# Downlink Non-Orthogonal Multiple Access (NOMA) in Poisson Networks

Konpal Shaukat Ali*, *Student Member, IEEE,* Martin Haenggi†, *Fellow, IEEE,* Hesham ElSawy*, *Senior Member, IEEE,* Anas Chaaban*, *Senior Member, IEEE,* and Mohamed-Slim Alouini*, *Fellow, IEEE*



**Abstract**

A network model is considered where Poisson distributed base stations transmit to $N$ power-domain non-orthogonal multiple access (NOMA) users (UEs) each that employ successive interference cancellation (SIC) for decoding. We propose three models for the clustering of NOMA UEs and consider two different ordering techniques for the NOMA UEs: mean signal power-based and instantaneous signal-to-intercell-interference-and-noise-ratio-based. For each technique, we present a signal-to-interference-and-noise ratio analysis for the coverage of the typical UE. We plot the rate region for the two-user case and show that neither ordering technique is consistently superior to the other. We propose two efficient algorithms for finding a feasible resource allocation that maximize the cell sum rate $\mathcal{R}_{\mathrm{tot}}$, for general $N$, constrained to: 1) a minimum rate $\mathcal{T}$ for each UE, 2) identical rates for all UEs. We show the existence of: 1) an optimum $N$ that maximizes the constrained $\mathcal{R}_{\mathrm{tot}}$ given a set of network parameters, 2) a critical SIC level necessary for NOMA to outperform orthogonal multiple access. The results highlight the importance in choosing the network parameters $N$, the constraints, and the ordering technique to balance the $\mathcal{R}_{\mathrm{tot}}$ and fairness requirements. We also show that interference-aware UE clustering can significantly improve performance.

## I. Introduction

The available spectrum is a scarce resource, and many new technologies to be incorporated into 5G aim at reusing the spectrum more efficiently to improve data rates and fairness. Traditionally,

*The authors are with the Computer, Electrical, and Mathematical Sciences and Engineering (CEMSE) Divison, King Abdullah University of Science and Technology (KAUST), Thuwal, Makkah Province, Saudi Arabia. (Email: {konpal.ali, hesham.elsawy, anas.chaaban, slim.alouini}@kaust.edu.sa)

† The author is with the Department of Electrical Engineering, University of Notre Dame, USA. (Email:mhaenggi@nd.edu)

Part of this work will be presented at the IEEE International Conference on Communications (ICC'18) [1].

M. Haenggi gratefully acknowledges the support of the U.S. National Science Foundation through grant CCF 1525904.

temporal, spectral, or spatial[1] orthogonalization techniques, referred to as orthogonal multiple access (OMA), are used to avoid interference among users (UEs) in a cell. They allow only one UE per time-frequency resource block in a cell. A promising candidate for more efficient spectrum reuse in 5G is non-orthogonal multiple access (NOMA), which allows multiple UEs to share the same time-frequency resource block. The set of UEs being served by a base-station (BS) via NOMA is referred to as the UE cluster. A UE cluster is served by having messages multiplexed either in the power domain or in the code domain. NOMA is therefore a special case of superposition coding [2]. Decoding techniques using successive interference cancellation (SIC) [3] for multiple-access channels have been studied from an information theoretic perspective for several decades [4], and they were implemented on a software radio platform in [5]. The focus of our work is on NOMA where messages are superposed in the power domain, i.e., by transmitting them at different power levels. Hence, NOMA allows multiple UEs to transmit/receive messages in the same time-frequency resource block by multiplexing them at different power levels. SIC techniques are then used for decoding.

With NOMA come a number of challenges, including:

1) Determining the size of the UE cluster, i.e., the number of UEs to be served by a BS.
2) Determining which UEs to include in a cluster, referred to as UE clustering.
3) Ordering UEs within a cluster according to some measure of link quality.
4) The objective of the cluster – prioritizing individual UE performance, total cluster performance, or a middle ground between the two.
5) Resource allocation (RA) for the UEs in a cluster.

Promising results for NOMA as an efficient spectrum reuse technique have been shown [6], [7]. In [8], [9] power allocation (PA) schemes are investigated for universal fairness by achieving identical rates for NOMA UEs. The idea of cooperative NOMA is investigated in [10], [11]. Most NOMA works order UEs based on either their distance from the transmitting BS [9], [11]–[13] or on the quality of the transmission channel [7], [8], [10], [14], [15]. A number of works in the literature focus on RA [8], [9], [15]–[17]. RA schemes for maximizing rates with constraints often focus on small NOMA clusters such as the two-user case [9], [16], [17], though works such as [8], [15] consider a general number of UEs in a NOMA cluster.

[1]Spatial separation of UEs with MIMO can be used with either OMA or NOMA.

The works in [6]–[11], [15]–[17] consider NOMA in single cells and therefore do not account for intercell interference, denoted by $I^{\emptyset}$, which has a drastic negative impact on the NOMA performance as shown in [18]. Stochastic geometry has succeeded in providing a unified mathematical paradigm to model large cellular networks and characterize their operation while accounting for intercell interference [19]–[22]. Using stochastic geometry-based modeling, a large uplink NOMA network is studied in [13], [23], a large downlink NOMA network in [23], [24], and a qualitative study on NOMA in large networks is carried out in [18]. However, [24] does not take into account the SIC chain in the signal-to-interference-and-noise ratio (SINR) analysis which overestimates coverage. In [23] two-user NOMA with fixed PA is studied. In the downlink, comparisons are made between random UE selection and selection such that the weaker UE's channel quality is below a threshold and the stronger UE's is greater than a second threshold. It ought to be mentioned that fixed RA doesn't allow the system to meet a defined cluster objective and makes comparison with other schemes such as OMA unfair.

In this work we analytically study a large multi-cell downlink NOMA system that takes into account intercell interference and intracell interference, henceforth called *intraference* denoted by $I^{\circ}$, error propagation in the SIC chain, and the effects of imperfect SIC for a general number of UEs served by each BS (i.e., a general cluster size). We discuss all of the challenges enumerated above. Our goal is to analyze the performance of such a large network setup using stochastic geometry. We introduce and study three different models to show the impact of location-based selection of NOMA UEs in a cluster, i.e., UE clustering, on performance. We analyze and compare the network performance using two ordering techniques, namely mean signal power- (MSP-) based ordering, which is equivalent to distance-based ordering, and instantaneous signal-to-intercell-interference-and-noise-ratio- (ISINR-) based ordering. In this context, the rate region for the two-user case is studied for both ordering techniques. To the best of our knowledge, an analytical work that compares both ordering techniques does not exist. We consider two main objectives: 1) maximizing the cell sum rate defined as the sum of the rates of all the UEs in a NOMA cluster of the cell, subject to a threshold minimum rate (TMR) constraint of $\mathcal{T}$ on the individual UEs 2) maximizing the cell sum rate when all UEs in a cluster have identical rates, i.e., maximizing the symmetric rate. Accordingly, we formulate optimization problems and propose algorithms for intercell interference-aware RA for both objectives. OMA is used to benchmark the gains attained by NOMA.

The contributions of this paper can be summarized as follows:

- We propose three models based on the Poisson cluster process. The UEs are distributed uniformly on a sector of the in-disk which is the largest disk centered at the BS that fits inside the cell (cf. Fig. 1). The sector has angle $\phi$. For Model 1, $\phi = 2\pi$; for Model 2, $\phi = \pi$; and for Model 3, $\phi = 0$ (i.e., the sector reduces to a line). Consequently, UE clustering in our models depends on the characteristics of the Voronoi cell. In contrast, using a standard Matern cluster process such as in [13] would lead to the unrealistic situation where UEs from another Voronoi cell may be part of a NOMA cluster.
- From the rate region for the two-user case we show that contrary to the expected result UE ordering based on ISINR, that takes into account information about not only the path loss but also fading, intercell interference and noise, is not always superior to MSP-based ordering. We discuss how RA and intraference impact this finding.
- We show that there exists an optimum NOMA cluster size that maximizes the constrained cell sum rate given the residual intraference (RI) factor $\beta$.
- We show the existence of a critical level of SIC $1 - \beta$ that is necessary for NOMA to outperform OMA.

The rest of the paper is organized as follows: Section II describes the system model. The SINR analysis and relevant statistics are in Section III. In Section IV the two optimization problems are formulated, the rate region for the two-user scenario is discussed, and algorithms for solving the problems are proposed. Section V presents the results, and Section VI concludes the paper.

*Notation:* Vectors are denoted using bold text, $\|\mathbf{x}\|$ denotes the Euclidean norm of the vector $\mathbf{x}$, $b(\mathbf{x}, r)$ denotes a disk centered at $\mathbf{x}$ with radius $r$, and $s(\mathbf{x}, r, \phi)$ denotes the sector of the disk centered at $\mathbf{x}$ with radius $r$ and angle $\phi$; when $\phi = \pi$, we denote the half-disk by $s(\mathbf{x}, r)$. $\mathcal{L}_X(s) = \mathbb{E}[e^{-sX}]$ denotes the Laplace transform (LT) of the PDF of the random variable $X$. The ordinary hypergeometric function is denoted by ${}_2F_1$.

