Power and Discrete Rate Adaptation in Wideband NOMA in Frequency-Selective Channels: A Systematic Approach

S. Sruthy, *Student Member, IEEE* and Neelesh B. Mehta, *Fellow, IEEE*

ICC 2023 - IEEE International Conference on Communications | 978-1-5386-7462-8/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICC45041.2023.10278816

***Abstract*—Non-orthogonal multiple access (NOMA) superimposes signals of multiple users and transmits them simultaneously. To be implementable in 5G and beyond cellular systems, it must adhere to the constraint imposed by the standard that the same modulation and coding scheme (MCS) and power must be used across all physical resource blocks (PRBs) assigned to the users. However, the channel gains of different PRBs are different in wideband channels and the MCSs must belong to a discrete, pre-specified set. We propose a method that uses the exponential effective signal-to-noise ratio mapping (EESM) to systematically determine whether a feasible power allocation exists for a given choice of MCSs, and to find the MCSs that maximize the weighted sum rate. We then propose a novel, lower complexity method called power-normalized EESM, which leads to explicit analytical criteria for the existence of a feasible power allocation. We prove that this method is a relaxation of the original problem under various conditions and is exact for narrowband channels. Wideband NOMA achieves a higher average weighted sum rate than orthogonal multiple access, which is employed by 5G.**

I. Introduction

Non-orthogonal multiple access (NOMA) enables a base station (BS) to serve multiple users simultaneously over the same frequency resource blocks. It has received considerable attention in the literature and standards due to its ability to address the demands for higher data rates, more connected devices, and improved user fairness despite requiring a higher decoding complexity [1], [2].

In downlink power-domain NOMA, the BS superimposes signals from two or more users with different transmit powers and transmits them simultaneously. One or more users employ successive interference cancellation (SIC) in their receivers to retrieve their data in the presence of interference from other users' signals. For example, in the SIC-stable regime of operation of two-user NOMA, the near user first decodes the far user's data, cancels it from its received signal, and decodes its own signal. On the other hand, the far user decodes in the presence of interference from the near user's signal [3, Ch. 6].

NOMA must operate in combination with orthogonal frequency division multiplexing (OFDM), which is the radio access technology in 4G and 5G New Radio (NR). OFDM is also a likely candidate for 6G systems. In OFDM, the system bandwidth is divided into physical resource blocks (PRBs), which are the smallest units of allocation to users. In NOMA, the scheduler at the BS assigns multiple contiguous PRBs to two or more users depending on the data payload.

5G systems are wideband in nature with bandwidths that span several MHz. Due to frequency-selectivity, the channel gain can vary from one PRB to another. However, the 5G standard mandates that the same modulation and coding scheme (MCS) and power must be used on all the PRBs assigned to a user even in wideband frequency-selective channels [4, Sec. 5.2.5]. We shall refer to this as the common MCS-power constraint. This is done to limit the uplink feedback and downlink control signaling overhead. Furthermore, the MCSs must belong to a discrete, pre-specified set. We shall refer to NOMA must operates under the above constraints as *wideband NOMA*.

The common MCS-power constraint implies that the choice of the transmit power and MCS of a user is a function of the vector of signal-to-interference-noise ratios (SINRs) of the PRBs assigned to the user. In wideband NOMA, these SINRs are a function of the vectors of channel gains of the PRBs of all the users since the users interfere with each other. As a result, rate and power adaptation are different in wideband NOMA compared to the adaptation models typically considered in the OFDM-NOMA literature, in which different rates and powers can be assigned to different PRBs [5]–[9].

Contributions: We present a novel theoretical framework for power and discrete rate adaptation in downlink wideband NOMA with two users. Each user is subject to a constraint on its block error rate (BLER), which is the probability that the codeword cannot be decoded by the user. We present an effective SINR-based approach that uses the exponential effective SINR mapping (EESM) to map the vector of SINRs of each user to a single equivalent flat-fading SINR. The optimal power and MCS for each user are systematically determined from the effective SINRs of the two users. EESM leads to decoding constraints that are non-linear functions of the powers of the users. We first propose a gradient descent algorithm based on a barrier function to numerically find a feasible power allocation for a given choice of MCSs. Among the MCSs for which a feasible power allocation exists, the one with the largest weighted sum rate is the optimal one.

We then propose a novel lower-complexity method called power-normalized EESM (PNEESM) in which the decoding constraints become linear inequalities in the user powers. This leads to explicit analytical criteria to determine whether a

This work was supported in part by the "Next Generation Wireless Research and Standardization on 5G and Beyond" project sponsored by the Ministry of Electronics and Information Technology, Govt. of India.

The authors are with the ECE department, Indian Institute of Science, Bengaluru, India (e-mails: {sruthys,nbmehta}@iisc.ac.in).

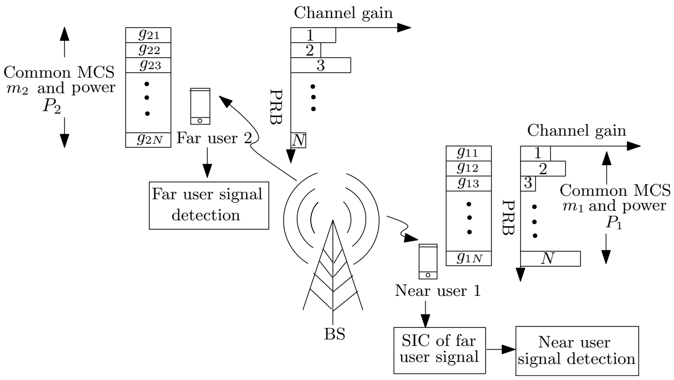

Fig. 1. System model for wideband NOMA with near and far users in the SIC-stable regime. Shown are the PRB gains of different users, their common MCSs, and common powers.

feasible solution exists. We show that the PNEESM method is a relaxation of the original optimization problem under various conditions. Thus, any feasible solution of the original optimization problem is also feasible for the relaxation. Furthermore, this relaxation is exact for narrowband channels.

