On the Multiplexing of Broadband Traffic and Grant-Free Ultra-Reliable Communication in Uplink

Abreu, Renato Barbosa; Jacobsen, Thomas; Berardinelli, Gilberto; Pedersen, Klaus I.; Mahmood, Nurul Huda; Kovacs, Istvan; Mogensen, Preben Elgaard

Published in:
2019 IEEE 89th Vehicular Technology Conference (VTC Spring)

DOI (link to publication from Publisher):
10.1109/VTCSpring.2019.8746589

Publication date:
2019

Document Version
Accepted author manuscript, peer reviewed version

On the Multiplexing of Broadband Traffic and Grant-Free Ultra-Reliable Communication in Uplink

Renato Abreu*, Thomas Jacobsen*, Gilberto Berardinelli*, Klaus Pedersen*†, Nurul H. Mahmood*,
István Z. Kovács†, Preben Mogensen*†
*Dept. of Electronic Systems, Aalborg University; †Nokia Bell Labs, Aalborg, Denmark
Email: {rba, tj, gb}@es.aau.dk

***Abstract*—5G networks should support heterogeneous services with an efficient usage of the radio resources, while meeting the distinct requirements of each service class. We consider the problem of multiplexing enhanced mobile broadband (eMBB) traffic, and grant-free ultra-reliable low-latency communications (URLLC) in uplink. Two multiplexing options are considered; either eMBB and grant-free URLLC are transmitted in separate frequency bands to avoid their mutual interference, or both traffic share the available bandwidth leading to overlaying transmissions. This work presents an approach to evaluate the supported loads for URLLC and eMBB in different operation regimes. Minimum mean square error receivers with and without successive interference cancellation (SIC) are considered in Rayleigh fading channels. The outage probability is derived and the achievable transmission rates are obtained based on that. The analysis with 5G new radio assumptions shows that overlaying is mostly beneficial when SIC is employed in medium to high SNR scenarios or, in some cases, with low URLLC load. Otherwise, the use of separate bands supports higher loads for both services simultaneously. Practical insights based on the approach are discussed.**

I. Introduction

The support for services with heterogeneous requirements is one of the goals of fifth generation (5G) new radio (NR). In particular, the enhanced mobile broadband (eMBB) and ultra-reliable low-latency communications (URLLC) service classes have distinct characteristics in terms of traffic type and key performance indicators. While eMBB tolerates a moderate reliability and focus on high data rates, URLLC targets highly reliable small packets transmissions with short latency deadlines, such as $1\,\mathrm{ms}$ with 99.999% reliability [1].

In uplink, the eMBB traffic can be dynamically scheduled using large block lengths. However, the scheduling request and grant procedure required for a packet transmission are source of delays and errors, which can jeopardize the latency and reliability [2]. Therefore grant-free access, which allows immediate access to the channel without the scheduling procedure, is considered for URLLC [3]. Multiple users can share the same grant-free allocation to improve the radio resource utilization [4]. In a 5G network, the same carrier may need to support both grant-free URLLC and scheduled eMBB traffic. One option is to split the available bandwidth between each service class. However, this may lead to poor spectral efficiency in case of sporadic URLLC transmissions. Sharing the same radio resources for grant-free URLLC and eMBB traffic, with overlaying allocations, might improve the spectral efficiency. The consequence is the mutual interference between the two service classes, which may compromise the reliability of URLLC or degrade the eMBB data rate. Power control schemes and multi-antenna receivers, including successive interference cancellation (SIC), are potential solutions to mitigate the interference [5]. Our interest is then to study whether separate bands or overlaying allocations is preferred for ensuring efficient multiplexing of both services, depending on the scenario, traffic load and receiver characteristics.

Previous works have formed the bases for studying the coexistence of multiple traffic. The capacity of multi-antenna systems with spatial multiplexing is provided in [6], with and without SIC. The work in [7] derives the reliability of the minimum mean square error (MMSE) receiver in Rayleigh channel including multiple interferers. In [8], the overlaying of broadband traffic and sporadic transmissions is studied using basic information theoretic tools. The dynamic multiplexing of URLLC and eMBB traffic is evaluated considering preemption [9] and superposition schemes [10], which can be applied for scheduled transmissions. The recent work in [11] investigates the potential of non-orthogonal multiple access (NOMA) for heterogeneous services, though collisions between URLLC transmissions are not considered. The achievable rates in collision prone resources is discussed in [12] for sporadic URLLC transmissions and linear receivers. Collisions between multiple URLLC transmissions and eMBB transmissions is not considered in the related works.