## II. System Model

### A. NOMA System Model

We consider a downlink cellular network where BSs transmit with a total power budget of $P = 1$. Each BS serves $N$ UEs in one time-frequency resource block by multiplexing the signals for each UE with different power levels; here $N$ denotes the cluster size. The BSs use fixed-rate transmissions (that can be different for each UE) for the transmission of each message leading to UE rates that are lower than the transmission target rate. The BSs are distributed according

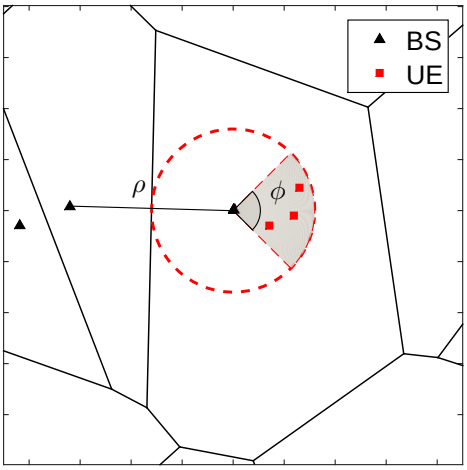


Fig. 1: A realization of the network with $N = 3$ based on a model with $\phi = \pi/2$. The UEs, in-disk (dashed circle) and sector (shaded) for the cell at $\mathbf{o}$ are shown.

to a homogeneous Poisson point process (PPP) $\Phi$ with intensity $\lambda$. To the network we add an extra BS at the origin $\mathbf{o}$, which, under expectation over the PPP $\Phi$, becomes the typical BS serving UEs in the typical cell. In this work we study the performance of the typical cell. Note that since $\Phi$ does not include the BS at $\mathbf{o}$, $\Phi$ is the set of the interfering BSs for the UEs in the typical cell. Denote by $\rho$ the distance between the BS at $\mathbf{o}$ and its nearest neighbor. Since $\Phi$ is a PPP, the distance $\rho$ follows the distribution

$$f_\rho(x) = 2\pi\lambda x e^{-\pi\lambda x^2}, \quad x \geq 0. \tag{1}$$

Consider a disk centered at the $\mathbf{o}$ with radius $\rho/2$. We refer to this as the in-disk as shown in Fig. 1. The in-disk is the largest disk centered at the serving BS that fits inside the Voronoi cell. UEs outside of this disk are relatively far from their BS and thus are better served in their own resource block (without sharing) or even using CoMP if they are near the cell edge [25], [26]. These UEs will not be discussed further in this work.

We consider three models. Each model is a Poisson cluster process with $N$ points distributed uniformly and independently in each cluster. Let $\mathbf{x}$ be the parent point, i.e., the BS, and $\rho_{\mathbf{x}}$ the distance to its nearest neighbor $\mathbf{y}_{\mathbf{x}}$; for brevity, $\rho_{\mathbf{o}}$ is denoted by $\rho$. The points in the cluster are:

- Model 1: distributed on the disk $b(\mathbf{x}, \rho_{\mathbf{x}}/2)$
- Model 2: distributed on the half-disk $s(\mathbf{x}, \rho_{\mathbf{x}}/2)$ such that all points in $s$ have distance at least $\rho_{\mathbf{x}}$ from $\mathbf{y}_{\mathbf{x}}$
- Model 3: distributed on the line segment $s(\mathbf{x}, \rho_{\mathbf{x}}/2) \cap \ell(\mathbf{x}, \mathbf{y}_{\mathbf{x}})$, where $\ell(\mathbf{x}, \mathbf{y}_{\mathbf{x}})$ is the line

through $\mathbf{x}$ and $\mathbf{y_x}$

More compactly, let $s(\mathbf{x}, \rho_{\mathbf{x}}/2, \phi) \subseteq b(\mathbf{x}, \rho/2)$ be the (closed) disk sector of angle $\phi$ whose curved boundary has midpoint $\mathbf{z_x} = (3\mathbf{x} - \mathbf{y_x})/2$. Then for Model 1, $\phi = 2\pi$, for Model 2, $\phi = \pi$, and for Model 3, $\phi = 0$. A realization of the cell at $\mathbf{o}$, its in-disk, and the surrounding cells are shown in Fig. 1. The sector $s(\mathbf{x}, \rho_{\mathbf{x}}/2, \phi)$ with $\phi = \pi/2$ is shaded in the figure.

For Model 1, the union of all the disks $\cup_{\mathbf{x}\in\Phi} b(\mathbf{x}, \rho_{\mathbf{x}}/2)$ is the so-called *Stienen model* [27]. The area fraction covered by the Stienen model is 1/4. This means that if all users form a stationary point process, 1/4 of them are served using NOMA in Model 1 and 1/8 in Model 2 (and 0 in Model 3). More generally, for arbitrary $\phi$, the area fraction is $\phi/(8\pi)$.

Focusing on the typical cell, the link distance $R$ between a UE uniformly distributed in the sector of the in-disk $s(\mathbf{o}, \rho/2, \phi)$ with $\phi > 0$ and the BS at $\mathbf{o}$, conditioned on $\rho$, follows

$$f_{R|\rho}(r \mid \rho) = \frac{8r}{\rho^2}, \quad 0 \le r \le \frac{\rho}{2}. \tag{2}$$

Since $\phi > 0$ for Models 1 and 2, the statistics of their link distances are according to (2). For Model 3, however, the sector becomes a line segment as $\phi \to 0$. Consequently, $R$, conditioned on $\rho$, in Model 3 follows the distribution

$$f_{R|\rho}(r \mid \rho) = \frac{2}{\rho}, \quad 0 \le r \le \frac{\rho}{2}. \tag{3}$$

***Remark 1:*** Model 3 gives an exact distance between the UE and the interferer nearest to $\mathbf{o}$, $z = R + \rho$.

***Remark 2:*** As there is no interfering BS inside $b(\mathbf{o}, \rho)$, a UE located at $\mathbf{u}$ in $s(\mathbf{o}, \rho/2, \phi)$, for any $\phi$, is $\rho - R$ away from the boundary of this disk. Hence, all three models guarantee that there is no interfering BS in $b(\mathbf{u}, \rho - R)$.

It makes sense to employ NOMA for UEs that have good channel conditions and thus can afford to share resource blocks with other UEs rather than those that cannot. Accordingly, any user close to a cell edge is worse off than the cell center users, on average. As $\phi$ decreases, users are located in the in-disk farther from any cell edge, particularly the edge closest to $\mathbf{o}$, and consequently have better intercell interference conditions. In this context, Model 2 can be used as a technique to improve the performance by selecting UEs for NOMA operation, i.e., UE clustering, more efficiently based on their locations, and Model 3 can be viewed as a method to upper bound the achievable performance via UE clustering.

A Rayleigh fading environment is assumed such that the fading coefficients are i.i.d. with a unit mean exponential distribution. A power law path loss model is considered where the signal power decays at the rate $r^{-\eta}$ with distance $r$, where $\eta > 2$ denotes the path loss exponent and $\delta = \frac{2}{\eta}$.

SIC is employed for decoding NOMA UEs. According to the NOMA scheme, the PA and transmission target rate are designed such that the $i^{\text{th}}$ strongest UE is able to decode the messages intended for all those UEs weaker than itself. This requires ordering of UEs based on the quality of the transmission link. We order UEs in such a way that the $i^{\text{th}}$ UE, referred to as UE$_i$, has the $i^{\text{th}}$ strongest transmission link. There are various ways to define what comprises the link quality. The link quality should include the effect of path loss (and therefore link distance), fading and intercell interference. The impact of the large-scale path loss is more stable and dominant than the fading effect which varies on a much shorter time scale. Additionally, accounting for intercell interference and fading necessitates very high feedback overhead. Since for small values of $N$ the path loss dominates the channel relative to fading, considering the quality of a channel to be based on the distance between a UE and its BS is often assumed to be a reasonable approximation [9], [11]–[13], [28], [29]. The link quality can be determined by ordering the UEs of the typical cell from strongest to weakest according to descending

- Mean signal power (MSP)[2]: this ignores fading and therefore orders UEs based on descending $R^{-\eta}$. Equivalently it can be viewed as ordering based on ascending link distance $R$.
- Instantaneous signal power (ISP): this includes fading and therefore orders UEs based on descending $hR^{-\eta}$.
- Mean signal-to-intercell-interference-and-noise ratio (MS$\mathring{\text{I}}$NR): this ignores fading in both the transmission from the serving BS and in the intercell interference and therefore orders UEs based on descending $\frac{R^{-\eta}}{\sum_{\mathbf{x}\in\mathbf{\Phi}} \|\mathbf{x}-\mathbf{u}\|^{-\eta}+\sigma^2}$ where $\|\mathbf{x}\|$ and $\|\mathbf{u}\|$ are the locations of the interfering BSs and UE, respectively and $\sigma^2$ is the noise power.
- Instantaneous signal-to-intercell-interference-and-noise ratio (IS$\mathring{\text{I}}$NR): this includes fading and therefore orders UEs based on descending $\frac{hR^{-\eta}}{I^{\circ}+\sigma^2}$.

