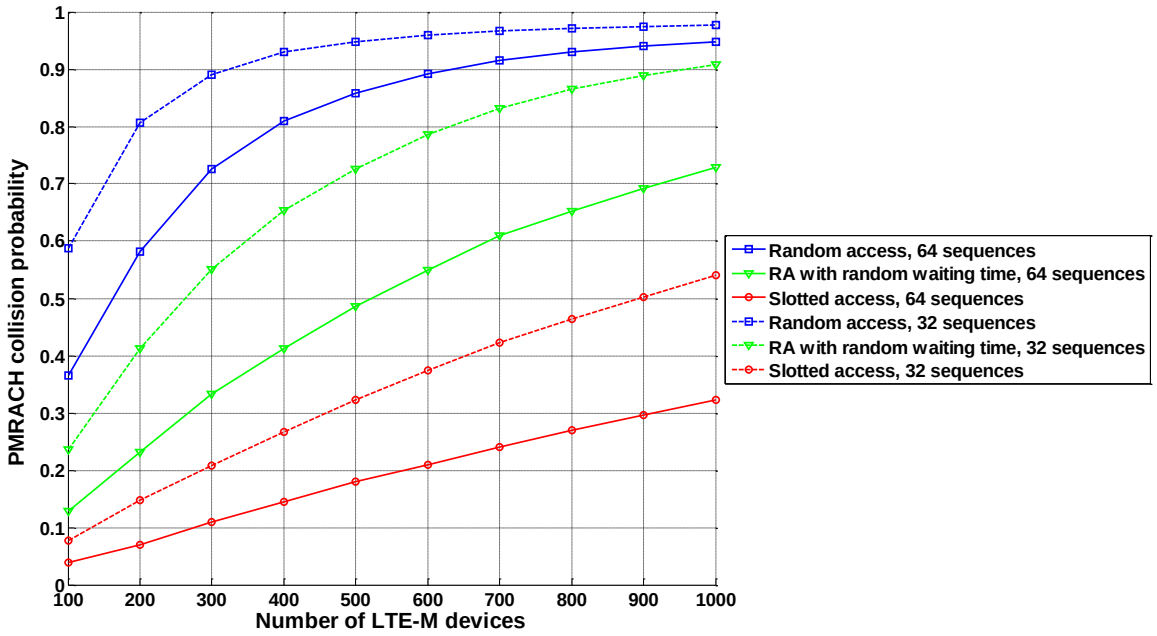


Figure 5-7: Collision probability of PMRACH resources depending on the number of LTE-M devices.

The collision probability for slotted access is clearly smaller compared with the conventional random access. Assuming a fixed collision probability, e.g. $P = 0.3$, the number of LTE-M devices fulfilling this criterion increases from below 100 (solid blue curve) to over 900 (solid red curve).

5.2.5 AGTI scheduler

The proposed solution comprises a semi-static scheduling algorithm for M2M traffic which is designed to operate over the existing LTE resource pool [39], [40], [41]. Unlike LTE where resources are redistributed among the competing terminals every TTI based on CQI and buffer status reporting (BSR) messages sent by each terminal to the scheduling entity, our solution calculates a scheduling rate for each M2M device, and then the devices access the frame at the particular periodic time-intervals. Such a scheduling policy minimizes the feedback signalling (that is the exchanged CQI and BSR messages, which are proportional to the number of terminals/devices) as well as the forward signalling needed for informing each M2M device at which RB-TTI should access the frame. Minimizing the total overhead signalling has a direct positive impact on the number of supported devices within a cell. In addition, unlike LTE-VoIP semi-persistent scheduling where the access period is fixed and predetermined, our solution is able to tune the scheduled access period for vastly diverse M2M applications, with different traffic intensity (average number of packet arrivals per TTI) and QoS requirement (delay thresholds, delay threshold violation probability) profiles. Such a flexible scheduler is able to deal with a wide heterogeneity of machines performance requirements and capabilities, compared to the conventional LTE cellular networks.

The AGTI-based scheduling solution is evaluated based on the assumptions of Scenario 2. A single-cell is simulated, 100 M2M devices are active, each one generating data packets

based on the Poisson traffic model with average packet rate of $\lambda = 0.02$ packets/TTI. No channel impairments are considered. Various scenarios are examined with different required real-time delay constraints (probability of exceeding a delay threshold or dropped packet rate due to delay threshold violation). For each scenario the associated scheduling period (T_g(TTIs)) has been calculated based on an analytical model proposed in D3.3 [1]. In Figure 5-8 we demonstrate the QoS performance levels achieved by our scheduling solution through simulation ("ACHIEVED-SIM") and theoretically predicted levels based on the analytical model ("QoS-THEORY"). Hence, our solution: (1) guarantees the required delay-based QoS requirements by tuning the scheduling access period of the M2M devices accordingly, and (2) avoids dynamic scheduling decisions which increase complexity and overhead signalling.

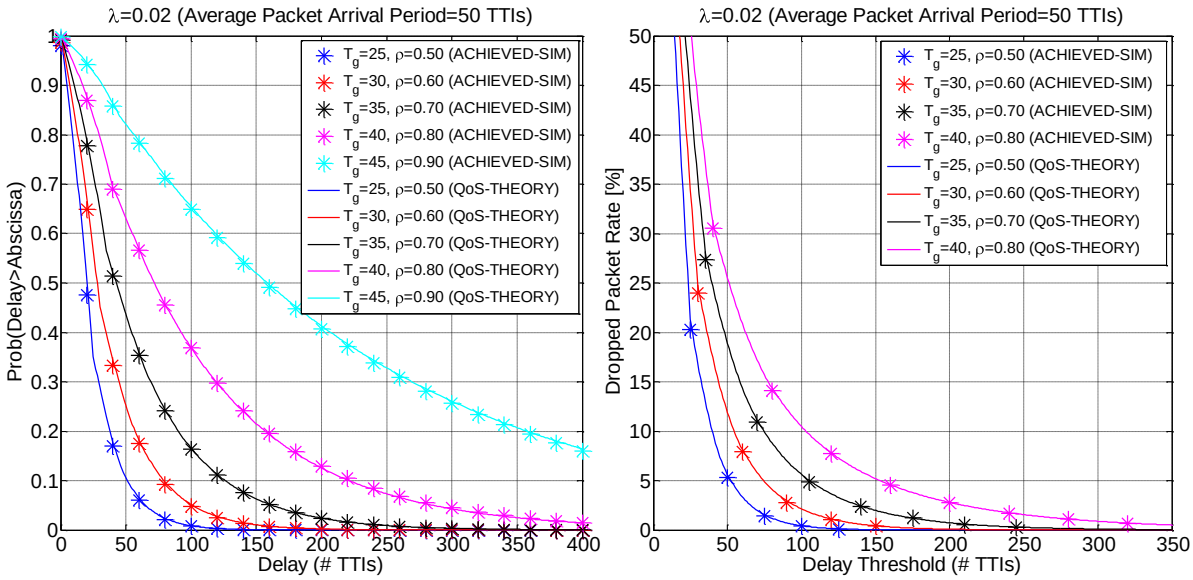


Figure 5-8: Delay-based QoS Performance Levels of the AGTI-based Scheduling Solution: Achieved Simulation-based & Theoretically-predicted Results for Periodic Access Scheduling with No Feedback Signalling.

5.2.6 QoS-based scheduler

In the following, using a LTE simulator, the performance of the LTE PS is examined when a number of M2M devices with specific requirements request to be served. The simulation model is based on the 3GPP LTE system model and it includes a single cell with uplink transmissions. The available bandwidth is 5MHz, which means that 25 RBs are available per TTI. The number of M2M devices ranges from 10 to 1500, with each one having an average SNR that uniformly ranges from 0 to 10 dBs. We assume no LTE users and that 5 MHz bandwidth is available for the M2M devices. In the simulation it is assumed that there is no interference between the subcarriers and detection errors occur due to the fading channel and the noise at the receiver. Moreover, it is assumed that the M2M devices can send their BSR signalling over the same bandwidth without interfering with each other, by using the Zado-Chu sequence generation. All M2M devices use binary phase shift keying (BPSK) modulation, since they are supposed to be low-rate low-complexity devices (e.g. sensors). The maximum delay tolerance of each device uniformly ranges from 10 ms to 10 minutes (TS 36.331 [18]).

In Figure 5-9 the performance of the LTE PS and that of the two proposed PS are compared under the same operation conditions [42]. Note that in this figure the optimum performance is shown. This means that the scheduler is assumed to know the exact delay tolerance of each

device. Under this assumption, we observe that knowing the exact delay tolerance of each device leads to significant increase in the number of satisfied devices. However, such a scheme is not practical, since the required feedback is prohibitive. Hence, a finite number of classes has to be found, covering a wide range of M2M QoS requirements.

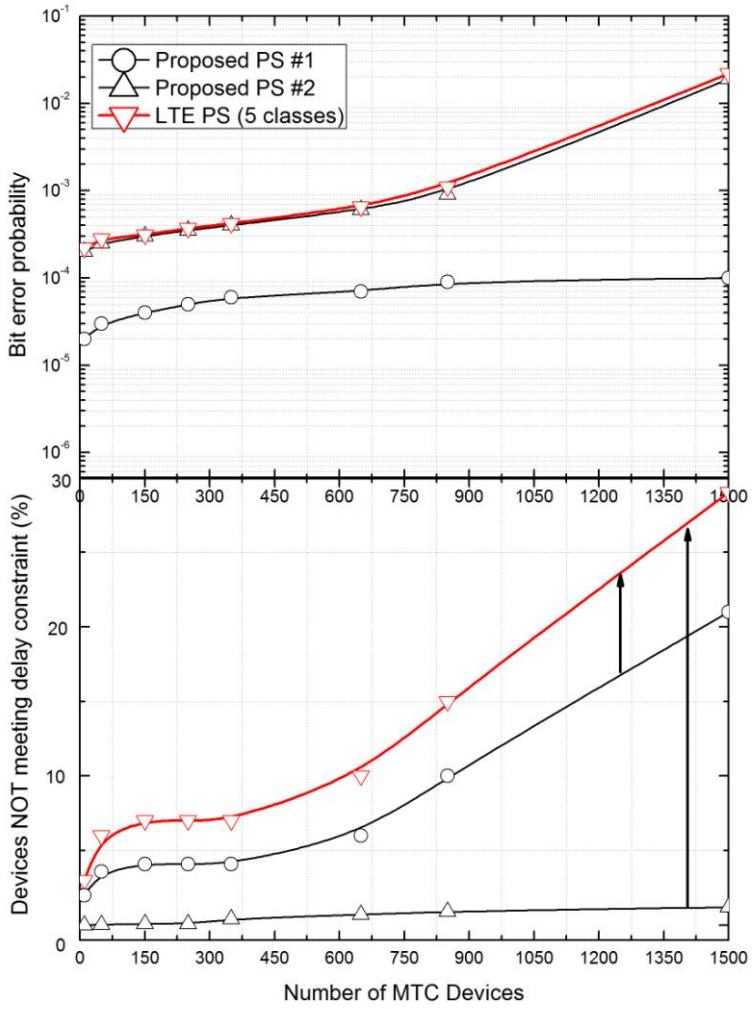


Figure 5-9: Performance of the proposed PS algorithms.

The number of delay classes, which lead to near optimum performance, heavily depends on the traffic statistics and the number of devices in each class. Given the number of classes that can be supported (e.g. depending on the available feedback bandwidth, a near optimum strategy is not to create classes uniformly (e.g. 1, 100, 1000, 10^4 ms, etc.), but according to the number of devices around a specific delay tolerance value. Moreover, the impact of the number of QoS classes is depicted in Figure 5-10.

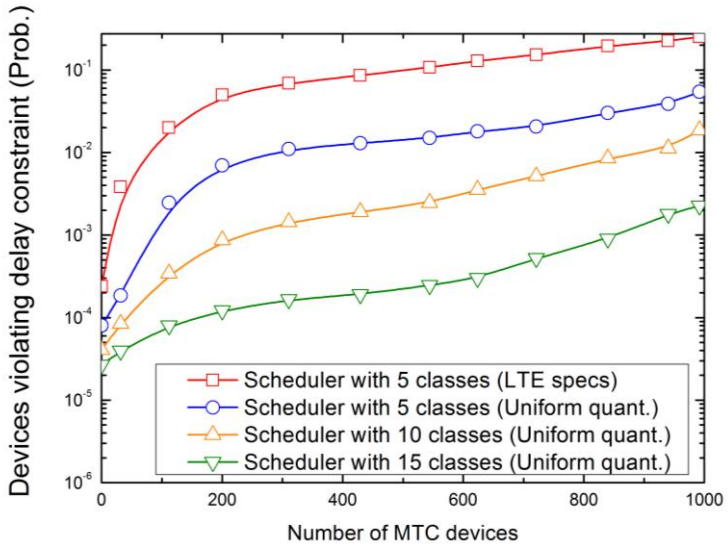


Figure 5-10: Dependency between the delay classes' range of values and the number of devices per class.

5.2.7 GFDM

With GFDM, potentially a higher number of devices can be served, when compared to traditional OFDM and SC-FDM approaches. This is due to two reasons that will be presented and evaluated in the following.

Consider the typical LTE FDD setup: Within a bandwidth of 30.72 MHz, there are 2048 potential subcarrier positions. From these, the 1200 centre carriers are in use which, leads to 30.72*1201/2048 = 18.015 MHz of occupied bandwidth, which are then transmitted in a 20 MHz band. The 2 MHz difference remains as guard band which are distributed on both sides. In GFDM, the flexible subcarrier pulse shaping enables to occupy more subcarriers while maintaining the spectral mask of the 20 MHz band. Filling these extra 2 MHz with useful data can increase the number of resources and thus also the number of supported devices by up to 10%. This is exemplarily shown in Figure 5-11.

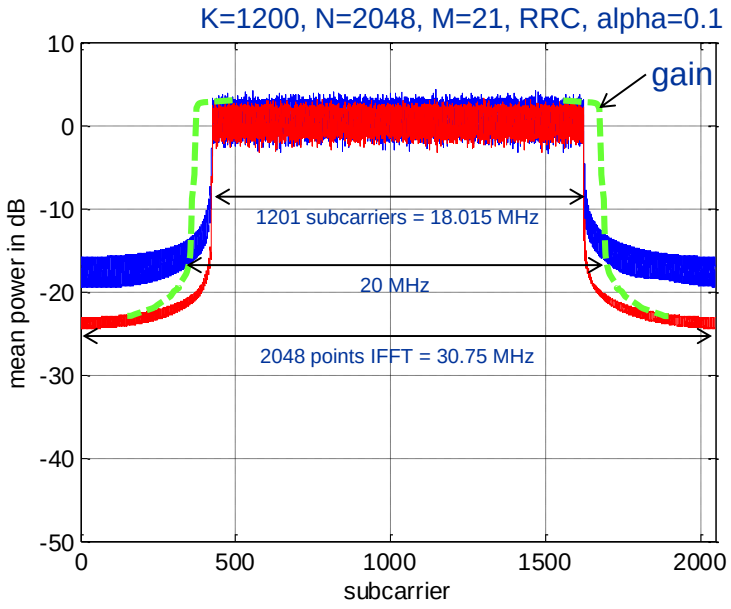


Figure 5-11: Potential frequency resource savings.

Further potential to increase the number of supported devices lies in the granulation of the resources in the time dimension. While an OFDM/SC-FDMA based system transmits symbols which are each attached with a cyclic prefix, in GFDM several symbols are combined to one block. See exemplarily Figure 5-12.

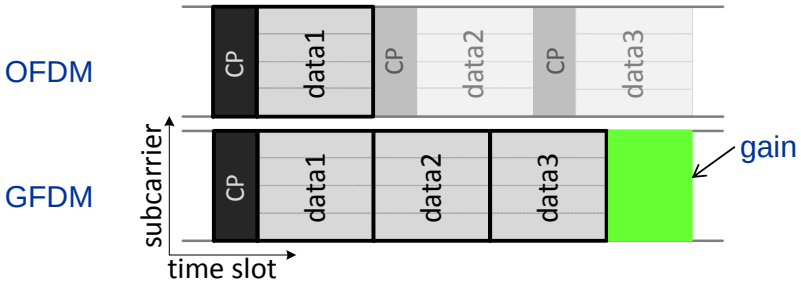


Figure 5-12: Potential time resource savings.

Suppose one GFDM block is chosen to be the length of two subframes. This corresponds to 14 OFDM symbols and 14 prefixes in contrast to M GFDM symbols in one GFDM block and 1 prefix. Depending on if normal or extended mode is used, the length of the prefix varies. In normal mode, for a duration of two slots, which is one subframe of 1 ms, a total of 14*66.67us are spent transmitting data, while 2*5.2+12*4.69us are spent transmitting prefix. A corresponding GFDM system could use 14.9*66.67us for data transmission and would require only 5.2us for the prefix. This gives additional 16.6% of additional resources. The gain is even larger for the extended mode, where OFDM/SC-FDM spends 12*66.67us on data and 12*16.67us on prefix, in contrast to 14.75*66.67us of data and 16.67us of prefix in an equivalent GFDM case. The gain here is up to 22.9%.

Even if transmitting a fraction of a symbol remains questionable, the potential for significant gains is evident. Exploiting gains both in time and in frequency dimensions together, the number of served devices could potentially be extended by up to 35% with GFDM.

5.2.8 Scheduling algorithm for heterogeneous traffics

The proposed method aims at scheduling the channel access in the downlink mode for devices of LTE and LTE-M systems that operate over the LTE downlink resources. Different RT and NRT traffic types are considered. More details concerning these traffic types are given in subsection 3.3.1. As described in TR 37.868 [30], the traffic for standalone LTE-M devices is generated according to a uniform law with a 60s period; this traffic will be referred to simply by “M2M”. On the other hand, the traffic flow for a capillary network behind a M2M gateway will be referred to by “M2M-GW”; parameters settings that have been adopted in the remainder of this section are as follows:

- 1’000 non LTE-M devices form the capillary network which is interfaced by the M2M gateway through the EPC;
- The gateway buffers all packets intended to the non LTE-M devices within the capillary network and then the gateway aggregates and forwards them in a ‘super-packet’ every T_{GW}=20s. It is out of the scope of this study to evaluate how these ‘super-packets’ are routed and disassembled through the right non LTE-M devices within the capillary network.

A system with homogeneous (single) traffic types is simulated, i.e. all active devices in the system receive data from the same traffic type.

Performance evaluation is assessed hereafter in the scope of the scenario 2 described in the section 3.3. The system is composed of a central cell surrounded by 2 rings of interfering cells; each cell is composed of 3 sectors. The simulations which were performed did just simulate the traffic flows for the devices within the 3 central sectors; meanwhile other sectors only generated inter-cell interference (19 cells, 57 sectors in total). At each TTI a maximum of 16 devices can be scheduled per sector, since there are 16 available control channels per TTI. Results are averaged over 10 independent dynamic runs, where at the beginning of each run devices are randomly uniformly located in the central cell. Each run simulates 120 seconds of network activity and at each TTI channel realisations are updated.

In this section 5 we prove how the investigated solutions permit to support a big number of devices. Thus Figure 5-13 presents the maximal number of devices that a sector can support with satisfaction, i.e. the achievable traffic load. Results were obtained by increasing the traffic load until satisfaction is not met any more.

More precisely, this maximum achievable sector traffic load is defined as the number of devices in the sector when more than 95% of the devices are satisfied. Devices are considered satisfied if their packet error rate is below 2% and their transfer delay is below the characteristic time-to-live (TTL) of the traffic type (50ms for VoIP, 100ms for NRTV, 1s for HTTP, 60s for both M2M and M2M-GW traffics).

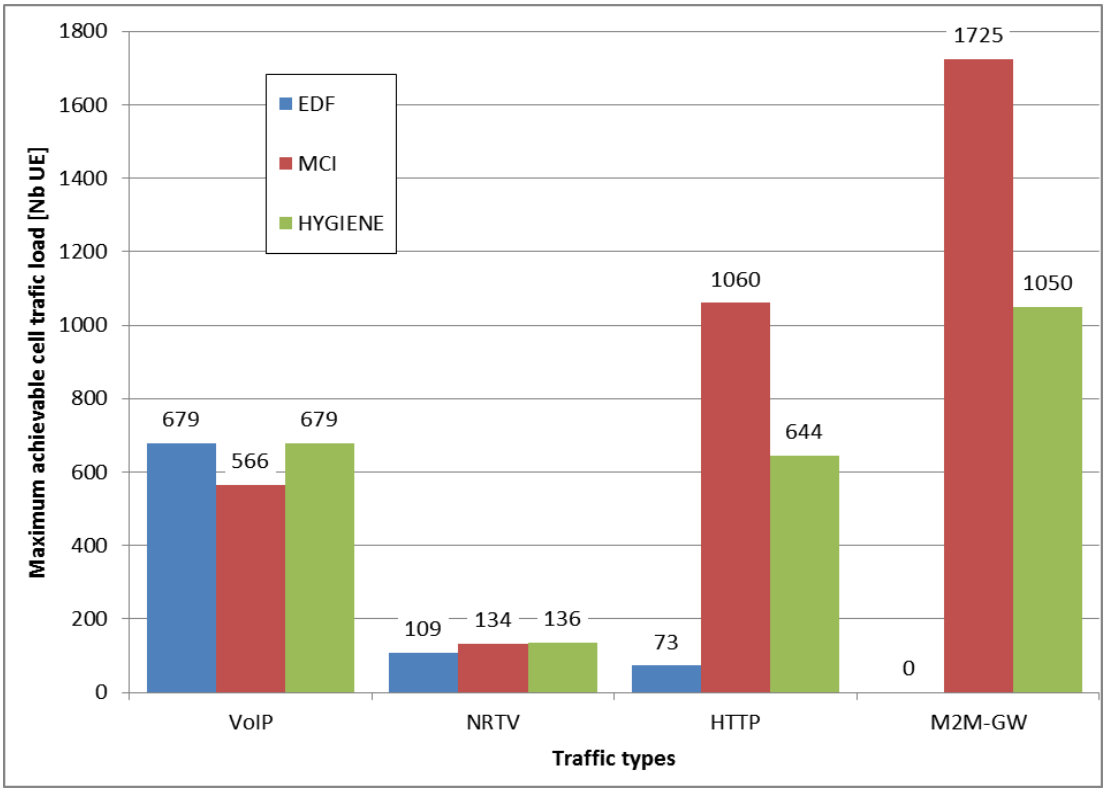


Figure 5-13: Maximum achievable sector capacity with EDF, MCI and HYGIENE schedulers for different single traffics.

Figure 5-13 compares the maximum achievable sector traffic load for all traffic types introduced previously (except the FTP traffic) for three different schedulers, respectively Earliest Deadline First (EDF), Maximum Channel to Interference ratio (MCI) and the proposed HYGIENE scheduler. EDF scheduler is designed to deal with RT QoS constraints regardless to the momentary device's channel quality, while MCI scheduler allocates resources to devices with the highest momentary instantaneous capacity, regardless to any traffic QoS constraints. HYGIENE tries to profit from benefits of both EDF and MCI without

suffering from their drawbacks. Consequently, by adjusting threshold according to the traffic type, HYGIENE can mimic EDF for RT traffic with 'rushing' entities or on the other hand HYGIENE behaves similar to MCI for NRT traffics with which few entities are 'rushing'. Results for EDF, MCI and HYGIENE are respectively shown on Figure 5-13 with the blue, red and green bar charts. Along the X-axis are displayed the results obtained for each homogeneous traffic type.

Some simulations have also been performed with the proportional fair scheduler that seeks to trade-off fairness and capacity maximization. This scheduler is known as being efficient for best effort traffics but not for RT traffics. For instance, the maximum achievable cell traffic load with HTTP traffic and the proportional fair scheduler is more than 1'600 UEs; then it serves better UEs than all other schedulers. Similarly, with a M2M-GW traffic, the proportional fair scheduler permits to achieve a cell traffic load greater than 1'400 gateways.

Results for the M2M traffic (i.e. for standalone LTE-M devices) are not shown since all simulations that we have run with more than 4'000 LTE-M devices permit to achieve a 100% satisfaction with all schedulers. Such schedulers but especially the system simulator is not designed to serve so many devices and the simulations have an amazing duration. So we propose not to focus so much on this traffic type.

Further simulations have been performed with a shorter period T_{GW} for the generation of messages at the gateways (M2M-GW traffic). The results presented above have been obtained with a period T_{GW}=20s. In case of a shorter period T_{GW}=1s, we have observed a huge reduction (around 80%) in the maximum achievable cell traffic load for MCI and HYGIENE schedulers with M2M-GW traffic. By reducing the period T_{GW}, the generated messages are shorter but more numerous, which leads to more requests for channel opportunity and then congestion. There is no change with the EDF scheduler that does not manage to serve UEs with MEM-GW traffic, especially with a shorted T_{GW}.

Lastly, simulations have been launched to evaluate the maximum achievable cell traffic load in case of heterogeneous traffics; i.e. when flows of two different types of traffic are simultaneously generated. For this purpose we have imposed to satisfy 200 gateways with M2M-GW traffic while trying to maximize the number of satisfied UEs served by HTTP traffic. The EDF scheduler fails in serving the least HTTP UE while managing to serve successfully the 200 M2M-GW gateways (just 180 gateways have been satisfied). In the other hand, MCI and HYGIENE schedulers both success in serving with satisfaction the 200 M2M-GW gateways and in addition the schedulers serve with satisfaction around respectively 400 and 50 HTTP UEs.

As it can be observed on Figure 5-13, HYGIENE permits to deal efficiently with all RT traffics by achieving the highest achievable sector traffic load. With NRT traffics, HYGIENE presents a reduction of the achievable traffic load in comparison to MCI but a huge gain in comparison to EDF. It seems that MCI outperforms HYGIENE for NRT traffics (both in case of homogeneous and heterogeneous traffics scenarios). This curious observation has not been explained but results most probably from an incompatible setting of threshold in the schedulers. Indeed HYGIENE has been designed to behave similarly to MCI with NRT traffics; the performance of successful scheduling should be then at least as well as those obtained with MCI.

Due to less stringent QoS constraints for the M2M traffics in comparison to the LTE traffics, the schedulers permit to support and satisfy more M2M devices than LTE devices. Recall that for the traffic 'M2M-GW', the maximum achievable sector traffic load refers to a number of M2M gateways that interface a capillary network with 1'000 non LTE-M devices. Results would be different if the number of non LTE-M devices within the capillary network changed.

5.3 Conclusion

This section evaluated the benefits LTE-M can gain with respect to the number of users served. Typical machine to machine communication focuses on applications that support a large number of devices and therefore the different parts of the LTE-M system design have been designed to fulfil this objective. Section 5.2 showed a detailed assessment of the different parts of the LTE-M system design regarding the support for a large number of devices and significant increases were shown.

The biggest gains were shown with solutions that take into account the specific nature of the M2M applications such as relaxed and diverse delay constraints or periodic traffic patterns. The semi-static scheduling option of LTE-M shows gain factors between 5 and 10 with realistic assumptions, which translates to an increase in the number of users of 500%-1000% only due to the exploitation of periodic messaging of LTE-M devices. The accompanying solution to calculate the periodicities for semi-static scheduling, the AGTI scheduling framework, can show that even when introducing such periodic scheduling rule, the delay constraints of the different nodes will be met while keeping the signalling low. The scheduling option that incorporates also the QoS of different M2M devices shows similar gains. At a delay violation probability of $2x10^{-3}$, the number of devices using this scheduling option can be increased from less than 100 up to 1000 which relates to a gain of about 1000%. Most of scheduling algorithms are not designed to support traffics whose QoS constraints may vary in a large range of values. In case of very heterogeneous traffics, an adaptive specific scheduling option can benefits from advantages of common scheduler approaches but without their drawbacks and shows best performance in terms of number of satisfied users for all different kinds of applications varying from VoIP communication and NRTV to M2M traffics, and yields a jointly optimal solution to achieve the support for large number of devices.