Since the optimal solution of a relaxation can be infeasible, we propose a PNEESM with backtracking (PB) algorithm that arrives at a feasible solution from an infeasible one. It entails far fewer numerical searches than a brute-force approach. Our numerical results that show that the average weighted sum rate, which is a performance metric often used in the literature, of PB is the same as that of the effective SINR-based approach. Wideband NOMA achieves a higher average weighted sum rate than orthogonal multiple access (OMA), which is employed by 5G.

Comparison with OFDM-NOMA Literature: Power allocation and continuous rate adaptation are done on a per-subcarrier basis in [5]–[7] and the references therein. Bit loading with variable number of bits per subcarrier is studied in [8] and [9] for fixed and continuous power allocation, respectively. While the common MCS-power constraint is considered in [10], the arithmetic average of the subband SINRs is taken to be the effective SINR. This leads to a higher than allowed BLER [11]. Furthermore, fractional power allocation is assumed. Joint user pairing, power allocation, and MCS selection using mutual information effective SINR mapping (MIESM) are studied in [12]. However, the intractable form of MIESM requires involved numerical techniques.

Outline: Section II describes the system model for downlink wideband NOMA. Section III presents the approach for power allocation and MCS selection. Section IV presents the numerical results, and our conclusions follow in Section V.

II. System Model and Problem Statement

We consider a BS that transmits to a near user 1 and a far user 2 simultaneously over N PRBs of an OFDM physical layer. Each PRB consists of S subcarriers. The transmit powers per subcarrier of users 1 and 2 are P_1 and P_2, respectively, and the total power per subcarrier is P.

Let g_{kn} denote the baseband channel power gain (with unit mean) between the k^{th} user and the BS on the n^{th} PRB, for $1 \le k \le K$ and $1 \le n \le N$, and $\mathbf{g}_k = [g_{k1}, g_{k2}, \cdots, g_{kN}]$. The channel is flat over a PRB. This is reasonable when the PRB bandwidth is of the order of 180 to 360 kHz, which is much less compared to the coherence bandwidth of most channels. The model is illustrated in Figure 1.

The SINR $\gamma_{12}^{(n)}$ of user 1 on subcarrier n when it decodes user 2's data is given by

$$\gamma_{12}^{(n)} = \frac{P_2 \ell_1 g_{1n}}{P_1 \ell_1 g_{1n} + N_0 B}, \tag{1}$$

where N_0 is the noise power spectral density, B is the subcarrier bandwidth, and ℓ_k is the pathloss for the k^{th} user.

The SINR $\gamma_{11}^{(n)}$ of user 1 on subcarrier n after it has cancelled user 2's interference and decodes its own data is $\gamma_{11}^{(n)} = P_1 \ell_1 g_{1n}/(N_0 B)$. And, the SINR $\gamma_{22}^{(n)}$ of user 2 on subcarrier n when it decodes its own data is

$$\gamma_{22}^{(n)} = \frac{P_2 \ell_2 g_{2n}}{P_1 \ell_2 g_{2n} + N_0 B}. \tag{2}$$

Let $\mathbf{\Gamma}_{kj} = [\gamma_{kj}^{(1)}, \gamma_{kj}^{(2)}, \ldots, \gamma_{kj}^{(N)}]$ denote the vector of SINRs of the k^{th} user when it decodes the j^{th} user's data, for $(k, j) \in \{(1,1), (1,2), (2,2)\}$.

Let $\Omega = \{0, 1, 2, \ldots, L\}$ denote the discrete set of L MCSs that can be used for transmission. The information rate of MCS $m \in \Omega$ is r_m. The MCSs are arranged in the increasing order of their rates: $0 = r_0 < r_1 < r_2 < \cdots < r_L$. Here, MCS 0, which has a rate $r_0 = 0$, means that no transmission occurs.

Problem Formulation: Our goal is to maximize the weighted sum rate by choosing the MCSs of users 1 and 2 and their powers. Let $m_1 \in \Omega$ and $m_2 \in \Omega$ be the MCSs of users 1 and 2, respectively. Let $\mathsf{BLER}_m(\mathbf{\Gamma}_{kj})$ denote the BLER of MCS m when it is transmitted over N PRBs whose vector of SINRs is $\mathbf{\Gamma}_{kj}$. The optimum MCSs and powers are the solution to the following constrained optimization problem:

$$\mathcal{S}_0 : \max_{\substack{m_1 \in \Omega, m_2 \in \Omega,\\ P_1 \ge 0, P_2 \ge 0}} \{r_{m_1} + w_2 r_{m_2}\}, \tag{3}$$

$$\text{s.t. } \mathsf{BLER}_{m_1}(\mathbf{\Gamma}_{11}) \le \epsilon, \tag{4}$$

$$\max\{\mathsf{BLER}_{m_2}(\mathbf{\Gamma}_{12}), \mathsf{BLER}_{m_2}(\mathbf{\Gamma}_{22})\} \le \epsilon, \tag{5}$$

