Physical Layer Network Slicing for eMBB and mMTC with Distributed Power Allocation

Luiz Eduardo Hupalo, Richard Demo Souza and João Luiz Rebelatto

Abstract— **This work addresses the network slicing of radio resources between enhanced mobile broadband (eMBB) and massive machine-type communication (mMTC) services, considering both heterogeneous orthogonal multiple access (H-OMA) and heterogeneous non-orthogonal multiple access (H-NOMA). For increasing the number of supported active mMTC devices, we adapt the transmission power of these devices in a decentralized fashion. The results indicate that the number of supported mMTC devices can be increased by up to 40% under H-NOMA, and that such gain achieves up to 60% in an orthogonal scenario.**

Keywords— **Beyond 5G, eMBB, mMTC, network slicing.**

I. Introduction

Each generation of mobile communications has progressed to provide new services for specific use cases and applications. From the 1G analog communications to the worldwide success of 4G, the primary focus has been connecting people [1]. The introduction of 5G brings a novel and disruptive concept of connecting both people and machines, by means of three main uses cases: massive Machine-Type Communications (mMTC), enhanced Mobile Broadband Communications (eMBB), and Ultra-Reliable Low-Latency Communications (URLLC) [2]. In 5G, the typical one-size-fits-all network solution is not viable. An alternative is network slicing, an approach that divides a network, at the physical and/or upper layers, into several virtual layers based on the specific service being provided. This enables different customizations to meet varying – and sometimes conflicting – service requirements [3].

Recent works consider the use of network slicing in 5G and beyond (B5G) networks. An overview of intra and interslice resource allocation methods inside a network slicing context is provided in [4]. Some approaches discuss the use of machine learning techniques in slices resource management [5] and to dynamically allocate, schedule and orchestrate resources [6]. Specifically in the physical layer, [7] provides the theoretical background regarding the slicing between the three aforementioned 5G services and introduces the concept of heterogeneous orthogonal multiple access (H-OMA) and heterogeneous non-orthogonal multiple access (H-NOMA), where the term *heterogeneous* comes from the different classes of services. Later, [8] elaborated on the model from [7] by considering the slicing between eMBB and URLLC services using the max-matching diversity (MMD) algorithm [9] to allocate channels to the eMBB devices. The use of the MMD algorithm introduces frequency diversity, simultaneously increasing the eMBB achievable rate and the URLLC reliability. In the context of eMBB and mMTC slicing, [10] considers the use of multiple receiving antennas in the uplink. The results show that the space diversity provided by the multiple antennas are more beneficial to H-NOMA than to H-OMA, and that such difference increases with the number of antennas.

This work focuses on the slicing between eMBB and mMTC. We propose a method to modify the power allocation of the mMTC devices in a distributed and uncoordinated fashion, aiming at spreading the set of power levels at the base station (BS), which improves the performance of the successive interference cancellation (SIC) technique. The proposed method is shown to achieve gains of up to 40% in terms of supported mMTC active devices for H-NOMA and up to 60% in the H-OMA slicing for slow eMBB traffic.

II. System Model

We consider the uplink of a single-cell where eMBB and mMTC devices transmit to a common BS through the same channel f. A single eMBB device access the channel, while the mMTC traffic is composed of a random number A_M of active devices, which follows a Poisson distribution with mean (arrival rate) equal to λ_M. We assume block Rayleigh fading, where the channel coefficients are constant during one time slot (TS), independently changing between time-slots. The eMBB channel coefficient is denoted by $H_B \in \mathcal{CN} \sim (0,\, \Gamma_B)$, where Γ_B is the average channel gain, which encompasses both path loss and transmission power. Likewise, the channel coefficient of the m-th mMTC device is denoted by $H_m \in \mathcal{CN} \sim (0,\, \Gamma_M)$ for $m \in \{1, 2, \ldots, A_M\}$.