Analyzing the SINR for ordering based on ISP and MS$\mathring{\text{I}}$NR is out of the scope of this work. Hence, we analyze and compare the following two schemes:

[2]It should be noted that this ordering is based on the total unit power transmission received at the UE.

- MSP-based: the UEs of the typical cell are indexed according to their ascending ordered distance $R_i$; the $i^{\text{th}}$ closest UE from $\mathbf{o}$ is referred to as $\text{UE}_i$, for $1 \leq i \leq N$.
- ISI$\mathring{\text{N}}$R-based: UEs of the typical cell are indexed with respect to their descending ordered ISI$\mathring{\text{N}}$R $Z_i$[3]; hence, $\text{UE}_i$ has the $i^{\text{th}}$ largest ISI$\mathring{\text{N}}$R, for $1 \leq i \leq N$.

The power for the signal intended for $\text{UE}_i$ is denoted by $P_i$, hence $P = \sum_i P_i$.

To successfully decode its own message, $\text{UE}_i$ must therefore be able to decode the messages intended for all UEs weaker than itself, i.e, $\text{UE}_{i+1}, \ldots, \text{UE}_N$. This is achieved by allocating higher powers and/or lower target rates to the data streams of the weaker UEs. Correspondingly, $\text{UE}_i$ is not able to decode any of the streams sent to UEs stronger than itself, i.e., $\text{UE}_1, \ldots, \text{UE}_{i-1}$ due to their smaller powers and/or higher target rates. Assuming perfect SIC, the intraference experienced at $\text{UE}_i$ when decoding its own message, $I_i^{\circ}$, is comprised of the powers from the messages intended for $\text{UE}_1, \ldots, \text{UE}_{i-1}$. Since in practice SIC is not perfect, our mathematical model additionally considers a fraction $0 \leq \beta \leq 1$ of RI from the UEs weaker than $\text{UE}_i$ in $I_i^{\circ}$. When perfect SIC is assumed, $\beta = 0$, while $\beta = 1$ corresponds to no SIC at all. Additionally, $\text{UE}_i$ suffers from intercell interference, $I_i^{\mathring{o}}$, arising from the power received from all the other BSs in the network, and noise power $\sigma^2$. For the NOMA network $2N-1$ parameters are to be selected, namely $N$ target rates and $N-1$ powers. Note that MSP-based ordering of UEs is agnostic to intercell interference and fading; however, our RA (choice of the $2N-1$ parameters) is not. For the case of ISI$\mathring{\text{N}}$R-based ordering, both ordering and RA are intercell interference- and fading-aware.

### *B. OMA System Model*

We compare our NOMA model with a traditional OMA network where only one UE is served by each BS in a single time-frequency resource block. We focus on time division multiple access (TDMA). For a fair comparison with the NOMA system, the BS serves $N$ UEs distributed uniformly at random in (part of) the in-disk as in the NOMA setup according to the model being employed. Each TDMA message is transmitted using full power $P = 1$ for a duration $T_i$. Without loss of generality, a unit time duration is assumed for a NOMA transmission and therefore $\sum_i T_i = 1$. Consequently, $2N-1$ parameters are to be selected for the OMA network,

[3]It ought to be noted that $Z_i$ is equivalent to $\text{SINR}_{\text{OMA}}^{i}$ in (28). We use the notation $Z_i$ for brevity and to differentiate between the context it is being used in.

namely $N$ target rates and $N-1$ fractions of the time slot. We compare both MSP-based UE ordering and ISINR-based ordering for the OMA model, too.

## III. SINR Analysis

### A. SINR in NOMA Network

In the NOMA network, the SINR at $\text{UE}_i$ of the message intended for $\text{UE}_j$ in the typical cell for $i \leq j \leq N$ is

$$\text{SINR}_j^i = \frac{h_i R_i^{-\eta} P_j}{\underbrace{h_i R_i^{-\eta}\left(\sum_{m=1}^{j-1} P_m + \beta \sum_{k=j+1}^{N} P_k\right)}_{I_{j,i}^{\circ}} + \underbrace{\sum_{\mathbf{x}\in\Phi} g_{\mathbf{y}_i}\|\mathbf{y}_i\|^{-\eta}}_{I_i^{\phi}} + \sigma^2},$$

where $\mathbf{y}_i = \mathbf{x} - \mathbf{u}_i$, $\mathbf{u}_i$ is the location of $\text{UE}_i$, $h_i$ ($g_{\mathbf{y}_i}$) is the fading coefficient from the serving BS (interfering BS) located at $\mathbf{o}$ ($\mathbf{x}$) to $\text{UE}_i$. The intraference experienced when $\text{UE}_i$ decodes the message for $\text{UE}_j$ is denoted by $I_{j,i}^{\circ}$. We use $I_i^{\circ}$ to denote $I_{i,i}^{\circ}$.

### B. Laplace Transform of the Intercell Interference

We analyze the LT of the intercell interference at both the unordered UE and the UE ordered based on MSP. Upon taking the expectation over the BS PPP and the unordered UE's (ordered UE's) location, the UEs in the cell with the BS at $\mathbf{o}$ become the typical unordered UEs (typical ordered UEs, from $\text{UE}_1$ to $\text{UE}_N$.).

***Lemma 1:*** The LT of $I^{\phi}$ ($I_i^{\phi}$) at the typical unordered UE (ordered $\text{UE}_i$) conditioned on $R$ ($R_i$) and $\rho$, where $u = \rho - R$ ($u_i = \rho - R_i$), in Model 1 is approximated as

$$\mathcal{L}_{I^{\phi}|R,\rho}(s) \approx \exp\left(-\frac{2\pi\lambda s}{(\eta-2)u^{\eta-2}}{}_2F_1\left(1, 1-\delta; 2-\delta; \frac{-s}{u^{\eta}}\right)\right)\frac{1}{1+s\rho^{-\eta}} \tag{4}$$

$$\overset{\eta=4}{=} e^{-\pi\lambda\sqrt{s}\tan^{-1}\left(\frac{\sqrt{s}}{u^2}\right)}\frac{1}{1+s\rho^{-4}} \tag{5}$$

$$\mathcal{L}_{I_i^{\phi}|R_i,\rho}(s) \approx \exp\left(-\frac{2\pi\lambda s}{(\eta-2){u_i}^{\eta-2}}{}_2F_1\left(1, 1-\delta; 2-\delta; \frac{-s}{u_i^{\eta}}\right)\right)\frac{1}{1+s\rho^{-\eta}} \tag{6}$$

$$\overset{\eta=4}{=} e^{-\pi\lambda\sqrt{s}\tan^{-1}\left(\frac{\sqrt{s}}{{u_i}^2}\right)}\frac{1}{1+s\rho^{-4}}. \tag{7}$$

*Proof:* Let $\mathbf{y} = \mathbf{x} - \mathbf{u}$, where $\|\mathbf{x}\|$ and $\|\mathbf{u}\|$ are the locations of the interfering BSs and the UE, respectively. The fading coefficient from the interfering BS at $\mathbf{x}$ to the UE is $g_{\mathbf{y}}$. The intercell interference experienced at the UE is

$$I^{\emptyset} = \sum_{\substack{\mathbf{x}\in\Phi \\ \|\mathbf{x}\|>\rho}} g_{\mathbf{y}}\|\mathbf{y}\|^{-\eta} + \sum_{\substack{\mathbf{x}\in\Phi \\ \|\mathbf{x}\|=\rho}} g_{\mathbf{y}}\|\mathbf{y}\|^{-\eta}. \tag{8}$$

The first term of the LT accounts for the first term in (8) corresponding to the non-nearest interferers from **o** lying at a distance at least $u$ ($u_i$) from the unordered UE (ordered UE$_i$). It is obtained from employing Slivnyak's theorem, the probability generating functional of the PPP, and MGF of $g_{\mathbf{y}} \sim \exp(1)$. However, this does not include the BS at distance $\rho$ from **o**, which is accounted for by the second term in (8) using the MGF of $g_{\mathbf{y}}$. Denote by $z$ the distance between this interferer and the typical UE. The exact expression is $\mathbb{E}_{z|\rho}\left[(1+sz^{-\eta})^{-1}\right]$. For simplicity we approximate it using the approximate mean of this distance. Since the average position of the typical UE is **o**, $\mathbb{E}[z \mid \rho] \approx \rho$. ■

*Note:* The first term of the LT of $I^{\emptyset}$ ($I_i^{\emptyset}$) is pessimistic since the interference guard zone in our model $u$ ($u_i$) is smaller than the actual one. For the second term, an exact evaluation shows that the difference between $\mathbb{E}[z \mid \rho]$ and $\rho$ is less than 3.2%.