$$P_1 + P_2 = P, \tag{6}$$

$$P_1 < P_2, \text{ if } m_2 > 0, \tag{7}$$

where $w_2 \ge 1$ is the weight for user 2. It is greater than 1 to ensure user fairness. This formulation encompasses the proportional fair scheduler, in which the weight is inversely proportional the average user throughput [13, Ch. 6]. The constraint in (4) requires that the BLER of user 1 when it decodes its data should always less than or equal to the target value ϵ. The constraint in (5) requires that the BLERs of users 1 and 2 when they decode user 2's data should be less than or equal to the target value ϵ. The total power constraint is captured in (6). The constraint in (7) ensures that error propagation does not happen during SIC. This is referred to as the SIC-stable regime in the NOMA literature [3], [14]. Given (6), (7) is equivalent to $P_1 < \frac{P}{2}$. However, when the BS does not transmit to user 2, such a constraint is not required.

For a given MCS pair (m_1, m_2), we say that a feasible power pair exists if it satisfies the constraints in (4), (5), (6), and (7). In such a case, we say that the MCS pair is feasible.

Comments and Extensions: The model above and the theory we develop below can be extended to support simultaneous transmissions to three or more users. The theory can also be extended to the information-theoretic regime of NOMA [13, Ch. 6], in which a lower power need not always be assigned to the near user. We focus on single-input-single-output systems.

III. Systematic Approach for Joint Power and Rate Adaptation

The problem $\mathcal{S}_0$ is intractable because the BLER of an MCS when it is transmitted over N PRBs with different SINRs is not available in closed-form. We address this by using EESM, which maps a vector of SINRs into an equivalent effective SINR over a flat-fading channel with the same BLER. EESM has been extensively used in 3GPP system simulations due to its accuracy [15].[1] For a vector of SINRs $\mathbf{x} = (x_1, x_2, \ldots, x_N)$, the EESM of MCS m, which we denote by $\mathsf{EESM}(\mathbf{x}, \beta_m)$, is defined as

$$\mathsf{EESM}(\mathbf{x}, \beta_m) \triangleq -\beta_m \ln\left(\frac{1}{N}\sum_{n=1}^{N} \exp\left(-\frac{x_n}{\beta_m}\right)\right), \quad (8)$$

where β_m is an MCS-dependent scaling constant that is tabulated in the literature [17, Table 1]. As a result, the effective SINR depends on the MCS. Let $\tilde{\Gamma}_{kj}(m)$ be the effective SINR of the k^{th} user when it decodes the j^{th} user's data that uses MCS m. From (8), we get

$$\tilde{\Gamma}_{kj}(m) = -\beta_m \ln\left(\frac{1}{N}\sum_{n=1}^{N} \exp\left(-\frac{\gamma_{kj}^{(n)}}{\beta_m}\right)\right). \quad (9)$$

Let T_m be the smallest SINR at which the BLER of MCS m in an additive white Gaussian noise channel is equal to ϵ. Then, (4) can be written as $\tilde{\Gamma}_{11}(m_1) \geq T_{m_1}$. And, (5) simplifies to $\min\{\tilde{\Gamma}_{12}(m_2), \tilde{\Gamma}_{22}(m_2)\} \geq T_{m_2}$. Furthermore, $\tilde{\Gamma}_{12}(m_2) \geq \tilde{\Gamma}_{22}(m_2)$ with high probability because the average channel power gain of the near user 1 is typically several dB greater than that of the far user 2. Hence, (5) simplifies to $\tilde{\Gamma}_{22}(m_2) \geq T_{m_2}$. Therefore, $\mathcal{S}_0$ simplifies to

$$\mathcal{S}_1: \max_{\substack{m_1\in\Omega, m_2\in\Omega,\\ P_1\geq 0, P_2\geq 0}} \{r_{m_1} + w_2 r_{m_2}\}, \quad (10)$$

$$\text{s.t. } \tilde{\Gamma}_{11}(m_1) \geq T_{m_1}, \quad (11)$$

$$\tilde{\Gamma}_{22}(m_2) \geq T_{m_2}, \quad (12)$$

$$P_1 + P_2 = P, \quad (13)$$

$$P_1 < \frac{P}{2}, \text{ if } m_2 > 0. \quad (14)$$

[1] We note that the MIESM has also been used as a link quality metric in 3GPP system simulations. We do not consider it as its involved form makes it intractable [16].

To solve $\mathcal{S}_1$, for every MCS pair, we first determine if a feasible power pair exists. Among the feasible MCS pairs, the one with the largest weighted sum rate is the optimal one.

A. Feasible Power Allocation For an MCS Pair (m_1, m_2)

The effective SINRs in (11) and (12) are non-linear functions of P_1 and P_2. To assess whether a feasible (P_1, P_2) exists, we design the following barrier function F. It consists of three exponential terms that are based on the constraints in (11), (12), and (14):

$$F(P_1) = \exp\left(-\left[\tilde{\Gamma}_{11}(m_1) - T_{m_1}\right]\right) + \exp\left(-\left[\tilde{\Gamma}_{22}(m_2) - T_{m_2}\right]\right) + \exp\left(-\left[\frac{P}{2} - P_1\right]\right). \quad (15)$$

Here, $\exp(-x)$ is an approximation to the indicator function [18, Ch. 11]. It is non-negative and increases rapidly if the inequality is not satisfied, i.e., when $x < 0$. However, for a feasible solution, F is small. We find a P_1 and $P_2 = P - P_1$ that minimize F using gradient descent and check if they satisfy all the constraints in $\mathcal{S}_1$.