Besides periodic message communication some applications are run in an event driven manner such as alarms. These devices need to have efficient and reliable access to the PMRACH even if the number of devices that simultaneously try to access the PMRACH is very high. To deal with the increased number of devices slotted access was proposed as an integral part of the LTE-M system design and simulations showed the gains of this scheme. When using slotted access, the number of users facing the same collision probability is increased from 100 to 900. This refers to an increase to 900%. To make the transmission of packets in the PMRACH more reliable against collisions, network coding was introduced within the LTE-M system design. The LTE-M RACH is therefore able to increase the throughput of the PMRACH up to 80% of normalized throughput compared to traditional solutions such as slotted ALOHA (only 37%) and CRDSA (up to 60%).

The optimization for resource usage shows also significant gains in terms of number of supported users. The introduction of GFDM as the radio access method of the LTE-M uplink gives benefits both in the usage of time and frequency resources. While by exploiting frequency localisation of GFDM, 10% more users can be served the specific block structure of GFDM in time gives rise to another 23% of users additionally served. Combined time and frequency gains yield 35% increase in number of supported users. Additionally, the HARQ scheme of LTE-M was optimized in a way to reduce resources used. When applying the optimized scheme, up to 30% more users can be served due to newly available resources for transmission. The introduction of new registering information can be seen as an enabling technique that further supports the efficient service for a large number of devices.

One of the key objectives of the LTE-M system design is the support of a large number of users as this is one drawback of currently used LTE Rel.8 solutions. This section showed the

different gains, that the mechanisms of the LTE-M system design could gain in various different applications yielding gains from 30% up to 1000%. The interplay of these mechanisms may further increase these gains. Finally, an increase in number of supported devices of an order of magnitude seems achievable for the LTE-M system design.

6. Objective 3: Provision of Wide Area Coverage

Provision of wide area coverage is a critical feature of LTE-M, since servicing embedded devices without network coverage may quickly become very expensive, a challenge that has been addressed in the 3GPP contribution [43]. This section evaluates techniques aiming at maximizing the coverage of LTE-M. An integrated look at the solutions provided highlights the advantages, which are then evaluated numerically in the EXALTED scenarios. Both uplink and downlink techniques are considered.

6.1 Proposed Solutions

Mechanisms proposed in EXALTED aiming at a reduction of complexity and cost of LTE-M devices, e.g. the restriction to one single antenna or a lower transmit power, may lead to impairments in the link budget and in consequence downsizing of the coverage area. Therefore, LTE-M must provide means for coverage extension. A variety of models exist for cellular propagation, all of which include a distance-dependent path loss component and a random attenuation known as shadow fading. In order to achieve very high signal availability levels, two approaches are provided. The first approach aims at improving the link budget and increasing the fading margin of the link by means of a CDMA overlay. The second approach aims at combating shadow fading by providing spatial macro-diversity through a collaborative retransmission approach. In statistical terms, the CDMA overlay approach improves mean received signal strength, while the collaborative retransmission approach changes the received signal strength distribution, reducing its variance. In both cases, an improvement in coverage can be achieved. The problem of coverage extension for M2M communications is discussed in 3GPP as well. Two contributions with potential solutions were submitted [44], [45].

In scenarios with isolated LTE-M devices or with severely changing radio propagation conditions, a CDMA-overlay is an effective approach; in particular for LTE-M devices with limited transmit power. The principle of the CDMA overlay is shown in Figure 6-1. One single subcarrier is considered and the frequency dimension is omitted in the following for the sake of simplification. The left figure depicts the consecutive transmission of three data symbols in the uplink. Without loss of generality, perfect synchronization is assumed.

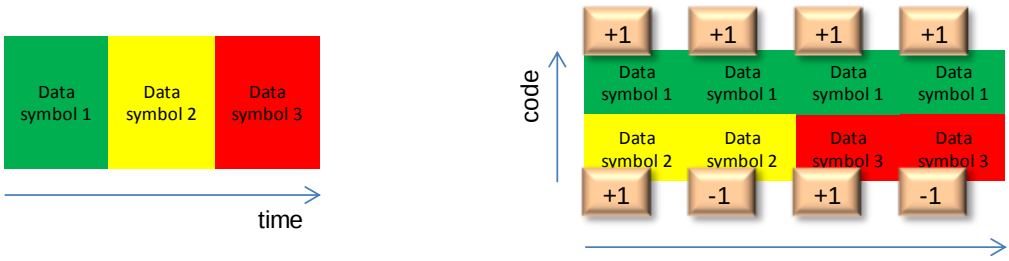


Figure 6-1: Transmission of data symbols (a) without and (b) with CDMA overlay.

In order to achieve the required link budget improvement for the green LTE-M device, its transmission time is extended in the right figure from one to four time steps, e.g. SC-FDMA or GFDM symbols. The eNodeB accumulates the received energy until the message can be decoded. To minimize the impact on the other two devices (yellow and red), which may be subject to a certain delay constraint, they are allowed to transmit their data in parallel to the green one. Thus, the flexibility for the allocation of radio resources is increased. User separation is done by multiplying the data symbols at the respective transmitters with orthogonal codes and by the usage of a correlation receiver at the eNodeB.

The codes {+1, +1} and {+1, -1} with code length 2 in Figure 6-1 are just an example for the explanation of the principle. In practice, it is most likely to use longer codes, which allow a bigger number of parallel transmissions, but also increase the minimal transmission time for one data symbol required for the user separation at the receiver. But for the required low data rates in LTE-M this is not critical. Assuming ideal conditions, there is a linear relationship between transmission time and accumulated energy at the receiver, i.e. expanding the transmission time by a factor of 2 yields a link budget improvement of 3 dB, given that the transmit power in both cases is the same. There is no limitation of the transmission time of one data symbol. As shown for the green device in Figure 6-1 the code sequence (here {+1, +1}) can be repeated arbitrarily.

The superposition of modulation symbols with orthogonal codes is fully compatible with the proposed LTE-M access scheme. The principle can be applied in both uplink and downlink, and the superimposed modulation symbols may belong to one single LTE-M device or to different LTE-M devices. However, the uplink is probably the bottleneck in the LTE-M system due to the assumed low transmit power and limited flexibility at the devices. Thus, it is probable to utilize CDMA overlay primarily in the uplink.

While the CDMA overlay approach increases the received signal strength, the resulting improvement may not be sufficient in locations suffering severe shadow fading. Such situations are more frequent in urban environments, where the density of devices is medium to high. An alternative approach consists in combating shadow fading with macro-diversity, i.e. retransmitting packets from different geographical locations. A collaborative approach is provided, based on random linear network codes. The system constitutes a distributed, binary encoding system, which adds the advantages of spatial diversity to those of time diversity which can be provided by rateless codes in the broadcast scenario. Unlike the previous technique, this one is primarily intended for the downlink [46].

In the proposed system, data packets (also “source packets”) are represented in a Galois field (GF) of size 2^q (GF(2^q)) and encoded at the eNodeB. An encoded packet is obtained as a linear combination of source packets in GF(2^q). Upon reception, intermediate nodes re-encode and re-transmit packets, which are finally decoded at the destination. In order for the destination node to recover the source packets, the coefficients of the linear combinations are sent along with each encoded packet. The extra overhead due to coefficients is negligible if source packets are long enough. Within the EXALTED system concept, network coding is applied over both LTE-M + capillary air interfaces in broadcast transmissions. The problem at hand is depicted in Figure 6-2.

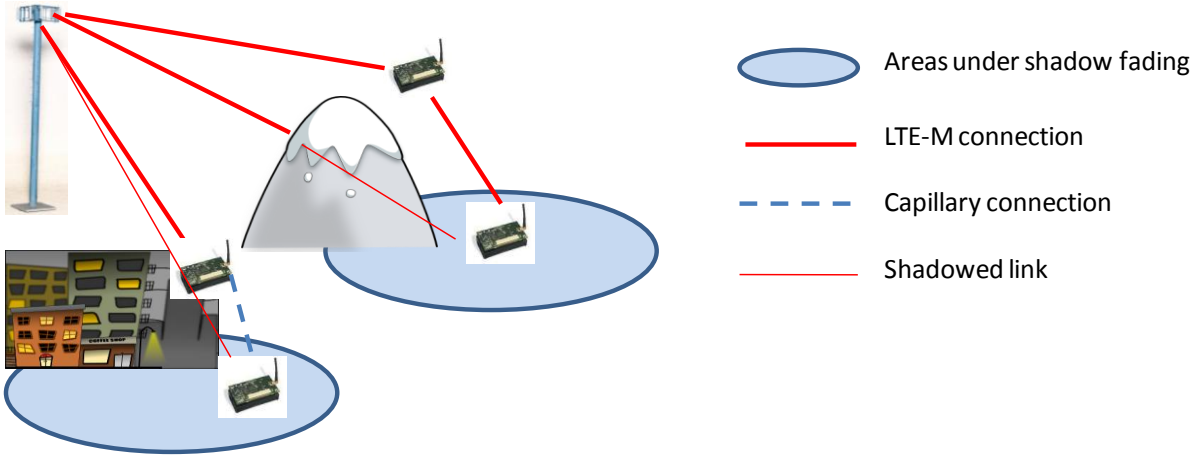


Figure 6-2: Descriptive diagram of the proposed collaborative broadcast architecture based on network coding.

Areas under shadow fading do not have sufficient coverage from the eNodeB, and rely on nearby devices to supply redundancy in order to decode the transmitted signal. Both LTE-M and capillary radio interfaces are combined at higher layers. In order to reduce the decoding complexity for nodes with good channel from the eNodeB, a rateless code (such as, e.g., a *Raptor code*) may be applied by the eNodeB, and encoded symbols are interpreted as linear combinations of source packets in GF(2^q) with coefficients equal to either 1 or 0.

6.2 Performance in Scenario 1

Both techniques were evaluated in Scenario 1 by means of simulations. Results on uplink improvement through a CDMA overlay and downlink improvement through the collaborative broadcast architecture are provided next. In both cases, scenario 1, which takes into account complexity and energy limitations of the terminals, is considered.

6.2.1 CDMA overlay

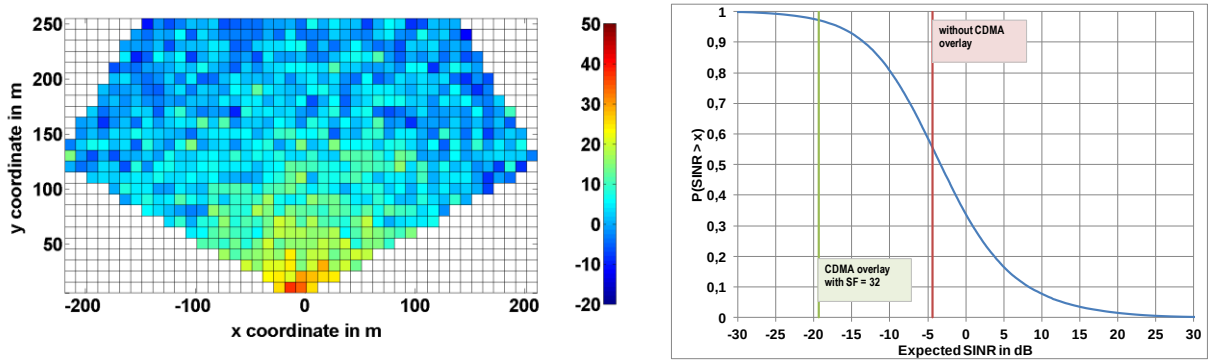
The intention of this subsection is to show to which extent the coverage can be enhanced by applying a CDMA overlay. The evaluation was carried out with a system level simulation tool. Among the alternative models introduced in section 3.2, a microcell environment with 500 m inter-site distance and 800 MHz carrier frequency was considered. Moreover, the greedy source traffic model was assumed. Other parameters were chosen as defined in section 3.2. The definition of coverage implies that the expected value of the SINR at one specific point in the cell must exceed -4.4 dB to belong to the coverage area. This value was derived from AWGN link level simulations with QPSK as modulation scheme and code rate 1/9.

The impact of the CDMA overlay on system level is modelled as an additional spreading gain. In other words, the perceived SINR is improved by the spreading factor. As an example, a spreading factor of 32 yields a SINR improvement of $G = 10 * \log 10(32) = 15\text{dB}$.

At the receiver (LTE-M enabled eNodeB) four antenna elements with λ/2-spacing are available. It is assumed that the combining algorithm adds up the amplitudes of the desired user signal, but the powers of noise and interference. This leads to an improvement of 6 dB in the link budget. Diversity effects and sophisticated interference rejection algorithms are not considered in this study.

Figure 6-3(a) shows the resulting area distribution of expectation values of the SINR in the cell in dB. Red colour, as visible close to the eNodeB at the bottom, indicates high values, whereas blue colour corresponds to low values. It is assumed that the transmitting LTE-M

device is located at the respective point in the cell and the colour coded SINR value shown at this point is measured at the eNodeB. The area distribution translates to the inverse cumulative density function depicted in Figure 6-3(b) (blue curve). The y-axis gives the probability that the perceived SINR at a random point in the cell exceeds the value on the x-axis. At $x = -4.4\text{dB}$ (red line), this corresponds to the achieved coverage without CDMA overlay, which is $C = 55.5\%$. The reason for this extremely bad value is primarily the reduced transmit power of the LTE-M device.



(a) (b)

Figure 6-3: (a) Distribution of expected SINR in the cell and (b) corresponding cumulative density function.

The capability of CDMA overlay to improve the link budget relaxes this problem considerably. Assuming a spreading gain of 15 dB, the improved coverage is now determined at $x = -19.4\text{dB}$, and it is extended to $C = 97.0\%$ of the cell area.

6.2.2 Collaborative broadcast architecture

The collaborative broadcast architecture was simulated according to the parameters defined in section 3. In the multicell layout, we considered a single cell broadcast only. It can be expected that a multicell broadcast approach similar to the Multimedia Broadcast Multicast Service (MBMS) defined in LTE will yield even better coverage results, since co-channel interference from adjacent cells or sectors is eliminated, and macro-diversity is provided when multiple cells broadcast the same signal. However, it may also be the case that broadcast transmissions only involve a limited geographic area and then the MBMS solution may not be adequate. Therefore, our results focus on the more limiting single cell broadcast case.

As explained in deliverable D3.3 [1], the Collaborative Broadcast Architecture (CBA) leverages on both LTE-M and capillary interfaces to improve the coverage provided by LTE-M. This architecture is particularly useful for messages addressed to a large number of devices, such as software or firmware upgrades. The eNodeB transmits a broadcast message encoded with a physical layer inner code and a Raptor outer code. The message blocks are then received by LTE-M nodes and those having a capillary interface retransmit a random linear combination of received packet locally, and with a certain probability. The resulting network code is decoded using standard decoding techniques by all LTE-M nodes.

Performance was evaluated in a vehicular environment where it can be assumed that **i)** the density of nodes is high, and **ii)** vehicle-mounted nodes are not energy-constrained and may afford to retransmit a large number of packets through their capillary interface. Nodes were arranged in a Manhattan-like street grid as shown in Figure 6-4. Two different node densities were considered: high, at 500 vehicles per sector, and low, at 100 vehicles per sector. Cell size and characteristics were set according to section 3, as well as all propagation parameters. Vehicle speed was set at 50 km/h. 802.11p DSRC was assumed as secondary

air interface, with a maximum range of 50m. A 400-byte message is transmitted, split into ten 40-byte packets, which constitute the basic unit (symbol) of the Raptor and network codes. The Raptor rate is 1/6, while the PHY layer code rate is 1/3. Physical layer abstraction is used to simulate the LTE-M link.

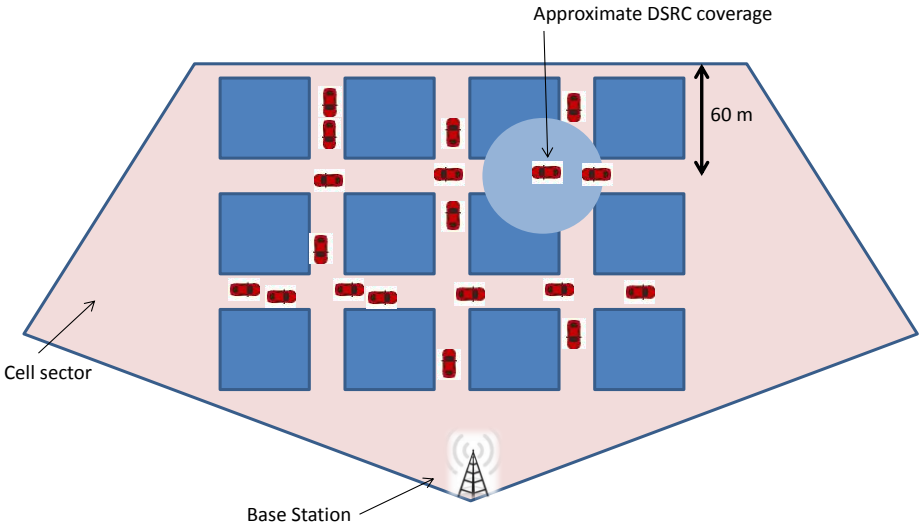


Figure 6-4: Manhattan grid used in the CBA simulations.

Broadcast coverage is shown in Figure 6-5 and Figure 6-6 for the high and low density scenarios, respectively. As it can be seen, in the high density scenario coverage is well above the 99.9% mark if the optimal average number of retransmissions ζ is used. On the other hand, the non-collaborative system and the simple relay system peak at around 98% and 98.2%, respectively. In addition, coverage for the CBA peaks at around ζ≈0.1, a relatively low value, which means it is not necessary to make extensive usage of the capillary interface in order to improve coverage significantly.

In the low density scenario it can also be seen that coverage of the CBA is above the 99.9% mark for low values of ζ, and that it clearly outperforms the non-cooperative and relay cases.

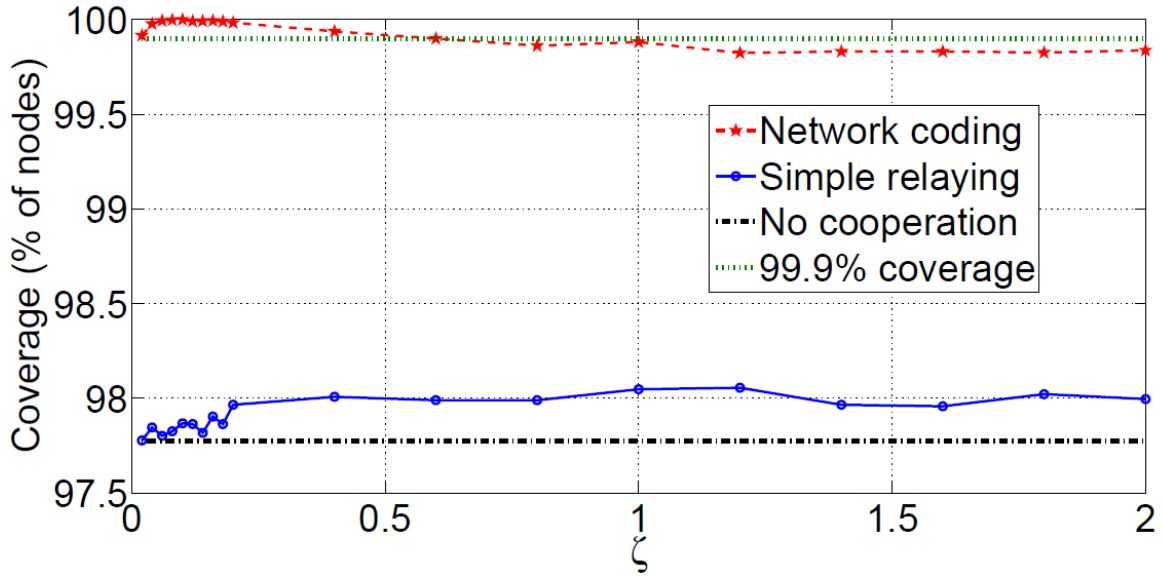


Figure 6-5: Broadcast coverage as a function of the average number of retransmissions per node, in the HIGH density scenario. Simple relaying and no cooperation are shown for comparison.

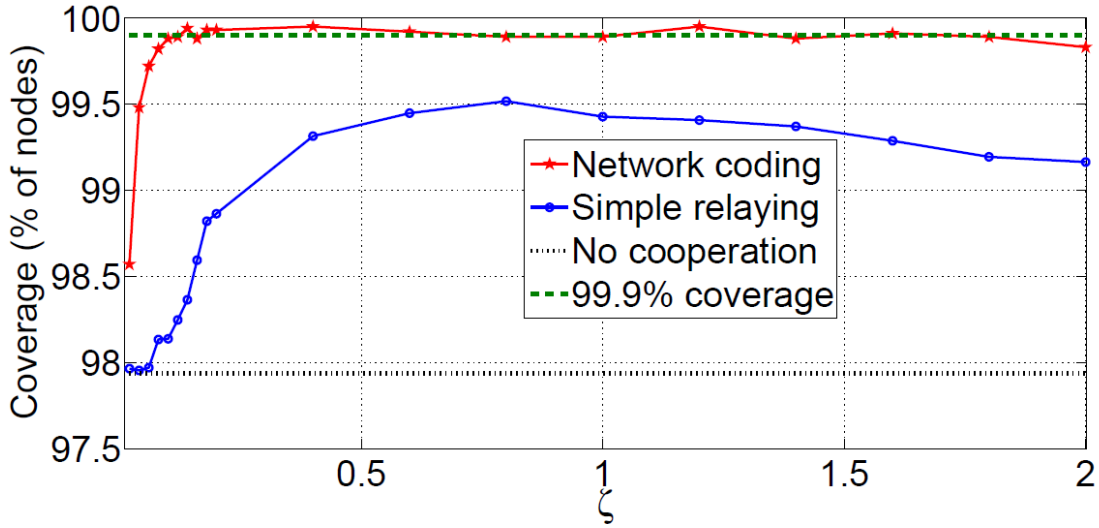


Figure 6-6: Broadcast coverage as a function of the average number of retransmissions per node, in the LOW density scenario. Simple relaying and no cooperation are shown for comparison.

6.3 Conclusion

In this section coverage improvements have been evaluated by two specific techniques designed in EXALTED for that purpose.

For the uplink, we can conclude that CDMA overlay solely, even with high spreading factor, is not sufficient to obtain full coverage. Either this technique has to be combined with other techniques or a sophisticated receiver capable to suppress interference must be applied.
For the downlink it can be seen that the proposed Collaborative Broadcast Architecture achieves very high level of coverage under the simulated conditions, and offers a considerable improvement with respect to the non-collaborative case. As a result, in some favourable situations, it may be possible to use higher coding and modulation rates to improve the throughput when transmitting large files.

7. Objective 4: Energy Efficiency

Energy efficiency is fundamental to the concept of EXALTED, as terminal devices may need to be autonomous for months or years. EXALTED aims to develop a cost, spectrum and energy efficient radio access technology for M2M applications, adapted to coexist within a high capacity LTE network. To facilitate this, new mechanisms are developed within the LTE extensions.

7.1 Proposed Solutions

Since this section is related to the energy efficiency objective, firstly the terminal power consumption breakdown model is presented. The model is used as the benchmark for further performance evaluations for LTE-M devices in comparison to LTE-M system. Following that detailed evaluations of proposed algorithms is done in comparison to the proposed model.

Battery power consumption breakdown:

Regarding terminal (device) battery power consumption basically two main modes are available, i.e. when terminal is in IDLE or in ACTIVE mode. Some preliminary calculations, that can be used for energy efficiency evaluations are presented:

1) In the case that terminal is powered by AAA battery, then let's assume capacity of the battery to be maximal i.e. 1200 mAh (typical capacity of AAA batteries is in the range 250-1200 mAh), and a nominal voltage of 1.5 V. That corresponds to theoretical energy of 6.5 kJ. Depends on the power consumption of device, standby lifetime will be shorter or longer.
2) For the calculation of used energy per bit transmitted, the following calculation model is proposed: Maximum output power in the EXALTED project is assumed to be 17dBm, which corresponds to 50mW. Let's also assume that transmission of the data is 100kbps. In this case, energy per transmitted bit is 0.5 µJ. If we consider a device that sends 10000 bit measurement report every hour, then it corresponds to 3bps, and a theoretical consumption of 1.5 µW.

The analysed model of terminal power consumption model is simplified, so power amplifier losses, power conversion losses are not taken into account. Also, since the most of the terminals should be very simple and cost efficient, we can assume that application part of the platform has a very small influence on the power consumption (what is contrary to regular LTE devices). In EXALTED it can be assumed that the terminal is dominated with radio platform consumption rather than application part (so there is no power demanding application processor and memory).

<u>Simplified M2M terminal should have:</u>

- Application platform (which is minimized or almost avoided);
- Clock which is always running in order to activate the terminal for scheduled events such as monitoring paging, performing measurements and performing periodic keep-alive messages.
- Receiver:
  a. Receiver front end (low noise amplifier) and mixer;
  b. Filter and ADC (Analogue to Digital conversion);
  c. Receiver baseband:
     - Detection and synchronization;
     - Measurements;

- Channel estimation;
- Demodulation;
- Decoding;
- Channel quality estimation;
- Protocol processing.

- Transmitter:
  a. Power amplifier;
  b. DAC (Digital to Analogue conversion);
  c. Transmitter baseband:
     - Modulation and pulse shaping;
     - Transmit buffering;
     - Encoding;
     - Protocol processing.

IDLE mode Power breakdown consumption:

Consumption of the device in IDLE mode is assumed to be 5 mW.
In Table 7-1 is given simplified model of power consumption for IDLE mode:

Table 7-1: Battery power consumption breakdown in IDLE mode.

	Application	Paging monitoring and measurements	Clock	Total power consumption
Power Consumption	4.1 mW	0.5 mW	0.4 mW	5 mW

ACTIVE mode Power breakdown consumption:

In ACTIVE mode power consumptions obviously depends on the traffic, i.e. the number of bits transmitted. Consumption of the device in ACTIVE mode is assumed to be 500 mW in average (in general it depends on used bandwidth and transmission rate).

Table 7-2: Battery power consumption breakdown in ACTIVE mode.

Parameter	Percentage of Total power consumption
Power Amplifier	21%
RF Tx part	24%
RF Rx part	17%
ADC + DAC	10%
Application + Clock	2%
Rx Baseband	17%
Tx Baseband	9%

Presented power consumption numbers in Table 7-2 should be seen as indicative number mainly showing the order of magnitude of power consumption and relationship between different parts in mobile platform.

Energy efficiency

The mechanisms that enable LTE-M to achieve the goal of energy efficiency of supported devices will be summarized shortly in the following. More detailed information can be found in D3.3 [1].