In the case of Model 2 the distance between the UEs and the interferer closest to **o** is larger than in the case of Model 1. This corresponds to a change in the impact of the second term of (8). The LT of intercell interference changes accordingly.

***Lemma 2:*** The LT of $I^{\emptyset}$ ($I_i^{\emptyset}$) at the typical unordered UE (ordered UE$_i$) conditioned on $R$ ($R_i$) and $\rho$, where $u = \rho - R$ ($u_i = \rho - R_i$), in Model 2 is approximated as

$$\mathcal{L}_{I^{\emptyset}|R,\rho}(s) \approx \exp\left(-\frac{2\pi\lambda s}{(\eta-2)u^{\eta-2}}{}_2F_1\left(1,1-\delta;2-\delta;\frac{-s}{u^{\eta}}\right)\right)\frac{1}{1+s(1.25\rho)^{-\eta}} \tag{9}$$

$$\overset{\eta=4}{=} e^{-\pi\lambda\sqrt{s}\tan^{-1}\left(\frac{\sqrt{s}}{u^2}\right)}\frac{1}{1+s(1.25\rho)^{-4}} \tag{10}$$

$$\mathcal{L}_{I_i^{\emptyset}|R,\rho}(s) \approx \exp\left(-\frac{2\pi\lambda s}{(\eta-2){u_i}^{\eta-2}}{}_2F_1\left(1,1-\delta;2-\delta;\frac{-s}{{u_i}^{\eta}}\right)\right)$$
$$\frac{1}{1+s(1.25\rho)^{-\eta}} \tag{11}$$

$$\overset{\eta=4}{=} e^{-\pi\lambda\sqrt{s}\tan^{-1}\left(\frac{\sqrt{s}}{u_i^2}\right)}\frac{1}{1+s(1.25\rho)^{-4}}. \tag{12}$$

*Proof:* The proof follows according to that in Lemma 1. However, in the second term, $\mathbb{E}[z \mid \rho] \approx 1.25\rho$. ■

In the case of Model 3 the distance between the UEs and the interferer closest to **o** is exactly $z = R + \rho$. This too corresponds to a change in the impact of the second term of (8). The LT of intercell interference changes accordingly.

***Lemma 3:*** The LT of $I^{\emptyset}$ ($I_i^{\emptyset}$) at the typical unordered UE (ordered UE$_i$) distributed according to Model 3, conditioned on $R$ ($R_i$) and $\rho$, where $u = \rho - R$ ($u_i = \rho - R_i$), $a_1 = \frac{(1.5\rho)^\eta}{s}$ and $a_2 = \frac{\rho^\eta}{s}$, is approximated as

$$\mathcal{L}_{I^{\emptyset}|R,\rho}(s) \approx \exp\left(-\frac{2\pi\lambda s}{(\eta-2)u^{\eta-2}}{}_2F_1\left(1, 1-\delta; 2-\delta; \frac{-s}{u^\eta}\right)\right)\left(1-3\ {}_2F_1\left(1, \frac{1}{\eta}; \frac{\eta+1}{\eta}; -a_1\right)+ 2\ {}_2F_1\left(1, \frac{1}{\eta}; \frac{\eta+1}{\eta}; -a_2\right)\right) \tag{13}$$

$$\overset{\eta=4}{=} e^{-\pi\lambda\sqrt{s}\tan^{-1}\left(\frac{\sqrt{s}}{u^2}\right)} \times \left(1-\frac{\tan^{-1}\left(a_1^{\frac{1}{4}}\right)+\tanh^{-1}\left(a_1^{\frac{1}{4}}\right)}{\frac{2}{3}a_1^{\frac{1}{4}}}+\frac{\tan^{-1}\left(a_2^{\frac{1}{4}}\right)+\tanh^{-1}\left(a_2^{\frac{1}{4}}\right)}{a_2^{\frac{1}{4}}}\right) \tag{14}$$

$$\mathcal{L}_{I_i^{\emptyset}|R,\rho}(s) \approx \exp\left(-\frac{2\pi\lambda s}{(\eta-2)u_i{}^{\eta-2}}{}_2F_1\left(1, 1-\delta; 2-\delta; \frac{-s}{u_i{}^\eta}\right)\right)\left(1-3\ {}_2F_1\left(1, \frac{1}{\eta}; \frac{\eta+1}{\eta}; -a_1\right)+ 2\ {}_2F_1\left(1, \frac{1}{\eta}; \frac{\eta+1}{\eta}; -a_2\right)\right) \tag{15}$$

$$\overset{\eta=4}{=} e^{-\pi\lambda\sqrt{s}\tan^{-1}\left(\frac{\sqrt{s}}{u_i^2}\right)} \times \left(1-\frac{\tan^{-1}\left(a_1^{\frac{1}{4}}\right)+\tanh^{-1}\left(a_1^{\frac{1}{4}}\right)}{\frac{2}{3}a_1^{\frac{1}{4}}}+\frac{\tan^{-1}\left(a_2^{\frac{1}{4}}\right)+\tanh^{-1}\left(a_2^{\frac{1}{4}}\right)}{a_2^{\frac{1}{4}}}\right). \tag{16}$$

*Proof:* The first term of (13) and (15) follows from the first term of the LTs in Lemma 1. The exact second term is $\mathbb{E}_{z|\rho}\left[\left(1+sz^{-\eta}\right)^{-1}\right]$. Since $z = R + \rho$, using (3), $f_{z|\rho}(y \mid \rho) = 2/\rho,\ \rho \le y \le 3\rho/2$,

$$\begin{aligned}\mathbb{E}_{z|\rho}\left[\left(1+sz^{-\eta}\right)^{-1}\right] &= \int_{\rho}^{1.5\rho} \frac{1}{1+sy^{-\eta}} f_{z|\rho}(y \mid \rho)dy \\ &= 1-3\ {}_2F_1\left(1, \frac{1}{\eta}; \frac{\eta+1}{\eta}; -a_1\right)+2\ {}_2F_1\left(1, \frac{1}{\eta}; \frac{\eta+1}{\eta}; -a_2\right).\end{aligned}$$

■

### C. UE Ordering

Since the order of the UEs is known at the BS, we use order statistics for the PDFs of the link quality. These are derived using the distribution of the unordered link quality statistics and the theory of order statistics [30].

*1) MSP-Based Ordering:* UEs are ordered based on the ascending ordered link distance $R_i$. Hence, $R_i$ is the distance between the $i^{\text{th}}$ nearest UE, i.e., $\text{UE}_i$ to its serving BS, given $\rho$, for $1 \leq i \leq N$. Using the distribution of the unordered link distance $R$ conditioned on $\rho$ in (2) for Models 1 and 2 we have

$$f_{R_i|\rho}(r \mid \rho) = \binom{N-1}{i-1} \frac{8rN}{\rho^2} \left(\frac{4r^2}{\rho^2}\right)^{i-1} \left(1-\frac{4r^2}{\rho^2}\right)^{N-i} \tag{17}$$

for $0 \leq r \leq \rho/2$, where $\binom{c}{d} = \frac{c!}{d!(c-d)!}$.

Similarly, using the distribution of the unordered link distance $R$ conditioned on $\rho$ in (3) for Model 3 we have

$$f_{R_i|\rho}(r \mid \rho) = \binom{N-1}{i-1} N \frac{2}{\rho} \left(\frac{2r}{\rho}\right)^{i-1} \left(1-\frac{2r}{\rho}\right)^{N-i} \tag{18}$$

for $0 \leq r \leq \rho/2$.

Note that MSP-based ordering guarantees that the nearest interfering BS from $\text{UE}_i$ is farther than $\rho - R_i$.

*2) $\mathring{\text{ISI}}\text{NR}$-Based Ordering:* UEs are ordered based on descending ordered $\mathring{\text{ISI}}\text{NR}$, $Z_i$. The unordered $\mathring{\text{ISI}}\text{NR}$ is $Z = \frac{hR^{-\eta}}{I^{\mathring{\text{o}}}+\sigma^2}$.

***Lemma 4:*** The CDF of the unordered $\mathring{\text{ISI}}\text{NR}$ $Z$ conditioned on $\rho$ is approximated as

$$F_{Z|\rho}(x) \approx 1 - \int_0^{\rho/2} \mathcal{L}_{I^{\mathring{\text{o}}}|R,\rho}(xr^\eta) \exp(-xr^\eta\sigma^2) f_{R|\rho}(r) dr, \tag{19}$$

where $\mathcal{L}_{I^{\mathring{\text{o}}}|R,\rho}(s)$ is approximated in Lemmas 1, 2, and 3 for Models 1, 2, and 3, respectively, and $f_{R|\rho}(r)$ is given in (2) for Models 1 and 2 and in (3) for Model 3.