Regarding multiple access, we adopt the H-OMA and H-NOMA strategies from [7]. While H-OMA considers that the slicing between mMTC and eMBB is orthogonal in time, in H-NOMA all devices transmit simultaneously, as illustrated in Fig. 1. It is important to mention that the mMTC devices always transmit in a non-orthogonal fashion; in H-OMA, the time orthogonality is between the services.

III. Performance Metrics without Slicing

Next, we define relevant performance metrics, considering that the resources are exclusively allocated to a single service.

A. eMBB

Following [7], the eMBB transmission relies on the assumptions that *i)* each transmission occurs at an exclusively allocated radio resource to the eMBB user; *ii)* the eMBB user

L. E. Hupalo e R. D. Souza are with the Federal University of Santa Catarina, Florianopolis-SC, Brazil. {l.hupalo@posgrad.ufsc.br, richard.demo@ufsc.br}. J. L. Rebelatto is with the Federal University of Technology - Parana, Curitiba-PR, Brazil. {jlrebelatto@utfpr.edu.br}.

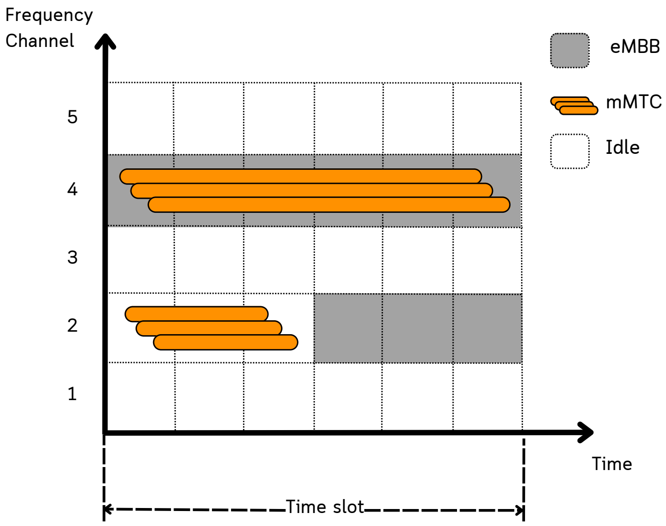

Fig. 1: Channel $f = 2$ illustrates H-OMA, being orthogonal in time. Channel $f = 4$, depicts H-NOMA, where both services transmit simultaneously during the entire time-slot.

is aware of the channel state information (CSI) G_B, which is used to properly adapt its transmission power $P_B(G_B)$. The goal is to maximize the eMBB rate r_B, which depends on the outage probability requirement of the service ϵ_B. This leads to the following optimization problem [7]:

$$\begin{aligned} \max \quad & r_B \\ \text{s.t.} \quad & \mathbb{P}\left[\log_2\left(1 + G_B P_B(G_B) \leqslant r_B\right)\right] \leqslant \epsilon_B \\ & \mathbb{E}\left[P_B(G_B)\right] = 1 \end{aligned} \tag{1}$$

The solution to (1) is the truncated power inversion. The transmission of the eMBB device then depends on the threshold signal-to-noise ratio (SNR) [7]

$$G_B^{\min} = \Gamma_B \ \ln\left(\frac{1}{1-\epsilon_B}\right), \tag{2}$$

which guarantees minimum conditions necessary for transmission and depends only on the eMBB reliability ϵ_B. Then, the eMBB user transmits only in case the SNR is above $G_B^{\min}$.

After the power inversion, the target SNR that is used to determine the required transmission power is [7]

$$G_B^{\text{tar}} = \frac{\Gamma_B}{\gamma\left(0, \frac{G_B^{\min}}{\Gamma_B}\right)}, \tag{3}$$

where $\gamma(\cdot, \cdot)$ is the lower incomplete gamma function. This implies that the eMBB achievable rate (in bps/Hz) is finally [7]

$$r_B^{\text{orth}} = \log_2\left(1 + G_B^{\text{tar}}\right). \tag{4}$$

B. mMTC

The mMTC devices transmit in the same channel while their traffic is sporadic, thus random and unknown. Assuming a fixed transmission rate r_M and a reliability constraint ϵ_M, our goal is to maximize the supported mMTC arrival rate λ_M. We also assume the use of a SIC decoder at the BS, so that more than one mMTC device may be served at the same time.