In this paper we study the multiplexing of eMBB and grant-free URLLC traffic using an analytical framework. The presented methodology is based on the findings in [7] and [8], where achievable rates in different interference scenarios and with different receiver types have been derived. The performance of both service classes is compared using overlaying allocations and separate bands. We describe the outage probability in each case, i.e. the complement of the reliability, considering linear MMSE receiver, and also MMSE with SIC for the case of overlaying transmissions. Numerical analysis is conducted considering NR requirements and numerology. The required rate for URLLC transmissions is obtained and the impact on the supported loads for eMBB and URLLC is evaluated with different settings. Further the paper discusses the implications when either of the multiplexing options are used and comes with concrete recommendations for 5G NR operation with heterogeneous services.

The rest of the paper is organized as follows. Section II describes the system model. Section III presents the outage and achievable load calculation. Numerical results are shown in Section IV and discussed in Section V. Finally, conclusions are drawn in Section VI.

II. System Model

We consider a scenario where users are connected and synchronized to one serving cell for uplink data transmission. N_e active users have eMBB service, while N_u users have URLLC service. The total available bandwidth W can either be split to each service class or be shared for overlaying transmissions, as illustrated in Fig. 1. The users transmit over a flat i.i.d Rayleigh fading channel with additive Gaussian noise. Users with a specific traffic type operate over the same resources.

For separate bands, we define a bandwidth split ratio R. With that, a bandwidth $W_u = WR$ is used for URLLC and a bandwidth $W_e = W(1-R)$ is used for eMBB, with $0 \leq R \leq 1$. For overlaying transmissions, it is assumed that both services use the full band W, so $W_u = W_e = W$. In this case, eMBB signals have an average interferer power relative to URLLC expressed as Ω, i.e. for URLLC users with average receive power $\bar{p_u}$ and eMBB with average receive power $\bar{p_e}$ over the same band, $\Omega = \bar{p_e}/\bar{p_u}$. It is assumed that the users from each service class are power controlled so that they are received with the same average signal-to-noise ratio (SNR). To meet strict latency requirements, the URLLC transmissions occur in a short transmission time interval (TTI) of duration T. Whereas eMBB transmissions use long TTIs which allows to benefit from larger coding gains [13].

The eMBB traffic is resource greedy, inducing an uninterrupted interference to other users that are transmitting simultaneously over the same band. $N_e > 1$ can be seen as the case of multi-user MIMO, in which multiple users are scheduled to transmit over the same time-frequency resources, exploiting the spatial dimension of a multi-antenna receiver [6]. The traffic from each URLLC user is assumed to follow a Poisson distribution with packet arrival rate λ per TTI and fixed payload size of D bits. The outage probability targeted for URLLC transmissions is ϵ_u, while for eMBB transmissions it is ϵ_e. For 5G NR use cases the value of ϵ_u should reach 10^{-5} in one or more transmission attempts, to satisfy the strict reliability requirement. Whereas, in cellular networks such as LTE the value of ϵ_e is in the order of 10^{-1}, for the sake of high throughput [14]. The effect of HARQ retransmissions is not considered in this work.

An MMSE receiver with M antennas is assumed. In the case that the URLLC transmissions overlay eMBB streams, we consider two different approaches: conventional MMSE receiver, and MMSE with SIC. For the latter, we assume that the URLLC transmissions should be identified, e.g. using a reference signal, and decoded first, considering the low latency requirement. Then SIC is employed, assuming that the interference of URLLC transmissions over the eMBB streams is completely canceled out.

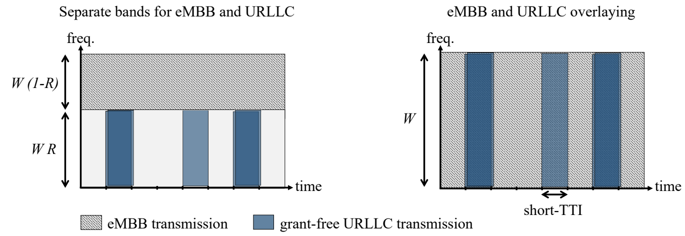


Fig. 1. Separate bands vs. overlaying transmissions for eMBB and URLLC.