*Proof:* By definition of $Z$,

$$\begin{aligned} F_{Z|\rho}(x) &= \mathbb{E}_{R,I^{\mathring{\text{o}}}} \left[\mathbb{P}\left(h \leq xR^\eta(I^{\mathring{\text{o}}} + \sigma^2) \mid R, I^{\mathring{\text{o}}}\right)\right] \\ &\overset{(a)}{=} \mathbb{E}_{R,I^{\mathring{\text{o}}}} \left[1 - \exp(-xR^\eta I^{\mathring{\text{o}}}) \exp(-xR^\eta\sigma^2)\right] \\ &\overset{(b)}{\approx} 1 - \int_0^{\rho/2} \mathcal{L}_{I^{\mathring{\text{o}}}|R,\rho}(xr^\eta) \exp(-xr^\eta\sigma^2) f_{R|\rho}(r) dr. \end{aligned}$$

Here (a) follows from $h \sim \exp(1)$ and (b) using the definition of the LT of $I^{\emptyset}$ conditioned on $R$ and $\rho$. ■

***Corollary 1:*** The CDF of the ordered ISI$\mathring{\emptyset}$NR $Z_i$ conditioned on $\rho$ is approximated as

$$F_{Z_i|\rho}(x) \approx \sum_{k=N+1-i}^{N} \binom{N}{k} \big(F_{Z|\rho}(x)\big)^k \big(1 - F_{Z|\rho}(x)\big)^{N-k}, \tag{20}$$

where $F_{Z|\rho}(x)$ is given in (19).

### *D. Coverage in NOMA Network*

In order to decode its intended message, $\text{UE}_i$ needs to decode the messages intended for all UEs weaker than itself. We use $\theta_j$ to denote the SINR threshold corresponding to the target rate associated with the message for $\text{UE}_j$. Coverage at $\text{UE}_i$ is defined as the event

$$C_i = \bigcap_{j=i}^{N} \left\{ \text{SINR}_j^i > \theta_j \right\} = \bigcap_{j=i}^{N} \left\{ h_i > R_i^{\eta} (I_i^{\emptyset} + \sigma^2) \frac{\theta_j}{\tilde{P}_j} \right\}, \tag{21}$$

where

$$\tilde{P}_j = P_j - \theta_j \left( \sum_{m=1}^{j-1} P_m + \beta \sum_{k=j+1}^{N} P_k \right).$$

***Remark 3:*** We observe that the impact of the intraference is that of a reduction in the effective transmit power to $\tilde{P}_j$; without intraference, $\tilde{P}_j$ in (21) would be replaced by $P_j$. This reduction and thus $\tilde{P}_j$ is dependent on the target rate of the message to be decoded.

We introduce the notion of *NOMA necessary condition* for coverage, which is coverage when only intraference, arising from NOMA UEs within a cell, is considered. By definition we can write the signal-to-intraference ratio (S$\mathring{\text{I}}$R) of the message for $\text{UE}_j$ at $\text{UE}_i$ as

$$\mathring{\text{SIR}}_j^i = \frac{h_i R_i^{-\eta} P_j}{\frac{h_i}{R_i^{\eta}} \left( \sum_{m=1}^{j-1} P_m + \beta \sum_{k=j+1}^{N} P_k \right)} = \frac{P_j}{\sum_{m=1}^{j-1} P_m + \beta \sum_{k=j+1}^{N} P_k}. \tag{22}$$

From (22), the S$\mathring{\text{I}}$R of the message for $\text{UE}_j$ is independent of the UE (i.e., $\text{UE}_i$) it is being decoded at; hence, it can be rewritten as $\mathring{\text{SIR}}_j$. In order for the message of $\text{UE}_j$ to satisfy the NOMA necessary condition for coverage, we require

$$\mathring{\text{SIR}}_j > \theta_j \quad \Leftrightarrow \quad \tilde{P}_j > 0. \tag{23}$$

The above condition constrains the target rate for the message of $\text{UE}_j$ to be less than a certain value that is a function of the power distribution among the NOMA UEs. If this condition is not satisfied, the message for $\text{UE}_j$ cannot be decoded since $\mathring{\text{SIR}}_j$ is an upper bound on $\text{SINR}_j^i$, $j \geq i$. Consequently $\text{UE}_i$ will be in outage as $\tilde{P}_j$ will not be positive. Note that for the particular case of $\text{UE}_1$ with perfect SIC (i.e., $\beta = 0$), there is no intraference and $\mathring{\text{SIR}}_1 = \infty$ implying $\text{UE}_1$ always satisfies the NOMA necessary condition for coverage when SIC is perfect; equivalently, when $\beta = 0$, $\tilde{P}_1 = P_1$. Hence, failing to satisfy the NOMA necessary condition for coverage guarantees outage for all UEs that need to decode that message simply because the target rate is too high for the given PA. This shows the importance of RA in terms of PA and target rate choice.

Using $M_i = \max\limits_{i \leq j \leq N} \frac{\theta_j}{\tilde{P}_j}$, $C_i$ in (21) can be rewritten compactly as

$$C_i = \{h_i > R_i^{\eta}(I_i^{\circledcirc} + \sigma^2)M_i\}. \tag{24}$$

Next, we derive the coverage probabilities for UEs using each ordering technique.

*1) Coverage for UEs Ordered Based on MSP:*

***Theorem 1:*** If $\tilde{P}_j > 0$, the coverage probability of the typical $\text{UE}_i$ ordered based on MSP, is approximated as

$$\mathbb{P}(C_i) \approx \int_0^{\infty} \int_0^{x/2} e^{-r^{\eta}\sigma^2 M_i} \mathcal{L}_{I_i^{\circledcirc}|R_i,\rho}\left(r^{\eta}M_i\right) f_{R_i|\rho}(r \mid x) dr f_{\rho}(x) dx, \tag{25}$$

where $f_{\rho}(x)$ is given in (1), $f_{R_i|\rho}(r \mid x)$ in (17) for Models 1 and 2 and in (18) for Model 3, and $\mathcal{L}_{I_i^{\circledcirc}|R_i,\rho}$ is approximated in Lemmas 1, 2, and 3 for Models 1, 2, and 3, respectively.

*Proof:*

$$\mathbb{P}(C_i) = \mathbb{E}_{\rho}\left[\mathbb{E}_{R_i}\left[e^{-R_i^{\eta}\sigma^2 M_i}\mathbb{E}\left[e^{-\left(R_i^{\eta}M_i\right)I_i^{\circledcirc}} \mid R_i, \rho\right]\right]\right],$$

as $h_i \sim \exp(1)$. The inner expectation is the conditional LT of $I_i^{\circledcirc}$ (given $R_i$ and $\rho$). From this we obtain (25). ■

*2) Coverage for UEs Ordered Based on $\text{ISI}\mathring{\text{N}}\text{R}$:*

***Theorem 2:*** If $\tilde{P}_j > 0$, the coverage probability of the typical $\text{UE}_i$ ordered based on $\text{ISI}\mathring{\text{N}}\text{R}$, is approximated as

$$\mathbb{P}(C_i) \approx \int_0^{\infty} \left(1 - F_{Z_i|\rho}(M_i \mid x)\right) f_{\rho}(x) dx, \tag{26}$$

where $f_\rho(x)$ is given in (1).

*Proof:* (26) follows using $\mathbb{P}(C_i) = \mathbb{P}\left(Z_i > M_i\right)$. ■

For a given SINR threshold $\theta_i$, corresponding to a target (normalized) rate of $\log(1+\theta_i)$, the rate of the typical UE$_i$ is

$$\mathcal{R}_i = \mathbb{P}(C_i)\log(1+\theta_i). \tag{27}$$

The cell sum rate is $\mathcal{R}_{\text{tot}} = \sum_{i=1}^{N} \mathcal{R}_i$.

### *E. OMA Network*

The SINR for UE$_i$ of the typical cell is

$$\text{SINR}^i_{\text{OMA}} = \frac{h_i R_i^{-\eta}}{\underbrace{\sum_{\mathbf{x}\in\Phi} g_{\mathbf{y}_i}\|\mathbf{x}-\mathbf{u}_i\|^{-\eta}}_{I_i^{\o}} + \sigma^2}. \tag{28}$$

where $\mathbf{u}_i$ is the location of UE$_i$, $h_i$ ($g_{\mathbf{y}_i}$) is the fading coefficient from the serving BS (interfering BS) located at $\mathbf{o}$ ($\mathbf{x}$) to UE$_i$. Coverage at UE$_i$ is defined as $\tilde{C}_i = \left\{\text{SINR}^i_{\text{OMA}} > \theta_i\right\}$.

