Joint Resource Configuration and MCS Selection Scheme for Uplink Grant-Free URLLC

Jacobsen, Thomas; Abreu, Renato Barbosa; Berardinelli, Gilberto; Pedersen, Klaus I.; Kovács, István; Mogensen, Preben Elgaard

Published in:
IEEE GlobeCOM 2018 Workshop proceedings

Publication date:
2018

Document Version
Early version, also known as pre-print

Link to publication from Aalborg University

Citation for published version (APA):

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

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

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

Downloaded from vbn.aau.dk on: December 01, 2018
Abstract—Ultra-reliable and low-latency communications (URLLC) addresses the most challenging set of services for 5G New Radio. Uplink grant-free transmissions is recognized as a promising solution to meet the ambitious URLLC target (1 ms latency at a 99.999% reliability). Achieving such a high reliability comes at the expense of poor spectral efficiency, which ultimately affects the load supported by the system. This paper proposes a joint resource allocation solution including multiple modulation and coding schemes (MCSs) and power control settings for grant-free uplink transmissions on shared resources. The scheme assigns smaller bandwidth parts and higher MCS to the UEs in good average channel conditions, reducing the probability of fully overlapping transmissions. The performance analysis shows that the scheme is capable of increasing the system outage capacity by ∼90%, compared to prior art solutions using a conservative single-MCS configuration with fully overlapping transmissions.

I. INTRODUCTION

One of the major goals of 5G New Radio (NR) is the support of ultra-reliable and low-latency communication (URLLC) to enable mission-critical applications. Meeting the strict URLLC requirements with a $10^{-5}$ packet failure probability within 1 ms is very challenging [1]. Many technology components towards achieving this have been investigated such as short transmission time intervals (TTIs) [2], semi-persistent scheduling (SPS) [3], fast hybrid automatic repeat request (HARQ) [4], and robust error correction coding [5].

For meeting the URLLC requirements in uplink, grant-free (GF) solutions have been found to be attractive, as time-consuming steps of grant-based scheduling and its potential errors are avoided [6], [7]. For 5G NR (Release-15) it has been agreed that GF transmissions happen according to a predefined configuration which includes power control settings, modulation and coding scheme (MCS), time-frequency resource allocation, among others. At most one GF configuration per bandwidth part is active at a time [8]. This is communicated to the user equipment (UE) by radio resource control (RRC) with possible activation via downlink control channel [9]. For GF transmissions, it is further assumed; that a configuration can be shared by multiple UEs [10], the MCS and transmission bandwidth is fixed [11], [12] and open loop power control is used [13].

It is known from numerous LTE uplink studies that dynamic link adaptation is beneficial. Using a combination of open and closed loop power control, and fast adaptive modulation and coding (AMC) based on channel state information (CSI) acquired by sounding brings clear benefits for mobile broadband traffic [14], [15]. This is found to be the case for dynamically scheduled transmissions, adjusting the MCS on a TTI basis. However, for GF URLLC cases, the situation is different. First, the URLLC traffic per UE is sporadic with small payloads appearing infrequently at the users for immediate uplink transmission. This means that there are no steady transmissions from the users that the base station nodes can utilize for CSI estimation. Secondly, as GF URLLC rely on fast uplink access without grant, there is no downlink signaling for conveying MCS and transmission bandwidth adjustments per transmission event. Finally, URLLC target transmissions where one URLLC packet is included in each transmission, as segmentation of URLLC payloads over multiple transmissions risks jeopardizing the latency targets of URLLC. Our hypothesis is therefore that a new joint MCS and transmission bandwidth selection method for GF URLLC transmission could help boosting the aggregated URLLC traffic that can be tolerated in the network.

We therefore propose a solution encompassing a hierarchical resource configuration that facilitates uplink transmissions of URLLC payloads (of fixed size) using different MCS schemes and transmission bandwidths. The idea is to allow partly overlapping transmissions with corresponding adjustments of the users MCS and power control settings. In short, we propose a solution where users are assigned to use different GF transmission settings according to a predefined resource grid, consisting of MCSs and different transmission sub-bands. The scheme allows to efficiently leverage the trade-offs between reducing the uplink collision probabilities by using lower transmission bandwidth per user versus the cost in terms of higher required signal-to-interference-plus-noise ratio (SINR) from using higher order MCS. The value of the proposed scheme is studied in a dynamic multi-user, multi-cell environment in line with the 3GPP NR assumptions.