In this scenario, each mMTC device transmission yields a different instantaneous SNR at the BS, being possible to order the devices as $G_{[1]} \geq G_{[2]} \geq \ldots \geq G_{[A_M]}$. In the absence of interference, the decoding of mMTC m_0 depends only on its instantaneous channel gain and on those of the remaining mMTC devices. To be correctly decoded, the following inequality must be satisfied:

$$\log_2\left(1 + \sigma_{[m_0]}^{\text{orth}}\right) \geq r_M, \tag{5}$$

with $\sigma_{[m_0]}^{\text{orth}}$ being the SNR[1] of the m_0–th device, given by

$$\sigma_{[m_0]}^{\text{orth}} = \frac{G_{[m_0]}}{1 + \sum_{m=m_0+1}^{A_M} G_{[m]}}. \tag{6}$$

In (6), the signals from all the mMTC devices yet to be decoded are treated as noise. Then, if the device is correctly decoded, its component is removed from the received signal and the procedure follows until all devices are decoded or an outage is declared. Let D_M be the number of mMTC devices in outage, then the error rate of the mMTC devices is [7]

$$\mathbb{P}(E_M) = \frac{\mathbb{E}[D_M]}{\lambda_M}, \tag{7}$$

where $\mathbb{E}[D_M]$ is the average number of users in outage. Finally, the maximum number of users that can be simultaneously supported under the reliability constraint $\mathbb{P}(E_M) = \epsilon_M$ can be numerically obtained from

$$\lambda_M^{\text{orth}}(r_M) = \max\left\{\lambda_M : \mathbb{P}(E_M) \leq \epsilon_M\right\}. \tag{8}$$

IV. Physical Layer Network Slicing

A. Orthogonal Slicing between eMBB and mMTC

In the H-OMA method, there is no interference between the two types of services using the same radio resource – *i.e.* when the eMBB device is active, the mMTC device is not transmitting. This can be modeled by time-sharing: let α be the fraction of time slot in which the resources are allocated to the eMBB device and $1-\alpha$ the fraction of time allocated to the mMTC devices, with $\alpha \in [0, 1]$. Then, the main objective of the coexistence between the two services is to maximize the pair $(r_B,\ \lambda_M)$, composed of the eMBB rate r_B given by $r_B = \alpha r_B^{\text{orth}}$ and the mMTC arrival rate given by [7]

$$\lambda_M = \lambda_M^{\text{orth}}\left(\frac{r_M}{1-\alpha}\right), \tag{9}$$

while meeting a rate r_M and error ϵ_M requirements.

B. Non-Orthogonal Slicing between eMBB and mMTC

In H-NOMA, both the eMBB and the mMTC users simultaneously transmit in the same radio resource, then the BS must decide whether to attempt to decode the eMBB or the mMTC devices first. In [7], it is assumed that at each decoding step, the BS first tries to decode the eMBB device – if it has not yet been decoded – or the next active mMTC device, in

[1]Strictly speaking, this is a signal to interference plus noise ratio (SINR). However, since the interference is treated as noise, and for the sake of simplicity, along this paper we simply refer to it as SNR.

a decreasing order of their instantaneous channel gains. This approach is guided by the idea that an mMTC device can have high channel gains, causing interference in the eMBB traffic and making it difficult to decode the latter.

For the non-orthogonal case, the eMBB rate is similar to the orthogonal case (4). However, now the target SNR does not depend only on the eMBB reliability constraint ϵ_B. In fact, in order to reduce the interference it can be appropriate to transmit at target SNR $G_B^{\text{tar}\prime}$ lower than that in (3).