The update equation in gradient descent at the $(k+1)^{\text{th}}$ iteration can be written as: $P_1^{(k+1)} = P_1^{(k)} - \eta \frac{\partial F}{\partial P_1^{(k)}}$, where $\eta > 0$ is the learning rate and $P_1^{(k)}$ is the power at the k^{th} step. Gradient descent terminates when the difference in the values of F after two consecutive iterations is less than a predetermined threshold or $P_1^{(k)}$ satisfies (11), (12), (13), and (14). We can show that F is a L-Lipschitz function. Hence, the gradient descent is guaranteed to converge to a stationary point so long as $\eta < (2/L)$ [19, Ch. 1].

B. Analytical Method that Avoids Numerical Search

We now propose an alternate, novel method that avoids the above numerical search. We first define the PNEESM $\tilde{G}_k(m)$ of the k^{th} user when it uses MCS m as

$$\tilde{G}_k(m) = -\beta_m \ln\left(\frac{1}{N}\sum_{n=1}^{N} \exp\left(-\frac{\alpha \ell_k g_{kn}}{N_0 B \beta_m}\right)\right), \quad (16)$$

where α is a pre-specified positive constant. $\tilde{G}_k(m)$ has the same form as EESM except that the power term is replaced with the constant α.

The following lemma connects the effective SINR $\tilde{\Gamma}_{22}(m_2)$ with the PNEESM $\tilde{G}_2(m_2)$ of user 2. Let $\overline{\mathsf{SNR}}_k = \frac{P_k \ell_2}{N_0 B N}\sum_{n=1}^{N} g_{2n}$ be the subcarrier-averaged signal-to-noise ratio (SNR) of the k^{th} user.

Lemma *1:* For large β_{m_2} and high $\overline{\mathsf{SNR}}_1$, the effective SINR $\tilde{\Gamma}_{22}(m_2)$ is upper bounded by

$$\tilde{\Gamma}_{22}(m_2) \leq \frac{\frac{P_2}{\alpha}\tilde{G}_2(m_2)}{\frac{P_1}{\alpha}\tilde{G}_2(m_2) + 1} + \frac{P_2}{P_1}\mathcal{O}\left(\frac{1}{\overline{\mathsf{SNR}}_1}\right). \quad (17)$$

When $\overline{\mathsf{SNR}}_1$ is small and $\frac{P_2}{\alpha} \geq 1$, the effective SINR $\tilde{\Gamma}_{22}(m_2)$ is upper bounded by

$$\tilde{\Gamma}_{22}(m_2) \leq \frac{\frac{P_2}{\alpha}\tilde{G}_2(m_2)}{\frac{P_1}{\alpha}\tilde{G}_2(m_2) + 1}. \quad (18)$$

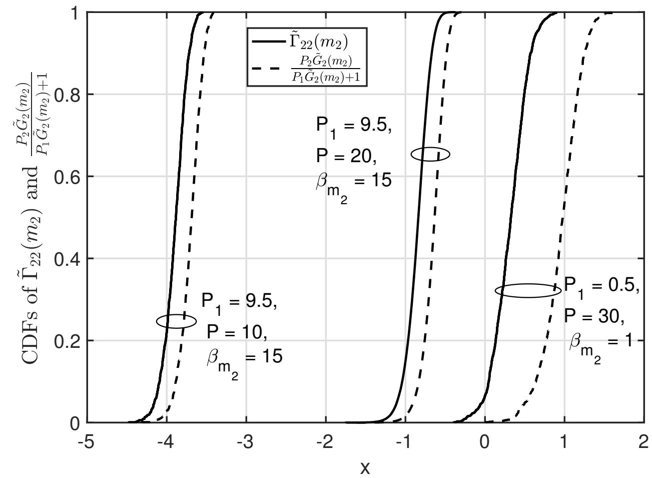

Fig. 2. CDFs of $\tilde{\Gamma}_{22}(m_2)$ and $\frac{P_2\tilde{G}_2(m_2)}{P_1\tilde{G}_2(m_2)+1}$ for small β_{m_2} and low $\overline{\mathsf{SNR}}_1$, large β_{m_2} and high $\overline{\mathsf{SNR}}_1$, and high $\overline{\mathsf{SNR}}_1$ and $P_2 < 1$ ($\alpha = 1$).

Proof: The proof is given in Appendix A. ∎

The next lemma connects $\tilde{\Gamma}_{11}(m_1)$ and $\tilde{G}_1(m_1)$ of user 1.

Lemma *2:* The effective SINR $\tilde{\Gamma}_{11}(m_1)$ is upper bounded by

$$\tilde{\Gamma}_{11}(m_1) \leq \frac{P_1}{\alpha}\tilde{G}_1(m_1), \text{ for } \frac{P_1}{\alpha} \geq 1. \quad (19)$$

Proof: The proof is similar to Appendix A. ∎

We shall refer to $\frac{P_k}{\alpha}\tilde{G}_k(m)$ as the scaled PNEESM of the k^{th} user for MCS m.

Comments: The above inequalities can be shown to be become equalities for the narrowband channel, which is equivalent to $N = 1$. And, α acts as a normalizing constant. Its value is inversely proportional to the ratio of the pathloss and the noise power.

It is of interest to understand whether and how often Lemma 1 holds even when the conditions specified in it do not hold. Figure 2 plots the empirical CDFs of $\tilde{\Gamma}_{22}(m_2)$ (in dB) and $\frac{P_2\tilde{G}_2(m_2)}{P_1\tilde{G}_2(m_2)+1}$ (in dB) for different values of β_{m_2} and powers. These CDFs are generated using 1000 realizations of the vector of the channel gains of users 1 and 2.