Due to the high degree of complexity of the system model, we rely on state-of-the-art system level simulations to preserve the high degree of realism, which would otherwise be jeopardized if imposing simplifications to allow analytical performance analysis. The simulations are based on the widely accepted models agreed in 3GPP for NR studies, and were also used for the works in [16], [17]. Finally, special care is given to ensure that statistically reliable performance results are generated, such that mature conclusions can be drawn.
The rest of the paper is structured as follows: Section II outlines the system model and objectives of the study. Section III presents the proposed resource configuration. Section IV outlines the simulation assumptions, while Section V presents the performance results. Section VI concludes the study.

II. SYSTEM MODEL AND PERFORMANCE METRICS

A. Network and transmission model

A multi-cell synchronous network is assumed, following the 3GPP guidelines as in [10], [16], [17]. A fixed number of U URLLC UEs are deployed in the cells and are assumed to be uplink synchronized and in connected state. Small packets of fixed size $B$ bytes are generated by each UE according to independent Poisson arrival processes with an average packet arrival rate $\lambda$. Grant-free uplink transmissions occur in a framed structure based on OFDM, frequency-division duplexing (FDD) and short-TTI [2]. The GF resources are shared by the $U$ UEs in the cell. In this sense, transmissions can occur simultaneously on the same time/channel resources (collisions). The successful reception of the packets depends on the used MCS and the post-processing SINR achieved after the receiver combining. Multi-user detection is assumed, therefore overlapping transmissions can be received depending on the resultant SINR [18]. If the reception fails the UE issues a HARQ retransmission after processing the feedback from the base station (BS) [17]. Chase-combining is used to improve the decoding performance after each retransmission.

B. Power control

Power control is utilized to regulate the transmit power in order to meet a target receive power and limit the generated interference in the network. We assume open-loop power control for the transmissions as in LTE [19], such that the UE transmit power is given by

$$P_{\text{dBm}} = \min\{P_{\text{max}}, P_0 + 10 \log_{10}(M) + \alpha P_L + \Delta_{\text{MCS}}\}, \quad (1)$$

where $P_{\text{max}}$ is the maximum transmit power, $P_0$ is the target receive power per resource block (RB), $M$ is the number of used RBs, $\alpha$ is the fractional pathloss compensation factor, $P_L$ is the slow faded pathloss and $\Delta_{\text{MCS}}$ is a power offset per RB that can be applied depending on the MCS. The $\Delta_{\text{MCS}}$ setting will be further discussed in this paper. As discussed in [13], we apply full pathloss compensation ($\alpha = 1$).

C. Performance metric

We adopt the performance target for URLLC defined by 3GPP [1]; a success probability of $1 - 10^{-5}$ to receive a small packet (32 bytes) in the radio interface with a maximum one-way latency of 1 ms.

The prior-art solutions use a conservative single-MCS, to meet the performance target [11], [13], [17]. In the baseline case, all UEs use the full band in an entire TTI, using QPSK1/8 as the conservative single-MCS. Our target is to improve the achievable load per cell ($L [\text{b/s}] = \lambda \cdot U \cdot B \cdot 8$) in the network, which meets the URLLC performance target, compared to the baseline. This load is referred to as the system outage capacity.