If the eMBB is inactive due to insufficient SNR, then the mMTC devices do not suffer from interference and are decoded in the order of their decreasing channel gains. But if there is an active eMBB device, the BS starts by evaluating the SNR from the m_{0-th} device using [7]

$$\sigma_{[m_0]}^{\text{non-orth}} = \frac{G_{[m_0]}}{1 + G_B^{\text{tar}\prime} + \sum_{m=m_0+1}^{A_M} G_{[m]}}. \qquad (10)$$

If the m_{0-th} mMTC device is correctly decoded, it is subtracted from the received signal and the procedure continues until an outage occurs or the last device is decoded. Then, the BS evaluates the eMBB instantaneous SNR as [7]

$$\sigma_B^{\text{non-orth}} = \frac{G_B^{\text{tar}\prime}}{1 + \sum_{m=m_0}^{A_M} G_{[m]}} \qquad (11)$$

and tries to decode it; if successful, the procedure runs as in the orthogonal slicing case between eMBB and mMTC – as now there is no more eMBB traffic on this time slot.

In order to maximize the achievable pair (r_B, λ_M), let $D_M \in \{1, \ldots, A_M\}$ and $D_B \in \{0, 1\}$ be the random variables denoting the number of correctly decoded mMTC and eMBB devices, respectively. The solution can be obtained from [7]

$$\begin{aligned} \lambda_M^{\text{non-orth}}(r_B) = \max \quad & \{\lambda_M \geq : \exists G_B^{\text{tar}} \text{ and } G_B^{\min}\} \\ \text{s.t.} \quad & \frac{\mathbb{E}\left[D_M\right]}{\lambda_M} \geq 1 - \epsilon_M, \\ & \mathbb{E}\left[D_B\right] \geq 1 - \epsilon_B. \end{aligned} \qquad (12)$$

V. Proposed mMTC Power Allocation Method

Recall that the set of mMTC devices transmit at the same channel even in H-OMA, and that their instantaneous received powers vary around Γ_M due to fading. Thus, the overall received signal is a sum of several independent components, which may allow the decoding of the strongest one and its cancellation from the received signal, thus latter enabling the decoding of the second strongest and so on. It is such difference in instantaneous received powers that enables SIC.

Our proposal is to generate more than one average SNR in the BS for the mMTC devices; this means that the instantaneous SNRs of part of these devices would fluctuate around a given average, while the SNRs of other devices would fluctuate around different values. This can be done without the need of coordination among the devices and the BS. For that sake, we propose that the devices randomly vary the average received SNR at the BS among a set of predefined values. In order to yield this average SNR at the BS, the devices need to set their transmission power accordingly, so that the knowledge of long-term path-loss is required. Let us say that, in the absence of the proposed method, the average SNR to be generated at the BS by the mMTC devices is Γ_M. Then, the new set of predefined SNR values is obtained by the multiplication of the previous average by the elements of a power allocation vector β. Thus, prior to transmission, the devices randomly choose one of the elements in β, lets say $\beta(i)$, and then adjust their transmission power to yield an average SNR of $\beta(i)\,\Gamma_M$ at the BS. The elements of β are constrained to have unitary mean so that the long-term average power consumption remains unchanged.