The minimization of signalling is one key prerequisite to support a large number of devices (more about that can be found in the section 5), but there is also another important reason to do so, namely the lifetime of battery in the device, i.e. more generally spoken the energy efficiency of the system. While signalling in LTE was designed for broadband applications with a large payload size, in EXALTED it is focused on the optimization of signalling with respect to sporadic short messages in order to minimize the energy required in the device to process and transmit such a short message. In LTE, random access is used to acquire a reserved time slot for transmission, and is managed by the eNodeB. In LTE-M, proposal is to use a slotted random access channel to transmit very short data packets directly, which can potentially reduce the number of transmissions and listen periods for the devices with consequent energy savings. A novel collision recovery mechanism is proposed in which the decoder retrieves information from the collided signals using physical layer network coding. This permits to achieve a lower energy cost (i.e., number of transmissions) per successfully delivered message, thus increasing the energy efficiently with respect to previously proposed random access schemes.

A cooperative broadcast architecture for coverage enhancement based on network coding is proposed where the terminals send out linear combination in a finite field of previously received messages. If a node receives enough linearly independent combinations of the source messages (i.e., messages to be broadcasted by the eNodeB) than it is able to recover all the broadcasted data by applying common matrix manipulation techniques. One of the advantages of such an approach is that no signalling (such as packets requests) is assumed among the cooperating nodes, thus reducing the energy consumption and the radio resources occupancy. Results show that the full diversity gain of the proposed method is achieved for low values of medium access probability, which suggests that the enhancement in coverage is obtained with a relatively low energy consumption increase with respect to the non-cooperative case.

A significant proportion of M2M device are expected to rely on batteries, without direct access to a power supply, and without the possibility to replace or manual recharge the batteries. In such cases, the limited battery size translates into a finite lifetime of the device, and possibly of the system. Consequently, energy efficiency is a key challenge in the sustainable deployment of battery-powered communication systems. A complementary approach has recently been made possible by introducing rechargeable batteries that can harvest energy from the environment in order to extend the lifetime of the system. The objective pursued here is that of designing the system operations taking into account the nature of the energy harvesting process to increase the overall energy efficiency (in Joules per bit). Considering a point to point link, the proposed solution maximizes, through power management, the amount of data that is transmitted from the source to the destination within a given time interval and under various assumptions regarding the energy harvesting model and battery limitations. Figure 7-1 illustrates how the optimal transmit power (derivative of the cumulative transmit energy curve $E^{opt}(t)$) relates to the cumulative harvested energy curve $H(t)$, and minimum energy curve $M(t)$ (which is due to the battery size constraint).

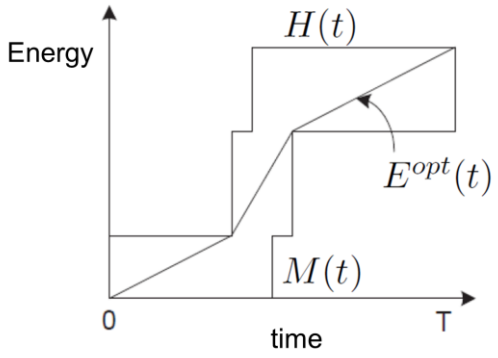


Figure 7-1: Illustration of the optimal transmission strategy for a discrete harvested energy flow.

RRC protocols and procedures are considered in order to create and define more efficient procedures for optimum utilization of radio resources. In order to enable the networks to differentiate between different terminals and their requirements, a new procedure is proposed that enables terminals to describe themselves and their specific capabilities, traffic generation patterns and service requirements. Information that is already stored in the network about the terminals is extended with new information. These additional descriptions are included in the existing signalling messages by extending some of the fields and stored in networks' HSS, which have to be extended as well. The proposed algorithm introduces extension of messages used by the terminals when registering to a network and extension of information that is supposed to be saved in the HSS will be investigated and proposed. By extending the system information messages, and using the knowledge about the type of terminals attached to a cell, the network will be in a position to specify different parameters for different groups of terminals thus enabling different treatment of LTE-M devices and making it possible to use modified protocols and procedures like the frequency of radio measurements. The proposed algorithm overcomes the main issues regarding monitoring of the paging channel, i.e. consumption of significant amount of energy in the IDLE mode. The LTE-M device determines the paging DRX cycle and other relevant paging parameters by processing the serving cell paging parameters broadcasted by the serving cell. Power consumption reduction during inactivity can be achieved by increasing the paging cycle, i.e. by introducing Paging flag (factor), and multiply DRX cycle by this factor (Paging factor should be defined in HSS per device). This factor will be downloaded on MME and delivered to LTE-M device during attach procedure.

Due to the characteristics of the M2M traffic (short transfers, mainly initiated from the devices towards remote servers), it is not required to have very smooth handovers and uninterrupted connections. The proposed algorithm is focused on optimization of the number of necessary radio measurements, and the possibility to allow the LTE-M devices to control the timing of measurements and to perform the required measurements only when needed. Frequency of radio measurements depends on DRX cycle, so it is proposed to Introduce Mobility flag (factor), and multiple DRX cycle by this factor that could be different for RRC_IDLE and RRC_CONNECTED states. Other proposed algorithm for minimization of network signalling regarding the monitoring paging channel of the LTE-M devices is focused on the number of paging messages sent towards LTE-M device during the paging procedure.

Multicast traffic is one important portion of the overall communication in machine-type environments. LTE-M devices are often grouped due to specific applications they are running. Such groups then often need to receive multicast messages due to configuration changes or software updates. Here LDPC codes, designed for incremental redundancy multicast are introduced. The channel codes are constructed [47] in a nested way and multicast transmission begins with a very short high-rate codeword which is initially sent out

and received by all LTE-M devices. After the initial transmission some LTE-M devices can already decode the multicast message and therefore need no longer to receive additional redundancy. They can go to sleep mode and preserve energy. Other nodes must listen longer to receive a sufficient amount of redundancy for decoding. The procedure is illustrated in Figure 7-2.

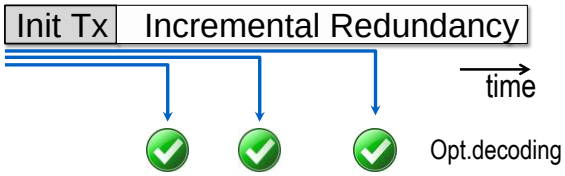


Figure 7-2: Incremental redundancy transmission with optimal decoding points.

Compared to the traditional solution energy savings can be accomplished on a per node basis as nodes with good channel qualities can go to sleep mode faster.

In LTE, downlink beamforming is supported resulting in considerably improved radio coverage. Considering a TDD-LTE system, the uplink and downlink take place on the same frequency, and hence based on channel reciprocity, the uplink sounding reference signals can be used directly to estimate the channel, which can then be used to derive the weighting for the downlink beamforming. A quite promising solution that has been also adopted in analysis is to use MU diversity with ZFBF at the transmitter together with a selection algorithm that is based on an orthogonality criterion. By using ZFBF, the weight vectors are appropriately selected in order to avoid interference among user streams. Additionally, a user selection algorithm that is based on orthogonality, the semi-orthogonal user selection (SUS), achieves near optimal sum rate, with, however, considerable less complexity, is also investigated.

Finally, MIMO configuration for transmit diversity is proposed, with 2 transmit antennas at the eNodeB and M>=1 receive antennas at the end device [48]. The targets of this algorithm are the following: reduce the number of CSI estimations at the LTE-M device by replacing the CSI estimator with a simple energy detector at each antenna element, reduce the number of required RF chains at the receiver, by selecting those antennas, which will utilize them and reduce the number of antenna switching operations, which lead to synchronization problems and larger delays. The basic feature of this transceiver is, in contrast to the conventional one it does not require to estimate the instantaneous channel gain for each of the receive antennas, but only for those that are eventually selected for the data decoding. In this way computational resources can be saved. Moreover, in addition to previous receiver, in order to mitigate the effect of excessive switching at the receiver, which may lead to synchronization problems and larger delays, an adaptive receive filter can be applied, according to the measured Doppler spread, please see figure below.

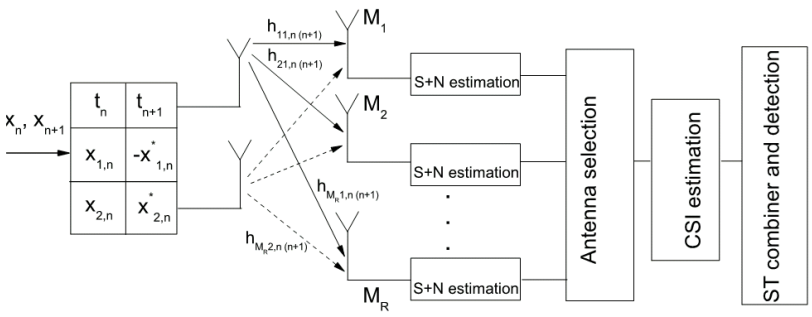


Figure 7-3: CSI and S+N-based MIMO transceivers.

The following sections evaluate the different gains introduced by the different parts of the LTE-M system design with respect to the objective of the support of energy efficiency.

7.2 Performance in Scenario 1

In the EXALTED project two different evaluation scenarios are defined in order to examine the performance of the LTE-M system with respect to the predetermined objectives and system requirements (more details can be found in section 3). The first scenario sets the context of low complexity and energy efficient M2M communications while the second one aims at supporting a large number of LTE-M devices with heterogeneous requirements and capabilities. This section is dedicated to objective Energy Efficiency which mostly corresponds to Scenario 1 case, i.e. low complexity and energy efficient M2M communications, where the battery lifetime is the most important metric. So, all analysed solutions for objective Energy Efficiency are related to Scenario 1.

7.2.1 Energy harvesting

The benefit of introducing energy harvesting aware transmission schemes is analysed here in order to minimize (through transmit power adaptation) the energy consumption of the LTE-M devices. A single point-to-point link is focused on, where the transmitter is equipped with energy harvesting capabilities. In the EXALTED system architecture, this transmitter could for instance be a LTE-M device or M2M gateway. The greedy source model is assumed as well as that a sub-band is permanently assigned to the device under investigation, leading to an interference-free scenario. Moreover, Shannon's formula is used to translate SNR values to spectral efficiency (or bit rate). With respect to the power breakdown presented earlier in this section, note that the analysis here focuses on the active mode, and considers only the power actually transmitted, i.e. a fraction (corresponding to the power amplifier efficiency) of the consumption appearing under the name "power amplifier" in Table 7-2. However, the study only considers the actual transmitted power, which constitutes only a fraction of the power amplifier consumption (depending on the efficiency of the power amplifier). As an illustrating example, for the harvested energy flow described in Figure 7-4 and the parameters provided in section 7.2.1, up to 25% transmit energy saving is achieved.

For the evaluation of the benefit of the proposed energy harvesting aware transmission scheme, the generic scenario depicted in Figure 7-4 is considered. It consists of a finite time interval *[0,T]*, during which two packets of harvested energy enter the system: the first one of E_0 joules at time $t=0$, and the second one of E_1 joules at time t_1. In particular, considering an energy harvesting rate (amount of energy which is harvested per unit of time) $R=16.65$ mJ/s (typically corresponding to a small size solar energy harvester), we have that $E_1 = t_1 R$ and $E_0 = T/4 R$ (assuming that the node has been inactive and harvesting energy during the time interval *[-T/4.0]*). Finally, for a 17 dBm transmission power (which corresponds to the maximal LTE-M device transmit power), we assume that the SNR at the receiver is equal to 15 dB.

The energy consumption associated with the optimal energy harvesting aware transmission scheme is compared (see green line in Figure 7-4) with that of a constant power transmission scheme (see red line in Figure 7-4) which would be the default strategy for a LTE UE. Figure 7-5 depicts the above-defined E_{energy} metric for different values of the ratio t_1/T. Remember that E_{energy} refers to the energy saving (in percentage) achieved by the proposed transmission strategy. Note that for a fair comparison, the energy consumption are compared for a same total amount of information (in bits) transmitted during the time interval *[0,T]*.

Let us analyse Figure 7-5. When t_1 is small, this corresponds to having all the harvested energy available at the beginning of the time interval *[0,T],* in which case the constant power transmission is optimal and no energy saving is achieved, as expected. However as t_1 increases, the energy saving increases, reaching for instance a value of 25% at $t_1/T=0.6$. Yet, let us stress that the value of the energy saving is strongly dependent on the actual flow of harvested energy into the system, and therefore should be considered as indicative only.

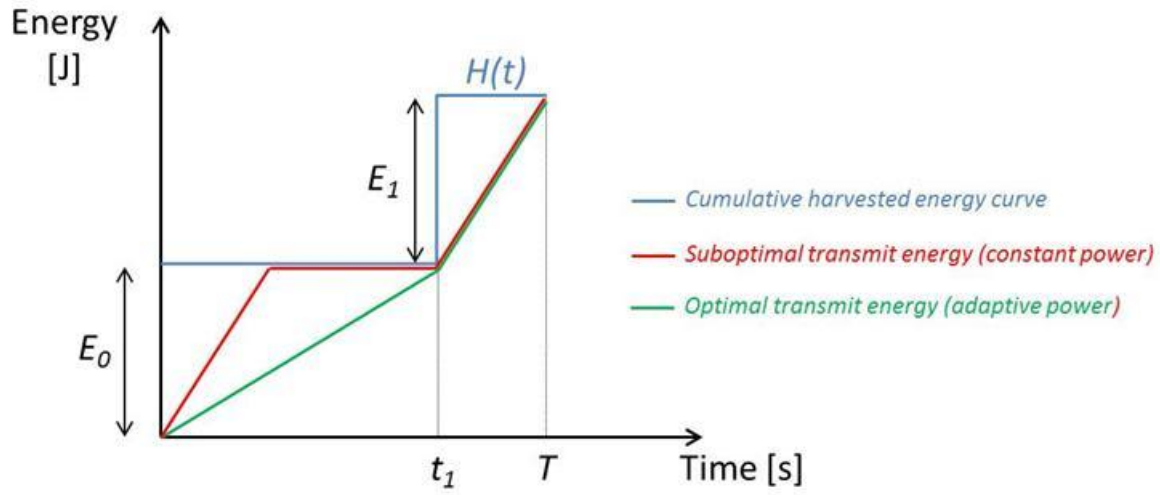


Figure 7-4: Generic energy harvesting flow.

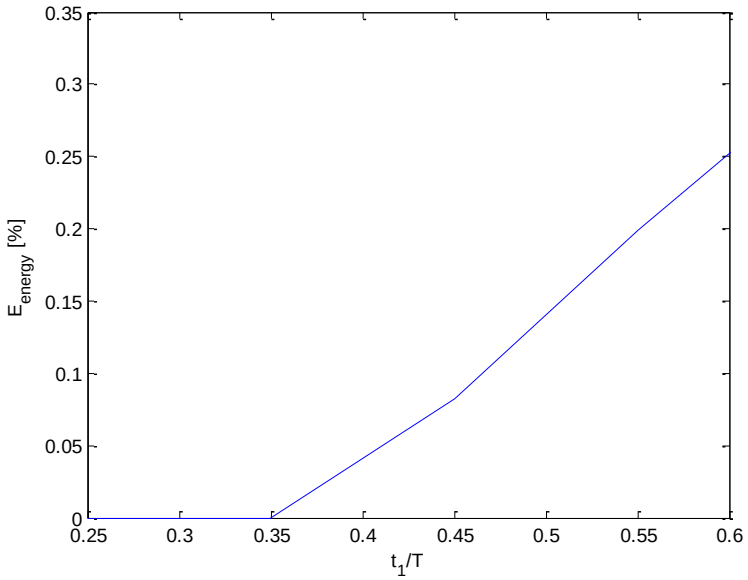


Figure 7-5: Energy saving E_{energy} for the generic energy harvesting scheme.

7.2.2 Collision recovery

The performance of the proposed LTE-M Random Access scheme with Collision Recovery, called NCDP protocol, was evaluated through simulations using the parameters defined in section 3. Details on the NCDP protocol definition can be found in EXALTED deliverable [1] and, to a lesser extent, in section 5 of this document. In order to assess the energy efficiency of the proposed solution, the number of packet transmissions required in order to deliver one successful packet is evaluated, which we define as η. With respect to the power breakdown presented earlier in this section, such number represents a multiplying factor of the terminal

active time for each received packet. For example, if η=1.5, then the device must be in active mode during 1.5 times the packet length in order to successfully transmit one packet. The proposed scheme presents no significant differences compared with benchmark Slotted Aloha, regarding idle time. In Figure 7-6 we plot the relationship between energy efficiency η and normalized throughput Φ, which is attained varying the normalized offered load. As it can be seen, Slotted Aloha is the most energy efficient solution for low traffic load, since each packet is transmitted little more than once on average. At higher traffic load, it becomes more efficient to use the proposed NCDP protocol or the benchmark CRDSA system equally, with 2 transmissions per frame. To attain higher throughput it is more efficient to increase the number of transmissions per frame to 3, or to use NCDP with uniform transmission probability in each slot p=0/0453. Therefore, as a function of the expected traffic load, different solutions should be chosen.

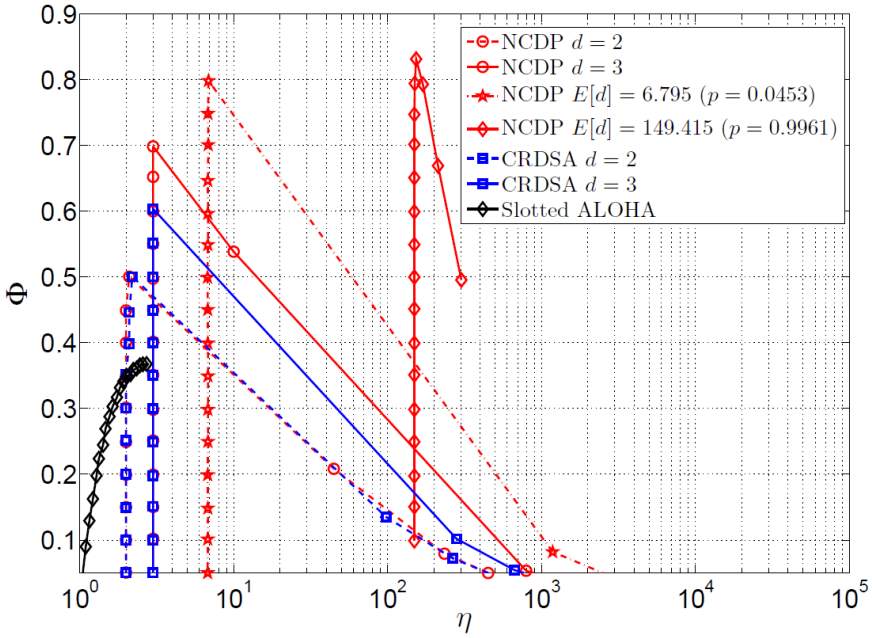


Figure 7-6: Throughput Φ versus energy efficiency η. Simulation parameters identical to those described in section 5.2.

For active mode, both transmitter (Tx) baseband and Tx RF, battery life can be improved by 33% with respect to Slotted Aloha for a normalized throughput of 0.36. The gains are smaller below that mark and for low load it becomes advisable to use Slotted Aloha. The NCDP protocol is able to attain higher throughput, up to 0.8, although battery life is reduced by 50% with respect to the 0.36 mark.

7.2.3 Collaborative broadcast architecture

The Collaborative Broadcast Architecture was simulated according to the parameters defined in section 3. Details on the CBA definition can be found in EXALTED Deliverable [1] and, which is also briefly outlined in section 6 of the present document. As in section 6.2.2 a single cell broadcast only and an identical simulation setup are considered. The CBA was evaluated in terms of energy consumption, measured as the average number of packets that must be transmitted by each node through the capillary air interface per received information packet (retransmissions/decoded messages). For this solution, increase in transmitted power cannot be related to the LTE-M device power breakdown presented earlier in this section. The reason is that the CBA increases active time and idle time of the capillary air interface, but leaves active time and idle time of the LTE-M air interface unchanged. Figure 7-7 and Figure 7-8 show the results for the high density and low density scenarios, respectively. In the figures, the vertical line shows the point where maximum coverage is attained. As it can

be seen, the number of retransmissions increases (and consequently the energy efficiency decreases) for higher values of ζ (which represents how many times a node retransmits every linearly independent packet it receives), so it does not make sense to go beyond that point. Therefore, maximum coverage is attained retransmitting 1.5 and 1 times each message, for high and low density scenarios, respectively.

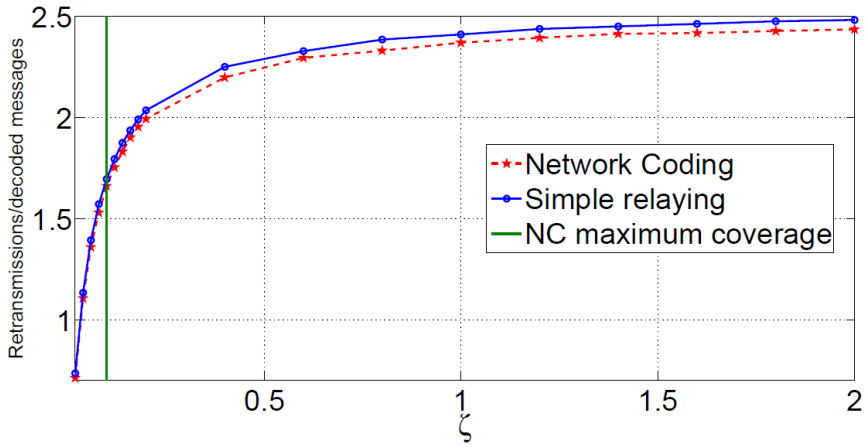


Figure 7-7: Retransmissions per decoded message as a function of the average number of messages sent through the secondary interface. HIGH density scenario (500 machine terminals / cell).

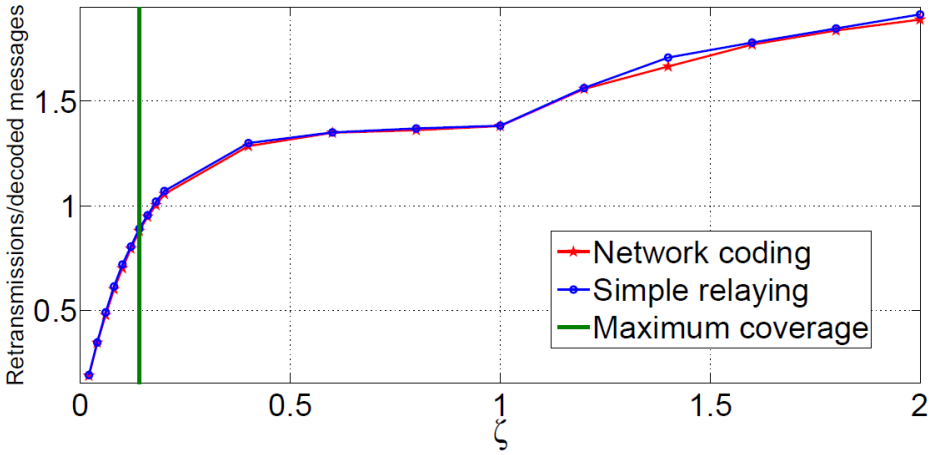


Figure 7-8: Retransmissions per decoded message as a function of the average number of messages sent through the secondary interface. LOW density scenario (100 machine terminals / cell).

7.2.4 Directional antennas

The main objective of the proposed algorithm is to reduce the signal overhead originated by the feedback load and this increases the nodes battery lifetime. In this context, a communication scenario is considered where a GW equipped with 2 transmitting antennas communicates with a large number of spatial distributed single antenna nodes. The wireless communication link between the GW and the nodes is considered to be affected by Rayleigh fading; additive white Gaussian noise is also present, while the SNR is considered to be equal to 10 dB. In the following figure the normalized feedback is plotted as a function of the number of users for several values of the normalized parameter β, which is used in order to optimize the throughput performance, considering different user selection algorithms. Specifically, we consider a) a user selection mechanism based on exhaustive search, b) proposed algorithm (based on zero forcing beamforming (ZFBF) with semi-orthogonal user selection (SUS)) with β=0.6 and c) proposed algorithm with β=1. In this figure, for all cases

feedback rates when ZFBF are employed are much lower, especially for the case of β=1. Furthermore, similar performances in terms of sum rates have been observed, with the algorithm 1 being always the optimum with the penalty of the highest complexity. Hence, it is reasonable to assume that based on the proposed algorithm the overall feedback bandwidth, could be significantly reduced, without considerable affecting the system performance, in terms of sum rates.

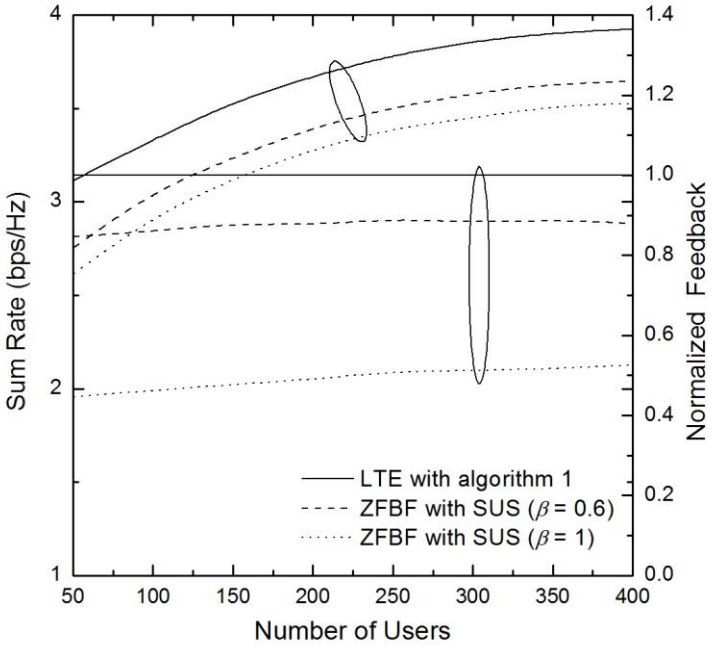


Figure 7-9: Normalized feedback as a function of the number of users.

7.2.5 LDPC codes for incremental redundancy multicast

To evaluate the performance of the incremental redundancy multicast solution with LDPC codes, extensive simulations according to the scenario described in section 3.2. The use of LDPC codes for incremental redundancy multicast aims at the reduction of energy consumption. Typically a code of rate 1/3 would be used to serve all devices in the cell. The proposed multicast scheme does not fix the code rate initially but starts the multicast transmission with a high code rate of R=0.75.

The assessment of the solution was done with respect to specific metrics, typical for communication systems:

- SNR (or Ec/N0): The average signal power to noise ratio is characterizing the quality of the given transmission link. In a cellular setting, the SNR is typically distributed in a specific manner over the whole coverage area. As the gains in terms of energy efficiency for the proposed multicast solutions are determined on a per node basis, considerations can be applied for specific SNRs
- Incremental redundancy packets: As the proposed multicast solution incorporates the subsequent transmission of several incremental redundancy packets, the number of such packets is the investigated metric. The fewer such packets have to be used for decoding in the receiver; the fewer energy will be spend at the receiving M2M device.

Simulations of this multicast scheme show (Figure 7-10), that for high SNRs, the initial transmission is suitable to serve the users. When the SNR starts to get worse, more incremental redundancy (IR) packets are needed at the receiver to recover the original message.