III. JOINT RESOURCE ALLOCATION AND MCS SELECTION

A. Resource allocation

The proposed hierarchical resource allocation scheme encompasses multiple transmission bandwidths and power control settings associated with the MCSs for grant-free transmissions. The scheme uses the resources within a bandwidth part of size $BW$. Each MCS is univocally associated to a specific sub-band size $\leq BW$. The supported set of MCS, $\mathbb{M}$, includes $N$ MCS options denoted by $M_{\text{CS}}(k)$, with index $n \in [1, N]$ and $k$ is the ratio between the bandwidth $BW$ and the sub-band size associated to the MCS. Shortened MCS notation can omit $k$. $\mathbb{M}$ is sorted such that $M_{\text{CS}}(1)$ has the lowest modulation and coding rate, i.e. the most conservative option and use the full bandwidth $BW$. Higher MCS options form a set $\mathbb{M}_{1\ldots n} \subset \mathbb{M}$ for $n > 1$, which are mapped to sub-bands of size $BW \cdot k^{-1}$ with $k > 1$. Considering the strict latency requirement for URLLC traffic, the MCS options and $k$ are chosen such that the URLLC payload can be fully transmitted in the corresponding sub-bands without segmentation. The UEs are pre-configured via RRC signaling with the resource allocation scheme, defining the sub-bands RBs, the set of corresponding MCSs and the power offsets.

Fig. 1 shows an example configuration of the resource grid, i.e. the sub-bands and MCS options, where the set $\mathbb{M} = \{M_{\text{CS}}(1), M_{\text{CS}}(2), M_{\text{CS}}(4)\} = \{\text{QPSK1/8}, \text{QPSK1/4}, \text{QPSK1/2}\}$ is supported. Each MCS has an associated $\Delta_{\text{MCS}}$. Transmissions with $M_{\text{CS}}$ use all the 48 RBs, while transmissions with $M_{\text{CS}}$ or $M_{\text{CS}}$ use sub-bands of size 24 and 12 RBs respectively. Fig. 2 illustrates examples of GF transmissions and their overlap which can occur using the configuration illustrated in Fig. 1. Fully overlapping transmissions can occur for transmissions using the same MCS whereas transmissions using different MCS can partially overlap.

The BS can estimate and decide, e.g. based on infrequent UE reports, the MCS and corresponding sub-band to be used and indicate it to the UE through downlink signaling. If multiple sub-bands are associated to the MCS, either the BS assigns one or allows the UE to randomly select. By knowing the possible combinations of transmitting UEs, $\mathbb{M}$ and the associated sub-bands, the blind decoding complexity at the receiver side is bounded. UEs in good average channel condition can be signaled to use one of the higher MCS options ($\mathbb{M}_{1\ldots n}$) instead of the conservative $M_{\text{CS}}$. Since
higher MCSs are leveraged through smaller bandwidth parts, the collision probability is reduced among the sub-bands, while UEs operating simultaneously with lower order MCSs are only partly overlapped. This can be of mutual benefit to the UEs in the network and potentially increase their achieved reliability and in the end the system outage capacity. The price to pay for UEs using $M_{1+}$ is that they need a corresponding higher power spectral density in order to maintain the reliability of their transmissions, which means that the interference in the used sub-band is increased. The power spectral density offset can be configured for the power control defined in (1), but due to the transmit power limitation $P_{\text{max}}$, it can not be guaranteed that $\Delta_{\text{MCS}}$ can be fully applied. For this reason, only UEs with sufficient transmit power headroom to fully apply $\Delta_{\text{MCS}}$ should use $M_{1+}$.

The choice of $\Delta_{\text{MCS}}$ should consider the higher SINR targets for $M_{1+}$, the power headroom, and the generated interference. Further, the values can be predetermined from the difference in required SINR to maintain a block error rate (BLER) target, which can be found using BLER/SINR curves obtained using extensive link-level simulations. As an initial setting we propose to use

$$\Delta_{\text{MCS},n}[\text{dB}] = 10\log_{10}(k), \quad (2)$$

such that the target transmit power is maintained, and apply fine-tuning based on the observed outage performance.

B. MCS selection scheme

We propose a simple MCS and correspondent bandwidth selection scheme which is defined using a set of $N-1$ coupling-gain thresholds $C = \{C_1, ..., C_{N-1}\}$ sorted in ascending order. The $MCS_n$ is selected according to $n = \arg \min \{ C_i | C \leq C_i \}$, where $C$ is the experienced coupling-gain which is inversely proportional to the pathloss. The selection is done such that the lower the coupling-gain is, the more conservative is the used MCS. For $C > C_{N-1}$, $MCS_N$ is used. Note that, the idea of grouping the UEs based on coupling-gain thresholds is similar to the one used in NB-IoT [20].