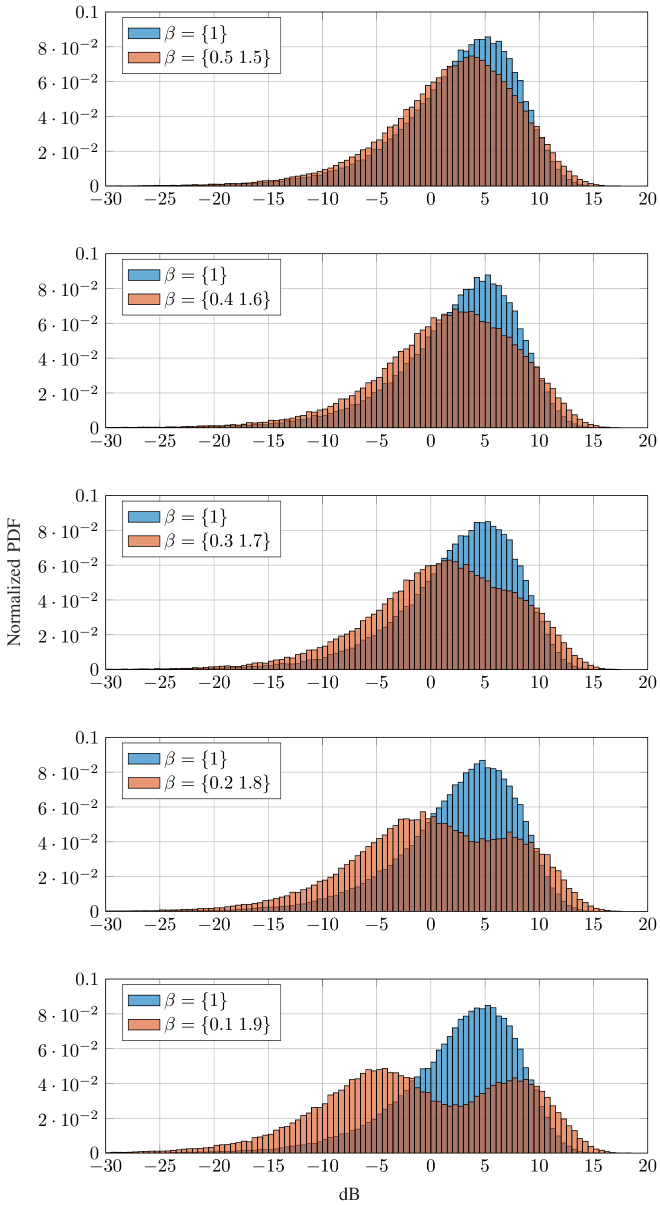

Fig. 2: Probability density function of mMTC devices SNRs at BS with and without the proposed method. When $\boldsymbol{\beta} = \{1.9\ 0.1\}$, one can clearly see the existence of "clusters of instantaneous SNR".

In order to illustrate the influence of β in the probability density function (PDF) of the SNR, let us consider the following set of power allocation vectors: $\boldsymbol{\beta} = \{1.5\ 0.5\}, \{1.6\ 0.4\}, \{1.7\ 0.3\}, \{1.8\ 0.2\}, \{1.9\ 0.1\}$. Fig. 2 shows the pdf of the mMTC devices instantaneous SNRs, considering those β vectors. One can see that the pdf considerably changes with β and that it becomes organized in clusters that are more visible when the elements in β are more distant.

The creation of the so-called clusters enables the occurrence of more events of successful SIC due to the larger difference in the instantaneous SNRs of the mMTC devices. Other power allocation vectors could be used to create more than two average SNRs at the BS. However, the constraint of having unitary mean into the elements of β makes it difficult to obtain well-separated clusters without the average SNR of one of them being very low, leading to good performance only in case the target rate is also very low. Thus, in this work we consider the cardinality of β to be two.

Finally, the power allocation vectors impact on the number of supported active mMTC devices as follows. For H-OMA, the SNR for the mMTC devices becomes

$$\sigma_{[m_0]}^{\mathrm{orth},\beta} = \frac{G_{[m_0]}^{\beta}}{1+\sum_{m=m_0+1}^{A_M} G_{[m]}^{\beta}}. \tag{13}$$

where $G_{[m]}^{\beta}$ is the SNR of the m-th mMTC device after the proposed power allocation. Similarly, for the H-NOMA case, the SNR for the mMTC devices becomes

$$\sigma_{[m_0]}^{\mathrm{non-orth},\,\beta} = \frac{G_{[m_0]}^{\beta}}{1+G_B^{tar\prime}+\sum_{m=m_0+1}^{A_M} G_{[m]}^{\beta}}. \tag{14}$$

VI. Numerical Results

Unless stated otherwise, the simulation results presented in this section consider the power allocation vectors utilized in Fig. 2, with the fixed simulations parameters being $\Gamma_M = 5$ dB, $\epsilon_M = 10^{-1}$, $\epsilon_B = 10^{-3}$ and $r_M = 0.04$. The benchmark is the model proposed in [7], which we refer to as standard H-OMA and standard H-NOMA. The results are obtained by solving (1), (8), and (12) through Monte Carlo simulations.