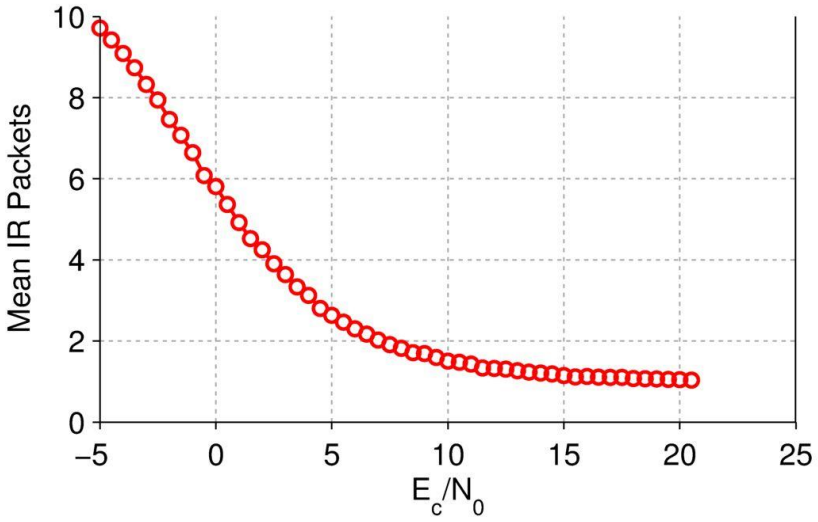


Figure 7-10: Mean number of sent redundancy packets.

The energy efficiency is obtained according to the metric defined in section 3.2. The assumption for the baseline solution of LTE is, that a multicast transmission is done with a channel code designed for servicing the worst channel conditions, e.g. at cell boarders. Typical code rate in the LTE setting is the Rate 1/3 turbo code. This code rate is approximately achieved by the proposed coding scheme with the last incremental redundancy transmission. If LTE-M devices can stop receiving and decoding earlier, energy is saved. Figure 7-11 shows the achieved energy savings.

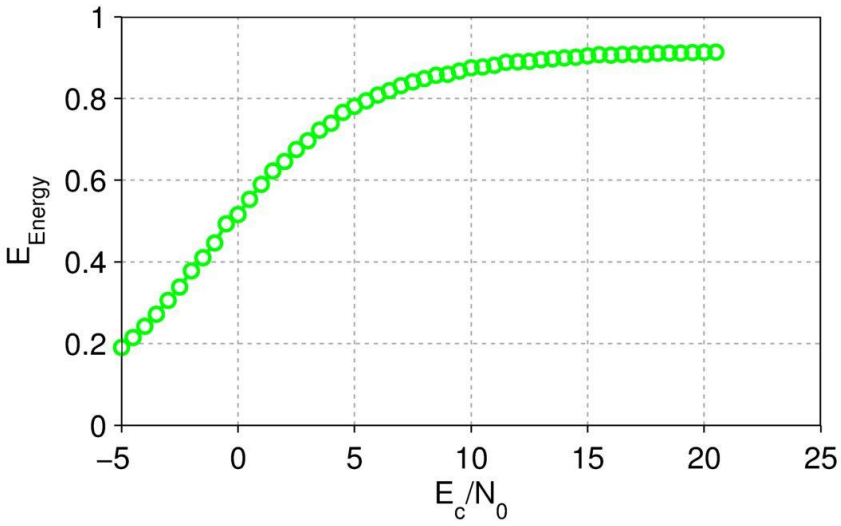


Figure 7-11: Mean energy savings when using incremental redundancy multicast instead of low rate code word multicast.

The highest energy savings of up to 91% can be achieved in the very high SNR regime where channel quality is good enough to be able to decode only with the initial transmission. For lower SNRs the savings decrease a little but still at the lowest considered SNR of -5dB the energy savings still remain at 20% compared to the low-rate solution. On average (over SNR), the proposed multicast scheme can achieve energy savings of approx. 74%.

7.2.6 Low complexity MIMO

Following the assumptions of the evaluation scenario 1 (i.e. a point to point link over Rayleigh fading, with accurate CSI estimations for the conventional receiver), the MIMO scheme will be evaluated using the Bit error rate at the output of the decoder and Consumed energy per message as KPIs.

As shown in Figure 7-12, the energy consumption at the receiver side is reduced significantly (the proposed receiver can offer energy savings with respect to the RF components up to 55% depending on the required SNR target compared to the conventional receiver), since the proposed receiver utilizes only one antenna at each time instant, and it requires no CSI estimations, which need extra power. The price for this reduction is slightly worse performance in terms of bit error rates, which however is not an issue for M2M applications, which are mainly characterized by their low bit rate requirements.

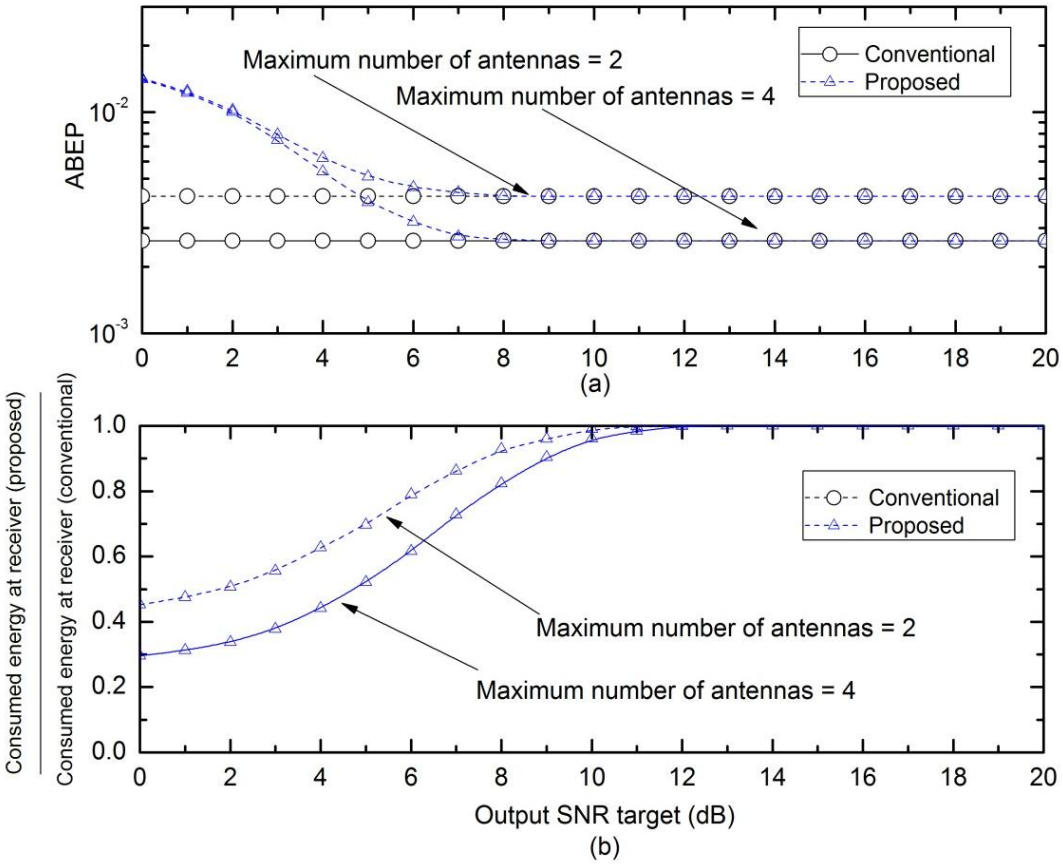


Figure 7-12: The energy reduction of the proposed low complexity MIMO and its performance.

Considering an operational environment where the SNR ranges between 0dB and 10dB (this is equal to a bit error rate that ranges from 0.08 to 0.008) the average energy saving is approximately 28%.

7.2.7 Registering information about terminals

To enable LTE-M network to differentiate between different terminals and their requirements, the following parameters are introduced and stored in networks' HSS as a part of user subscription:

Table 7-3: Additional registering information about terminals.

Field	Description	Format	Example
Paging Factor (PF)	Indicates IDLE mode paging expectations	**Unsigned integer:** Numeric string whose range is [0,65535]	• No paging at all PF=0 • Full paging support like in LTE. PF=1 • Paging factor which multiplies DRX cycle by this factor. PF>1
Mobility Management Factor (MF)	Indicates the level of needed mobility in IDLE mode	**Unsigned integer:** Numeric string whose range is [0,65535]	• No mobility needed. MF=0 • Full mobility like in LTE. MF=1 • Reduced mobility, Mobility factor which multiplies DRX cycle by this factor: MF>1
Traffic generation patterns	Expected traffic generation patterns. MODE (IDLE, ACTIVE), Traffic Type packet distribution (CONSTANT, UNIFORM, GAUSSIAN, LOGNORMAL), Traffic type inter-arrival times, (CONSTANT, UNIFORM, GAUSSIAN, LOGNORMAL),	**Structured attributes:** Each part of the structure is an string separated with ":" character	ACTIVE:CONSTANT:UNIFORM

Based on 3GPP TS 23.401 [49] document, the number of fields needed per subscriptions is as following:

- HSS, 39 values related to subscription and authentication data. The data held in the HSS is defined in Table 5.7.1 [49].
- MME, The MME maintains MM context and EPS bearer context information for UEs in the ECM-IDLE, ECM-CONNECTED and EMM-DEREGISTERED states. Table 5.7.2 [49] shows the context fields for one UE, i.e. subscription profile – 80 values.
- The UE (terminal) maintains the 25 context information (Table 5.7.5 [49]).

So, by introducing 3 new fields in HSS, the number of values will be increased from 39 to 42, which is about 8%. Regarding MME values will be increased from 80 to 83, which is about 4 %, and finally UE, i.e. terminal will have 27 instead 25 values which is exactly 8%. Traffic generation patterns are not information used by the terminals, but rather information used by the network regarding the terminals. Of course these all fields are not of the same length, but in average we can assume that they are equal. It is not expected that these extensions will dramatically affect provisioning time and HSS memory consumption. Provisioning time and memory consumption per user will be slightly increased, but with rather minor effect in comparison to the ordinary LTE subscriptions.

Table 7-4: Number of fields per subscription.

Node	Nr. of fields per subscription in LTE	Nr. of fields per subscription in LTE-M	Average increase of the number of fields length
HSS	39	42	7.6%
MME	80	83	3.75%
Terminal	25	27	8%

There is not direct battery power saving from this procedure which is actually rather used as an enabler for power savings in the other procedures. For example by using the first and second introduced parameters, power consumption in IDLE mode will be reduced (During attach procedure these parameters are downloaded on MME as a part of subscription and delivered to LTE-M device), while from the third parameter access schemes and resource allocations for LTE-M could use additional terminal information. Further, based on the solution some new fields also could be introduced like differentiation whether or not a device supports IP.

7.2.8 Adaptive paging

Adaptive paging algorithm aims for minimization of network signalling regarding the monitoring the paging channel of LTE-M devices. Two new tables are introduced and defined on MME (namely paging profile selection and paging profile tables). When MME initiates the Paging procedure towards a LTE-M device, firstly paging profile selection table will be checked. Paging profile selection table is used for identifying paging profile that should be used for specific LTE-M device. The paging profile selection table should be configured to fit the network configuration as well as the LTE-M device behaviour. The paging profile that is selected is based on the configured parameter values. The current values of the parameters are compared with the table values, and the paging profile for which all the fields in the row match is selected. If no paging profile selection table has been configured by the operator, the default paging profile selection is used. The following information can be used for selecting a paging profile: Access Point Name (APN), Number series of International Mobile Subscriber Identities (IMSIs), Number series of International Mobile Equipment Identities (IMEIs), etc. and different priorities can be defined. Below is given one example of the Paging profile selection table.

Table 7-5: Paging profile selection.

Priority	APN	IMSI	IMEI	Comment	Paging Profile
1	M2M.mnc.mcc.gprs	-	-	-	1, 2, 3, 4 …
2	-	12345	-	-	1, 2, 3, 4 …
3	-	-	12345	-	1, 2, 3, 4 …
4	-	-	-	Default	1

After going through this table, paging profile is selected for device. In the other table, i.e. the paging profile table, the number of paging attempts to perform is provided, using the specified paging target area. Different paging profiles corresponds to different demands related to latency, i.e. if short latency is critical or not. The first row where all criteria match those of the LTE-M device and its bearers determines which paging profile to use. The paging profile selection table can be used to select a paging profile starting with a narrow paging width for certain LTE-M device, i.e. the last visited eNodeB up to the whole Tracking Area (TA), so that the amount of paging messages and consequently the load on the radio network is minimized. Also default value for the number of paging attempts is 2, and it could be extended. TAI (Tracking Area Identity) list contains a list of tracking areas served by the same MME.

The paging profile, as shown in the table below, specifies the number of paging attempts to perform using the specified paging target area.

Table 7-6: Paging profile.

Paging Profile	Paging area	Number of page attempts
1	TAI list	4
2	TA	3
3	Last eNodeB+6 surrounding	2
4	Last eNodeB	2

Different paging profiles corresponds to different demands related to latency, i.e. if short latency is critical or not. Paging profiles are also configurable. The table should be defined in the manner that if we for example choose profile 4 and device is not found, then proceed with profile 3 and so on, i.e. go to a wider area of paging. Paging profile 1 should be default paging profile. The paging profile selection table can be used to select a paging profile starting with a narrow paging width for certain LTE-M device, so that the amount of paging messages and consequently the load on the radio network is minimized.

Here two extreme cases are considered. In the first case LTE-M device is on the fixed position, so it is necessary to perform paging only in the last visited eNodeB, thus allowing a huge save in the number of paging messages.

In the evaluation the following assumptions are taken:

- TA can include 10, 50 and 100 cells. In the practical LTE realizations at the moment 10 represents empirical minimum of eNodeBs and 100 is maximum number. Number of 50 is taken as a mean value.
- TAI list contains 4 TAs.
- Number of page attempts is 2.

Table 7-7: Number of paging messages when terminal stayed in the same eNodeB.

Paging Profile	Paging Area	Paging messages TA=10	Paging messages TA=50	Paging messages TA=100
1	TAI list (4 TAs)	40	200	400
2	TA (10, 50, 100)	10	50	100
3	Last visited eNodeB+6 surrounding	7	7	7
4	Last visited eNodeB	1	1	1

In the case that we expect that the LTE-M device (i.e. sensor) will stay within the same eNodeB it is desirable to use profile 4; reduction of the number of messages, under the given assumptions is going up to 400 times.

On the opposite side is illustrated the worst case scenario, when device is moved from TA, is in the list of TAI.

Table 7-8: Number of paging messages when terminal move to other TA (when the size of TA is 10 eNodeBs).

Paging Profile	Paging Area	Paging messages	Total number of paging messages
1	TAI list	40	40
2	TA x 2	10x2	60
	TAI list	40	
3	Last visited eNodeB+6 surrounding x 2	7x2=14	74
	TA x 2	10x2=20	
	TAI list	40	
4	Last visited eNodeB x 2	1x2=2	76
	Last visited eNodeB+6 surrounding x 2	7x2=14	
	TA x 2	10x2=60	
	TAI list	40	

Based on the calculations shown in the Table 7-8 the same calculation is conducted for the cases when TA contains 50 and 100 eNodeBs, and results are shown in the Table below.

Table 7-9: Number of paging messages when terminal move to other TA.

Paging Profile	TA =10, Total number of paging messages	TA =50, Total number of paging messages	TA=100, Total number of paging messages
1	40	200	400
2	60	300	600
3	74	314	614
4	76	316	616

On Figure 7-13 are presented obtained results.

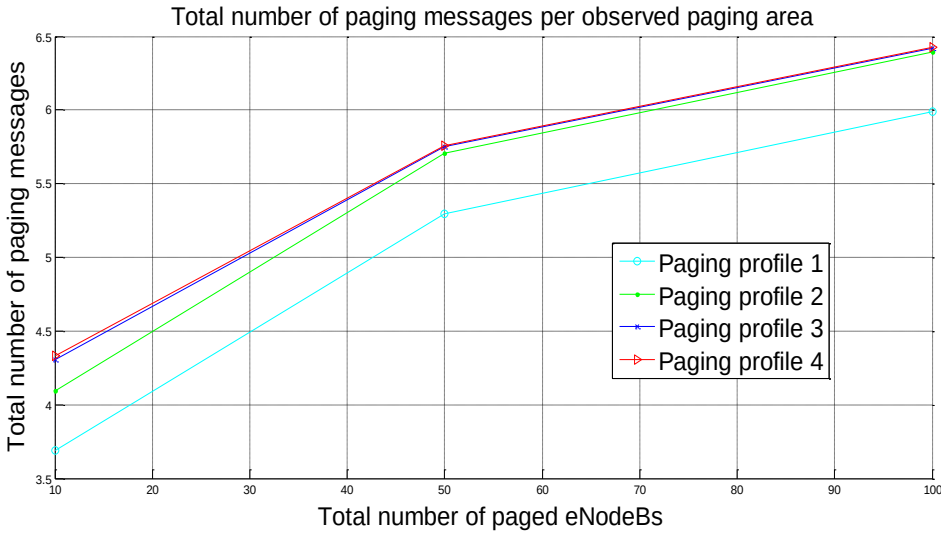


Figure 7-13: Number of paging messages when terminal move to other TA for different sizes of TA.

Since the values on the “y” axis are pretty closed especially for paging profile 1 and 2, the results are presented using logarithmic scale. It can be seen that now in comparison with the case when there is no adaptive paging, than when TA =10 we have increased the number of paging messages for 76/40=90%, for TA =50 we have 316/200=58% and finally for TA=100 there will be 616/400=54% of increase. Obviously the larger is tracking area, the smaller is increase of paging messages.

To conclude, proposed algorithm is highly appreciated in the case when device is on the fixed position when reduction of the number of paging messages could be up to 400 times in ideal case, for the given assumptions. From the other side in the worst case scenario presented in the Table 7-9 we will have 90% more paging messages than with regular paging procedure. So, the main challenge is to understand and predict behaviour of the LTE-M device so that we can properly define paging profile and achieve a significant savings in the numbers of paging messages. By reducing the number of paging messages, we contribute to a reduction in the network load. The available resources can be used for handling more users. Less paging also reduces the signalling in the radio access network. More cases will be analysed and presented soon.

7.2.9 Monitoring paging channel and mobility support

To enable LTE-M network to differentiate between terminals and their requirements Paging Factor (PF) and Mobility Management Factor (MF) are introduced. During attach procedure these parameters are downloaded on MME as a part of subscription and delivered to LTE-M device. More details about this can be found in description of registering information about terminals procedure in subsection 7.2.7. The LTE-M device determines the paging DRX cycle and other relevant paging parameters by processing the serving cell paging parameters broadcasted by the serving cell. The paging message includes the identity of the terminal(s) being paged and a terminal not finding its identity will discard the received information and sleep according to the DRX cycle. The terminal may use DRX in idle mode in order to reduce power consumption. DRX is operator defined parameter and it cannot be influenced. Therefore we have introduced PF (defined per device), which actually multiply DRX delivered by the system with device specific PF, so that terminal will listen for the paging at every PF*DRX ms (milliseconds). For the case when sensors are used for applications that do not require network initiated traffic, monitoring paging could be almost avoided i.e. PF=0, but this is an extreme case. For the applications that can tolerate long delays for network initiated traffic, paging procedure could be longer. Obviously there is a compromise between terminal battery power consumption and the time when the terminal has to wake-up and listen for paging.

As already mentioned DRX is set by the operator, and could have values based on document 3GPP TS 36.331 [18], up to 2560 subframes (sf - 10, 20, 32, 64, 80, 128, 160, 256, 320, 512, 640, 1024, 1280, 2048 and 2560). Since this value is the same for the whole cell it must be compromise for both LTE and LTE-M terminals in the cell, having on mind that LTE devices cannot tolerate huge latency in paging procedure.

Here few values of DRX are analysed, i.e. sf10, sf64 and sf256 are adopted values for DRX for the whole cell, that corresponds to 256 frames, i.e. 0.256 second (PO subframe will come every 256 ms), i.e. terminal has to wake-up every 256 ms, and to observe PO subframe. The same logic is for sf64, i.e. here terminal has to wake-up every 64 ms, and sf10.

If PF is set for example for specific device to be 50, than PF*sf256=50*256ms=12800ms, i.e. terminal will now wake up every 12800 ms instead 256ms. It has also to be assumed that terminal can’t immediately wake up, but need some time, for example few ms, thus getting even more battery savings in the proposed case.

From the other side, network is not aware that terminal has effectively longer DRX, and it could happen that if network needs to page the terminal that is still sleeping, terminal will miss paging (which is highly undesirable case). The solution for that situation is usage of adaptive paging procedure, described earlier where we can assume that terminal is in the same cell all the time, and we can allow to increase the number of paging attempts to 50 in order to cover the worst case for the number of paging attempts needed. Due to an adaptive paging procedure this is efficient since without it, and for the case that TA has 50 cells, then the number of paging attempts will be 2500 which is not acceptable.

As already discussed the power consumption of LTE-M terminal in IDLE mode is dominated also by performing mobility measurements. The proposed solution is to perform optimization of the number of necessary radio measurements, and the possibility to allow the LTE-M terminals to control the timing of measurements and to perform the required measurements only when needed. Frequency of radio measurements depends on DRX cycle, so it is proposed to Introduce Mobility flag (factor), and multiple DRX cycle by this factor, like for the paging case. Mobility factor (MF) is defined in HSS per device as the subscription parameter, and should be delivered to M2M device during the registration process.

This factor could be different for
- RRC_IDLE;
- RRC_CONNECTED.

Since in the RRC_CONNECTED mode it is rather complicated to change the procedures for mobility measurements and handover, here it is focused only on RRC_IDLE state. One idea for further work when terminal is in RRC_CONNECTED state, is to allow terminals to perform cell reselection instead of handover, i.e. terminal controlled mobility, what is not analysed here due to large impact on the overall system.

The terminal shall measure the RSRP and RSRQ level of the serving cell and evaluate the cell selection criterion as defined in 3GPP TS 36.304 [50] for the serving cell at least every DRX cycle. The terminal shall filter the RSRP and RSRQ measurements of the serving cell using at least 2 measurements. Within the set of measurements used for the filtering, at least two measurements should be taken. This requirement actually coincides with monitoring paging channel, i.e. this procedure could be performed even with the larger time span, or at least every paging cycle defined by MF. It is desirable obviously that MF=A*PF, where A is integer. So we can assume that in general MF=PF.

Assumptions for evaluation case are the following:
- DRX = sf10, sf64, sf256;
- PF = 50;
- MF = 50;
- Observed terminal is on the fixed location, i.e. under the coverage of one eNodeB;
- Number of paging attempts for specific terminal is increased by PF number, i.e. 50 times;
- TA has 50 cells;
- Due to the latency in the power up of terminal, it is necessary to start a little bit earlier the wake-up procedure, which covers totally with the desired subframe, 3 subframes, i.e. 3 ms;
- Power consumption in IDLE mode can be 1.7mW, 2mW, 4.7mW and 5mW.

Also, different types of terminals that can be used as LTE-M device will be observed.

Table 7-10: Analysis of terminal power breakdown in IDLE mode.

Terminal	Total power consumption in IDLE mode	Application	Paging monitoring and measurements	Clock
Smart Phone	5 mW	4.1 mW	0.5 mW	0.4 mW
M2M terminal with optimized application	2 mW	1.1 mW	0.5 mW	0.4 mW
M2M terminal with improved clock consumption	4.7 mW	4.1 mW	0.5 mW	0.1 mW
M2M terminal with optimized application and improved clock consumption	1.7 mW	1.1mW	0.5 mW	0.1 mW

So in IDLE mode these 3 factors are dominated for power consumption, i.e. application, paging monitoring and measurements, and clock. In order to understand influence of these factors, different assumptions are made regarding improvements in M2M terminal power consumption, namely application optimization and clock consumption.

Here an example is given below to show how we calculate the total power consumption.

Table 7-11: Analysis of power consumption for Smartphone.

DRX	DRX*PF	Power consumption	Number of paging attempts	Max. number of paging messages (adaptive paging is activated)
No DRX	-	**5mW**	50	-
256 ms	256 ms	(3/256)*5mW+(1-3/256)*0.4mW =0.0117*5mW+0.9883*0.4mW =0.058mW+0.39mW =**0.453mW**	50	-
256 ms	12800ms	(3/12800)*5mW+(1-3/12800)*0.4mW =0.000234*5mW+0.999766*0.4mW =0.00117mW+0.399mW =**0.40017mW**	-	• up to 50 if terminal is in fixed cell, in average 25 (paging profile 4) • up 50*7 if terminal is in the zone of 7 neighbouring cells =350, in average 175 (paging profile 3)

In the calculation it is assumed that terminal is “awake” 3ms every 256 ms for sf256 DRX used, or 3ms every 12800ms when PF=50 is applied, and consumption during that mode is 5mW, and in the rest of the time the terminal is switched off, so only clock battery consumption is present.

In Table 7-12 to Table 7-14 are presented results for different M2M terminals and DRXs:

Table 7-12: Analysis of power consumption for different M2M terminals sf256.

Terminal	Power consumption when no DRX is applied	Power consumption when DRX sf256 is applied / percentage of power savings	Power consumption when DRX sf256 is applied and PF=50 / percentage of power savings
Smart Phone	5 mW	0.453 mW / 91%	0.400 mW / 92%
M2M terminal with optimized application	2 mW	0.418 mW / 79%	0.400 mW / 80%
M2M terminal with improved clock consumption	4.7 mW	0.154 mW / 96.7%	0.101 mW / 97.8%
M2M terminal with optimized application and improved clock consumption	1.7 mW	0.119 mW / 93%	0.100 mW / 94.1%

Table 7-13: Analysis of power consumption for different M2M terminals sf64.

Terminal	Power consumption when no DRX is applied	Power consumption when DRX sf64 is applied / percentage of power savings	Power consumption when DRX sf64 is applied and PF=50 / percentage of power savings
Smart Phone	5 mW	0.615 mW / 87.7%	0.405 mW / 91.9%
M2M terminal with optimized application	2 mW	0.475 mW / 76.3%	0.401 mW / 80%
M2M terminal with improved clock consumption	4.7 mW	0.315 mW / 93.3%	0.104 mW / 97.7%
M2M terminal with optimized application and improved clock consumption	1.7 mW	0.175 mW / 89.7%	0.101 mW / 94%

Table 7-14: Analysis of power consumption for different M2M terminals sf10.

Terminal	Power consumption when no DRX is applied	Power consumption when DRX sf10 is applied / percentage of power savings	Power consumption when DRX sf10 is applied and PF=50 / percentage of power savings
Smart Phone	5 mW	1.780 mW / 64.4%	0.427 mW / 91.4%
M2M terminal with optimized application	2 mW	0.880 mW / 56%	0.409 mW / 79.5%
M2M terminal with improved clock consumption	4.7 mW	1.480 mW / 68.5%	0.127 mW / 97.2%
M2M terminal with optimized application and improved clock consumption	1.7 mW	0.580 mW / 65.9%	0.109 mW / 93.5%

The shorter value of DRX is set in the system, the more benefit will be obtained by proposed procedure. It is obvious that the introduction of DRX on the system level gives huge power savings, but from the other side the longer DRX is set, it will affect all terminals in the cell, what is highly undesirable for the cases, where quick response time is needed. Proposed procedure allows definition of DRX specified per terminal, not influencing other terminals in the cell.

Different system values of DRX as analysed. When DRX is set to be sf10, the smallest one, and PF=50, then there is about 25-30% more battery savings with the proposed algorithm. For sf64 it is about 4%, and finally for sf256 the power saving is about 1%. For the maximal value of DRX sf2560, battery power saving is about 0.15 %. It is also obvious that when DRX is longer, then the total consumption is influenced by clock consumption, so reducing of clock power consumption in the future devices will significantly improve battery lifetime.

Comparison of different devices, namely Smartphone and improved M2M device (application and clock) is presented in the figures below, for different values of DRX. Also, for the other two observed type of terminals, the figures are similar.