III. Analysis of overlaying and separate bands

In this section we present an analytical approach to evaluate the multiplexing of eMBB and sporadic URLLC traffic. The approach builds on top of closed-form solutions that models the reliability for an ideal MMSE receiver with additive interference channels. The model presented in [7] allows to consider each signal source with a different average interferer power relative to a desired source. The outage probability with randomly active sources with the same power characteristics are described and numerically validated in [12]. In this work, we distinguish two classes which can possibly have different average receive SNR, from a total of $v + w$ interferers. v of them have an average interferer power relative to the desired source given by Γ_v. And w interferers have an average interferer power relative to the desired source denoted by Γ_w. We later relate the v interferers as the URLLC ones, and the w interferers as the eMBB ones. The desired source can be either an eMBB or an URLLC signal, that can suffer with interference coming from users of the same or different class. The outage probability for the transmissions subject to interference is calculated as follows [7]:

$$P_f(\bar{\gamma}, v, w, \Gamma_v, \Gamma_w) = 1 - e^{\psi/\bar{\gamma}} \sum_{n=1}^{M} \frac{A_n}{(n-1)!} \left(\frac{\psi}{\bar{\gamma}}\right)^{n-1}, \quad (1)$$

where $\bar{\gamma}$ is the average SNR of the desired source signal at the receiver input, and ψ is the post-combining SINR required for receiving with an outage probability P_f. With the two classes of interferers, we have that

$$A_n = \begin{cases} 1 & \text{if } v + w \leq M - n \\ \dfrac{1 + \sum_{i=1}^{M-n} C_i \psi^i}{(1+\psi\Gamma_v)^v (1+\psi\Gamma_w)^w} & \text{if } v + w > M - n \end{cases}, \quad (2)$$

where C_i is the coefficient of ψ^i in the expansion of $(1+\psi\Gamma_v)^v(1+\psi\Gamma_w)^w$.

In a collision prone scenario the resultant outage probability, can be calculated by combining the collision probability and the outage probability for the given number of interferers [12]. This outage probability can be interpreted as a long term error rate. The probability of having x simultaneous transmissions generated by other y users that are randomly active is

$$P_c(x, y) = \binom{y}{x} P_a^x (1 - P_a)^{y-x}, \quad (3)$$

where P_a is the probability of each user to transmit. In the case of Poisson arrival traffic with arrival rate λ, as we assume for the URLLC users, $P_a = 1 - e^{-\lambda}$.

From that, we describe the outage probability for eMBB and URLLC transmissions for the case of separate bands and for overlaying transmissions.

A. MMSE receiver and separate bands

In the case that a separate band is reserved for each service class, URLLC and eMBB transmissions do not interfere with each other, and their outage probabilities can be derived independently. However, sporadic URLLC transmissions can still collide with each other within the URLLC band. With power control, all the URLLC interferers are assumed to have the same average power at the receiver input as the desired URLLC source. Given that, we assign $v = N_u - 1$ and $\Gamma_v = 1$, while $w = 0$ and $\Gamma_w = 0$ since there is no other type of interferer in the same band. The outage probability for the URLLC transmissions is then given by

$$P_{f,u} = \sum_{z=0}^{N_u-1} P_c(z, N_u - 1) P_f(\bar{\gamma}_u, z, 0, 1, 0), \tag{4}$$

where $\bar{\gamma}_u$ is the average SNR of the URLLC users. Note that (4) is equivalent to the result obtained in [12].

For eMBB, transmission streams from different users can mutually interfere when they are scheduled in the same time-frequency resources, as in the case of multi-user MIMO. Assuming that the eMBB users have the same power control configuration, which leads to the same average power at the receiver as the desired eMBB source, we set $\Gamma_w = 1$. Assuming that all the available resources are simultaneously used by the N_e active users, we have that $w = N_e - 1$. The outage probability of eMBB without URLLC interference can be expressed as

$$P_{f,e} = P_f(\bar{\gamma}_e, 0, N_e - 1, 0, 1), \tag{5}$$

where $\bar{\gamma}_e$ is the average SNR of the eMBB users.

B. MMSE receiver and overlaying transmissions

When URLLC and eMBB have overlaying allocations, the reliability of the URLLC transmissions is not only affected by collisions with sporadic URLLC interferers, but also by the continuous eMBB interferers. Hence, we set $w = N_e$ and $\Gamma_w = \Omega$, besides $\Gamma_v = 1$. With that, the outage probability for the URLLC transmissions is calculated as

$$P_{f,u} = \sum_{z=0}^{N_u-1} P_c(z, N_u - 1) P_f(\bar{\gamma}_u, z, N_e, 1, \Omega). \tag{6}$$