***Lemma 5:*** In the OMA network, the coverage probability of the typical UE$_i$ ordered based on MSP is approximated as

$$\mathbb{P}(\tilde{C}_i) \approx \int_0^\infty \int_0^{x/2} e^{-\theta_i r^\eta \sigma^2} \mathcal{L}_{I_i^{\o}|R_i,\rho}(\theta_i r^\eta) f_{R_i|\rho}(r|x)dr\, f_\rho(x)dx, \tag{29}$$

where $f_\rho(x)$ is given in (1), $f_{R_i|\rho}(r \mid x)$ in (17) for Models 1 and 2 and (18) for Model 3, and $\mathcal{L}_{I_i^{\o}|R_i,\rho}$ is approximated in Lemmas 1, 2, and 3 for Models 1, 2, and 3, respectively.

*Proof:* Using the exponential distribution of $h_i$ and the LT of $I_i^{\o}$ conditioned on $R_i$ and $\rho$ we obtain (29). ■

***Lemma 6:*** In the OMA network, the coverage probability of the typical UE$_i$ ordered based on ISI$^{\o}$NR is approximated as

$$\mathbb{P}(\tilde{C}_i) \approx \int_0^\infty \left(1 - F_{Z_i|\rho}(\theta_i \mid x)\right) f_\rho(x)dx, \tag{30}$$

where $F_{Z_i|\rho}(y \mid \rho)$ is given in (19) and $f_\rho(x)$ in (1).

*Proof:* (30) follows from $\mathbb{P}(\tilde{C}_i) = \mathbb{P}\left(Z_i > \theta_i\right)$. ■

For a given SINR threshold $\theta_i$ and corresponding target (normalized) rate $\log(1+\theta_i)$, the rate of the typical UE$_i$ is

$$\mathcal{R}_i = T_i \, \mathbb{P}(\tilde{C}_i) \log(1+\theta_i), \tag{31}$$

where $T_i$ is the fraction of the time slot allotted to UE$_i$.

## IV. NOMA Optimization

### A. Problem Formulation

Determining the optimization objective of the NOMA framework can be complicated. The objective of NOMA is to provide coverage to multiple UEs in the same time-frequency resource block. Naturally we are interested in maximizing the cell sum rate. It is well known that the cell sum rate is maximized by allocating all resources (power in the NOMA network) to the best UE [31]. However, this comes at the price of a complete loss of fairness among NOMA UEs, which is one of the main motivations behind serving multiple UEs in a NOMA fashion. Hence, we constrain the objective of maximizing cell sum rate to ensure multiple UEs are served. In addition to the power constraint we consider two kinds of constraints: 1) constraining resources so that each of the typical UEs achieves at least the threshold minimum rate (TMR), 2) constraining resources so that the typical UEs achieve symmetric (identical) rates. Formally stated, these objectives are:

- $\mathcal{P}1$ - Maximum cell sum rate subject to the TMR $\mathcal{T}$:

$$\begin{aligned} &\max_{(P_i,\theta_i)_{i=1,\ldots,N}} && \mathcal{R}_{\text{tot}} \\ &\text{subject to} && \sum_{i=1}^{N} P_i = 1 \text{ and } \mathcal{R}_i \geq \mathcal{T}. \end{aligned}$$

  Because this problem is non-convex, an optimum solution, i.e., choice of $P_i$ and $\theta_i$ that result in the maximum constrained $\mathcal{R}_{\text{tot}}$, cannot be found using standard optimization methods. However, from the rate region for static channels we know that a RA that results in all UEs achieving the TMR $\mathcal{T}$, while all of the remaining power being allocated to the nearest UE, i.e., UE$_1$, to maximize its rate is the optimum solution for that problem. An example of this for the two-user case is presented in [16].

- $\mathcal{P}2$ - Maximum symmetric rate:

$$\max_{(P_i,\theta_i)_{i=1,\ldots,N}} \quad \mathcal{R}_{\text{tot}}$$
$$\text{subject to} \quad \sum_{i=1}^{N} P_i = 1 \text{ and } \mathcal{R}_1 = \ldots = \mathcal{R}_N.$$

  This is equivalent to maximizing the smallest UE rate. Solving this results in a RA that achieves the largest symmetric rate (universal fairness), i.e., $\mathcal{R}_1 = \ldots = \mathcal{R}_N$. Since this problem is also non-convex, an optimum solution cannot be found using standard optimization methods.

***Remark 4:*** The maximum symmetric rate is the largest TMR that can be supported.

***Remark 5:*** Due to outage, the typical UEs that achieve the same effective rate do not necessarily have the same individual target rates (and corresponding $\theta_i$'s).

The same objectives hold for OMA networks. The constrained resource allocated to the UEs, however, is time for TDMA instead of power for NOMA, i.e., $\sum_i T_i = 1$. The OMA UEs enjoy full power in their transmissions. Optimization over target rate is done similarly to NOMA.

### *B. Case Study:* $N = 2$

In this subsection we focus on the two-user case for which we can plot the maximum rate for each UE subject to any power distribution for NOMA. This gives us the rate regions for the $N = 2$ scenario as shown in Figs. 2 and 3. We use Model 1, $\lambda = 10$, $\sigma^2 = -90$ dBm, $\eta = 4$ in this subsection. The rate regions in Fig. 2 are using different $\beta$ values and MSP-based ordering, while Fig. 3 uses both MSP- and ISINR-based UE ordering with perfect SIC (i.e., $\beta = 0$). Since the OMA scheme employed is TDMA, the RA in this case is not in terms of power but of time. We use the optimal $\theta_i$ for a given power (time) distribution between the two NOMA (TDMA) UEs.

In the rate regions in Figs. 2 and 3 each point on the curve is obtained from optimal target rate allocation that maximizes rate given a power (time) distribution for the two NOMA (TDMA) UEs. A zero rate of $\text{UE}_1$ (the intersection of the curves with the y-axis) corresponds to all the power being allocated to $\text{UE}_2$ in the case of NOMA and all the time being allotted to $\text{UE}_2$ in the case of TDMA and vice versa for zero rate of $\text{UE}_2$ (the intersection of the curves with the x-axis). The rest of the points in each NOMA curve (TDMA curve) are made of all possible power (time) distributions between the two UEs. Each curve is the boundary of the corresponding

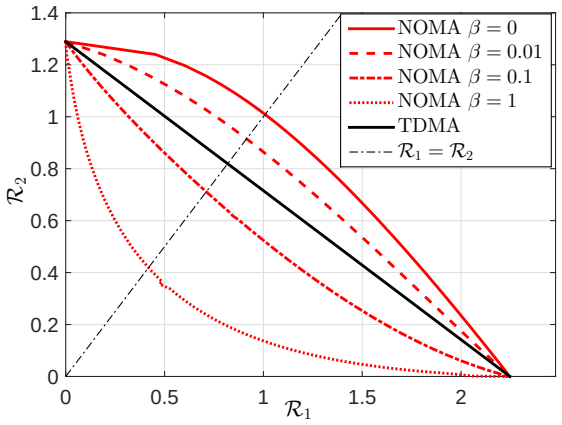


Fig. 2: Rate region for NOMA and TDMA with MSP-based UE ordering for Model 1 using different $\beta$ and $N = 2$.

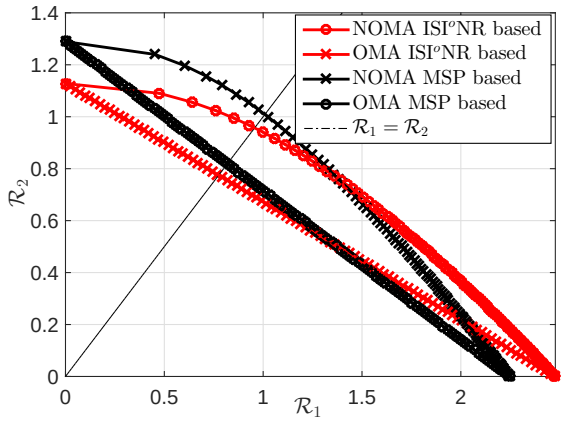


Fig. 3: Rate region for NOMA and TDMA with MSP and ISI$^{\mathring{o}}$NR-based UE ordering for Model 1 with $\beta = 0$ and $N = 2$.