A. *Results for H-OMA Slicing Between mMTC and eMBB*

Fig. 3 shows the achieved pairs (r_B, λ_M) for each power allocation vector β. When α tends to zero (*i.e.*, r_B tends to zero and the mMTC devices occupy more of the channel) there is an increase in the SIC performance, allowing a larger number of active devices when compared to the standard H-OMA. As the eMBB device increases its channel usage time (α moving closer to 1), the number of mMTC active devices ends up decreasing until reaching that of standard H-OMA. Moreover, the proposed method supports up to 60% more active mMTC devices than standard H-OMA, in the case of low channel usage by the eMBB device.

B. *Results for H-NOMA Slicing Between mMTC and eMBB*

In Fig. 4 we consider the H-NOMA case, whose results can be divided in two regions: *i)* The region where the eMBB traffic interference is small (on the left side) and *ii)* The region where the interference from eMBB traffic compromises the proposed method gains (on the right). In the first region, when r_B is low, the behavior is very similar to the H-OMA case, and the eMBB interference can be removed by

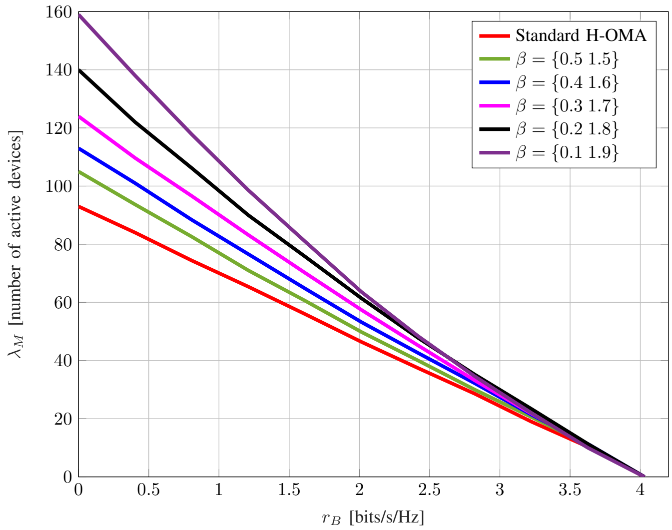

Fig. 3: Arrival rate $\lambda_M^{\mathrm{orth}}$ for H-OMA case, as a function of eMBB rate r_B, with $\Gamma_B = 20$ dB.

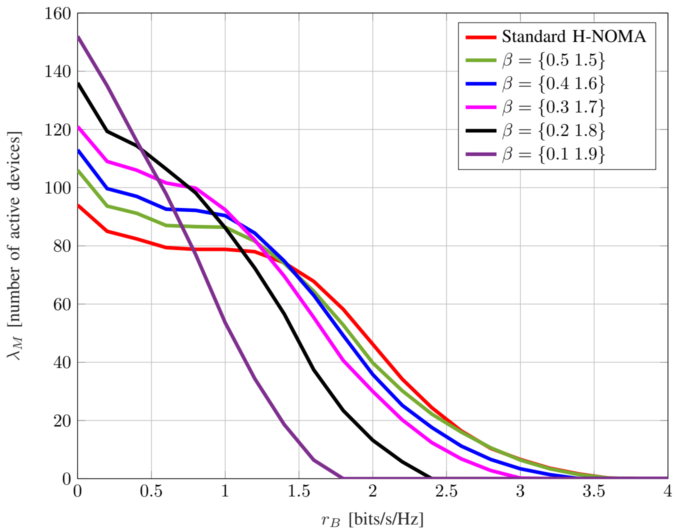

Fig. 4: Arrival rate $\lambda_M^{\mathrm{non-orth}}$ vs r_B for H-NOMA for different power allocation vectors, with $\Gamma_B = 20$ dB.