Likewise, eMBB is also affected by the transmissions from the N_u URLLC users in the same band. Given that $\Omega = \bar{p_e}/\bar{p_u}$ as described in Section II, the average URLLC interferer power relative to the desired eMBB source is the inverse of Ω. Hence, we set $\Gamma_v = 1/\Omega$ and $\bar{\gamma} = \bar{\gamma}_e = \bar{\gamma}_u \Omega$. At the same time, with other eMBB streams present with the same average interferer power, we have that $w = N_e - 1$ and $\Gamma_w = 1$. Then, the outage probability of the eMBB transmissions is given by

$$P_{f,e} = \sum_{z=0}^{N_u} P_c(z, N_u) P_f(\bar{\gamma}_u \Omega, z, N_e - 1, 1/\Omega, 1). \tag{7}$$

C. MMSE with SIC receiver and overlaying transmissions

With SIC we assume that URLLC traffic has to be decoded first, due to its strict latency. Then its interference contribution is removed from the receive signal. This means that only eMBB actually benefits from SIC. Given that, the outage probability of URLLC transmissions in this case can be also expressed by (6).

Assuming that $\epsilon_u << \epsilon_e$, the interference from failing URLLC transmissions, which cannot be canceled by SIC, is negligible. With eMBB not suffering from URLLC interference, the outage probability of the eMBB transmissions can be calculated with (5).

D. Achievable rate and load calculation

Using the described outage probability for each case, we can calculate numerically the minimum value for the SINR ψ to meet a given requirement. Here, we find ψ that satisfy $P_{f,u} = \epsilon_u$ for the URLLC cases, and $P_{f,e} = \epsilon_e$ for the eMBB cases. For a certain rate r in bps/Hz, the outage probability is expressed as $Prob[\log_2(1 + \psi) < r]$. From this relation we can obtain the maximum rate corresponding to the outage probability requirement as

$$r[\text{bps/Hz}] = \log_2(1 + \psi). \tag{8}$$

The achievable eMBB load, which corresponds to the maximum throughput with a given ϵ_e, is calculated as

$$L_e[\text{bps}] = r W_e N_e (1 - \epsilon_e). \tag{9}$$

For URLLC transmission of a packet of size D in a bandwidth W_u and in a TTI of duration T, the transmission rate is given by

$$r_u[\text{bps/Hz}] = D/T/W_u. \tag{10}$$

With the correspondent SINR for this rate, i.e. $2^{r_u} - 1$, we calculate numerically the maximum arrival rate $\hat{\lambda}$ that is allowed for a given number of URLLC users meeting the outage probability requirement. Then, the achievable URLLC load can be calculated as

$$L_u[\text{bps}] = D \hat{\lambda} N_u / T. \tag{11}$$

Given that ϵ_u is very low, the impact of transmission failures in the resultant load is considered negligible.

IV. Numerical analysis

In this section we first present the achievable rate for URLLC transmissions overlaying a eMBB stream. We then find the achievable load for both kind of services, considering NR assumptions. Finally, a comparison between the allocation approaches is provided for different operation regimes.

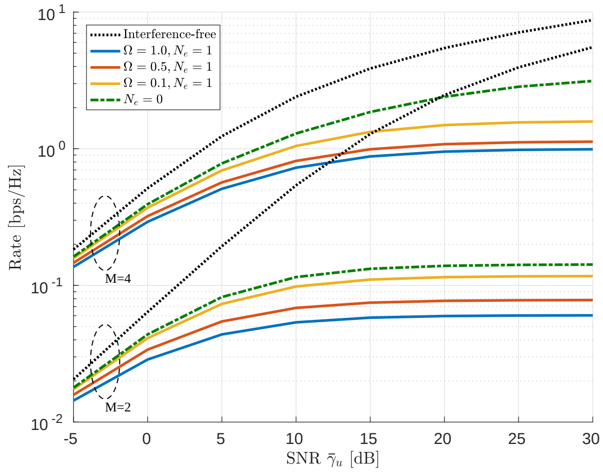


Fig. 2. Achievable rates for URLLC overlaying one eMBB stream with different Ω, considering $N_u = 50$, $\lambda = 10^{-2}$, and MMSE with 2 and 4 antennas. For the interference-free curve it is assumed dedicated resources.

A. Achievable rates for URLLC

For eMBB we consider $\epsilon_e = 10^{-1}$, whereas $\epsilon_u = 10^{-3}$ for URLLC. These values are usual block error rate targets for the initial transmission of these services, considering that a higher reliability is more efficiently achieved after retransmission [2]. We consider the case of MMSE with $M = 2$ and $M = 4$ receive antennas. A URLLC load is imposed with $N_u = 50$ users and packet arrival rate $\lambda = 10^{-2}$ per TTI for each user. Different relative receive power of eMBB with respect to the URLLC signals are assumed with $\Omega = \{1, 0.5, 0.1, 0\}$. Setting $\Omega = 0$ is equivalent to no eMBB, i.e. $N_e = 0$.