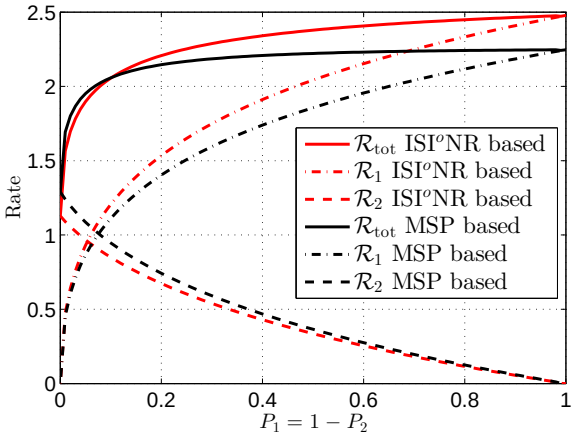


Fig. 4: Optimum cell sum rate and individual UE rates vs. $P_1$ for NOMA with MSP and ISI$^{\mathring{o}}$NR-based ordering for Model 1 with $\beta = 0$ and $N = 2$.

rate region. Optimal RA allows us to operate on the boundary of the rate region. This sort of graph also reveals what areas of rate operation result in higher cell sum rate given a TMR constraint on the UEs. Additionally, if a symmetric rate is required, the rate region shows us the maximum rate possible. Obtaining the rate region for larger $N$, however, is impractical as it requires exhaustively going through the $2N - 1$ parameters for RA.

With perfect SIC (i.e., $\beta = 0$), if RA is optimum, i.e., if we operate at the boundary of the rate region, NOMA outperforms TDMA for both the symmetric-rate objective ($\mathcal{P}2$) and given any TMR ($\mathcal{P}1$) as shown in Figs. 2 and 3. In Fig. 2 we observe that increasing $\beta$ deteriorates performance by pushing the boundary of the rate region inward. Also, if $\beta$ is too high, with

optimum RA, TDMA always outperforms NOMA. Additionally, the rate region graphs shed light on the importance of RA; suboptimum RA can result in significant deterioration in performance as one could be operating inside the rate region far from the boundary. Thus, appropriate RA is very important to fully exploit the potential of NOMA.

In Fig. 4 we plot the optimum cell sum rate and individual UE rates for $N = 2$ against increasing $P_1$ (decreasing $P_2$) for NOMA with the two UE ordering techniques. Intuitively, UE ordering that incorporates more information about the channel is more accurate and should result in superior performance given any power distribution. Accordingly, one may anticipate that ISINR-based ordering, which takes into account path loss, fading, intercell interference and noise, to always be superior in terms of $\mathcal{R}_{\text{tot}}$ to MSP-based ordering, which only accounts for path loss. Contrary to this expectation, we observe that ISINR-based ordering is not always superior. In particular, below a certain $P_1$, MSP-based ordering outperforms ISINR-based ordering in terms of cell sum rate. ISINR-based ordering exceeds in performance when $P_1$ is increased. In Fig. 3 we observe that this holds for TDMA, too.

ISINR-based ordering does, in fact, incorporate more information about the channel; the weakest (strongest) UE in this case on average is weaker (stronger) than the weakest (strongest) UE in MSP-based ordering. This can be seen in Fig. 4 where the weak (strong) UE of ISINR-based ordering consistently underperforms (outperforms) its MSP counterpart. Additionally for NOMA, $\text{UE}_N$ is unable to cancel SI for the messages of any other UE and therefore suffers the largest intraference. In the case of ISINR-based ordering, unlike its MSP counterpart, $\text{UE}_N$ may not necessarily be the farthest UE from the BS making the impact of residual intraference larger; this further deteriorates the SINR of the ISINR-based $\text{UE}_N$. Hence, when $P_1$ is small, the impact on $\mathcal{R}_{\text{tot}}$ of the larger $\mathcal{R}_2$ in MSP-based ordering is more signifcant than the impact on $\mathcal{R}_{\text{tot}}$ of the larger $\mathcal{R}_1$ in ISINR-based ordering. When $P_1$ increases the impact of the significance is reversed.

From Fig. 3 we observe that for higher TMR values (including the symmetric rate), MSP-based ordering outperforms ISINR-based ordering in terms of $\mathcal{R}_{\text{tot}}$. This will become obvious in the results section as well.

### *C. Algorithm for Solving $\mathcal{P}1$*

Since standard optimization techniques cannot be employed for any $N$, the optimum solution to $\mathcal{P}1$ can only be found exhaustively by searching over all combinations of power and target

rate for each of the $N$ NOMA UEs. This, however, is an extremely tedious approach, particularly as $N$ increases. In this subsection we propose an efficient algorithm that, while not guaranteeing an optimum solution, finds a feasible solution, i.e., a solution that satisfies the constraints (but there is no guarantee that the cell sum rate is close to the global maximum).

Given a certain power, $\text{UE}_1$ is able to achieve a larger rate from this resource than any other UE. It therefore makes sense to solve $\mathcal{P}1$ by first ensuring that all UEs other than the strongest achieve TMR with the smallest powers possible. This will leave the largest $P_1$ for $\text{UE}_1$. $\text{UE}_1$ can then maximize the cell sum rate by maximizing $\mathcal{R}_1$ with this power by finding the appropriate target rate.

We tackle this problem by decoupling the choice of power and target rate; our algorithm finds the minimum possible power and corresponding smallest target rate[4] that achieve $\mathcal{T}$ for $\text{UE}_2$ to $\text{UE}_N$ and allocates the remaining power to $\text{UE}_1$. $\text{UE}_1$ then optimizes its target rate (and therefore $\theta_1$) with the remaining power to maximize its rate. If a UE cannot attain $\mathcal{T}$, the available power is insufficient and the algorithm is unable to find a feasible solution as the cluster size $N$ is too large to attain this TMR for all UEs. This can be remedied by either decreasing $\mathcal{T}$ or $N$. Formally, we state the working of the algorithm in Algorithm 1.

Since SIC requires knowledge of RA for the weaker UEs in the decoding chain, we start with $\text{UE}_N$ in line 1. In lines 3 to 16, starting with zero power, we search for the smallest corresponding target rate that can meet the TMR constraint. If for a given power no target rate can achieve TMR, the power is gradually increased. Once the minimum power that can meet the TMR is found, this power and the corresponding minimum target rate is saved. We then move on to the next weakest UE, using the stored power and target rate for the stronger UEs. This process is repeated for the $N-1^{th}$ strongest UE to $\text{UE}_2$. If the TMR cannot be met for $\text{UE}_i$, $i \in \{2,\ldots,N\}$, even when all of the remaining power $1-\sum_{k=i+1}^{N} P_k$ is allocated to it, the power budget is not sufficient for the current TMR and we exit the algorithm in line 13. Otherwise, the rate achieved in line 7 is $\mathcal{R}_i = \mathcal{T}$. After the minimum required powers to achieve the TMR have been allocated to $\text{UE}_2,\ldots\text{UE}_N$, we use the remaining power in lines 17 to 34 to maximize $\mathcal{R}_1$ by finding the appropriate $\theta_1$. If the remaining power is 0 or if $\mathcal{R}_1 < \mathcal{T}$ in lines

[4]For an $i \in \{1,\ldots,N\}$, the function $\mathcal{R}_i(\theta_i)$ is monotonically increasing from zero and then monotonically decreasing to zero, with a unique maximum at a finite $\theta_i > 0$. This is because for small $\theta_i$, $\mathbb{P}(C_i)$ is close to 1, hence $\mathcal{R}_i$ increases linearly with $\log(1+\theta_i)$, while for large $\theta_i$, $\mathbb{P}(C_i)$ goes to zero more quickly than $\log(1+\theta_i)$ grows. Hence, each $\mathcal{R}_i$ (except the maximum) can be satisfied by two $\theta_i$ values. We select the smaller value since it increases the coverage probability for all UEs that are to decode the $i^{\text{th}}$ message.

**Algorithm 1** RA of a feasible solution to $\mathcal{P}1$

Begin with UE$_N$, $i = N$, $P = [\ ]$, $\theta = [\ ]$, $\mathcal{R} = [\ ]$
**while** $i > 0$ **do**
  **if** $i > 1$ **then**
    **for** $0 \leq P_i \leq 1 - \sum_{k=i+1}^{N} P_k$ **do**
      **for** $\theta_i > 0$ **do**
        Calculate $\mathcal{R}_i$ using (27) with (25) for MSP-based (with (26) for ISINR-based) UE ordering
        **if** $\mathcal{R}_i \geq \mathcal{T}$ **then**
          Update: $P = [P_i;\ P]$; $\theta = [\theta_i;\ \theta]$; $\mathcal{R} = [\mathcal{R}_i;\ \mathcal{R}]$; $i = i-1$
          Go to 2
        **end if**
      **end for**
      **if** $P_i = 1 - \sum_{k=i+1}^{N} P_k$ **then**
        TMR cannot be met for all UEs;
        **exit**
      **end if**
    **end for**
  **else**
    $P_1 = 1 - \sum_{k=2}^{N} P_k$
    **if** $P_1 > 0$ **then**
      $\mathcal{R}_1^{\text{vec}} = [\ ]$
      **for** $\theta_1 > 0$ **do**
        Calculate $\mathcal{R}_1$ using (27) with (25) for MSP-based (with (26) for ISINR-based) UE ordering
        Update $\mathcal{R}_1^{\text{vec}} = [\mathcal{R}_1^{\text{vec}}; \mathcal{R}_1]$
      **end for**
      Update: $\mathcal{R}_1 = \max(\mathcal{R}_1^{\text{vec}})$ and corresponding $\theta_1$
      **if** $\mathcal{R}_1 \geq \mathcal{T}$ **then**
        Update: $P = [P_1;\ P]$; $\theta = [\theta_1;\ \theta]$; $\mathcal{R} = [\mathcal{R}_1;\ \mathcal{R}]$; $i = 0$
        Go to 2
      **else**
        TMR cannot be met for all UEs;
        **exit**
      **end if**
    **else**
      TMR cannot be met for all UEs;
      **exit**
    **end if**
  **end if**
**end while**

32 and 29, respectively, the TMR cannot be met for all UEs and we exit the algorithm.