The choice of $C_T$ depends on the scenario, $M$ and the power control settings. For this reason an expression valid for all deployment scenarios is not straightforward. We propose that $C_T$ is chosen based on outage statistics computed using one-way latency measurements collected at the BS, prior to applying the joint resource and MCS selection scheme, and sorted into coupling-gain intervals. Good candidates for threshold values are found between intervals where the outage probability increases significantly.

C. Example of partly overlapping transmissions

In this section we give an example of how a resource configuration with $M_{1+}$ can give SINR improvements compared to a single-MCS configuration. Consider the simple example illustrated in Fig. 3, where two UEs transmit with fully overlapping transmissions on the left and the alternative configuration on the right. For simplicity, this example does not consider the effect of fading.

In the first case, $UE_a$ and $UE_b$ use $MCS_1$ in full band with $w$ RBs. In the alternative configuration, $UE_b$ is configured to use a higher MCS $MCS_2 \in M_{1+}$ and hence uses a smaller bandwidth of $m$ RBs, ensuring that when both UEs transmit simultaneously their transmissions only partly overlap. $UE_b$ use $\Delta_{\text{MCS}_2}$ to increase its power spectral density. The post detection SINRs of the used RBs are averaged per RB for computation of the effective SINR of the data stream.

The resultant SINR of the two fully overlapping transmissions for $UE_a$ and $UE_b$ can be expressed by $\gamma_a = P_a/(N_0 + P_b)$ and $\gamma_b = P_b/(N_0 + P_a)$ respectively, where $N_0$ is the Gaussian noise spectral density, $P_a$ and $P_b$ are the power spectral density (PSD) from $UE_a$ and $UE_b$ respectively, giving $\gamma_a = \gamma_b$ for $P_a = P_b$. With the partial overlapping configuration, the transmission from $UE_b$ uses a higher spectral density $P_b = P_b \cdot 10^{\Delta_{\text{MCS}_2}/10}$, resulting in an SINR expressed by $\gamma_b = P_b/(N_0 + P_a)$. The SINR for $UE_a$ maintaining $MCS_1$ and $P_a$ can be expressed by

$$\gamma_a = \frac{w-m}{w} \frac{P_a}{N_0 + m} + \frac{m}{w} \frac{P_a}{N_0 + 10^{\Delta_{\text{MCS}_2}/10}}. \quad (3)$$

An evaluation of the SINR gain $\gamma_a/\gamma_a$ using (3) is shown in Fig. 4 considering different PSDs $P_b/P_a$ and sub-band size ratios $m/w$. It is assumed $w = 48$ RBs, $N_0 = -126$ dBm/RB and $P_a = -131$ dBm/RB. At a given power density ratio, the respective SINR gain for $UE_a$ decreases with the increase of the overlapping ratio. The dashed line follows the performance when $\Delta_{\text{MCS}_2}$ is selected according to (2). An SINR gain for
The simulator models the 5G NR design, adopting the topology with 500 meters of inter-site distance (ISD), using each having 3 sectors composing a regular hexagonal grid of a single layer urban macro network consisting of 7 sites, studies published in [4], [13], [17]. The network layout is as mentioned in Section III-A. The same simulator was also used in the earlier URLLC evaluation [10], [22]. Ideal channel estimation of both the BS receiver type MMSE-IRC, 1 TX x 2 RX UL

\[ \Delta_{\text{MCS}} = \frac{12}{k-1} = 12/48 = 0.25 \text{ meaning } k = 4 \text{ gives an initial } \Delta_{\text{MCS,a}} = 10\log_{10}(4) \approx 6 \text{ dB marked in the point X.} \]