The achievable rate for URLLC depending on the SNR $\bar{\gamma}_u$ is shown in Fig. 2. The interference-free curve denotes a benchmark assuming dedicated resources for each user. It is observed that the rate practically saturates after $\bar{\gamma}_u = 10\,\text{dB}$ for $M = 2$, i.e. a higher SNR does not yield on higher URLLC capacity. This is due to the eMBB interference and collisions with the imposed URLLC load. The achievable rate obviously increases with lower values of Ω, since the SINR of URLLC increases. This means that, for guaranteeing high URLLC capacity, the power of URLLC signals should be higher than the ones of eMBB in the overlaying band. It is evident that $M = 4$ allows the highest rates due to the better interference rejection capability of the receiver. At $\bar{\gamma}_u = 10\,\text{dB}$ and $\Omega = 1$, it allows a rate just 3.3 times lower than the interference-free benchmark, compared to the 10 times lower with $M = 2$. The higher number of receive antennas allows higher URLLC rates and gives possible room for multiple eMBB streams.

B. Achievable loads

Now we compare the resource allocation options for multiplexing URLLC and eMBB traffic, considering particular NR assumptions [4]. For that, we calculate the achievable load for each service according to the receiver type, average SNR, average interferer power relative to source, and allocated band. We consider a bandwidth $W = 10\,\text{MHz}$. For separate bands, we assume $R = \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$, corresponding to full band for eMBB until full band for URLLC. For overlaying transmissions, we assume $\Omega = \{0, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$, which corresponds to no eMBB until eMBB with same average receive power as URLLC. Given the higher priority of URLLC, we do not consider the option of eMBB with higher average receive power than URLLC.

URLLC users transmit payloads of $D = 256\,\text{bits}$ using a short-TTI of $0.143\,\text{ms}$. This may represent the case of a NR mini-slot numerology with 4 symbols per TTI and $30\,\text{kHz}$ subcarrier spacing. The eMBB users transmit large volume of data exploiting capacity-achieving codes. In the following examples we assume $M = 4$ and $N_e = 2$, i.e. two eMBB streams are simultaneously active in the same band, as in MU-MIMO.

Four operation modes are considered:

- Separate bands and equal SNR: the average SNR is $\bar{\gamma}_u = \bar{\gamma}_e = \bar{\bar{\gamma}}$ for URLLC and eMBB, where $\bar{\bar{\gamma}}$ is the average SNR over the bandwidth W. It refers to a system in which users keep the same power spectral density (PSD) regardless of the operational bandwidth.
- Separate bands and scaled SNR: $\bar{\gamma}_u = \bar{\bar{\gamma}}/R$ for URLLC and $\bar{\gamma}_e = \bar{\bar{\gamma}}/(1-R)$ for eMBB, i.e. the average SNR is increased as much as the associated bandwidth decreases. It refers to a system where users maintain the same output power regardless of the operational bandwidth.
- Overlay with SIC: overlaying transmissions considering MMSE with ideal SIC and different values of Ω.
- Overlay without SIC: overlaying transmissions with MMSE receiver and different values of Ω.

Fig. 3 and Fig. 4 show the achievable loads for eMBB and URLLC in a low SNR scenario ($\bar{\gamma}_u = 0\,\text{dB}$ in full band) and medium SNR scenario ($\bar{\gamma}_u = 10\,\text{dB}$ in full band), respectively. Each line delimits the maximum load that can be achieved depending on R or Ω, while meeting the requirements given by ϵ_e and ϵ_u. The region to the left of the line represents lower load combinations that can be supported. The maximum supported URLLC load is denoted by $\hat{L}_u$. At 20% of $\hat{L}_u$ is indicated a low URLLC load regime, and at 80% of $\hat{L}_u$ is indicated a high URLLC load regime. The maximum gain of overlaying allocation relative to using separate bands in terms of eMBB throughout is denoted by $G_{o,e}$.