We see that the CDF of $\tilde{\Gamma}_{22}(m_2)$ is left to the CDF of $\frac{P_2\tilde{G}_2(m_2)}{P_1\tilde{G}_2(m_2)+1}$ not just for small β_{m_2} and low $\overline{\mathsf{SNR}}_1$ but also for high $\overline{\mathsf{SNR}}_1$ and $P_2 < 1$. Therefore, the upper bound holds with high probability for all realizations of powers, channel gains, and β_m.[2]

When Lemma 1 holds, the constraint $\tilde{\Gamma}_{22}(m_2)$ in (12) implies $\frac{P_2}{\alpha}\tilde{G}_2(m_2) \geq T_{m_2}\left(\frac{P_1}{\alpha}\tilde{G}_2(m_2)+1\right)$. Similarly, when Lemma 2 holds, the constraint in (11) implies $\frac{P_1}{\alpha}\tilde{G}_1(m_1) \geq T_{m_1}$. This leads to the following relaxation of $\mathcal{S}_1$ in which the constraints are linear functions of the powers P_1 and P_2:

$$\mathcal{S}_2 : \max_{\substack{m_1\in\Omega, m_2\in\Omega,\\ P_1\geq 0, P_2\geq 0}} \{r_{m_1} + w_2 r_{m_2}\}, \quad (20)$$

$$\text{s.t. } \frac{P_1}{\alpha}\tilde{G}_1(m_1) \geq T_{m_1}, \quad (21)$$

$$\frac{P_2}{\alpha}\tilde{G}_2(m_2) \geq T_{m_2}\left(\frac{P_1}{\alpha}\tilde{G}_2(m_2)+1\right), \quad (22)$$

$$P_1 + P_2 = P, \quad (23)$$

$$P_1 < \frac{P}{2}, \text{ if } m_2 > 0. \quad (24)$$

The following result explicitly specifies when a feasible (P_1, P_2) exists for $\mathcal{S}_2$ for a given MCS pair (m_1, m_2).

Result *1:* For an MCS pair (m_1, m_2), a feasible power allocation for the PNEESM method exists if and only if

$$\tilde{G}_1(m_1) > \max\left\{\frac{\alpha T_{m_1}\tilde{G}_2(m_2)(T_{m_2}+1)}{\tilde{G}_2(m_2)P - \alpha T_{m_2}}, \frac{2\alpha T_{m_1}}{P}\right\}, \quad (25)$$

$$\tilde{G}_2(m_2) \geq \frac{\alpha T_{m_2}}{P}. \quad (26)$$

Proof: The proof is given in Appendix B. ∎

Thus, determining whether a feasible power allocation exists no longer requires a numerical search. The proof in Appendix B also shows that multiple solutions for (P_1, P_2) exist. We note that no such analysis is available for the MIESM-based method [12].

The solution obtained for $\mathcal{S}_2$, which is a relaxation when both lemmas hold, might be infeasible for $\mathcal{S}_1$. We present below an approach called PB to find a feasible solution from the infeasible solution so obtained. Let $\mathcal{Q} = \{(m_1, m_2) \in \Omega \times \Omega : (m_1, m_2) \text{ is feasible}\}$ be the set of all MCS pairs that are feasible solutions of $\mathcal{S}_2$ for a given vector of channel realizations. $\mathcal{Q}$ can be easily determined by applying Result 1 to each MCS pair. Let $\mathbf{m}_{\text{opt}} = (m_1^*, m_2^*)$ be the MCS pair in $\mathcal{Q}$ with the largest weighted sum rate. If $\mathbf{m}_{\text{opt}}$ is a feasible solution for $\mathcal{S}_1$, then we are done. Its feasibility can be checked by applying the barrier function method of Section III-A. Otherwise, we eliminate this $\mathbf{m}_{\text{opt}}$ from $\mathcal{Q}$ and select the MCS pair with the highest weighted sum rate from $\mathcal{Q}\backslash\{\mathbf{m}_{\text{opt}}\}$. We then check its feasibility, and so on. The algorithm terminates when a feasible $\mathbf{m}_{\text{opt}}$ is found. Since $\mathcal{Q}$ is finite, the algorithm is guaranteed to terminate. Thus, the gradient descent-based numerical search needs to be done only for a handful number of MCS pairs.

IV. Numerical Results

We now present Monte Carlo simulation results to benchmark the weighted sum rate of wideband NOMA. We set $N = 15$, $S = 12$, $w_2 = 4$, $\epsilon = 0.1$, $\alpha = 1$, and $B = 15$ kHz. The BS has 16 MCSs available to it, as specified in [20, Table 5.2.2.1-2]. Their rates range from 0.15 to 5.55 bits/symbol. The results are averaged over 1000 independent realizations of the PRB channel gains of the users.

Figure 3 plots the average weighted sum rate of wideband NOMA using PB and the effective SINR-based approach as

[2]We have observed that the upper bound is violated in only 5 out of 10^5 channel realizations.