SIC and the mMTC performance remains high when using the proposed power allocation. As r_B increases, the eMBB device can only be decoded after the mMTC devices with better channels are decoded and canceled by SIC. In the standard case, in which the proposed power allocation method is not used, after decoding the eMBB device, the process fails due to interference from mMTC devices that had not yet been decoded. When using the proposed method, after eMBB decoding, the remaining mMTC devices have channels with different enough instantaneous SNRs for SIC to be able to decode and cancel the best ones for a few more steps, explaining the better performance in this initial r_B range.

In the second region, as r_B increases, the mutual interference between mMTC and eMBB devices imposes a poor performance to both services. Regarding the proposed scheme,

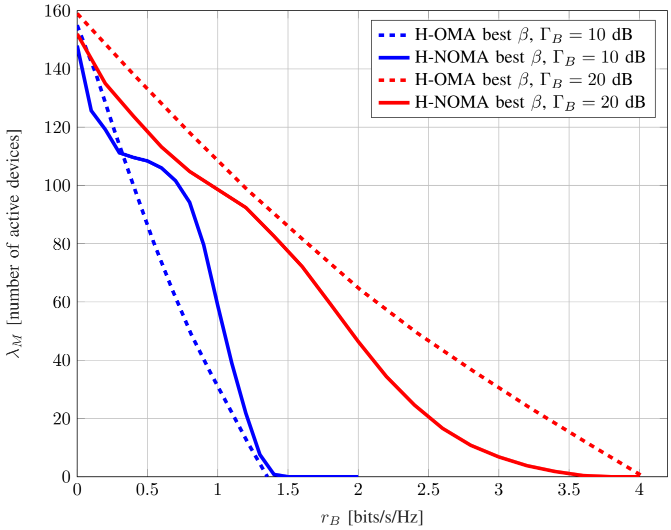

Fig. 5: Arrival rate $\lambda_M^{\text{non-orth}}$ for both H-OMA and H-NOMA case, with the envelope containing the β values that allows the best performance, as a function of eMBB rate r_B.

the more unbalanced the gains of the mMTC devices, the sooner the degradation of the two services occurs. Fig. 4 shows that there is a trade-off between the channel usage by the eMBB service and the imbalance SNR level of the mMTC devices. In a high imbalanced scenario – for example, $\beta = \{0.1\ 1.9\}$ – the minimum channel usage by the eMBB service causes the two services to degrade by r_B around 1.5 bits/s/Hz. On the other hand, one can forego having a high amount of λ_M for low values of r_B and thus extend the point where service degradation occurs, just using $\beta = \{0.5\ 1.5\}$.

C. Comparison between H-OMA and H-NOMA

To compare the performance between H-OMA and H-NOMA using the proposed power allocation method, two average SNR scenarios are simulated for the eMBB service, with the results shown in Fig. 5. The curves are the envelope of the sets of curves in Fig. 3 and 4, containing the best power allocation vectors for each (r_B, λ_M) pair.

For $\Gamma_B = 10\,\text{dB}$, H-OMA outperforms H-NOMA only when the channel is almost fully allocated for mMTC devices. As time is shared until the moment when only the eMBB service uses the channel, H-NOMA is the alternative that allows reaching the largest pairs (r_B, λ_M). However, this same behavior does not occur when $\Gamma_B = 20\,\text{dB}$. In this scenario, the use of the proposed method allows greater gains to H-OMA when compared to H-NOMA. In fact, there is a lower bound for the H-NOMA pairs causing saturation of the result as Γ_B increases [7]. For small values of r_B, the envelope of H-NOMA stretches upwards using the power allocation method, but even so, it cannot outperform H-OMA.