In the low SNR scenario as it is shown in Fig. 3, we observe that the separate bands and equal SNR operation (dashed red line) shows the lowest achievable loads. For example with $R = 0.5$, only up to $1\,\text{Mbps}$ can be reliably supported for URLLC, and up to $11\,\text{Mbps}$ for eMBB. This performance can happen when same power control settings are used for both services. On the other hand, for separate bands and service SNR scaling with R (solid red line), the performance is generally better. For overlay without SIC (dashed blue line), a lower achievable load is experienced for both services compared to the use of separate bands as in the previous case. For example with $\Omega = 0.8$ and $2\,\text{Mbps}$ URLLC load, up to $14\,\text{Mbps}$ can be reliably supported for eMBB, while $17\,\text{Mbps}$ can be reached if traffic

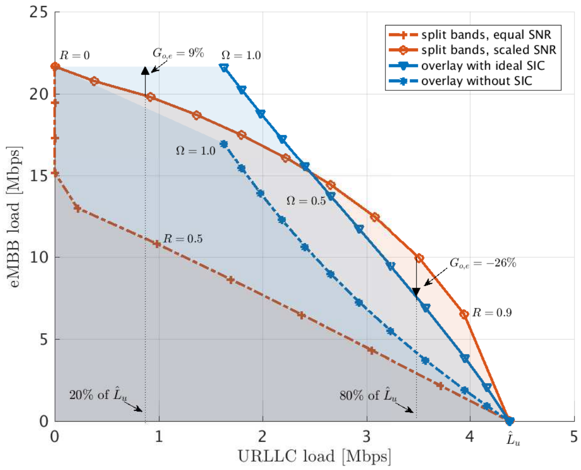


Fig. 3. Achievable loads for URLLC and eMBB considering different receive strategies and low average SNR $\bar{\gamma}_u = 0\,\mathrm{dB}$. $W = 10\,\mathrm{MHz}$, $D = 256\,\mathrm{bits}$, $N_u = 50$, $N_e = 2$ and $M = 4$.

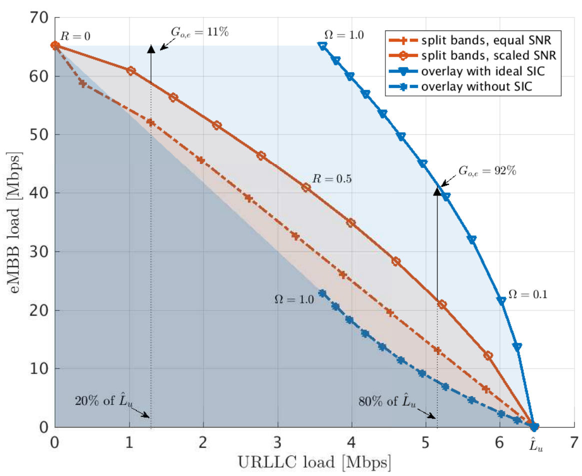


Fig. 4. Achievable loads for URLLC and eMBB considering different receive strategies and moderate average SNR $\bar{\gamma}_u = 10\,\mathrm{dB}$. $W = 10\,\mathrm{MHz}$, $D = 256\,\mathrm{bits}$, $N_u = 50$, $N_e = 2$ and $M = 4$.

is conveyed in separate bands. While for overlay with SIC (solid blue line), there is an advantage of overlaying when the URLLC load is lower than $2.4\,\mathrm{Mbps}$, due to the reduced interference in this condition. Anyway, it can be noted that overlaying is generally not a good option in low SNR cases.

For the medium SNR scenario in Fig. 4, there is a clear advantage of overlaying when MMSE with SIC is used. Without noise limiting and canceled URLLC interference, the antenna combining can strength the eMBB signal boosting its throughput. However, without SIC the achievable load for both services is higher if separate bands are allocated. This avoids that the mutual interference between the traffic penalizes the performance of each other. Given that the URLLC rate saturates, the result for a high SNR scenario is omitted here, though the same observations as for medium SNR are valid.

C. Comparison for different regimes

Fig. 5 shows the gain $G_{o,e}$ of overlaying relative to separate bands allocation in terms of eMBB throughput, for low and high URLLC load regimes. Two packet sizes, $D = 256\,\mathrm{bits}$ and $D = 1600\,\mathrm{bits}$, are assumed for URLLC. Besides, we also assume two values for the outage probability targeted for URLLC. $\epsilon_u = 10^{-3}$ refers to a system in which a higher reliability can be achieved after a retransmission, and $\epsilon_u = 10^{-5}$ refers to a system where the reliability target should be achieved with a single shot transmission. The absolute values of the maximum supported URLLC load $\hat{L}_u$ for each case are shown on the top of the plots.