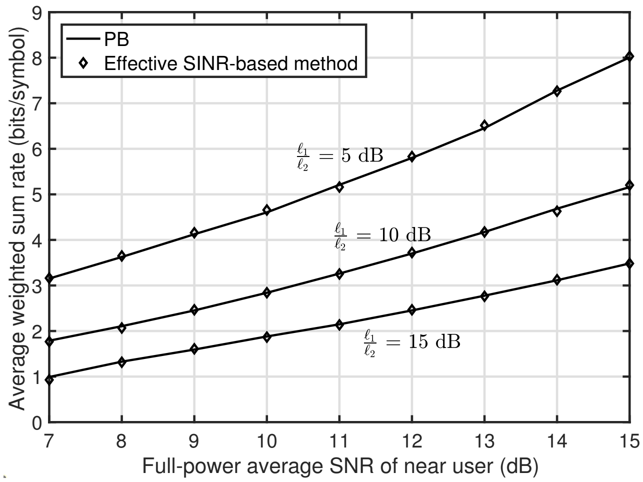

Fig. 3. Average weighted sum rate of wideband NOMA using PB and the effective SINR-based approach as a function of $\frac{P\ell_1}{N_0 B}$.

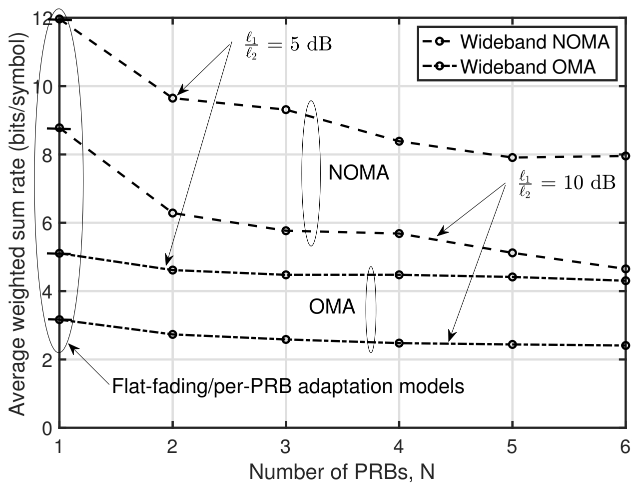

Fig. 4. Comparison of average weighted sum rates of wideband NOMA and wideband OMA as a function of the number of PRBs $\left(\frac{P\ell_1}{N_0 B} = 10 \text{ dB}\right)$. The ticks on the y-axis correspond to the flat-fading/per-PRB adaptation models.

a function of $\frac{P\ell_1}{N_0 B}$, which we shall refer to as the full-power average SNR of the near user. PB's average weighted sum rate is indistinguishable from that of the effective SINR-based approach for all SNRs and pathloss ratios $\frac{\ell_1}{\ell_2}$. As $\frac{\ell_1}{\ell_2}$ increases and the near user's SNR is kept fixed, the average weighted sum rate decreases because the far user's SNR decreases.

Complexity: PB has a much lower complexity than the effective SINR-based approach as it requires at most 4 to 5 searches to identify the optimal MCS pair. It achieves the same BLER target.

Figure 4 plots the average weighted sum rates of wideband NOMA and wideband OMA as a function of the number of PRBs N for $\frac{\ell_1}{\ell_2} = 5$ and 10 dB. In wideband OMA, the BS transmits to only one user at a time and uses the same MCS for all the PRBs assigned to the user in compliance with the common MCS-power constraint. As N increases, the average weighted sum rate decreases. This is because the rate and power adaptation are not allowed to be done on a per-PRB basis. For $N \geq 5$, the decrease is negligible. Also shown as benchmarks are the results for the flat-fading and per-PRB adaptation [21]–[24]. We see that these overestimate the average weighted sum rate. Furthermore, wideband NOMA achieves a substantially higher average weighted sum rate than wideband OMA for all N and $\frac{\ell_1}{\ell_2}$.

V. CONCLUSIONS

We presented a novel approach based on effective SINRs to determine the MCSs and powers of the users whose signals were superimposed in wideband NOMA. Unlike several approaches in the literature that assumed flat-fading or per-PRB adaptation, our approach enabled implementation of NOMA in the 5G OFDMA-based standard by adhering to its common MCS and power constraint. We also proposed a lower complexity analytical approach called PB that exploited properties of the effective SINR.

The average weighted sum rate of PB was indistinguishable from that of the effective SINR-based approach. Wideband NOMA had a higher average weighted sum rate than wideband OMA employed by the standard. We also saw that the per-PRB adaptation approaches significantly overestimated the average weighted sum rate. An interesting avenue for future work is to extend this work to multiple-input-multiple-output systems.

APPENDIX

A. Proof of Lemma 1

a) Proof of (17)*:* As $\beta_{m_2} \to \infty$, the right hand side of (17) can be written as

$$\lim_{\beta_{m_2}\to\infty} \frac{\frac{P_2}{\alpha}\tilde{G}_2(m_2)}{\frac{P_1}{\alpha}\tilde{G}_2(m_2)+1} = \frac{\frac{P_2}{\alpha}\frac{1}{N}\sum_{n=1}^{N}\frac{\alpha\ell_2 g_{2n}}{N_0 B}}{\frac{P_1}{\alpha}\frac{1}{N}\sum_{n=1}^{N}\frac{\alpha\ell_2 g_{2n}}{N_0 B}+1},$$
$$= \frac{\overline{\mathsf{SNR}}_2}{\overline{\mathsf{SNR}}_1 + 1}. \quad (27)$$