Following the dotted line for \( \Delta_{\text{MCS,a}} > 6 \text{ dB, the SINR of } UE_a \text{ increases together with the ratio } P_b/P_a, \text{ however the SINR gain of } UE_a \text{ reduces. It should be observed that, for low overlapping } m/w \text{ ratios, the increase of } P_b \text{ in relation to } P_a \text{ has lower impact on the SINR gain of } UE_a. \text{ However, for ratios such as } \frac{m}{w} = 0.5 \text{ or higher, there is not much room to adjust } \Delta_{\text{MCS,a}} \text{ without causing a loss in SINR for } UE_a. \text{ Notice that this example does not include the effect of intra sub-band interference, as only 1 UE is considered per MCS, which would affect the observed gains. For this reason, after applying the initial } \Delta_{\text{MCS,a}}, \text{ fine-tuning it can be beneficial, as mentioned in Section III-A.} \]

IV. SIMULATION METHODOLOGY

An advanced system-level simulator is used for assessing the performance of the proposed resource allocation scheme. The simulator models the 5G NR design, adopting the commonly agreed mathematical models in 3GPP for radio propagation, traffic models, key performance indicators, etc [10]. The same simulator was also used in the earlier URLLC studies published in [4], [13], [17]. The network layout is a single layer urban macro network consisting of 7 sites, each having 3 sectors composing a regular hexagonal grid topology with 500 meters of inter-site distance (ISD), using wrap-around [21]. UEs are random distributed (all outdoor), following a spatial uniform distribution. The traffic per UE follows a Poisson arrival process in line with system model in Section II. The offered URLLC traffic load is adjusted by varying the number of users \( U \) per macro-cell area, while keeping \( \lambda = 10 \text{ packets per second (PPS)} \) and \( B = 32 \text{ bytes} \) fixed. The time-granularity of the simulator is one OFDM symbol, and the frequency resolution is one sub-carrier. The main simulation assumptions are described in Table I.

![Fig. 4. SINR gain \( \gamma_a/\gamma_a \) in dB of \( UE_a \) using the MCS1 as a function of \( m/w \) and \( P_b/P_a \) ratios](image)

For each GF transmission from a UE to a BS, the received post detection SINR is calculated (accounting for both inter- and intra-cell interference) assuming a two-antenna receiver and Minimum Mean Square Error Interference Rejection Combining (MMSE-IRC) which is the baseline detector for NR evaluation [10], [22]. Ideal channel estimation of both the desired and the interfering signals is assumed. Based on [23], [24], the SINR values are mapped to the mutual information domain, taking the applied modulation scheme into account. Given the mean mutual information per coded bit (MMIB) and the used coding rate of the transmission, the error probability of the transmission is determined from look-up tables that are obtained from extensive link level simulations.

The simulations of the GF URLLC transmissions are in line with the presented system model; including open loop power control, HARQ with chase combining, queuing, etc. Results from the simulator have been benchmarked against calibration results shared in 3GPP for the NR macro simulation scenario, confirming a good match. To ensure statistical reliable results, information is collected from at least \( 5 \cdot 10^6 \) completed URLLC payload transmissions. With this amount of independent samples the outage probability can be said to be within a 27% error margin around the 10^-3 quantile with 95% confidence using the interval estimation of a binomial proportion [25].
V. RESULTS

This section evaluates a two MCS resource allocation configuration $M = \{MCS_1(1), MCS_2(4)\} = \{QPSK1/8, QPSK1/2\}$. QPSK1/8 is used as the conservative MCS option (as in [13], [17]) and QPSK1/2 as the higher MCS option. We set the initial power spectral density offset $\Delta_{MCS_2} = 6$ dB by following (2).

Fig. 5 shows the outage probability at 1 ms per coupling-gain interval for the baseline and for the proposed scheme. The offered load is 486.4 kbps per cell. To get high accuracy per coupling-gain interval, $50 \cdot 10^6$ transmission latency samples have been collected in the network for this result. The percentage of samples per interval is $\approx 6\%$. Each marker is placed on the maximum coupling-gain of the interval. This means, for example, that the marker on coupling gain $-110$ dB represents the outage in the interval ($-113$ dB, $-110$ dB]. The MCS selection threshold set $C_T$ is defined based on outage probability statistics of one-way latency measurements calculated per coupling-gain interval. The threshold $C_T = C_{T1} = -110$ dB is chosen by observing that below this value the outage probability increases significantly for the baseline configuration, as indicated in the figure.