In many cases marked with ”x”, we note that no URLLC load can be supported. This is observed in most cases for $M = 2$ in low SNR scenarios, independent of the allocation scheme. As can be seen in Fig. 5a and Fig. 5c, for small packet size there is a significant gain of overlaying at high SNR, specially for 4 receive antennas and high URLLC load regime (up to $+260\%$). In case of large packets as shown in Fig. 5b and Fig. 5d, overlaying allocation may lead to losses, while minor gains appears only in case of $M = 4$ antennas and $N_e = 1$ eMBB stream, at high SNR. For stricter reliability such as 10^{-5}, the gain of overlaying is reduced, and losses get more evident with the $1600\,\mathrm{bits}$ packets.

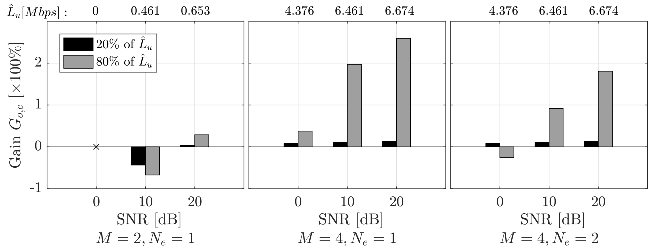
V. Discussion

In many cases the allocation of separate bands for each service class shows to be more efficient, specially when SIC is not employed. In practice, it implies that the bandwidth needs to be reconfigured for all grant-free users whenever the target supported load changes. This results in additional control signaling overhead. To avoid this issue, for instance in a scenario where the URLLC load varies very often, it would be recommended to proactively allocate a larger share of the bandwidth for URLLC to cope with the load variation, to the detriment of the eMBB capacity.

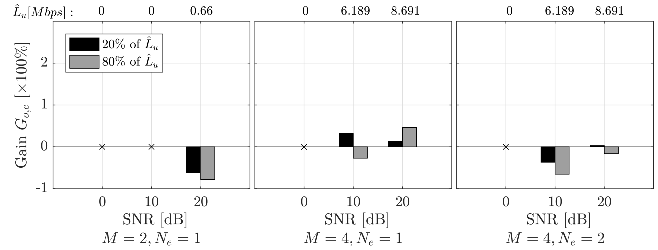
For scenarios with low average SNR, e.g. macro deployments, the gains of overlaying transmission using SIC are insignificant compared to operating with a simple MMSE receiver. Besides, even when SIC is available, the crossing regions indicate that it is beneficial to switch between separate bands and overlaying mode depending on the load aimed for each service. On the other hand, in a dense deployment with medium/high SNR, the application of a more complex receiver with SIC is more relevant, given the higher achievable loads.

It is important to note also that, for a network with users that have multiple traffic types, as for eMBB and URLLC services, it is beneficial to use different transmission parameters for each kind of service. This means, for example, that one user should be configured with a power control setting for eMBB and another for URLLC.

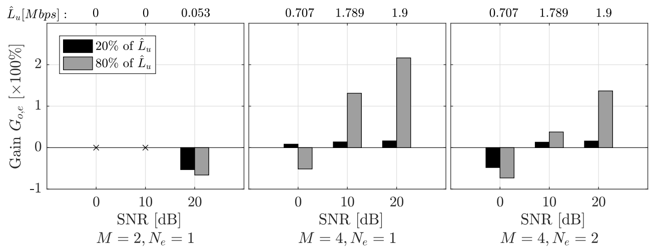
The proposed approach presented in this paper can be also relevant for feasibility analysis and decision making. For example, by assigning costs to each traffic, one can find the optimal load balance policy that results in the highest profit,



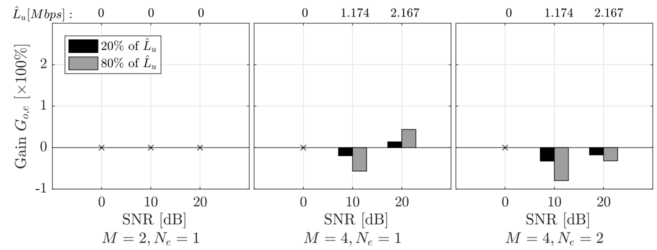
(a) $D = 256\,\text{bits}, \epsilon_u = 10^{-3}, T = 0.143\,\text{ms}$.



(b) $D = 1600\,\text{bits}, \epsilon_u = 10^{-3}, T = 0.143\,\text{ms}$.



(c) $D = 256\,\text{bits}, \epsilon_u = 10^{-5}, T = 0.250\,\text{ms}$.