For large $\overline{\mathsf{SNR}}_1$, we have $\frac{\overline{\mathsf{SNR}}_2}{\overline{\mathsf{SNR}}_1+1} = \frac{P_2}{P_1} - \frac{P_2}{P_1}\mathcal{O}\left(\frac{1}{\overline{\mathsf{SNR}}_1}\right)$, since $\overline{\mathsf{SNR}}_2/\overline{\mathsf{SNR}}_1 = (P_2/P_1)$. Hence,

$$\lim_{\beta_{m_2}\to\infty} \frac{\frac{P_2}{\alpha}\tilde{G}_2(m_2)}{\frac{P_1}{\alpha}\tilde{G}_2(m_2)+1} = \frac{P_2}{P_1} - \frac{P_2}{P_1}\mathcal{O}\left(\frac{1}{\overline{\mathsf{SNR}}_1}\right). \quad (28)$$

We know that $e^{-\frac{P_2\ell_2 g_{2n}}{(P_1\ell_2 g_{2n}+N_0 B)\beta_{m_2}}} > e^{\frac{-P_2}{P_1\beta_{m_2}}}$. Summing over n on both sides, we get $\frac{1}{N}\sum_{n=1}^{N} e^{-\frac{P_2\ell_2 g_{2n}}{(P_1\ell_2 g_{2n}+N_0 B)\beta_{m_2}}} > e^{\frac{-P_2}{P_1\beta_{m_2}}}$. Taking log on both sides and rearranging terms yields $\tilde{\Gamma}_{22}(m_2) < \frac{P_2}{P_1}$. Combining this inequality with (28) yields (17).

b) Proof of (18)*:* First, we prove that $\tilde{\Gamma}_{22}(m)$ is an increasing function of β_m. The derivative of $\tilde{\Gamma}_{kj}(m)$ with respect to β_m can be shown to be equal to

$$\frac{d\tilde{\Gamma}_{22}(m)}{d\beta_m} = \ln(N) - \sum_{n=1}^{N} v_n \ln(v_n), \quad (29)$$

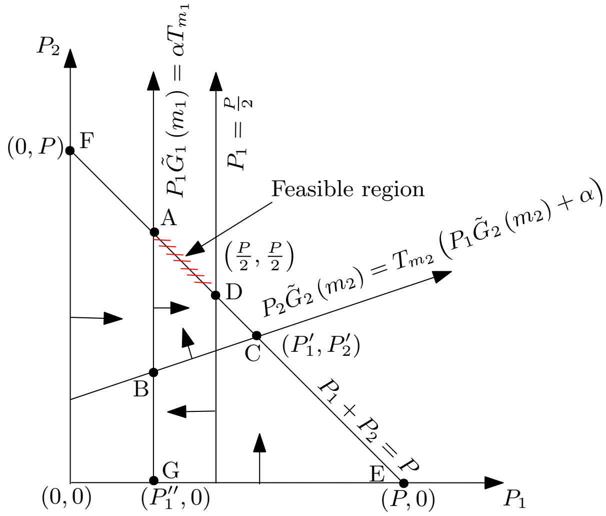

Fig. 5. Illustration of the feasible power region for an MCS pair using the PNEESM method. Arrows represent the feasible regions for the constraints in (21), (22), and (24). The shaded line segment shows the feasible region for (23).

where $v_n = e^{-\frac{\gamma_{22}^{(n)}}{\beta_m}} / \left(\sum_{n=1}^{N} e^{-\frac{\gamma_{22}^{(n)}}{\beta_m}}\right)$. Note that v_n lies between 0 and 1, and $\sum_{n=1}^{N} v_n = 1$. Thus, $\{v_n\}_{n=1}^{N}$ is a probability mass function. From Jensen's inequality, it follows that $\sum_{n=1}^{N} v_n \ln(v_n) \leq \ln(N)$. Thus, $\frac{d\tilde{\Gamma}_{22}(m)}{d\beta_m} \geq 0$. Hence, $\tilde{\Gamma}_{22}(m)$ increases as β_m increases.

For small $\overline{\mathsf{SNR}}_1$, $P_1\ell_2 g_{2n} + N_0 B \approx N_0 B$. Hence, using (9) the expression for $\tilde{\Gamma}_{22}(m_2)$ simplifies to

$$\tilde{\Gamma}_{22}(m_2) = -\beta_{m_2} \ln\left(\frac{1}{N}\sum_{n=1}^{N} \exp\left(-\frac{P_2\ell_2 g_{2n}}{N_0 B \beta_{m_2}}\right)\right). \quad (30)$$

Since $\frac{\alpha\beta_{m_2}}{P_2} \leq \beta_{m_2}$ for $\frac{P_2}{\alpha} \geq 1$, the monotonicity of $\tilde{\Gamma}_{22}(m)$ implies that

$$\begin{aligned}\tilde{\Gamma}_{22}(m_2) &= -\frac{P_2}{\alpha}\frac{\beta_{m_2}\alpha}{P_2} \ln\left(\frac{1}{N}\sum_{n=1}^{N} \exp\left(-\frac{\alpha\ell_2 g_{2n}}{N_0 B \frac{\beta_{m_2}\alpha}{P_2}}\right)\right), \\ &\leq \frac{P_2}{\alpha}\tilde{G}_2(m_2). \quad (31)\end{aligned}$$

Furthermore, for small $\overline{\mathsf{SNR}}_1$, we have $\frac{\frac{P_2}{\alpha}\tilde{G}_2(m_2)}{\frac{P_1}{\alpha}\tilde{G}_2(m_2)+1} \approx \frac{P_2}{\alpha}\tilde{G}_2(m_2)$. Hence, (18) follows.