With the chosen $C_{T_1}$, fine-tuning of $\Delta_{MCS_2}$ is performed. Fig. 5, also shows the performance of the proposed scheme with $\Delta_{MCS_2} = \{6$ dB, $10$ dB, $20$ dB$\}$. Increasing $\Delta_{MCS_2}$ from the initial setting improves the reliability for the UEs using $MCS_2$, while also degrading the reliability for the UEs using $MCS_1$. For $\Delta_{MCS_2} = \{6$ dB, $10$ dB$\}$ the reliability in the intervals using $MCS_2$ are comparable, which indicates that the UEs in these intervals are able to apply the full PSD offset through power control. For a very high PSD offset ($\Delta_{MCS_2} = 20$ dB) the variation on reliability indicates that not all coupling-gain intervals are capable of applying the full offset and reaching the reliability requirement.

The reliability statistics per coupling-gain interval in Fig. 5 does not show the systems overall reliability when combining all latency samples. For that, the latency CCDF for the system is shown in Fig. 6, for both the baseline and the considered scheme with $\Delta_{MCS_2} = \{6$ dB, $10$ dB, $20$ dB$\}$. The staircase behavior comes from HARQ retransmissions [17]. From the figure, it can be seen that the option with $\Delta_{MCS_2} = 10$ dB is capable of reaching the target outage probability of $10^{-5}$ within 1 ms. The baseline is only capable of reaching an outage probability of $3.7 \cdot 10^{-5}$ at the 1 ms latency deadline. Considering the fine-tuning of $\Delta_{MCS_2}$ it can be seen that $\Delta_{MCS_2} = 10$ dB is the best option, indicating that further increasing the offset does not improve the performance.

Fig. 7 shows a sensitivity study of $C_{T_1}$ impact on the outage probability. The threshold that gives the lowest outage for both $\Delta_{MCS_2} = \{6$ dB, $10$ dB$\}$ is $C_{T_1} = -110$ dB, confirming the earlier choice. This coupling-gain threshold value corresponds to $12\%$ of all transmissions using the $MCS_1$ and $88\%$ using $MCS_2$.

Fig. 8 summarizes the achieved overall outage probability at 1 ms comparing the baseline with the proposed joint resource allocation and MCS selection scheme with $\Delta_{MCS_2} = \{6$ dB, $10$ dB$\}$. The maximum supported offered load for the baseline is $256.0$ kbps/cell, which aligns with previous work done in [13]. Using the proposed scheme the supported
load increases to 358.4 kbps/cell using $\Delta_{MCS_2} = 6$ dB and 486.4 kbps/cell using $\Delta_{MCS_2} = 10$ dB. The proposed scheme is capable of increasing the system outage capacity up to 40% using the initial $\Delta_{MCS}$ and a further 35% by fine-tuning it.

VI. CONCLUSION

In this paper we have proposed a joint resource allocation and MCS selection scheme for uplink grant-free URLLC. The scheme allows to pre-define a set of MCSs, transmissions.

and MCS selection scheme for uplink grant-free URLLC. The scheme allows to pre-define a set of MCSs, transmis-

sions.

VI. CONCLUSION

In this paper we have proposed a joint resource allocation and MCS selection scheme for uplink grant-free URLLC. The scheme allows to pre-define a set of MCSs, transmission bandwidths and power offsets. The MCS selection is based on the coupling-gain of the UEs. UEs in good average channel condition have reduced collision probability at the expense of eventual higher interference power in the sub-bands, while UEs in poor average channel conditional have lower degradation with partial overlapping. Compared with a conservative single-MCS configuration, the proposed scheme shows that the system outage capacity can be increased by 90%, up to 486.4 kbps per cell, while still fulfilling the URLLC requirements.

Future work will focus on the potential of multi-site reception and receiver diversity together with the proposed joint resource allocation and MCS selection scheme to further enhance the system capacity for uplink grant-free URLLC transmissions.

ACKNOWLEDGMENT

This research is partially supported by the EU H2020-ICT-2016-2 project ONE5G. The views expressed in this paper are those of the authors and do not necessarily represent the project views.

REFERENCES