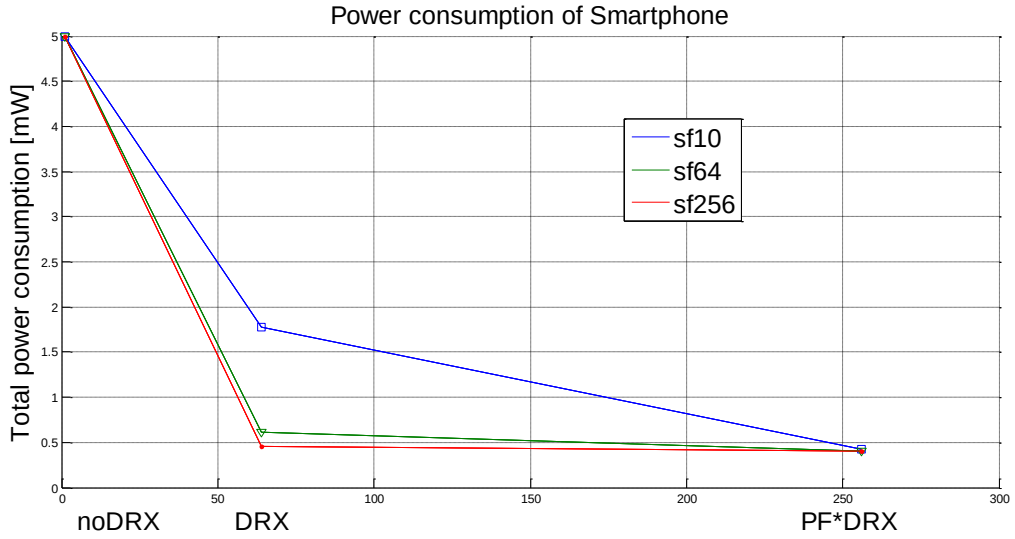


Figure 7-14: Power consumption of Smartphone for different length DRX.

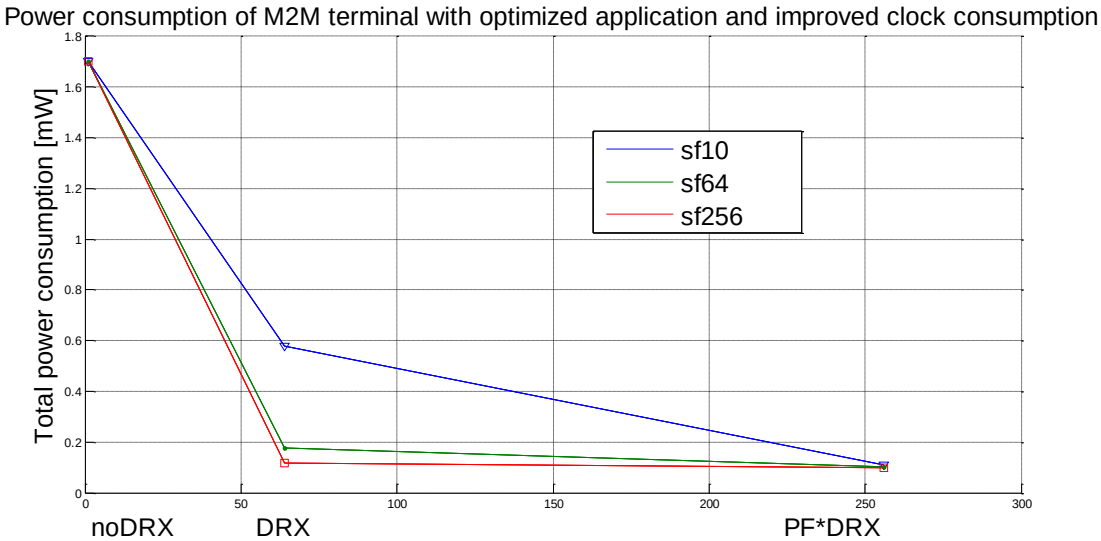


Figure 7-15: Power consumption of M2M terminal with optimized application and improved clock consumption for different length DRX.

It is obvious that increase of DRX influenced all terminals in the same way of power reduction.

When PF is used, also adaptive paging is used. Then, if paging profile 4 from Table 7-5 is chosen in average there will be also less paging messages (in the worst case it will be the same). In the case that paging profile 3 is chosen, then in average we will have 175/50=350% more paging messages, i.e. in the worst case 7 times more paging messages. It is unlikely that terminal which is on the fixed location can move out of the 7 observed cells, so the case when the whole tracking area has to be paged is not considered. The number of paging messages is the same for both analysed cases.

Also it can be assumed that for example, the terminal is operated with AAA battery which has a capacity of up to 1200 mAh which corresponds to theoretical energy of 6.5 kJ, what corresponds to lifetime of the battery if terminal is all the time in IDLE mode, standby lifetime of 15 days. With DRX and proposed algorithms battery lifetime in IDLE mode could be prolonged for up to 15-20 times, i.e. to 300 days for the given assumptions.

7.3 Conclusion

In section 7, evaluation of the new proposed energy efficiency algorithms in the EXALTED is presented. At the beginning of the section 7.1, terminal battery power consumption breakdown model is given, and the algorithms are described briefly. In the section 7.2 the algorithms are evaluated in detail and the improvements are presented.

The proposed algorithms cover IDLE and ACTIVE mode of terminal. The algorithms also influence different parts of the terminal.

Registering information about the terminal enables LTE-M network to differentiate between different terminals and their requirements by adding new parameters related to terminal subscription data. There is no battery power saving directly from this, instead the proposed algorithm is rather used as an enabler for power savings in the other procedures. For example by using the first and second introduced parameters, power consumption in IDLE mode will be reduced.

Two algorithms are proposed for IDLE mode of work, i.e. Adaptive paging, and Monitoring paging channel and mobility support. Adaptive paging algorithm aims for minimization of network signalling regarding the monitoring paging channel of the M2M devices. Proposed algorithm is highly appreciated in the case when device is on the fixed position when reduction the number of paging messages could be up to 400 times in ideal case, for the given assumptions. In average we can expect that reducing of paging messages will be 20-50 times depends on radio network planning. By reducing the number of paging messages, we contribute to a reduction in the network load.

Monitoring paging channel and mobility support algorithm is concentred on the reduction of monitoring paging channel activity and mobility measurements which use significant power consumption in IDLE mode. Different system values of DRX are analysed. For the small values of DRX battery savings are about 25-30%. For medium values it is about 4%, and finally for higher values the power saving is about 1%. For the maximal value of DRX, battery power saving is about some part of %. When DRX is longer than consumption is influenced by clock consumption, so reducing of clock power consumption in the future devices will significantly improve battery lifetime. With proposed algorithms battery lifetime in IDLE mode could be prolonged for up to 15-20 times, i.e. to 300 days for the given assumptions.

The use of LDPC codes for incremental redundancy multicast algorithm aims at the reduction of energy consumption. Decoding of LDPC encoded frame is related to the demodulation and other procedures of Rx Baseband. In the terminal power model it is assumed that 17% of total power is going to Rx Baseband part. By using LDPC decoding part will be reduced variably between 20-91%, in average 74%, which means that now Rx Baseband power consumption will be reduced for 74%, i.e. from 17% to 4.5% when receiving multicast messages (12.5% reduction of total power consumption).

In low complexity MIMO algorithm, the baseline receiver is assumed to have two RF chains and utilizes one or both of them according to the desired output SNR target. Compared to this baseline, the proposed receiver can offer energy savings with respect to the RF components up to 55% depending on the required SNR target. Considering an operational environment where the SNR ranges between 0dB and 10dB (this is equal to a bit error rate that ranges from 0.08 to 0.008) the average energy saving is approximately 28%. In the terminal power model it is assumed that 17% of total power is going to Rx Baseband part, so by applied algorithm we will have decrease of power consumption for 28%, which leads to reduction from 17% to 12.25% of power used by RX Baseband (i.e. 4.75 % reduction of total power consumption).

The proposed approach for user selection that is based on zero forcing beamforming (ZFBF) with semi-orthogonal user selection (SUS) provides a considerable reduction on the required feedback rate (up to 60%) as compared to the baseline user selection algorithm. Specifically, considering a communication scenario with SNR equal to 10 dB the average feedback reduction is between 15%-60% depending on the parameter related with the user channels orthogonality. Proposed algorithm influencing Tx baseband operations which cover 9% of total battery power consumption. Energy efficiency of algorithm depends actually on the number (ratio) of feedback bits with all bits (data plus signalling sent by the terminal). With the obtained feedback reduction of 15-60% (depends on total number of bits transmitted) can lead up to 1% of Tx baseband power reduction, i.e. reducing the total Tx Baseband power consumption from 9% to 8%.

The proposed energy harvesting aware transmission schemes is able to reduce the amount of transmit energy needed to send a given amount of data in a given time period. The approach is based on adapting the transmit power to the battery state, and therefore to the flow of harvested energy. The amount of energy saved strongly depends on the actual flow of harvested energy into the system. As an illustrating example, for the harvested energy flow described in Figure 7-4 and the parameters provided in section 7.2.1, up to 25% transmit energy saving is achieved in active mode, which corresponds to up to 6% of total battery power.
The proposed collision recovery algorithms for random access improve energy efficiency by taking advantage of packet collisions in the decoding process, and therefore requiring a smaller amount of retransmissions. In the energy analysis it was observed that the optimal configuration of the algorithm, in terms of energy consumption, depends on the expected traffic load. For active mode, both Tx baseband and Tx RF, battery life can be improved by 33% with respect to Slotted Aloha for a normalized throughput of 0.36, which corresponds to 11% of total battery power save. The gains are smaller below that mark and for low load it becomes advisable to use Slotted Aloha. The NCDP protocol is able to attain higher throughput, up to 0.8, although battery life is reduced by 50% with respect to the 0.36 mark.

The collaborative broadcast architecture has a negative impact on energy consumption, using retransmissions to improve coverage. However, such an impact, which is related to the capillary air interface, was shown to be modest, with a small number of retransmissions achieving most of the attainable coverage gain, therefore making efficient usage of the energy.

In this section we presented the different gains i.e. terminal power consumption reductions obtained by new proposed LTE-M algorithms. Proposed algorithms, depends on the network scenario and circumstances provide significant power reductions, allowing longer terminal battery lifetime. Algorithms are covering different mode of terminal work, and different parts of the terminal.

8. Objective 5: Cost Efficiency

One of the key challenges of M2M communication is the provision of low-cost LTE-M devices. This issue on has been extensively studied in the 3GPP community (see TR 36.888 [4]). Essential aspects have been investigated to reduce the OPEX and CAPEX of network operators. As the M2M market is expanding massively the reduction of the overall cost for network maintenance becomes one of the main issues for network operators. To this end, a number of means have been discussed that can bring potentially significant cost savings. For instance, the reduction of the maximum transmission bandwidth, or the application of a single receiver chain for both RF and baseband processing at the device are potential solutions. Further cost reductions are also achievable by decreasing the peak rate and the transmit power as well as half-duplex operation.

According to the 3GPP study [4] we assume the device cost breakdown of an LTE UE in Table 8-1 as baseline.

Table 8-1: Device cost breakdown based on [4].

Functional block	Recommendation for evaluation
Ratio of RF to baseband cost	40:60
RF	
Power amplifier	25%
Filters	5%
RF transceiver	45%
Duplexer /Switch	20%
Other	5%
Baseband	
ADC / DAC	10%
FFT/IFFT	5%
Post-FFT data buffering	10%
Receiver processing block	25%
Turbo decoding	10%
HARQ buffer	10%
DL control processing & decoder	5%
Synchronization / cell search block	10%
UL processing block	5%
MIMO specific processing blocks	10%

The following technical contributions to this 3GPP study were submitted to and noted at 3GPP meetings.

- Support of reduced maximum bandwidth for low-cost MTC UEs [51]. It was concluded that the main challenge of bandwidth reduction lies in the position of the physical channels within the LTE frame. This finally led to the definition of LTE-M specific physical channels (see section 2.2.3) and for the design of the LTE-M super-frame structure in (see section 2.2.4) that solve this problem
- On single receive RF chain for low-cost MTC UEs [52]. The contribution mainly addresses the problem of coverage loss and led to the coverage extension study presented in section 6.

- On reduction of maximum transmit power for low-cost MTC UEs [53]. Also in this contribution the coverage loss problem is treated.
- On half duplex operation for low-cost MTC UEs [54]. This contribution discusses the potential benefits of half duplex operation and comes to the conclusion that a small cost reduction can be achieved.

The achievable cost savings for different cases and the expected performance impact are summarized in [4]. For suitable combinations of the discussed means device cost savings of up to 59% are claimed.

Additionally, a special low complexity MIMO scheme for M2M was analysed in EXALTED. It is described and evaluated in the following sections.

8.1 Proposed Solution: Low Complexity MIMO for M2M

In LTE and LTE-A systems, MIMO configurations are supported for both downlink and uplink connections, with two different modes, i.e. spatial multiplexing and transmit diversity. MIMO configurations provides a more efficient usage of available spectrum and offer higher data rates (spatial multiplexing) or improved link robustness (transmit diversity). Nevertheless, these schemes increase the system's complexity and device cost, since they require more Radio Frequency (RF) chains, Low Noise Amplifiers (LNA) and converters at the transmitter or/and the receiver, while they also require full CSI estimations at each antenna element with increased baseband signal processing.

Considering M2M scenarios, where complexity is an issue, MIMO may be on first sight not an attractive technology. On the other hand, taking advantage of MIMO, while keeping the complexity within affordable limits, would be an appealing option. To this context, two issues are identified regarding the applicability of LTE MIMO to M2M communications: The baseband complexity due to the CSI estimations at the receiver, which also require the transmission of pilot bits, and the RF costs at the receiver side, because of the utilization of more than one RF chains.

Towards reducing the receiver costs a transceiver scheme (Figure 8-1) is introduced with the basic features being the absence of CSI estimations at each of the receive antennas and the requirements of only one RF chain, resulting in the following cost savings [48].

- **Baseband cost**
  The baseband signal processing at the receiver is reduced because of the absence of CSI estimations. In the proposed signal-plus-noise-based MIMO scheme, the antennas that participate at the detection stage are determined according to the energy of the received signal. In other words, only those L antennas with the highest sum-of-amplitudes of the received signal are selected, i.e.

$$s_{Cj} = |r_{j,n}| + |r_{j,n+1}|$$

  where $r_{j,n}, r_{j,n+1}$ are the received signals enveloped at the receiver in two subsequent time instances, n and n+1. Moreover, in order to mitigate the effect of excessive switching at the receiver, which may lead to synchronization problems and larger delays, an adaptive receive filter can be applied, according to the measured Doppler spread.

- **RF cost**
  The proposed scheme requires a single RF chain at the receiver, which switches among two or more antennas, in order to increase the diversity gain and hence the performance

in terms of the bit error rate. The selection of the antenna to which the RF chain is attached is determined by the value of the measured received signal plus noise signal.

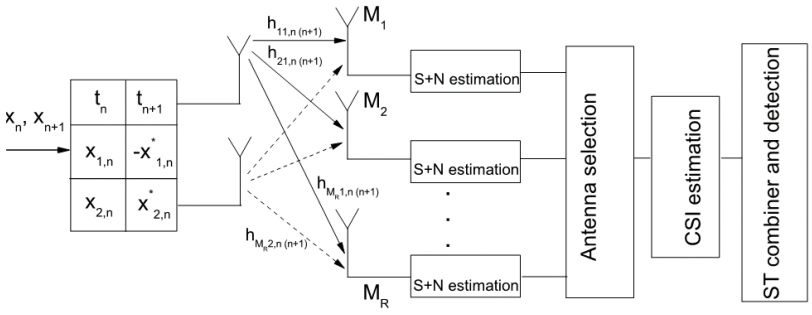


Figure 8-1: CSI and S+N-based MIMO transceivers.

8.2 Performance in Scenario 1

Following the assumptions of the evaluation scenario 1, the MIMO scheme will be evaluated using the following KPI's:

- **(K1) - BER**: Bit error rate at the output of the decoder.
- **(K43) - Number of active antennas**: number of activated antennas compared to the available ones.

As shown Figure 8-2, only one antenna is active at each time instant, compared to the conventional receiver. Note, that only one RF chain is used which switches between the antennas. According to the cost breakdown proposed in the 3GPP document [4] summarized in Table 8-1, the overall cost reduction of the proposed scheme compared to a receiver with 2 RF chains is up to 9%[1]. Moreover, as it can be seen in Figure 8-2, using only one RF chain is sufficient in terms of bit error rate performance (uncoded), taking also into account that the requirements of most M2M services are not as demanding as the conventional ones (e.g. video streaming, or VoIP).

[1] The cost reduction is calculated as (% RF:BB ratio) * (% fraction of RF transceiver cost in the RF breakdown) * (% RF chains reduction), i.e. 40% * 45% * 50% = 9%

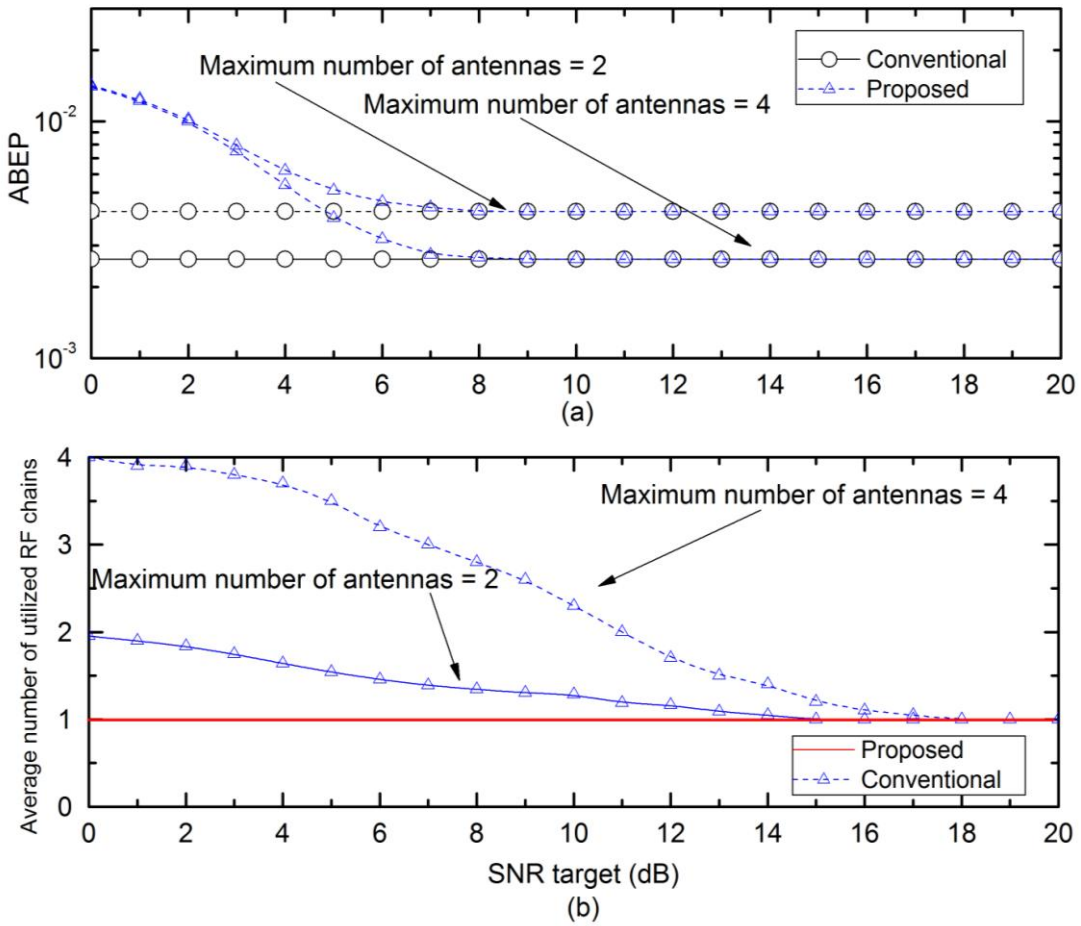


Figure 8-2: The cost efficiency of the proposed low complexity MIMO and its performance.

9. Conclusion and Recommendations

The first part of this report (section 2) presented the composition of the LTE-M system design. Although, the radio protocol architecture is based on the existing LTE stacks, it is possible to implement algorithms and protocols independent from LTE and tailored for M2M by the usage of LTE-M specific physical channels. This system design is the conceptual framework for the algorithms and protocols presented afterwards.

In the remaining sections it was shown to what extent the objectives of the work package can be achieved by applying these novel techniques. For each objective, the evaluation was done with respect to one single, meaningful system KPI. For this, two different evaluation scenarios are defined in section 3, one of them more tending towards energy efficiency, and the other towards spectrum efficiency.

At first, section 4 provided the evidence that LTE-M is a system coexisting with LTE. The separation of LTE-M radio resources in MBSFN subframes guarantees backward compatibility to previous LTE releases. A set of proposals, e.g. the HARQ scheme, underline the re-use of existing functional units in the eNodeB hardware. Also beneficial properties, like the spectral behaviour of GFDM, support the co-existence of two systems in the same frequency band without disturbing the performance. Table 9-1 summarizes the proposed solutions, indicates how they can be applied beneficially, and points out possible interactions with other EXALTED solutions.

Table 9-1: Proposed solutions and their recommended usage.

Solution	Expected performance	Recommended usage	Interactions with other EXALTED solutions
Registering information about terminals	Co-existence achieved with respect to backward compatibility	General usage in LTE-M systems	Enabler of other solutions that exploit this information, e.g. scheduling, optimisation or paging
Slotted access	Co-existence achieved with respect to avoidance of network overload situations,	Event-driven applications with a huge number of devices	Combination with ‘Random Access with Collision Recovery’ is recommended.
HARQ for LTE-M	Co-existence achieved with respect to re-use of existing hardware components	Generally applicable, but tailored for applications with short messages	It is required to adapt the LTE-M rate matching algorithm according to the proposed HARQ scheme.
Innovative scheduling techniques	Co-existence achieved with respect to maintaining the performance of LTE UEs in the presence of LTE-M devices	Generally applicable, but particularly beneficial in heterogeneous environments with mixed applications and QoS classes.	It can be used in all cases, where ‘Semi-persistent scheduling’ cannot be applied. Both approaches complement each other.
GFDM	Co-existence achieved with respect to maintaining the performance of LTE UEs in the presence of LTE-M devices	Generally applicable, but tailored for applications with short messages	GFDM replaces SC-FDMA in LTE-M uplink. A combination with CDMA-overlay is possible.

Section 5 handled the question by which factor the number of devices can be increased with LTE-M compared to LTE. It was found out that there are two possible ways to support a huge number of devices, either by reducing or even avoiding the feedback and signalling information, which can be realized e.g. by semi-persistent scheduling, or by applying spectrum efficient techniques on the payload itself. Examples for the latter are GFDM and the HARQ scheme. An optimization of the random access procedure and specialized scheduling techniques showed additional improvements. Table 9-2 summarizes the proposed solutions, indicates how they can be applied beneficially, and points out possible interactions with other EXALTED solutions. Benefits between some ten and some hundred percent were observed. All in all, with a suitable combination of the proposed methods the number of supported devices can be increased by one order of magnitude.

Table 9-2: Proposed solutions and their recommended usage.

Solution	Expected performance	Recommended usage	Interactions with other EXALTED solutions
Random access with collision recovery	80% throughput improvement on PMRACH	Applications with a huge number of devices, but rare transmissions	Combination with ‘Slotted access’ is recommended.
HARQ for LTE-M	Up to 30% more LTE-M devices	Generally applicable, but tailored for all applications with short messages	It is required to adapt the LTE-M rate matching algorithm according to the proposed HARQ scheme.
Semi-persistent scheduling	500% - 1000% more LTE-M devices	Applications with frequent time-driven transmissions	This solution complements the proposed innovative scheduling concepts.
Slotted access	900% more LTE-M devices	Event-driven applications with a huge number of devices	Combination with ‘Random Access with Collision Recovery’ is recommended.
AGTI scheduler	Up to 1000% more LTE-M devices	Beneficial if applications with different delay constraints are mixed.	It can be used in all cases, where ‘Semi-persistent scheduling’ cannot be applied. Alternatives are ‘QoS based scheduler’ and ‘Scheduling algorithm for heterogeneous traffics’.
QoS based scheduler	Up to 1000% more LTE-M devices	Beneficial if applications with different delay constraints are mixed.	It can be used in all cases, where ‘Semi-persistent scheduling’ cannot be applied. Alternatives are ‘AGTI scheduler’ and ‘Scheduling algorithm for heterogeneous traffics’.
GFDM	Up to 35% more LTE-M devices	Generally applicable, but tailored for applications with short messages	GFDM replaces SC-FDMA in LTE-M uplink. A combination with CDMA-overlay is possible.
Scheduling algorithm for heterogeneous traffics	Up to 1000% improvement based on definition of satisfied users	Beneficial if different LTE and LTE-M traffic types are mixed	It can be used in all cases, where ‘Semi-persistent scheduling’ cannot be applied. Alternatives are ‘AGTI scheduler’ and ‘QoS based scheduler’.

Coverage extension is essential to overcome performance losses caused by complexity- and cost reduction. Section 6 presented two possible solutions, CDMA-overlay and a collaborative broadcast architecture. Table 9-3 summarizes the proposed solutions, indicates

how they can be applied beneficially, and points out possible interactions with other EXALTED solutions.

Table 9-3: Proposed solutions and their recommended usage.

Solution	Expected performance	Recommended usage	Interactions with other EXALTED solutions
CDMA-overlay	97% coverage in the considered scenario	LTE-M uplink for applications with power-limited devices if the radio channel quality is bad	It can be combined with GFDM, but also with SC-FDMA or OFDMA.
Collaborative broadcast architecture	100% broadcast coverage	One message is addressed to multiple devices in the LTE-M downlink	Combination with E2E solutions for capillary networks studied in work package 4.

In section 7 a broad set of solutions targeting at energy efficiency and battery lifetime extension was discussed. Based on a breakdown of the energy consumption of the device components, it was assessed how much percentage of the energy can be saved compared to LTE, and to which battery lifetime this reduction finally translates. Thereby, it was distinguished between IDLE mode and ACTIVE mode. The analysed solutions range from RRC protocol optimization till unconventional techniques like energy harvesting. Table 9-4 summarizes the proposed solutions, indicates how they can be applied beneficially, and points out possible interactions with other EXALTED solutions. The final conclusion is that it is possible, depending on the characteristics of the application, to achieve battery lifetimes in the range of one year.

Table 9-4: Proposed solutions and their recommended usage.

Solution	Expected performance	Recommended usage	Interactions with other EXALTED solutions
Energy Harvesting	Energy reduction up to 6% for the evaluated case	Isolated power limited devices	No interaction with other solutions known.
Random access with collision recovery	Energy reduction up to 11% for the evaluated case	Applications with a huge number of devices, but rare transmissions	Combination with ‘Slotted access’ is recommended.
Collaborative broadcast architecture	Must be considered together with capillary networks.	One message is addressed to multiple devices in the LTE-M downlink	Combination with E2E solutions for capillary networks studied in work package 4.
Directional antennas	Energy reduction 1%	Beneficial in scenarios with limited feedback capacity	It is useful to combine the solutions with the methods aiming at the support of a big number of users.
LDPC Codes for incremental redundancy multicast	Average energy reduction 12.5%	To be used in the LTE-M downlink if the same message shall be delivered to a huge number of devices	No interaction with other solutions known. It is exclusively applied in the PMDMCH.
Low complexity MIMO	Average energy reduction 4.75 %	All LTE-M uplink scenarios with sufficient coverage	No interaction with other solutions known.