(d) $D = 1600\,\text{bits}, \epsilon_u = 10^{-5}, T = 0.250\,\text{ms}$.

Fig. 5. Gain of overlaying relative to separate bands allocation in terms of eMBB throughput for different settings.

and select the corresponding bandwidth shares or the power control settings for that.

VI. Conclusion

In this work we studied how to efficiently multiplex grant-free URLLC and eMBB services in the uplink. Two possible options of multiplexing are considered, namely, separate bands and overlaying transmissions. We describe the outage probability for each service and for each multiplexing option considering MMSE receiver and MMSE with SIC. With this approach we can compare the achievable load that can be supported for each traffic. The resource allocation considers different shares of the bandwidth for each traffic in separate bands, or different relative receive power when the transmissions are overlaying. Numerical analyses considering NR assumptions are carried out. The results show that overlaying provides better performance generally using MMSE with SIC either in high SNR or for low URLLC loads. Separate bands for each service class is better when a SIC processing is not employed, the URLLC packet size is large and higher reliability levels are required for URLLC. Future work should consider traffic bursts and the effect of power limitation for overlaying transmissions.

VII. Acknowledgments

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

[1] ITU-R, "Report ITU-R M.2410-0 - Minimum requirements related to technical performance for IMT-2020 radio interface(s)," International Telecommunication Union (ITU), Tech. Rep., Nov. 2017.

[2] G. Pocovi, H. Shariatmadari, G. Berardinelli, K. Pedersen, J. Steiner, and Z. Li, "Achieving Ultra-Reliable Low-Latency Communications: Challenges and Envisioned System Enhancements," *IEEE Network*, vol. 32, no. 2, pp. 8–15, Mar. 2018.

[3] P. Popovski, J. J. Nielsen, C. Stefanovic, E. d. Carvalho, E. Strom, K. F. Trillingsgaard, A. S. Bana, D. M. Kim, R. Kotaba, J. Park, and R. B. Sørensen, "Wireless Access for Ultra-Reliable Low-Latency Communication: Principles and Building Blocks," *IEEE Network*, vol. 32, no. 2, pp. 16–23, Mar. 2018.

[4] 3GPP TR 38.802 v14.0.0, "Study on New Radio Access Technology," Mar. 2017.

[5] R1-1803659, "UL multiplexing between URLLC and eMBB," Apr. 2018.

[6] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[7] H. Gao, P. J. Smith, and M. V. Clark, "Theoretical reliability of MMSE linear diversity combining in Rayleigh-fading additive interference channels," *IEEE Transactions on Communications*, vol. 46, no. 5, pp. 666–672, May 1998.

[8] G. Berardinelli and H. Viswanathan, "Overlay transmission of sporadic random access and broadband traffic for 5G networks," in *2017 International Symposium on Wireless Communication Systems (ISWCS)*, Aug. 2017, pp. 19–24.

[9] C.-P. Li, J. Jiang, W. Chen, T. Ji, and J. Smee, "5G ultra-reliable and low-latency systems design," in *2017 European Conference on Networks and Communications (EuCNC)*, Jun. 2017, pp. 1–5.

[10] A. Anand, G. de Veciana, and S. Shakkottai, "Joint Scheduling of URLLC and eMBB Traffic in 5G Wireless Networks," *ArXiv e-prints*, Dec. 2017. [Online]. Available: http://arxiv.org/abs/1712.05344

[11] P. Popovski, K. F. Trillingsgaard, and G. D. Osvaldo Simeone, "5G Wireless Network Slicing for eMBB, URLLC, and mMTC: A Communication-Theoretic View," *CoRR*, vol. abs/1804.05057, 2018. [Online]. Available: http://arxiv.org/abs/1804.05057

[12] G. Berardinelli, R. Abreu, T. Jacobsen, N. H. Mahmood, K. Pedersen, I. Z. Kovács, and P. Mogensen, "On the Achievable Rates over Collision-Prone Radio Resources with Linear Receivers," in *2018 IEEE 29th PIMRC*, Sep. 2018.

[13] K. I. Pedersen, G. Berardinelli, F. Frederiksen, P. Mogensen, and A. Szufarska, "A Flexible 5G Frame Structure Design for Frequency-Division Duplex Cases," *IEEE Communications Magazine*, vol. 54, no. 3, pp. 53–59, Mar. 2016.

[14] P. Wu and N. Jindal, "Coding versus ARQ in Fading Channels: How Reliable Should the PHY Be?" *IEEE Transactions on Communications*, vol. 59, no. 12, pp. 3363–3374, Dec 2011.