B. Brief Proof of Result 1

In Figure 5, a feasible (P_1, P_2) that satisfies the constraints in (21), (22), and (24) exists if the minimum of vertices C and D is to the right of the vertical line that joins A and B. From (21), the x co-ordinate of A and B can be shown to be equal to $\frac{\alpha T_{m_1}}{\tilde{G}_1(m_1)}$. Similarly, from (6) and (22), we can show that the x-coordinate of C is $\frac{\tilde{G}_2(m_2)P - \alpha T_{m_2}}{\tilde{G}_2(m_2)\left(T_{m_2}+1\right)}$. Comparing these two x-coordinates yields the desired result.

References

[1] 3GPP, "Study on non-orthogonal multiple access (NOMA) for NR," 3rd Generation Partnership Project (3GPP), TR 38.812 (v16.0.0), Dec. 2018.
[2] S. M. R. Islam, N. Avazov, O. A. Dobre, and K.-s. Kwak, "Power-domain non-orthogonal multiple access (NOMA) in 5G systems: Potentials and challenges," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 721–742, 2nd Qtr. 2017.
[3] M. Vaezi, Z. Ding, and H. Poor, *Multiple Access techniques for 5G Wireless Networks and Beyond*. Springer Intl. Publishing, 2018.
[4] 3GPP, "NR and NG-RAN overall description; stage 2 (release 17)," 3rd Generation Partnership Project (3GPP), TS 38.300 (v17.2.0), Sept. 2022.
[5] S. Mcwade, M. F. Flanagan, J. Mao, L. Zhang, and A. Farhang, "Resource allocation for mixed numerology NOMA," *IEEE Wireless Commun. Lett.*, vol. 10, no. 10, pp. 2240–2244, Oct. 2021.
[6] C. Yuan, X. Tao, W. Ni, N. Li, A. Jamalipour, and R. P. Liu, "Optimal power allocation for superposed secrecy transmission in multicarrier systems," *IEEE Trans. Veh. Technol.*, vol. 70, no. 2, pp. 1332–1346, Feb. 2021.
[7] X. Li, C. Li, and Y. Jin, "Dynamic resource allocation for transmit power minimization in OFDM-based NOMA systems," *IEEE Commun. Lett.*, vol. 20, no. 12, pp. 2558–2561, Dec. 2016.
[8] T. Assaf, A. Al-Dweik, M. S. E. Moursi, and H. Zeineldin, "Efficient bit loading algorithm for OFDM-NOMA systems with BER constraints," *IEEE Trans. Veh. Technol.*, vol. 71, no. 1, pp. 423–436, Jan. 2022.
[9] J.-H. Tseng, Y.-F. Chen, and C.-L. Wang, "User selection and resource allocation algorithms for multicarrier NOMA systems on downlink beamforming," *IEEE Access*, vol. 8, pp. 59 211–59 224, Feb. 2020.
[10] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "System-level performance evaluation of downlink non-orthogonal multiple access (NOMA)," in *Proc. IEEE PIMRC*, Sept. 2013, pp. 611–615.
[11] S. N. Donthi and N. B. Mehta, "An accurate model for EESM and its application to analysis of CQI feedback schemes and scheduling in LTE," *IEEE Trans. Wireless Commun.*, vol. 10, no. 10, pp. 3436–3448, Oct. 2011.
[12] Q.-T. Thieu, C.-H. Wang, and H.-Y. Hsieh, "A wideband scheduling method for non-orthogonal multiple access in the Vienna LTE-A downlink system-level simulator," in *Proc. IEEE Globecom*, Dec. 2016, pp. 1–6.
[13] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge Univ. Press, 2005.
[14] Z. Ding, M. Peng, and H. V. Poor, "Cooperative non-orthogonal multiple access in 5G systems," *IEEE Commun. Lett.*, vol. 19, no. 8, pp. 1462–1465, Jun. 2015.
[15] S. Lagen, K. Wanuga, H. Elkotby, S. Goyal, N. Patriciello, and L. Giupponi, "New radio physical layer abstraction for system-level simulations of 5G networks," in *Proc. ICC*, Jun. 2020, pp. 1–7.
[16] J. Bonafé, S. Boqué, M. Garcia-Lozano, and D. Martin-Sacristan, "Link abstraction models based on mutual information for LTE downlink," COST 2100, Tech. Rep.TD(10)11052, Jun. 2010.
[17] J. Fan, Q. Yin, G. Y. Li, B. Peng, and X. Zhu, "MCS selection for throughput improvement in downlink LTE systems," in *Proc. ICCCN*, Aug. 2011, pp. 1–5.
[18] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge Univ. Press, 2004.
[19] B. T. Polyak, *Introduction to Optimization*. New York : Optimization Software Inc., 1987.
[20] 3GPP, "NR; physical layer procedures for data," 3rd Generation Partnership Project (3GPP), TS 38.214 (v16.11.0), Sept. 2022.
[21] Z. Ding, P. Fan, and H. V. Poor, "Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6010–6023, Aug. 2016.
[22] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, "On the performance of non-orthogonal multiple access in 5G systems with randomly deployed users," *IEEE Signal Process. Lett.*, vol. 21, no. 12, pp. 1501–1505, Jul. 2014.
[23] P. Xu, Z. Ding, X. Dai, and H. V. Poor, "NOMA: An information theoretic perspective," Available on arXiv:1504.07751 [cs.IT], Apr. 2015.
[24] Z. Yang, Z. Ding, P. Fan, and N. Al-Dhahir, "A general power allocation scheme to guarantee quality of service in downlink and uplink NOMA systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7244–7257, Nov. 2016.