Registering information about terminals	No gain as stand-alone solution	General usage in LTE-M systems	Enabler of other solutions that exploit this information, e.g. scheduling, optimisation or paging.
Adaptive paging	Reduction of paging messages by factor 20-50 in average	All applications with fixed devices	Enabled by 'Registering information about terminals'.
Monitoring paging channel and mobility support	Up to 30% energy reduction in IDLE mode	All application with long periods in IDLE mode	Enabled by 'Registering information about terminals'.

Section 8, dealing with cost efficiency must be understood as an extension of the outcome of the 3GPP study item on provision of low-cost M2M UEs based on LTE [4]. The expected performance of the low-complex MIMO scheme according to the device cost breakdown adopted from the 3GPP study, its recommended usage and possible interactions with other EXALTED solutions is given in Table 9-5.

Table 9-5: Proposed solution and its recommended usage.

Solution	Expected performance	Recommended usage	Interactions with other EXALTED solutions
Low complex MIMO	9% cost reduction	All LTE-M uplink scenarios with sufficient coverage	No interaction with other solutions known.

Overall, it is claimed that a plenty of solutions aiming at all work package objectives have been studied. Significant performance benefits with respect to the two main challenges, namely support of a huge number of devices in the network and considerable extension of the battery lifetime, were observed. Moreover, expected drawbacks due to complexity- and cost reduction are fully compensated through coverage extension techniques. The basic conclusion is that the project objectives are thoroughly achieved.

It is pointed out that the performance of combinations of solutions, including those originating from different work packages, that are suitable for particular scenarios or use cases, is presented in the project report D2.4 [55].

Appendix

A1. KPIs and Subordinated Metrics

The following list summarizes the KPIs evaluated in this report and gives corresponding subordinated metrics:

- Co-existence achieved (yes or no) → Objective 1
  - (K1) Bit Error Rate
  - (K2) Packet Error Rate
  - (K3) Packet Loss Rate
  - (K4) Frame Error Rate
  - (K6) Peak-to-Average Power Ratio
  - (K7) Out-of-band radiation
  - (K12) Throughput
  - (K13) Average packet call throughput
  - (K14) Spectral efficiency
- Maximum number of active devices → Objective 2
  - (K5) Outage Probability
  - (K8)/(K9) Average number of retransmissions
  - (K10) Feedback bandwidth
  - (K11) Redundancy overhead
  - (K14) Spectral efficiency
  - (K15) Average packet delay per sector
  - (K17) Access delay
  - (K18) Bandwidth delay product
  - (K20) Number of addresses mapped
  - (K22) Percentage of satisfied users
  - (K23) User per cell capacity
  - (K24) CDF of number of served multicast users
  - (K27) PHY Control channel and pilot overhead
  - (K28) Paging efficiency
  - (K29) Mobility management efficiency
  - (K30) Transmission payload size
  - (K31) Payload encoding
  - (K32) Actual payload size
- Range and coverage, LTE-M can achieve → Objective 3
  - (K1) Bit Error Rate
  - (K2) Packet Error Rate
  - (K3) Packet Loss Rate
  - (K4) Frame Error Rate
  - (K8) Average number of retransmissions
  - (K25) Range
  - (K26) Coverage
  - (K33) Mean power per signalling bit per user
- Consumed energy per message / battery lifetime → Objective 4
  - (K8)/(K9) Average number of retransmissions
  - (K10) Feedback bandwidth
  - (K11) Redundancy overhead
  - (K27) PHY Control channel and pilot overhead
  - (K28) Paging efficiency
  - (K29) Mobility management efficiency

- (K30) Transmission payload size
- (K31) Payload encoding
- (K32) Actual payload size
- (K33) Mean power per signalling bit per user
- (K34) Ratio between transmitted power and achieved throughput
- (K35) Consumed energy per message
- (K36) Standard deviation of node energy levels
- (K37) Average node energy levels
- (K38) Coefficient of variation
- (K39) Network lifetime

- Relative cost reduction with respect to. LTE → Objective 5
  - (K40) Complexity of encoding and decoding
  - (K41) Distortion
  - (K42) Number of CSI estimation
  - (K43) Number of active antennas

A2. Improved Raptor Codes Aided HARQ

In recent years, fountain codes (papers [56] [57] [58]) have attracted researchers' eyes, since they have shown a capacity-achieving capability on Binary Erasure Channels (BECs). At present, most publications about fountain codes focus on protecting packet transmissions over BECs. However, people expect desirable performance beneficial from their natural rateless property, if they may act as a channel code for bit transmissions in the physical layer.

The definition of `rateless' indicates that such channel codes may provide a codeword with any coding rates, where the coding rate R is defined as the ratio of the number of information bits K over the number of encoded bits N, namely R = K/N. Namely, N in the expression of R = K/N may be taken to any value. If channel codes have this rateless property, they may achieve the highest throughput, as well as approach potentially zero outage probability. The authors of [59] studied a complex construction of rateless codes based on traditional fixed-rate codes for Gaussian channels. Due to their inherent rateless feature, the authors of [60] [61] [62] [63] [64] [65] studied the performance of fountain codes employed as a channel code in the physical layer over noisy fading channels. More explicitly, [60] and [61] revealed that Luby Transform (LT) codes suffer from error floors, while Raptor codes may achieve lower Bit Error Ratio (BER) than that of LT codes. The authors of [62] showed the requirements of an optimal degree distribution over Binary Memoryless Symmetric Channels (BIMSCs). Furthermore, [65] demonstrated that Raptor codes may fail to achieve capacity, when the channel Signal to Noise Ratio (SNR) lies outside of an interval with a lower and higher SNR bounds.

However, the performance of fountain codes over noisy channels is not as appealing as expected in the above studies, even though they have considered sufficiently long block length. Moreover, the performance degrades dramatically, when they are applied for transmitting finite-length blocks. In this appendix, we analysed that the unattractive performance of fountain codes is not only caused by the channel conditions, but also by the tanner graph structure. Tanner graph is an efficiently visual way to characterize the encoding and decoding process of fountain codes. We found there are three structural flaws in the tanner graph, referred to as `no degree-one check nodes', `no emerging `degree-one' check nodes' and `uncovered variable nodes', which may significantly decrease the performance. Aiming for finite-length blocks in this appendix, we propose a method by transmitting several special encoded bits, which are deliberately generated to repair these three structural flaws of the tanner graph. A simple degree distribution may be adopted with the aid of these special encoded bits. As a result, the Packet Loss Ratio (PLR) is significantly decreased,

while the decoding complexity drops down to a half, compared to normal systematic fountain codes.

We also combine the Improved Raptor (IRaptor) codes with HARQ. In our proposed IRaptor coded HARQ, the receiver may stop transmissions at the early stage, if the channel is detected in the deep fading. Otherwise, the receiver continues receiving, until it has received a sufficient number of encoded bits which can guarantee a low PLR. The organization of this section is as follows. In section A2.1, we briefly introduce the encoding algorithm of fountain codes and its soft decoding process based on belief propagation. By analysing the tanner graph corresponding to their encoding and decoding, a novel method is proposed in section A2.2.2 to improve the performance for transmissions of finite-length blocks. Simulation results are presented in section A2.3 to compare the performance of improved fountain codes to normal fountain codes. Additionally, the improved Raptor codes will be combined with HARQ in section A2.4, where the turbo coded HARQ in the Long Term Evolution (LTE) standard is chosen as a benchmarker. Finally, section A2.5 draws a conclusion.

A2.1. Fountain Codes

All traditional codes have had a coding rate limit, until fountain codes were invented, where the terminology of 'fountain' is a metaphor of indicating that the encoder of this kind of codes is capable of potentially producing an endless supply of encoded bits, like a fountain spraying drops of water.

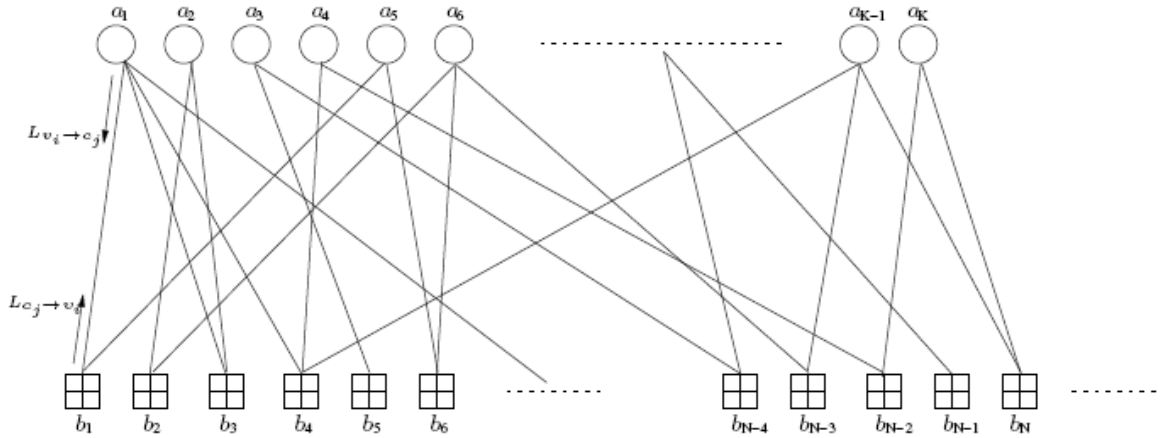


Figure A2-1: The tanner graph representing fountain codes, where the hollow circles denote source bits. Each square represents an encoded bit generated by the bitwise exclusive-OR sum of its several connected source bits.

Fountain codes may be characterized by a tanner graph having sparse edges, as seen in Figure A2-1, where circles denote the so-called variable nodes and squares denote check nodes. Sparsity implies that fountain codes basically have a low encoding and decoding complexity. As shown in Figure A2-1, the encoding process of fountain codes reveals that the number of encoded bits N may reach any value, even theoretically infinity, which justifies their remarkable rateless property. More explicitly, fountain code encoder outputs a sequence of encoded bits b_1, b_2, …, b_N for a block of source bits a_1, a_2, …, a_K. Each encoded bit b_j is generated by modulo-2 adding several randomly chosen source bits $a_{d_i} \in \Omega$, where Ω is the set formed by all source bits connected with this b_j, as seen in Figure A2-1. Mathematically, each encoded bit may be expressed by:

$$b_j = a_{d_1} \oplus a_{d_2} \cdots \oplus a_{d_v} \qquad (a_{d_1}, a_{d_2}, \cdots, a_{d_v} \in \Omega) \ . \tag{A2-1}$$

The number of source bits v in the set Ω is referred to as the 'degree' of that encoded bit, which is a random variable following a specific degree distribution. Furthermore, the positions (represented by d_i) of these source bits involved in generating each encoded bit are evenly distributed in the block region [1, K].

LT codes, as the first practical fountain codes, utilize belief propagation to simplify the decoding for packet transmissions over BECs. When they act as channel codes in the physical layer, sum product algorithm [66] based on belief propagation is also feasible for bit transmissions over noisy channels, which may be found in many papers [67] [68] [69]. In more details, the receiver may know the same tanner graph as Figure A2-1 at the transmitter, by using the pseudo-random generator or the side information passed along with information bits to reconstruct the same degree v and the same connection set Ω for each received encoded bit. Then, soft information will be iteratively exchanged along the edges between check nodes and variable nodes in Figure A2-1, where $L_{c_j \to v_i}^{(m)}$ represents the Log Likelihood Ratios (LLRs) passed from check nodes c_j to variable nodes v_i at the m^{th} iteration, while $L_{v_i \to c_j}^{(m)}$ represents the LLRs reversely passed from variable nodes v_i to check nodes c_j at the m^{th} iteration. The following two equations are calculated in turn during the iterative decoding [66] [70]:

$$L_{c_j \to v_i}^{(m)} = \tilde{\mathbf{b}} \boxplus \left(\sum_{i \in \Psi, i \neq j} \boxplus (L_{v_i \to c_j}^{(m-1)}) \right), \quad \text{(A2-2)}$$

$$L_{v_i \to c_j}^{(m)} = \sum_{j \in \Omega, j \neq i} L_{c_j \to v_i}^{(m-1)}. \quad \text{(A2-3)}$$

In Equation (A2-2), $\tilde{\mathbf{b}}$ represents the received LLRs corresponding to the encoded bit sequence b, and the set Ψ similar to the Ω defined before is formed of all variable nodes connected to this single check node c_j. Furthermore, the operation of $\boxplus$ is referred to as box-plus. The box-plus of two LLR values may be specifically expressed as (paper [71]):

$$L_1 \boxplus L_2 = sign(L_1) sign(L_2) min\{|L_1|, |L_2|\} + f(\cdot), \quad \text{(A2-4)}$$

where f(·) guarantees the precision of box-plus calculation. f(·) may obtain the values from a look-up table, or it may take the accurate expression of $\log\left(1 + e^{(-|L_1 + L_2|)}\right) - \log\left(1 + e^{(-|L_1 - L_2|)}\right)$. Furthermore, the box-plus operator of Equation (A2-2) for two LLRs can be extended to more LLRs by exploiting the associativity of the box-plus operator.

Iterative decoding starts from Equation (A2-2), calculating the extrinsic LLRs $L_{c_j \to v_i}^{(m)}$ on each edge for all check nodes in Figure A2-1. In the first computation of Equation (A2-2), the initial *a priori* LLRs of $L_{v_i \to c_j}^{(0)}$ are set to zeros. The output $L_{c_j \to v_i}^{(1)}$ will be input as the *a priori* LLRs into Equation (A2-3). The next invoking of Equation (A2-3) calculates the corresponding extrinsic LLRs $L_{v_i \to c_j}^{(1)}$ for all variable nodes. This may provide increased *a priori* LLRs $L_{v_i \to c_j}^{(1)}$ back to Equation (A2-2) for the second round calculation. Iterations continue until reaching a successful recovery or a convergence. Finally, the hard decision may be made on the *a posteriori* LLRs of $\tilde{\mathbf{a}}$, each element of which may be computed by:

$$\tilde{a}_i = \sum_{j \in \Omega} L_{c_j \to v_i}. \quad \text{(A2-5)}$$

A2.2. Improved Fountain Codes

The degree distribution of check nodes in fountain codes has a significant influence for the performance (papers [62] [65]). However, it depends on statistical characteristics, which implies that its property may only be held as long as the block length is sufficiently long, while not for short, even medium lengths. Furthermore, the tanner graph like Figure A2-1 is known at the transmitter after an amount of encoded bits has been generated. In this section, we will investigate the errors caused by the structure of tanner graph for the finite-length blocks, and conceive solutions in order to improve the fountain codes' performance.

A2.2.1.Tanner graph analysis

No degree-one check nodes

As explained in section A2.1, the iterative decoding starts from calculating the extrinsic LLRs of check nodes, each of which is related to a channel LLR. By examining Equation (A2-4), its calculating result can be a valid non-zero value, only when L1 and L2 are all non-zeros. Therefore, only the channel LLRs $\tilde{b}_j$ connected to degree-one check node may be initially passed through into the edges of the tanner graph. Taking a simple tanner graph shown in Figure A2-2 as an example, three extrinsic LLRs $L_{c_1 \to v_1}$, $L_{c_1 \to v_2}$ and $L_{c_2 \to v_2}$ are output for two check nodes. Only $L^{(1)}_{c_2 \to v_2}$ is not zero, equal to $\tilde{b}_2$ after the first invoking of Equation (A2-2).

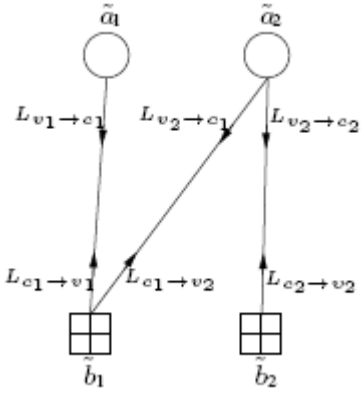


Figure A2-2: The tanner graph including one check node with degree-2, and the other check node with degree-1.

The continuous iteration becomes meaningful due to this non-zero $L^{(1)}_{c_2 \to v_2}$. On the contrary, if there is no degree-one check node, the output extrinsic LLRs at all edges are zeros. The exchanged information between check nodes and variable nodes in the following iterations never increases, always being zero. As a result, 50% bit errors may happen in the final decoded bits.

Although the degree distribution tends to ensure a specific percentage of the degree-one among all check nodes, this may not guarantee there must be a check node having degree one in once encoding, especially for a higher coding rate. Then, the performance of fountain codes suffers significantly because of this situation.

No emerging ‘degree-one' check nodes

Effective decoding may be triggered by an initial degree-one check node according to the above analysis. Nevertheless, continuous iterations depend on new ‘degree-one' check nodes emerging at each iteration. This ‘degree-one' check node is not a real degree-one

check node. Instead, it indicates such a check node having all non-zero *a priori* LLRs inputs for calculating the extrinsic LLR at one edge. Still using the example in Figure (A2-2), $L^{(1)}_{v_2 \to c_1}$ becomes a non-zero value after calling Equation (A2-3) in the first iteration. Then, it will act as an *a priori* LLR for calculating $L^{(2)}_{c_1 \to v_1}$ in the second iteration based on Equation (A2-2), namely $L^{(2)}_{c_1 \to v_1} = \tilde{b}_1 \boxplus L^{(1)}_{v_2 \to c_1}$. Since all operands are non-zeros here, an increased non-zero $L^{(2)}_{c_1 \to v_1}$ may be obtained. In this way, the channel LLR $\tilde{b}_1$ may go into the tanner graph and help the rest iterative decoding. If the edge on which there is a non-zero *a priori* LLR of $L^{(1)}_{v_2 \to c_1}$ is cancelled from the check node 1 in Figure A2-2, its degree becomes one. Therefore, we refer to this phenomenon as the appearance of a new 'degree-one' check node.

There exist a specific tanner graph structure, where no new 'degree-one' check node appears at some stage, and hence the rest of check nodes without 'degree-one' and their connected channel LLRs are helpless for the decoding. Figure (A2-3) shows an example of this structure, where the decoding becomes useless after the first iteration.

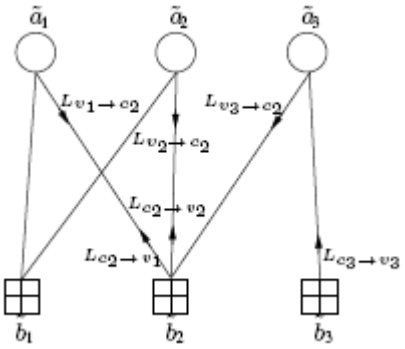


Figure A2-3: The tanner graph where no new 'degree-none' check node will emerge at the second iteration.

More explicitly, except $L^{(1)}_{v_3 \to c_2} = L^{(1)}_{c_3 \to v_3} = \tilde{b}_1$, all of other LLRs $L^{(1)}_{c_j \to v_i}$ and $L^{(1)}_{v_i \to c_j}$ on the edges of tanner graph in Figure A2-3 are zeros after the first iteration. When calling Equation (A2-2) to calculate all $L^{(2)}_{c_j \to v_i}$ values in the second iteration, only the calculations of $L^{(2)}_{c_2 \to v_1}$ and $L^{(2)}_{c_2 \to v_2}$ for the check node 2 are related to the non-zero $L^{(1)}_{v_3 \to c_2}$. Unfortunately, the check node 2 has a degree of 3 corresponding to three connected edges, as seen in Figure A2-3. Only one edge has a non-zero *a priori* LLR, e.g. $L^{(1)}_{v_3 \to c_2}$, while the other two edges have the zero *a priori* LLRs of $L^{(1)}_{v_1 \to c_2}$ and $L^{(1)}_{v_2 \to c_2}$. If the non-zero edge is cancelled from the check node 2, its degree becomes 2, not 1. As discussed before, the computing result of Equation (A2-2) retains zero, as long as one of its operands is zero. Based on the expression of $L^{(2)}_{c_2 \to v_1} = \tilde{b}_2 \boxplus \left(L^{(1)}_{v_2 \to c_2} \boxplus L^{(1)}_{v_3 \to c_2} \right)$, here $L^{(2)}_{c_2 \to v_1}$ may not be increased in the second iteration, keeping zero. Likewise, the resultant $L^{(2)}_{c_2 \to v_2}$ is zero too. Therefore, we may find that no new non-zero LLRs may be obtained during the second iteration and even during the following iterations. In this example, the channel LLRs $\tilde{b}_1$ and $\tilde{b}_2$ never go into the graph and become the wasteful transmissions. The error rate suffers due to this tanner graph structure.

Uncovered variable nodes

An important objective of degree distribution design is to make the average degree as small as possible, since the computational complexity of fountain codes scales with the number of

edges in the coding graph of Figure A2-1. However, in order to avoid encountering decoding failures potentially caused by the total absence of some source bits amongst the encoded bits, each variable node in Figure A2-1 should be connected with a check node at least. This requires the degree distribution may produce a small fraction of high-degree check nodes, which actually increases the fountain codes complexity.

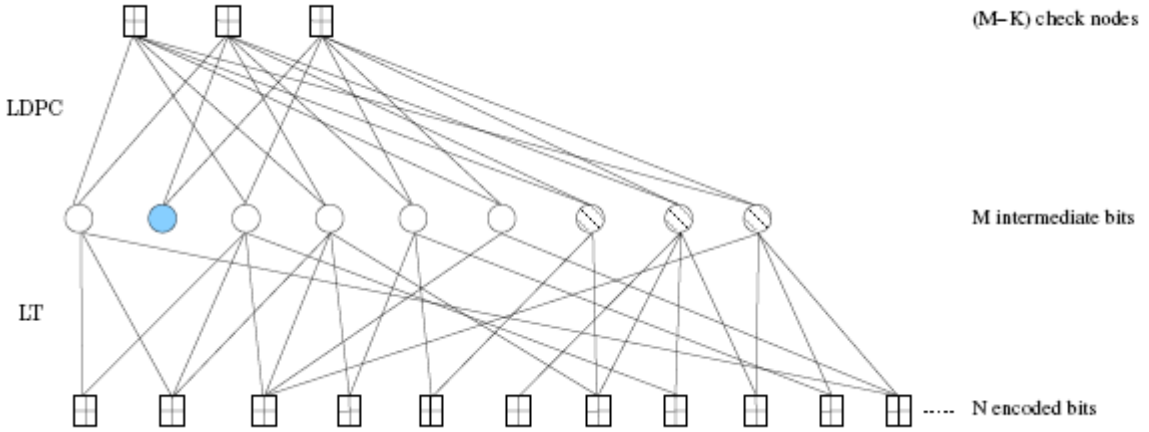


Figure A2-4: Raptor Codes: K source bits are pre-encoded to M > K intermediate bits using systematic LDPC codes; then the having a reduced average degree LT codes process the M intermediate bits to obtain N encoded bits, which will be transmitted to the receiver. The node denoted by the filled circle in the set of intermediate bits is not protected by the LT codes having a reduced average degree. Furthermore, the nodes with filled pattern at the end of intermediate bit set are the redundant bits for the prior K systematic bits.

That is why the author of [56] introduced the robust solution distribution including a number of high degree fractions. Even with these high-degree check nodes, the phenomenon of uncovered variable nodes cannot be completely eliminated due to the random selection process, especially for short block lengths. Furthermore, the resultant average degree becomes relatively high, so does the decoding complexity. Hence, Raptor codes [58] combines Low Density Parity Check (LDPC) code as an outer code in order to decode those uncovered source bits, as seen in Figure A2-4, where the filled circle is excluded by all encoded bits in the inner LT encoding, while it may be recovered by the second-step LDPC decoding. With the aid of outer LDPC code, the degree distribution of inner LT code may adopt the expression having a single-digit average degree, such as the following one proposed in [58] and widely employed:

$$\begin{aligned}\delta(x) = {} & 0.007969 + 0.493570x + 0.166220x^2 + 0.072646x^3 \\ & + 0.082558x^4 + 0.056058x^7 + 0.037229x^8 \\ & + 0.055590x^{18} + 0.025023x^{64} + 0.003135x^{65}\end{aligned} \quad . \text{(A2-6)}$$

In spite of the lower complexity of inner LT codes, present studies about Raptor codes ignore the LDPC decoding cost. Moreover, the degree distribution of Equation (A2-6) also contains several items of high degree, for example, the degree of 65 and 66. Therefore, the complexity of Raptor codes is not yet attractive.

A2.2.2. Solutions

Systematic fountain codes which transmit the whole information bits before fountain encoded bits may naturally solve the problems of `no degree-one check nodes' and `uncovered variable nodes' in the tanner graph, as analysed in section A2.2.1. However, the

performance of systematic fountain codes is not satisfactory, since they may not provide sufficient parity redundancy as non-systematic fountain codes. In the meanwhile, the problem of `no emerging `degree-one' check nodes' in section A2.2.1 still exists in systematic fountain codes. In section A2.3, we will enclose systematic fountain codes in our simulations as a benchmarker for comparison.

For non-systematic fountain codes, no matter how good the channel conditions are, the performance may be degraded by the above three attributes of tanner graph. However, these structural factors can be eliminated by the transmitter, since it may rebuild the tanner graph after sending a specific number of encoded bits. In order to realize this reconstruction, we allocate a list for each variable node and each check node to record the indices of its connected counterparts. For the lists of check nodes, they are simply formed once an encoded bit is created, while the lists of variable nodes are dynamically filled during the generation of each encoded bit. For example, noting the tanner graph in Figure A2-3, the list of check node 1 is like {1, 2} storing the indices of its connected variable nodes when the check node 1 is generated. Simultaneously, this check node's index 1 is reversely appended to the lists of those variable nodes that it is connected to, namely here the variable node 1 and the variable node 2. Finally, after three encoded bits are sent out, all recorded lists are as shown in Table A2-1.

Table A2-1: Recorded lists for the tanner graph in Figure A2-3.

check node index	lists
1	{1, 2}
2	{1, 2, 3}
3	{3}
variable node index	lists
1	{1, 2}
2	{1, 2}
3	{2, 3}

With the aid of these lists, it's not difficult to solve the problems caused by the tanner graph structure. After a number of encoded bits being sent out, the transmitter may easily find out ‘no degree-one check nodes' and ‘uncovered variable nodes' by respectively going through the recorded lists of check nodes and variable nodes. In order to get rid of the case of ‘no degree-one check nodes’, the encoder will generate a new degree-one check node in a usual way. For banishing those ‘uncovered variable nodes', the encoder produces the corresponding number of check nodes with a random degree plus 1, where the random degree indicates those variable nodes evenly randomly selected, while the 1 corresponds to the one among those uncovered variable nodes.

‘No emerging ‘degree-one' check nodes' may also be found by cancelling the nodes and their connected edges from the tanner graph one by one. More explicitly, this cancelation commences from any of the current degree-one check nodes. Using the illustrative lists in Table A2-1 as an example, the degree-one check node 3 will be first removed, namely deleting its list {3} to be an empty list. Nonetheless, its list {3} before cancelation implies that the belief may be propagated along the edge to variable node 3 during the decoding. Continuously, based on the list {2, 3} of variable node 3, the belief is then conveyed back to the check node 2. We refer to this as a belief propagation round, starting from a degree-one check node, and flowing back to several other check nodes whose degree will be minus 1. More explicitly, after this round of belief propagation, the list of variable node 3 is also set to NULL, and all edges connected to variable node 3 need to be removed from its connected check nodes. In this case, its connected check node 2 will remove the index 3 from its list, changing from {1, 2, 3} to {1, 2}. In other words, the check node 2's degree has been

decreased from 3 to 2, since the number of indices in a check node's list indicates that check node's degree.