VII. Conclusions

This work addressed the coexistence of eMBB and mMTC services in the physical layer, based on the model from [7]. Through the concept of heterogeneity between services, two forms of coexistence can be used: H-OMA and H-NOMA, for orthogonal and non-orthogonal methods, respectively. A decentralized method of changing the transmission powers of mMTC devices through power allocation vectors is proposed. In this way, the instantaneous SNR of the mMTC devices, seen at the BS, fluctuates around more than one average, improving SIC performance. It is shown that for H-OMA, for reduced eMBB traffic, there is an average gain of approximately 60% of active mMTC devices when using the same frequency resource of an eMBB device in time-sharing, in contrast to the standard procedure without the proposed method. For H-NOMA, an average gain of approximately 40% is verified for eMBB rates of up to $r_B = 1.5$ bits/s/Hz, that is, for low eMBB traffic; as traffic increases, interference between services becomes too high, harming performance. To obtain these gains metrics, it is evaluated the average ratio between the curves under comparison over all values of r_B.

Aknowledgements

This work was partially supported by CNPq.

References

[1] Z. Zhang, Y. Xiao, Z. Ma, M. Xiao, Z. Ding, X. Lei, G. K. Karagiannidis, and P. Fan, "6G wireless networks: Vision, requirements, architecture, and key technologies," *IEEE Vehicular Technology Magazine*, vol. 14, no. 3, pp. 28–41, Sep. 2019. [Online]. Available: https://doi.org/10.1109/mvt.2019.2921208

[2] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. D. Silva, F. Tufvesson, A. Benjebbour, and G. Wunder, "5G: A tutorial overview of standards, trials, challenges, deployment, and practice," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 6, pp. 1201–1221, Jun. 2017. [Online]. Available: https://doi.org/10.1109/jsac.2017.2692307

[3] S. Zhang, "An overview of network slicing for 5G," *IEEE Wireless Communications*, vol. 26, no. 3, pp. 111–117, Jun. 2019. [Online]. Available: https://doi.org/10.1109/mwc.2019.1800234

[4] F. Debbabi, R. Jmal, L. C. Fourati, and R. L. Aguiar, "An overview of interslice and intraslice resource allocation in B5G telecommunication networks," *IEEE Transactions on Network and Service Management*, vol. 19, no. 4, pp. 5120–5132, Dec. 2022. [Online]. Available: https://doi.org/10.1109/tnsm.2022.3189925

[5] W. Wu, C. Zhou, M. Li, H. Wu, H. Zhou, N. Zhang, X. S. Shen, and W. Zhuang, "AI-native network slicing for 6G networks," *IEEE Wireless Communications*, vol. 29, no. 1, pp. 96–103, Feb. 2022. [Online]. Available: https://doi.org/10.1109/mwc.001.2100338

[6] J. A. H. Sánchez, K. Casilimas, and O. M. C. Rendon, "Deep reinforcement learning for resource management on network slicing: A survey," *Sensors*, vol. 22, no. 8, p. 3031, Apr. 2022. [Online]. Available: https://doi.org/10.3390/s22083031

[7] P. Popovski, K. F. Trillingsgaard, O. Simeone, and G. Durisi, "5G wireless network slicing for eMBB, URLLC, and mMTC: A communication-theoretic view," *IEEE Access*, vol. 6, pp. 55 765–55 779, 2018. [Online]. Available: https://doi.org/10.1109/access.2018.2872781

[8] E. J. dos Santos, R. D. Souza, J. L. Rebelatto, and H. Alves, "Network slicing for URLLC and eMBB with max-matching diversity channel allocation," *IEEE Communications Letters*, vol. 24, no. 3, pp. 658–661, Mar. 2020. [Online]. Available: https://doi.org/10.1109/lcomm.2019.2959335

[9] B. Bai, W. Chen, Z. Cao, and K. Letaief, "Max-matching diversity in OFDMA systems," *IEEE Transactions on Communications*, vol. 58, no. 4, pp. 1161–1171, Apr. 2010. [Online]. Available: https://doi.org/10.1109/tcomm.2010.04.080478

[10] E. N. Tominaga, H. Alves, O. L. A. Lopez, R. D. Souza, J. L. Rebelatto, and M. Latva-aho, "Network slicing for eMBB and mMTC with NOMA and space diversity reception," in *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*. IEEE, Apr. 2021. [Online]. Available: https://doi.org/10.1109/vtc2021-spring51267.2021.9448974