Then, we need to check if there emerges a new degree-one check node. If a new degree-one check node appears, the same cancelling actions will be repeated. Otherwise, a new check node ought to be generated, attempting to overcome this structural flaw. The degree of this new check node is not randomly chosen as a normal fountain encoded bit. Instead, it is set to be a degree of 2. Namely, the new check node will have two edges, one of which connecting to the just cancelled variable node, the other connecting to any one of variable nodes which are connected to the least-degree check node in the rest tanner graph. For the example of Figure A2-3, the new degree-2 check node is connected to the just-cancelled variable node 3, as well as the variable node 2 which is connected to the current least-degree check node 1, as shown in Figure A2-5. It may be seen that the belief propagation may continue to flow, after producing this new check node 4, which is supposed to have a list of {2, 3}. Yet, its cut list {2} without the index of 3 will be added into the rest tanner graph to trigger the next cancelations, since the variable node 3 has been cancelled at this moment.

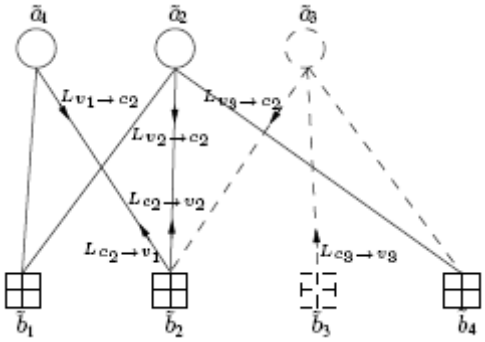


Figure A2-5: The enhanced tanner graph of Figure A2-3 by adding an extra check node 4 which has a degree of 2 connecting the just-cancel variable node 3 and the variable node 2. The variable node 2 is connected to the check node 1, which has the least degree of 2 among the rest check nodes. Here, the dashed parts indicate they are cancelled from the tanner graph during the process of checking and resolving the structural flaw of 'no emerging 'degree-one' check nodes'.

In a conclusion, the general process of solving the flaw of 'no emerging 'degree-one' check nodes' works as follows:

- Step 1: Attempt to find a degree-one check node. If none exists, go to step 3.
- Step 2: Cancel the found degree-one check node, its connected variable node and all edges of that variable node. Then, go to step 4.
- Step 3: Generate a new degree-two check node, connecting to the just-cancelled variable node and any one of variable nodes connecting to the check node which has the least degree in the rest tanner graph.
- Step 4: Loop from step 1, until the tanner graph becomes empty.

When the above proposed strategies are imported into the fountain codes to solve the structural flaws of tanner graph, the high degrees in the degree distribution may be left out to reduce the average degree, since there is no need to worry about uncovered variable nodes. As explained in section A2.2.1, the high degree fraction in the degree distribution is to reduce the probability of uncovered variable nodes as less as possible for non-systematic fountain codes. Therefore, in our simulations in section A2.3, we adopt the following degree distribution for our improved fountain codes:

$$\delta(x) = 0.091717 + 0.493570x + 0.166220x^2 + 0.072646x^3 + 0.082558x^4 + 0.056058x^7 + 0.037229x^8 \quad \text{, (A2-7)}$$

which is straightforward formed by moving the fractions of high degrees in the degree distribution of (A2-6) into the fraction of degree 1.

A2.3. Simulation results

In order to demonstrate the performance of our improved fountain codes, the PLR and complexity will be compared to those of normal fountain codes. We chose non-systematic LT codes, systematic and non-systematic Raptor codes, and our improved non-systematic LT codes, improved non-systematic Raptor codes to be compared, by collecting the statistical results after a relevant number of blocks were transmitted, using the Binary Phase-Shift Keying (BPSK) modulation scheme over an Additive White Gaussian Noise (AWGN) channel. The finite block lengths of 100 bits and 1000 bits are considered in our simulations.

The normal LT codes adopted the degree distribution of Equation (A2-6), while our improved LT codes employed the degree distribution of Equation (A2-7). For all normal and improved Raptor codes in our simulations, the coding rate of outer LDPC codes was set to a high value of 0.9. We employed regular systematic LDPC codes with the parameter set of (M, 3, 30), where M is the number of intermediate bits in Figure A2-4, and 3 is the variable node's degree, 30 is the check node's degree in the LDPC tanner graph in Figure A2-4. LDPC codes also employ sum-product algorithm based on the tanner graph for the decoding ([66]).

We investigated the performance of these five fountain codes with a fixed coding rate, as well as with gradually decreasing coding rates. For the sake of a fair comparison, the overall coding rate of five fountain codes ought to be kept the same. Hence, the coding rate for the sole LT codes was set to 0.45; whereas the coding rate for the inner LT codes in Raptor codes was set to 0.5, which is times the outer 0.9-rate LDPC to have the same overall coding rate of 0.45. Following the normal fountain encoding process, our improved fountain codes transmit a series of encoded bits, until it reaches this pre-set coding rate. Then, the encoder reconstructs the tanner graph using the recorded lists and generates several special encoded bits to solve its structural flaws, as discussed in section A2.2.2. Compared to normal fountain encoded bits, these extra encoded bits are special, since their degree and the variable nodes involved in the XOR operations are designated, instead of being randomly selected. Nevertheless, simulation results demonstrated that the influence to the coding rate by these extra encoded bits can be ignored, highly close to 0.45. Specifically, the coding rate of improved fountain codes is 0.4498 in our simulations.

For all LT codes, except performing the Cyclic Redundancy Check (CRC), an early stopping strategy proposed in paper [72] was employed to stop the iterative soft decoding. More specifically, instead of a fixed number of iterations, the decoder will halt the decoding once the increment of Mutual Information (MI) on the extrinsic LLRs of variable nodes becomes less than 0.0001, represented by $MI(L_{v_i \to c_j}^{(m)}) - MI(L_{v_i \to c_j}^{(m-1)}) < 0.0001$, since more iterations cannot improve the performance any more. Here, the function MI(·) which may be found in section 2.3 of paper [73] estimates the MI of the extrinsic LLR sequence $L_{v_i \to c_j}$. There are two tanner graphs in Raptor codes, as seen in Figure A2-4, reflecting the inner LT decoding and the outer LDPC decoding. More explicitly, after a number of iterations performed in the inner LT decoding, the outputs *a posteriori* LLRs are input for a number of sum-product iterations performed by the outer LDPC decoder. The above early stopping strategy is also employed to stop the inner LT iterative decoding, while we take 5 iterations for the outer LDPC iterative decoding, in order to control the whole complexity. Nevertheless,

the LDPC iterative decoding may not be invoked at all, if the systematic bits have already satisfied the CRC after the inner LT iterative decoding.

Each edge in a tanner graph is related to the calculations of plus and box-plus. The complexity of fountain codes is reasonably to be defined as the total number of operations of addition and box-plus per source bit, respectively denoted by N_{+} and $N_{\boxplus}$. Note that the total number of operations indicates the number of additions and box-pluses performed during all iterations, for not only LT codes, but also for LDPC codes if this statistic is for the complexity of Raptor codes. Considering the look-up table implementation of f(·) in Equation (A2-4), the complexity of box-plus is approximately equal to $(5 + 2 * \log_2(T))$ additions, where T is the number of entries in the loop-up table. The number of clock cycles performing a single box-plus operation in Equation (A2-4) includes 5 cycles for comparison operations for obtaining the minimum value between L1 and L2; as well as includes twice searching of the look-up table, each of which consumes log2(T) cycles. Referring to conference paper [74], when the number of entries becomes T≥4, the performance improvement of belief propagation decoding may hardly be obtained. The least T=4 entries for the look-up table may ensure the decoding accuracy, while keeping the low complexity. Therefore, the expression of $(9 \times N_{\boxplus} + N_{+})$ may generally reflect the complexity of fountain codes. This makes it possible to compare the complexity of LT codes and Raptor codes together.

With these simulation parameters, Figures A2-6-(a) and A2-6-(b) respectively show the PLR performance for the 100-bit block length and the 1000-bit block length. Furthermore, Figures A2-7-(a) and A2-7-(b) show the corresponding complexity performance for both block lengths, where the values of complexity for all schemes have been divided by 1000, since their original values are too large to be well displayed.

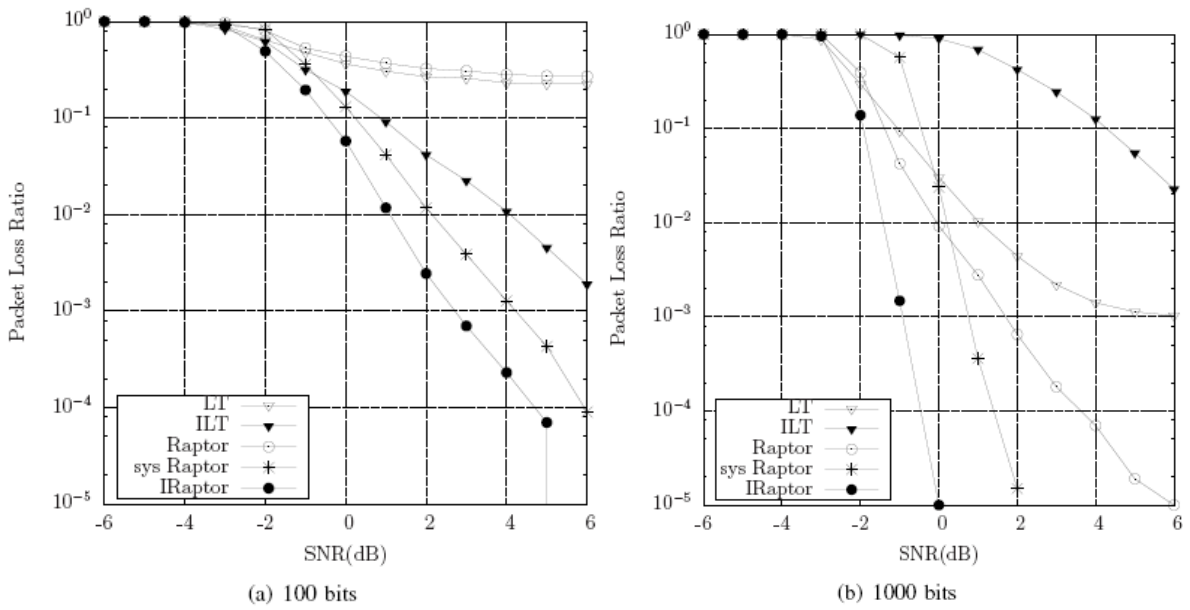


Figure A2-6: PLR versus the AWGN channel SNR for the block length: a) 100 bits, b) 1000 bits.

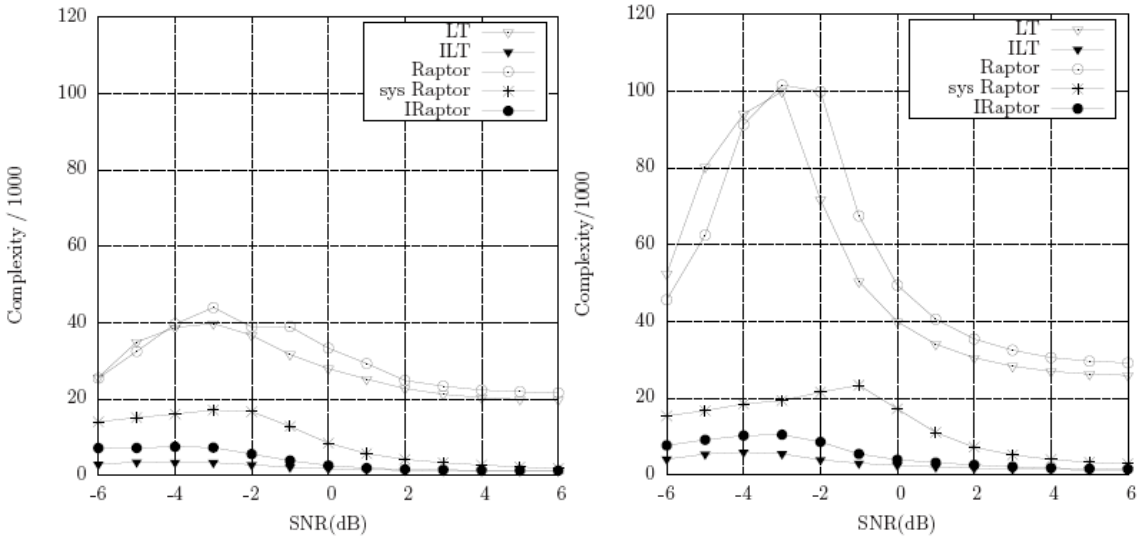


Figure A2-7: Complexity versus the AWGN channel SNR for the block length: a) 100 bits, b) 1000 bits.

For the block length of 100 bits, it may be observed in Figure A2-6-(a) that our improved Raptor codes achieve the lowest PLR, while having a significantly lower complexity compared to the normal fountain codes. More specifically, the PLR of the improved Raptor codes descends rapidly, approaching zero along the increasing SNRs. This PLR is 1.5dB better than that of the systematic Raptor codes, which have the best performance among the normal fountain codes. Despite 2dB worse than the PLR of the systematic Raptor codes, the improved LT codes outperform the non-systematic fountain codes, which have a PLR of higher than 0.2 even at high SNRs. As for the complexity seen in Figure A2-7-(a), the improved LT codes achieve the lowest one among the five codes. The complexity of the improved Raptor codes is slightly higher than it, since the LDPC decoding cost is counted. However, the improved Raptor codes still dramatically decrease the complexity to less than a half of the systematic fountain codes, beneficial from the simple degree distribution of Equation (A2-7).

Although the improved LT codes cannot work attractively for the 1000-bit block length in Figure A2-6-(b), the improved Raptor codes exhibit the lowest PLR performance among five fountain codes. This is because the straightforward design of degree distribution of Equation (A2-7) may not be an optimal one. A small number of errors remain in each block after the decoding of improved LT codes. Nevertheless, these remaining errors can be corrected by the outer 0.9-rate LDPC code. The improved Raptor codes attain a PLR of 10^{-5} at the SNR of 0dB, which is 2dB lower than the normal systematic Raptor codes and 6dB lower than the non-systematic Raptor codes. Simultaneously, as seen in Figure A2-7-(b), the complexity of improved Raptor codes reaches a half of the normal systematic Raptor codes. Compared to the normal non-systematic fountain codes, the complexity has been significantly reduced by 10 times most.

In an addition, we explored the candidates' behaviour at different coding rates. Figure A2-8 illustrates the PLR performance versus the reverse of coding rate for the 100-bit block length at the SNR of -3.0dB. It showed that the improved Raptor codes have a significantly lower PLR than the other fountain codes. Specifically, as seen in Figure A2-8, when the improved Raptor codes achieve a PLR of 0.1, the overall coding rate is required to about 1/3.3 = 0.303. By contrast, the systematic Raptor codes which present the best performance among normal fountain codes approximately need the coding rate of 1/3.65 = 0.274. In the meanwhile, Figure A2-9 reveals that the complexity of improved Raptor codes maintains a lower one

than the systematic Raptor codes by a reduction of three times, and by a further reduction of 10 times when compared to the non-systematic fountain codes.

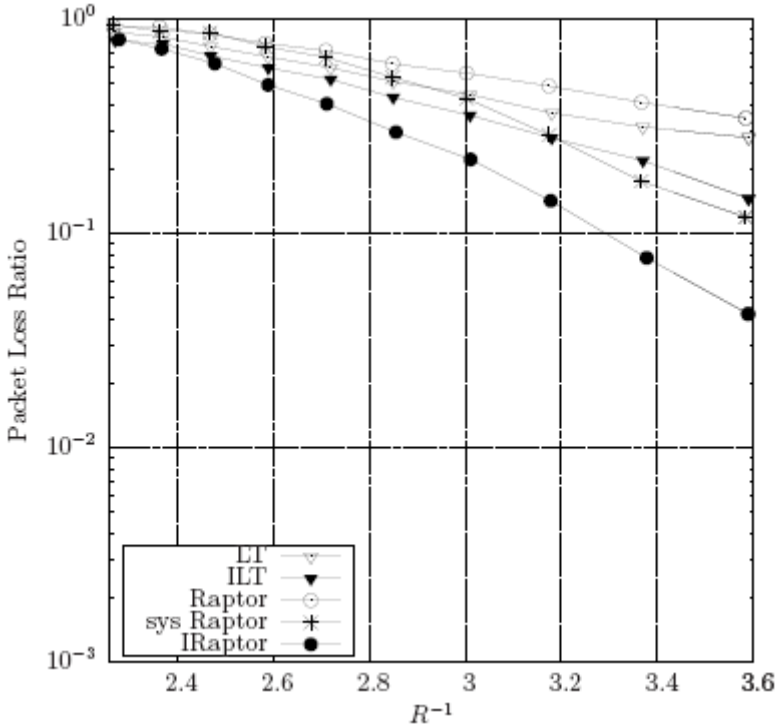


Figure A2-8: PLR versus the reverse of coding rate for the 100-bit block length at the SNR of -3.0dB.

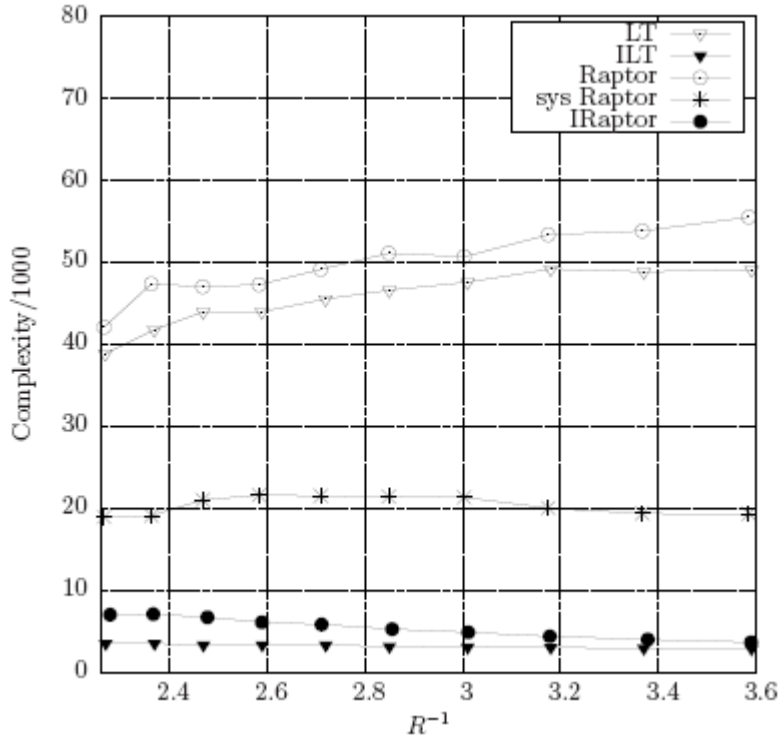


Figure A2-9: Complexity versus the reverse of coding rate for the 100-bit block length at the SNR of -3.0dB.

Simulation results in this section have demonstrated that the improved Raptor codes achieve the lowest PLR, as well as having a lower complexity than the normal fountain codes. This is achieved at the cost of a small number of memories and simple operations during the encoding at the transmitter. More specifically, these memories are used for storing the lists of all variable and check nodes, while the operations involve going through these lists to generate several special encoded bits, as discussed in section A2.2.2.

A2.4. Improved Raptor (IRaptor) Codes Aided HARQ

Since our IRaptor codes have shown attractive performance in section A2.3, we combine them with Hybrid Automatic Repeat reQuest (HARQ) to reach a higher performance. In a straightforward scheme, the transmitter may continuously send IRaptor encoded bits, until it receives an Acknowledgement (ACK) message from the receiver. The receiver will activate the iterative decoding after the reception of each encoded bit, and feedback the ACK message once the CRC is satisfied. In such a scheme, the decoding complexity may become significantly high, since the iterative decoding is performed for each reception.

Table A2-2: The coding rates of IRaptor codes, aiming for a PLR of 0.1 when the packet length is 104 bit.

	104-bit
-12.0dB	0.05
-11.0dB	0.07
-10.0dB	0.08
-9.0dB	0.10
-8.0dB	0.13
-7.0dB	0.16
-6.0dB	0.2
-5.0dB	0.24
-4.0dB	0.29
-3.0dB	0.34
-2.0dB	0.41
-1.0dB	0.47
0.0dB	0.54
1.0dB	0.61
2.0dB	0.68
3.0dB	0.75
4.0dB	0.82
5.0dB	0.89
6.0dB	1

In order to solve the above issue, we design an IRaptor coded HARQ scheme based on a Look-Up Table (LUT), which pre-stores the coding rates of IRaptor codes for different SNRs. Using these coding rates, a PLR of 0.1 may be approximately achieved for transmissions over an AWGN channel. This LUT may be generated by an off-line training process, by transmitting a sufficient number of packets, which are IRaptor encoded with various coding rates as described in section A2.3. Table A2-2 illustrates an example of the LUT from the SNR of -12.0dB to 6.0dB for the packet length of 104 bits.

Generally, the transmission gap between (re)transmissions is a main factor of delay in a HARQ scheme, while the propagation time for each (re)transmission can be ignored. Hence, the maximum number of retransmissions is determined by the tolerable maximum delay of a system. During the period of this maximum delay, the channel quality varies due to different

fading coefficients. However, the fading coefficient may be considered as the same for each (re)transmission because of the short propagation time. That is why the quasi static channel model is a reasonable assumption for simulations in a large number of publications. More explicitly, the channel is equivalent to an AWGN channel having a specific SNR during the period of each (re)transmission, while the SNR fluctuates from one by one transmission.

According to the LUT in Table A2-2, a low coding rate is required for guaranteeing a high probability of successful reception, if the transmission happens in a deep fading. This may significantly degrade the throughput. If a HARQ scheme avoids transmissions in the deep fading, instead it chooses a relatively better channel within the maximum delay allowed, it will achieve a high throughput, as well as a desirable PLR performance.

Based on the above analysis, our proposed IRaptor aided HARQ stops the transmission if the transmitter has not received a Channel State Information (CSI) indicator message at the early stage; or continues transmitting until the number of encoded bits can guarantee a low PLR corresponding to the CSI. Figure A2-10 illustrates the flow charts of the transmitter and the receiver in our proposed IRaptor coded HARQ scheme. In detail, in our proposed IRaptor coded HARQ, the original packet with a length of K bits is firstly appended by the CRC check bits and then encoded by the high-rate outer LDPC. The resultant M-bit length of packet will be input to the IRaptor encoder, which may continuously generate the encoded bits. The transmitter commences sending out several encoded bits and pauses. If the signals are able to be detected by the receiver, it carries out channel estimation after receiving these several bits. A short CSI message containing the detected SNR will be immediately fed back to the transmitter. When the transmitter receives this CSI, it will check if the detected channel SNR is larger than a pre-set SNR, which is a system parameter indicating a deep fade, as seen in Figure A2-10. If the detected SNR becomes larger than that pre-set SNR, the transmitter searches the LUT to find out a corresponding coding rate. Then, the transmitter continues transmitting until that coding rate has been reached. After that, it will reconstruct the Tanner graph and send out several special encoded bits for overcoming its potential structural flaws, as described in section A2.2.2. On the other side, the receiver will execute the iterative decoding after receiving all encoded bits. If the CRC is satisfied, it sends an ACK message back; otherwise it discards the packet. As seen in Figure A2-10, several situations will lead to a retransmission during the process of a single transmission. For example, if no CSI message is fed back, which implies that the signals may not be distinguished due to a deep fade; or if the detected SNR is found less than the pre-set SNR; or if the transmitter times out after sending all encoded bits out. The same transmission process will be repeated until a retry limit is reached.

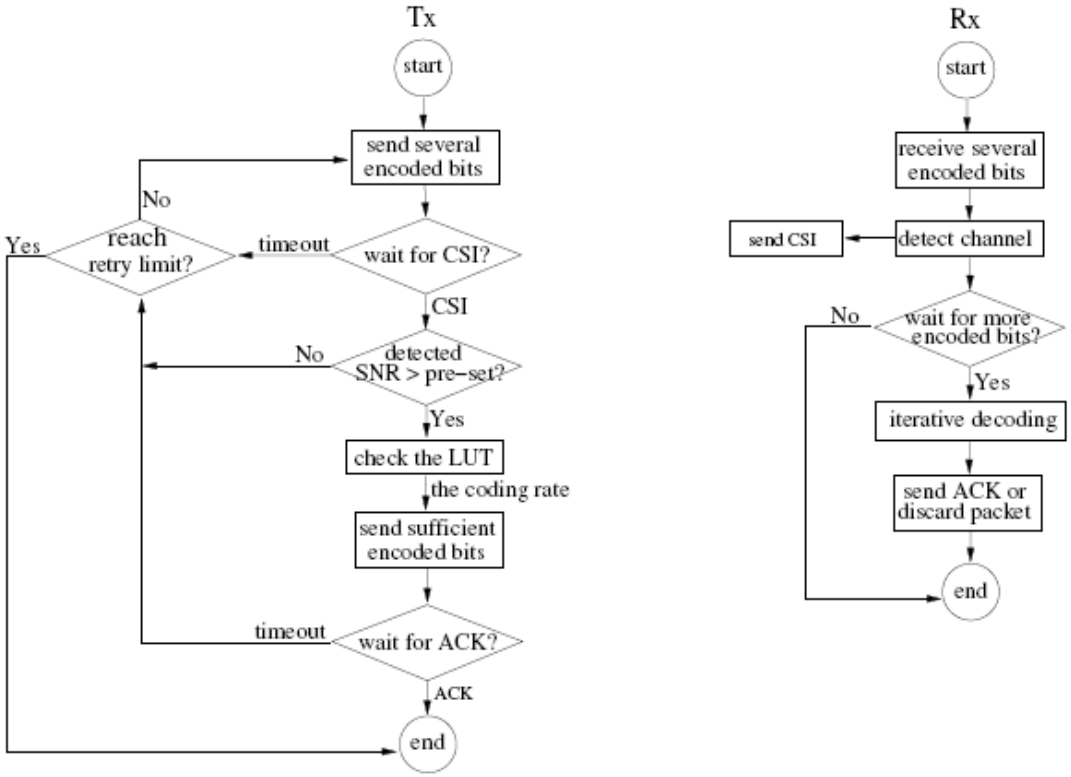


Figure A2-10: The flow chart of our proposed IRaptor coded HARQ.

The pre-set SNR value determines a trade-off between the throughput and the PLR , where the throughput is defined as a ratio of the number of successful information bits over the number of all transmitted bits. If it is set too low, the packet may be successfully transmitted in a deep fading, using a low coding rate. Hence, the throughput suffers, while guaranteeing a low PLR. By contrast, if it is high, transmissions on the mild fading channels may be given up, which leads to a relatively high PLR, but a high throughput. We developed simulations for investigating the throughput and PLR trends of our proposed IRaptor coded HARQ scheme for four pre-set SNR values in a set of {-12.0dB, -10.0dB, -8.0dB, -6.0dB}, which were illustrated in Figures A2-11 and A2-12.

In our simulations, the retry limit of the proposed IRaptor HARQ is set to 4, same as the turbo coded HARQ in the LTE standard [11], which was chosen as a benchmarker for comparison. Furthermore, a normal IRaptor coded HARQ was also enclosed. It employs the same strategy as a normal HARQ, transmitting IRaptor encoded bits in four (re)transmissions, each of which has the same length of M bits and several special bits for overcoming the tanner graph structural flaws. In all HARQ schemes, a sufficient number of packets with an original length of 104-bit were BPSK modulated and transmitted over a quasi-static Rayleigh fading channel. In the implementation of our proposed IRaptor HARQ, when mapping a decimal detected SNR to the coding rate at the receiver, the smaller integer SNR is adopted to check the LUT. For example, the coding rate for the detected SNR of 0.65dB is 0.54 corresponding to the SNR of 0.0dB, as seen in Table A2-2. For the sake of achieving a stably lower PLR performance, the transmitter in our simulations sends extra (0.1*M) bits on the basis of that coding rate.

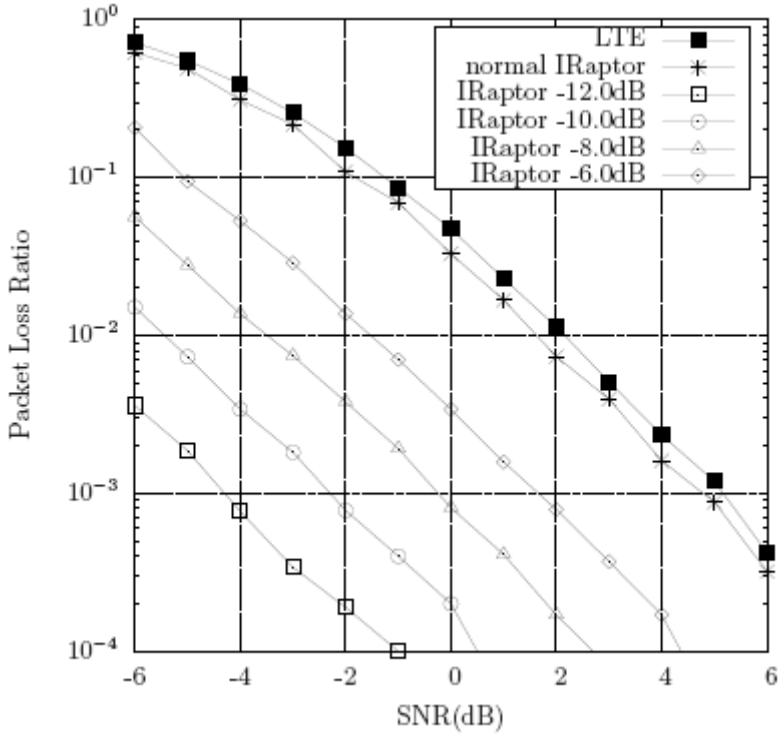


Figure A2-11: PLR versus SNR for all candidate HARQ schemes over a quasi-static Rayleigh fading channel.

Observe from Figures A2-11 and A2-12, although the normal IRaptor coded HARQ has a similar throughput to that of the LTE turbo coded HARQ, the PLR is slightly lower than it, namely about 0.5dB lower. However, Figure A2-11 demonstrated that a significantly low PLR may be achieved by our proposed IRaptor HARQ even at low SNRs. More specifically, when the deep fading parameter is set to -12.0dB, the PLR may be achieved below 0.001 from the SNR of -4.0dB. The PLR performance has a 2dB degradation if the deep fading is set to 2dB higher, while the throughput observed in Figure A2-12 has an improvement of about 4%. Considering the deep fading parameter of -6.0dB, with which our proposed IRaptor HARQ has an overall higher throughput than that of the LTE HARQ in Figure A2-12, the PLR also has a 3.5dB gain than the PLR of the LTE HARQ, as seen in A2-11.

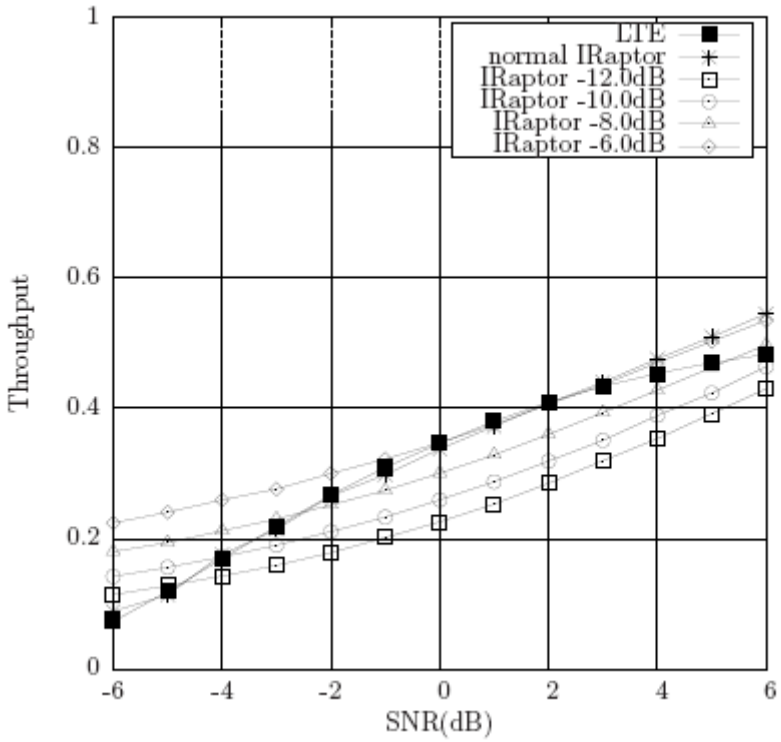


Figure A2-12: Throughput versus SNR for all candidate HARQ schemes over a quasi-static Rayleigh fading channel.

A2.5. Conclusion

Fountain codes naturally own a fantastic rateless property. They are widely used for protecting packet transmissions over BECs. However, unsatisfactory performance is expected when they act as channel codes in the physical layer. In this appendix, we analysed the reason of their undesirable performance is not only caused by the channel conditions, but also by the tanner graph structure, especially for the transmissions of finite-length block. Aiming for three structural flaws, we conceived a method by transmitting several special encoded bits which are deliberately generated to remedy these flaws, based on the rebuilt tanner graph at the transmitter. Simulation results have demonstrated that the improved Raptor codes exhibit an excellent PLR and complexity performance for the short and medium block lengths, only requiring less memories and simple operations.

Furthermore, a novel IRaptor coded HARQ based on the LUT was proposed in this appendix, beneficial from the rateless property of IRaptor codes. The basic principle is that an appropriate number of IRaptor encoded bits will be transmitted over a channel in the mild fading, while only a small number of bits have occupied the transmission for the deep-fading channel. Here, the appropriate number of encoded bits is determined according to the coding rate pre-stored in the LUT, which guarantees a low PLR of IRaptor codes. Simulation results showed that our proposed IRaptor HARQ may achieve a significantly low PLR performance, as well as remain a high throughput.

List of Acronyms

Acronym	Meaning
3GPP	3rd Generation Partnership Project
16-QAM, 64-QAM	16- Quadrature Amplitude Modulation / 64-QAM
ACK	Acknowledge
AGTI	Access Grant Time Interval
AMF	Address Mapping Function
APN	Access Point Name
ARQ	Automatic Repeat reQuest
ASN	Abstract Syntax Notation
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BLER	Block Error Rate
BPSK	Binary Phase Shift Keying
BSR	Buffer Status Reporting
CA	Carrier Aggregation
CAPEX	Capital Expenditure
CBA	Collaborative Broadcast Architecture
CDMA	Code Division Multiple Access
CH	Cluster Head
CMAS	Commercial Mobile Alert Service
CP	Cyclic Prefix
CQI	Channel Quality Indicator
CRC	Cyclic Redundancy Check
CRDSA	Contention Resolution Diversity Slotted Aloha
C-RNTI	Cell Radio Network Temporary Identifier
CRS	Cell specific or Common Reference Signals
CSI	Channel State Information
DFT	Discrete Fourier Transform
DL	Downlink
DL-RS	DR Reference Signal
DL-SCH	DL Shared Channel
DM-RS	Demodulation Reference Signal
DRX	Discontinuous Reception
eNodeB	Enhanced Node B
EPC	Evolved Packet Core
ETWS	Earthquake and Tsunami Warning System
E-UTRA(N)	Evolved UMTS Terrestrial Radio Access (Network)
EXALTED	EXPanding LTE for Devices
FDD	Frequency Division Duplex
FEC	Forward Error Correction
FFT	Fast Fourier Transform
FTP	File Transfer Protocol
GBR	Guaranteed Bit Rate
GF	Galois Field
GFDM	General Frequency Division Multiplexing
GPRS	General Packet Radio Service
GSM	Global System for Mobile communications
GW	Gateway

H-ARQ	Hybrid Automatic Repeat reQuest
HSS	Home Subscriber Servers
HTTP	HyperText Transfer Protocol
HYGIENE	HurrYGuided-Irrelevant-Eminent-NEeds
IFFT	Inverse Discrete Fourier Transform
IMEI	International Mobile Equipment Identities
IMSI	International Mobile Subscriber Identities
ID	Identification
IP	Internet Protocol
ITS	Intelligent Transport System
IRaptor	Improved Raptor
KPI	Key Performance Indicator
LDPC	Low Density Parity Check
LOS	Line Of Sight
LTE	Long Term Evolution
LTE-A	LTE-Advanced
LTE-M	LTE for MTC
LUT	Look-Up Table
M2M	Machine-to-Machine
M2M-GW	M2M traffic for M2M Gateway
MAC	Medium Access Control
MBCCH	MTC Broadcast Control Channel
MBCH	MTC Broadcast Channel
MBMQ	Multimedia Broadcast Multicast Services
MBSFN	Multicast-Broadcast Single Frequency Network
MCCCH	MTC Common Control Channel
MCS	Modulation Coding Scheme
MDCCH	MTC Dedicated Control Channel
MDL-SCH	MTC Downlink Shared Channel
MDTCH	MTC Dedicated Traffic Channel
MF	Mobility Factor
MIMO	Multiple Input Multiple Output
MMCCH	MTC Multicast Control Channel
MMCH	MTC Multicast Channel
MME	Mobility Management Entity
MMTCH	MTC Multicast Traffic Channel
MPCCH	MTC Paging Control Channel
MPCH	MTC Paging Channel
MRACH	MTC Random Access Channel
MRATCH	MTC Random Access Traffic Channel
M-SI-RNTI	MTC System Information Radio Network Temporary Identity
MTC	Machine Type Communications
MTCH	MTC Traffic Channel
M-UL-SCH	MTC Uplink Shared Channel
MUX	Multiplexing
NACK	Non-Acknowledge
NAS	Non-Access Stratum protocol
NCDP	Network Coded Diversity Protocol
NRT	Non-Real-Time
NRTV	Near Real Time Video
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access

OPEX	Operational Expenditure
PAPR	Peak to Average Power Ratio
PBCH	Physical Broadcast Channel
PCell	Primary Cell
PCFICH	Physical Control Format Indicator channel
PDCCH	Physical Downlink Control Channel
PDCP	Packet Data Convergence Protocol
PDN	Packet Data Network
PDP	Packet Data Protocol
PF	Paging Factor
P-GW	PDN Gateway
PHY	Physical Layer
PL	Path Loss
PLR	Packet Loss Ratio
PMCH	Physical Multicast Channel
PMDBCH	Physical MTC Downlink Broadcast Channel
PMDCCH	Physical MTC Downlink Control Channel
PMDMCH	Physical MTC Downlink Multicast Channel
PMDSCH	Physical MTC Downlink Shared Channel
PMI	Precoding Matrix Indicator
PMPCH	Physical MTC Paging Channel
PMRACH	Physical MTC Random Access Channel
PMUCCH	Physical MTC Uplink Control Channel
PMUSCH	Physical MTC Uplink Shared Channel
PSS	Primary Synchronization Signal
PO	Paging Ocasion
PUSCH	Physical Uplink Shared Channel
QCI	QoS Class Indicator
QoS	Quality of Service
QPSK	Quadrature Phase Shift Keying
RA	Random Access
RACH	Random Access Control Channel
RA-RNTI	Random Access Radio Network Temporary Identifier
RAT	Radio Access Technology
RB	Resource Block
RF	Radio Frequency
RI	Rank Indicator
RLC	Radio Link Control
RLC-AM	RLC Acknowledge Mode
RLC-UM	RLC Unacknowledge Mode
RM	Rate Matching
RRC	Radio Resource Control
RSRP	Reference Signal Received Power
RSRQ	Reference Signal Received Quality
RSSI	Received Signal Strength Indicator
RT	Real-Time
Rx	Receiver
SCell	Secondary Cell
SC-FDMA	Single Carrier Frequency Division Multiple Access
SCM	Spatial Channel Model
SDU	Service Data Unit

SIB	System Information Bloc
SMM	Smart Metering and Monitoring
SNR	Signal to Noise Ratio
(S)SS	(Secondary) Synchronization Signal
TA	Tracking Area
TAI	Tracking Area Identity
TB	Transport Block
TR	Technical Requirements
TS	Technical Specifications
TTI	Transmission Time Interval
TTL	Time-To-Live
Tx	Transmitter
UE	User Equipment
UE-EUTRA	UE Evolved Universal Terrestrial Radio Access
UL	Uplink
UL-SCH	UL Shared Channel
VoIP	Voice over IP
WP	WorkPackage

References

[1] FP7 EXALTED consortium, “D3.3 – Final report on LTE-M algorithms and procedures,” project report, Jul. 2012.
[2] FP EXALTED consortium, “D2.3 – The EXALTED system architecture (Final),” project report, Aug. 2012.
[3] 3rd Generation Partnership Project (3GPP), http://www.3gpp.org.
[4] 3GPP, “TR 36.888 – Study on provision of low-cost Machine-Type Communications (MTC) User Equipments (UEs) based on LTE,” Rel.11, Jun. 2012.
[5] FP7 EXALTED consortium, “D2.1 - Description of baseline reference systems, scenarios, technical requirements & evaluation methodology,” project report, v2.0, Jan. 2012.
[6] A. S. Lioumpas, A. Alexiou, C. Antón-Haro and P. Navaratnam “Expanding LTE for Devices: Requirements, Deployment Phases and Target Scenarios” European Wireless 2011, Vienna, Austria, Apr. 2011.
[7] FP EXALTED consortium, “D4.1 – M2M Packet Data Protocols between LTE-M and Capillary Networks,” project report, Jun. 2012.
[8] FP EXALTED consortium, “D4.2 – IP Networking System for M2M communications for EXALTED use cases,” project report, v2.0, Oct. 2012.
[9] 3GPP, “TS 36.201 – LTE physical layer; General description,” Rel.10, V10.0.0, Dec. 2010.
[10] 3GPP, “TS 36.211 – Physical Channels and Modulation,” Rel.10, V10.5.0, Jun. 2012.
[11] 3GPP, “TS 36.212 – Multiplexing and channel coding,” Rel.10, V10.6.0, Jun. 2012.
[12] FP7 EXALTED consortium, “D3.1 – First Report on LTE-M Algorithms and Procedures,” project report, v2.0, Jan. 2012.
[13] 3GPP, “TS 36.214 – Physical Layer Measurements,” Rel.10, V10.1.0, Mar. 2011.
[14] 3GPP, “TS 36.321 – Medium Access Control (MAC) protocol specification,” Rel.10, V10.6.0, Sep. 2012.
[15] C. Qian, H. Chen, Y. Ma and R. Tafazolli, “A Novel Adaptive Hybrid-ARQ Protocol for Machine to Machine Communications,” IEEE Vehicular Technology Conference, VTC-Spring, Dresden, June 2013.
[16] 3GPP, “TS 36.222 - Radio Link Control (RLC) protocol specification,” Rel.10, v10.0.0, Dec. 2010.
[17] 3GPP, “TS 36.323 – Packet Data Convergence Protocol (PDCP) specification,” Rel.10, v10.1.0, Mar. 2011.
[18] 3GPP, “TS 36.331 – Radio Resource Control (RRC); Protocol specification,” Rel.10, V10.7.0, Sep. 2012.
[19] 3GPP, “TS 36.306 – Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) radio access capabilities,” Rel.11, V11.1.0, Sep. 2012.
[20] 3GPP, “TS 36.213 – Physical layer procedures,” Rel.10, V10.7.0, Sep. 2012.
[21] 3GPP, “TS 45.005 – Radio transmission and reception,” Rel.11, V11.1.0, Aug. 2012.
[22] Alcatel-Lucent, Alcatel-Lucent Shanghai Bell, “R1-130463 – Channel characteristics and channel estimation for extended coverage MTC,” 3GPP TSG RAN WG1 #72, St. Julians, Malta, Jan. 2013.
[23] 3GPP, “TR 25.996 - Spatial channel model for Multiple Input Multiple Output (MIMO) simulations,” Rel.10, Mar. 2011.
[24] 3GPP, “TR 36.814 – Further advancements for E-UTRA physical layer aspects,” technical report, Rel.9, V9.0.0, Mar. 2010.
[25] 3GPP, “TS 36.101 - User Equipment (UE) radio transmission and reception,” Rel.10, V10.8.0, Sep. 2012.
[26] Iversen V.B., “Tele-traffic Engineering and Network Planning 2011,” URL http://oldwww.com.dtu.dk/education/34340.

[27] 3GPP, “TS 23.203 – Policy and charging control architecture,” Rel.11, V11.8.0, Dec. 2012.
[28] 3GPP, TSG RAN1 Meeting 48, Orange et al, “R1-070674, LTE physical layer framework for performance verification,” Feb. 2007.
[29] 3GPP, “TS 25.892 – Feasibility Study for Orthogonal Frequency Division Multiplexing (OFDM) for UTRAN enhancement”, Rel.6, Jun. 2004.
[30] 3GPP, “TR 37.868 – Study on RAN Improvements for Machine-type Communications,” Rel.11, Sep. 2011.
[31] N. Michailow, S. Krone, M. Lentmaier, G. Fettweis, “Bit Error Rate Performance of Generalized Frequency Division Multiplexing,” Vehicular Technology Conference, VTC-Fall, Quebec City (Canada), Sep. 2012.
[32] Alcatel-Lucent, Alcatel-Lucent Shanghai Bell, “R2-113183 - Further Study of Access Performance for MTC,” 3GPP TSG RAN WG2 #74, Barcelona, Spain, May 2011.
[33] Alcatel-Lucent, Alcatel-Lucent Shanghai Bell, “R2-114391 - Integrated Slotted Access with EAB for MTC,” 3GPP TSG RAN WG2 #75, Athens, Greece, Aug. 2011.
[34] G. Cocco, C. Ibars, D. Gündüz, O. d. R. Herrero, “Collision resolution in slotted ALOHA with multi-user physical-layer network coding,” in Proceedings of IEEE Vehicular Technology Conference (VTC 2011-Spring), 15-18 May 2011, Budapest (Hungary).
[35] G. Cocco, C. Ibars, D. Gündüz, O. d. R. Herrero, "Collision resolution in multiple access networks with physical-layer network coding and distributed fountain coding", in Proceedings of the IEEE 36th International Conference on Acoustics, Speech and Signal Processing (ICASSP 2011), 22-27 May 2011, Prague (Czech Republic).
[36] G. Cocco, N. Alagha, C. Ibars, S. Cioni, "Practical Issues in Multi-User Physical Layer Network Coding", in Proceedings of the 6th IEEE Advanced Satellite Mobile Systems Conference (ASMS 2012), 5-7 September 2012, Baiona (Spain).
[37] N. Michailow, R. Datta, S. Krone, M. Lentmaier and G. Fettweis, “Generalized Frequency Division Multiplexing: A Flexible Multi-Carrier Modulation Scheme for 5th Generation Cellular Networks,” GeMiC 2012: the 7th German Microwave Conference, Ilmenau University of Technology, Germany, Mar. 2012.
[38] E. Casini, R. De Gaudenzi, and O. del Rio Herrero, “Contention resolution diversity slotted ALOHA (CRDSA): an enhanced random Access scheme for satellite access packet networks,” IEEE Trans. on Wireless Communications, vol. 6, no. 4, pp. 1408-1419, Apr. 2007.
[39] A. Gotsis, A. S. Lioumpas and A. Alexiou, “Challenges and New Perspectives on Machine-to-Machine Scheduling over LTE Networks,” 28th meeting of the Wireless World Research forum, Piraeus, Greece.
[40] A. G. Gotsis, A. S. Lioumpas and A. Alexiou, "M2M Scheduling over LTE: Challenges and new Perspectives," IEEE Vehicular Technology Magazine, vol. 7, no. 3, pp. 34-39, Sep. 2012
[41] A. Gotsis, A. S. Lioumpas and A. Alexiou, “Evolution of Packet Scheduling for Machine-Type Communications over LTE: Algorithmic Design and Performance Analysis,” IEEE Globecom 2012 Workshops, California, U.S.A, Dec. 2012.
[42] A. S. Lioumpas and A. Alexiou, “Uplink Scheduling for Machine-to-Machine Communications in LTE-based Cellular Systems,” IEEE Globecom Workshops, Houston, U.S.A, Dec. 2011.
[43] Alcatel-Lucent, Alcatel-Lucent Shanghai Bell, “R1-125205 – Requirements for coverage improvement to MTC devices,” 3GPP TSG RAN WG1 #71, New Orleans, USA, Nov. 2012.
[44] Alcatel-Lucent, Alcatel-Lucent Shanghai Bell, “R1-124883 – Potential solutions for improved coverage for MTC UEs,” 3GPP TSG RAN WG1 #71, New Orleans, USA, Nov. 2012.
[45] Alcatel-Lucent, Alcatel-Lucent Shanghai Bell, “R1-130462 – Feasibility of coverage extension of physical channels for MTC devices,” 3GPP TSG RAN WG1 #72, St. Julians, Malta, Jan. 2013.

[46] G. Cocco, C. Ibars, N. Alagha, "Cooperative Coverage Extension in Heterogeneous Machine-to-Machine Networks", in proceedings of IEEE Global Communications Conference (GLOBECOM 2012): Second International Workshop on Machine-to-Machine Communications 'Key' to the Future Internet of Things (GC'12 Workshop-IWM2M), 3-7 December 2012, Anaheim, California (USA).
[47] W. Nitzold, M. Lentmaier and G. P. Fettweis, “Spatially Coupled Protograph-Based LDPC Codes for Incremental Redundancy,” 7th International Symposium on Turbo Codes & Iterative Information Processing, Sweden, Aug. 2012.
[48] A. S. Lioumpas and A. Alexiou, “On the Switching Rate of ST-MIMO Systems with Energy-based Antenna Selection,” EuCap, Rome, Italy, Apr. 2011.
[49] 3GPP, “TS 23.401 – General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access,” Rel.10, V10.8.0, Jun. 2012.
[50] 3GPP, “TS 36.304 – Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) procedures in idle mode," Rel.10, V10.5.0, Mar. 2012.
[51] Alcatel-Lucent, Alcatel-Lucent Shanghai Bell, “R1-120510 – Support of reduced maximum bandwidth for low-cost MTC UEs,” 3GPP TSG RAN WG1 #68, Dresden, Germany, Feb. 2012.
[52] Alcatel-Lucent, Alcatel-Lucent Shanghai Bell, “R1-121255 – On single receive RF chain for low-cost MTC UEs, “3GPP TSG RAN WG1 #68bis, Jeju, South Korea, Mar. 2012.
[53] Alcatel-Lucent, Alcatel-Lucent Shanghai Bell, “R1-121256 – On reduction of maximum transmit power for low-cost MTC UEs, “3GPP TSG RAN WG1 #68bis, Jeju, South Korea, Mar. 2012.
[54] Alcatel-Lucent, Alcatel-Lucent Shanghai Bell, “R1-121257 – On half duplex operation for low-cost MTC UEs, “3GPP TSG RAN WG1 #68bis, Jeju, South Korea, Mar. 2012.
[55] FP7 EXALTED consortium, “D2.4 - The EXALTED system concept and its performance,” project report, to appear in Feb. 2013.
[56] D.J.C. MacKay, “Fountain codes,” IEE Proceedings-Communications, vol. 152, no. 6, pp. 1062-1608, Dec. 2005.
[57] M. Luby, “LT codes,” in Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science, pp. 271-280, 2002.
[58] A. Shokrollahi, “Raptor codes,” IEEE Trans. on Information Theory, vol. 52, no. 5, pp. 2033–2051, May 2006.
[59] U. Erez, M. Trott, and G. Wornell, “Rateless coding for Gaussian channels,” IEEE Trans. on Information Theory, vol. 58, no. 2, pp. 530 –547, Feb. 2012.
[60] R. Palanki and J. Yedidia, “Rateless codes on noisy channels,” in Proceedings of International Symposium on Information Theory, p. 37, Jun. 2004.
[61] J. Castura and Y. Mao, “Rateless coding over fading channels,” IEEE Communications Letters, vol. 10, no. 1, pp. 46 – 48, Jan. 2006.
[62] O. Etesami and A. Shokrollahi, “Raptor codes on binary memoryless symmetric channels,” IEEE Trans. on Information Theory, vol. 52, no. 5, pp. 2033 – 2051, May 2006.
[63] K. Hu, J. Castura, and Y. Mao, “Performance-complexity tradeoffs of raptor codes over Gaussian channels,” IEEE Communications Letters, vol. 11, no. 4, pp. 343 –345, Apr. 2007.
[64] B. Sivasubramanian and H. Leib, “Fixed-rate Raptor codes over Rician fading channels,” IEEE Trans. on Vehicular Technology, vol. 57, no. 6, pp. 3905 –3911, Nov. 2008.
[65] Z. Cheng, J. Castura, and Y. Mao, “On the design of raptor codes for binary-input Gaussian channels,” IEEE Trans. on Communications, vol. 57, no. 11, pp. 3269 –3277, Nov. 2009.
[66] F. Kschischang, B. Frey, and H.-A. Loeliger, “Factor graphs and the sum-product algorithm,” IEEE Trans. on Information Theory, vol. 47, no. 2, pp. 498 –519, Feb. 2001.

[67] H. Jenka_c and T. Mayer, “Soft decoding of LT-codes for wireless broadcast,” in Proceedings of IST Mobile Summit, Jun. 2005.
[68] X. Liu and T. J. Lim, “Fountain codes over fading relay channels,” IEEE Trans. on Wireless Communications, vol. 8, no. 6, pp. 3278–3287, Jun. 2009.
[69] N. Bonello, R. Zhang, S. Chen, and L. Hanzo, “Reconfigurable rateless codes,” IEEE Trans. on Wireless Communications, vol. 8, no. 11, pp. 5592 –5600, Nov. 2009.
[70] J. Hagenauer, E. Offer, and L. Papke, “Iterative decoding of binary block and convolutional codes,” IEEE Trans. on Information Theory, vol. 42, no. 2, pp. 429 –445, Mar. 1996.
[71] S. Tong, P. Wang, D. Wang, and X. Wang, “Box-minus operation and application in sum-product algorithm,” Electronics Letters, vol. 41, no. 4, pp. 197–198, Feb. 2005.
[72] H. Chen, R. G. Maunder, and L. Hanzo, “Low-complexity multiple component turbo decoding aided hybrid ARQ,” IEEE Trans. on Vehicular Technology, vol. 60, pp. 1571–1577, May 2011.
[73] J. Hagenauer, “The EXIT chart - introduction to extrinsic information transfer in iterative processing,” in Proceeding of 12th European Signal Processing Conference, pp. 1541–1548, 2004.
[74] X. Zuo, R. Maunder, and L. Hanzo, “Design of fixed-point processing based LDPC codes using EXIT charts,” in IEEE Vehicular Technology Conference, VTC-Fall, Sep. 2011.