The same algorithm is employed for OMA, except that rates are calculated using using (31) with (29) for MSP-based (with (30) for ISINR-based) UE ordering, and the contending resource is $T$ instead of $P$. Note that since our problem includes intercell interference, our RA is intercell interference-aware.

We compared the solutions of Algorithm 1 with those found using an exhaustive search for the case of $N = 2$ and different values of $\mathcal{T}$. It turns out that our solutions are, in fact, optimum. It is of course not possible to compare the results of Algorithm 1 with those from an exhaustive search for larger $N$ as it is computationally too expensive.

### D. Algorithm for Solving $\mathcal{P}2$

Since $\mathcal{P}2$, like $\mathcal{P}1$, is non-convex, the optimum solution cannot be found using standard optimization techniques. As mentioned in the previous subsection, doing an exhaustive search over all combinations of power and target rate for each of the $N$ NOMA UEs is extremely tedious. We propose an algorithm which, while not guaranteeing an optimum solution, finds a feasible solution.

We start with a target rate, which we will denote by $\mu$, that each UE must achieve. As in Algorithm 1, starting with $\text{UE}_N$, we aim to find the smallest power that can achieve $\mu$. Unlike Algorithm 1 which does this upto $\text{UE}_2$ only, in this case we do it until $\text{UE}_1$, i.e., for all the UEs so that the UEs have symmetric rate $\mu$. If the total power consumed is less than the power budget, we increase $\mu$. However, if each UE cannot achieve $\mu$, the target rate is too high and needs to be reduced. In this way we update $\mu$ until the highest $\mu$ that can be achieved by all UEs while consuming the full power budget is found. Formally, the algorithm is stated in Algorithm 2.

**Algorithm 2** RA of a feasible solution to $\mathcal{P}2$

1: Begin with $\mu = 0.3$, $\mu_H = \infty$, $\mu_{\text{L}} = 0$, $\zeta = 0$, $a = 1.3$, $n = 1$
2: **while** $n$ **do**
3:  **if** $\zeta = 0$ **then**
4:  **if** $\mu_{\text{H}} = \infty$ **then**
5:  $\mu = a\mu$
6:  **else**
7:  $\mu = \frac{\mu_{\text{H}}+\mu}{2}$
8:  **end if**
9:  **else**
10:  $\mu = \frac{\mu_{\text{L}}+\mu}{2}$
11:  **end if**
12:  Begin with $\text{UE}_N$, $i = N$, $P = [\ \ ]$, $\theta = [\ ]$
13:  **while** $i > 0$ **do**
14:  **for** $0 \leq P_i \leq 1 - \sum_{k=i+1}^{N} P_k$ **do**
15:  **for** $\theta_i > 0$ **do**
16:  Calculate $\mathcal{R}_i$ using (27) with (25) for MSP-based (with (26) for ISINR-based) UE ordering
17:  **if** $\mathcal{R}_i \geq \mu$ **then**
18:  Update: $P = [P_i;\ P]$; $\theta = [\theta_i;\ \theta]$; $\zeta = 0$; $i = i - 1$
19:  Go to 13
20:  **end if**
21:  **end for**
22:  **if** $P_i = 1 - \sum_{k=i+1}^{N} P_k$ **then**
23:  $\mu$ cannot be met for all UEs; update: $\zeta = 1$
24:  Go to (28)
25:  **end if**
26:  **end for**
27:  **end while**
28:  **if** $\zeta = 1$ **then**
29:  $\mu_{\text{H}} = \mu$
30:  **else**
31:  $\mu_{\text{L}} = \mu$
32:  **end if**
33:  **if** $\mu_{\text{H}} - \mu_{\text{L}} < 0.01\mu_{\text{H}}$ **then**
34:  Algorithm has converged, update: $n = 0$
35:  **end if**
36: **end while**

In Algorithm 2 all UEs must achieve the target rate of $\mu$, which is achieved in lines 12 to 27. This is done by starting with $\text{UE}_N$ to find the smallest power and its corresponding smallest target rate that can attain $\mu$; once found, these are stored. We then move on to the next weakest UE, using the stored power and target rate for the stronger UEs. This process is repeated until $\text{UE}_1$. If there isn't sufficient power for a UE to attain $\mu$, the flag $\zeta$ in line 23 is updated from 0 to 1 denoting that the current target rate $\mu$ is too high and we exit the while loop to update $\mu$; otherwise, the flag $\zeta = 0$. We begin the algorithm assuming the last $\mu = 0.3$ and $\zeta = 0$. The upper bound on the target rate (which not all of the UEs can attain at once), $\mu_{\text{H}}$, is initially set to $\infty$ and the lower bound on the target rate (which all of the UEs can attain), $\mu_{\text{L}}$, is set to 0. We update $\mu_{\text{H}}$ in line 29 when a smaller value of $\mu$ is found which all of the UEs fail to attain, i.e., when $\zeta = 1$. Similarly, $\mu_{\text{L}}$ is updated in line 31 when a larger value of $\mu$ is found which all of the UEs can attain, i.e., when $\zeta = 0$. This way we iteratively update $\mu$ to be the average of the most updated upper and lower bounds in lines 7 and 10. When the difference between $\mu_{\text{H}}$ and $\mu_{\text{L}}$ is smaller than a certain value, such as 1% in line 37, we assume the algorithm has converged. This way we are able to find the largest symmetric rate. It should be noted that we use the coefficient $a$ such that $a > 1$; this allows us to update $\mu$ when we do not have available a finite $\mu_{\text{H}}$ in line 5. Also, note that although the algorithm begins with $\mu = 0.3$ and $\zeta = 0$, the choice of these parameters is arbitrary; even if $\mu = 0.3$ is not achievable by all of the UEs, since $\mu_{\text{H}}$ will be updated in the next iteration, the algorithm will not function incorrectly.

The same algorithm is employed for OMA, except that rates are calculated using using (31) with (29) for MSP-based (with (30) for ISI̊NR-based) UE ordering, and the contending resource is $T$ instead of $P$.

## V. Results

We first show that the approximations in Theorems 1 and 2 are tight. To this end, we simulate using BS intensity $\lambda = 10$, noise power $\sigma^2 = -90$ dB and $\eta = 4$. The results are shown in Fig. 5 which considers a system with $N = 3$ employing Model 1, a fixed PA scheme where $P_1 = 1/6$, $P_2 = 1/3$ and $P_3 = 1/2$, and both MSP-based and ISI̊NR-based UE ordering. For clarity of presentation we choose the same target rate for all three UEs in both cases and plot coverage of each UE against the corresponding SINR threshold. The figure verifies the accuracy of our SINR analysis as the coverage expressions for both types of UE ordering match the simulation closely. We observe that ISI̊NR-based UE ordering is superior for all UEs other than $\text{UE}_N$. As

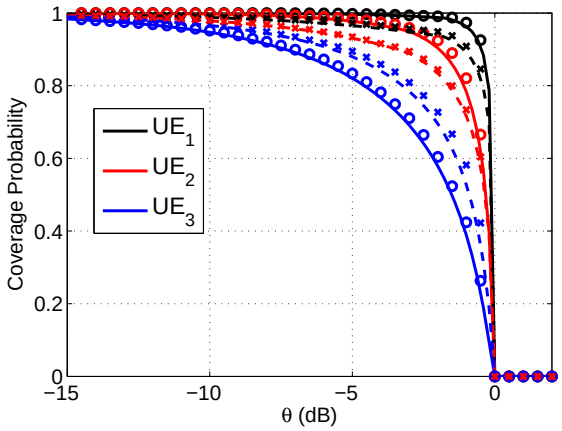


Fig. 5: SINR coverage vs. $\theta$ (identical target rate for all UEs) for Model 1 with $N = 3$ employing the fixed PA $P_1 = 1/6$, $P_2 = 1/3$ and $P_3 = 1/2$. Solid (dashed) lines show the analysis for ISINR-based (MSP-based) UE ordering. Markers show the Monte Carlo simulations.

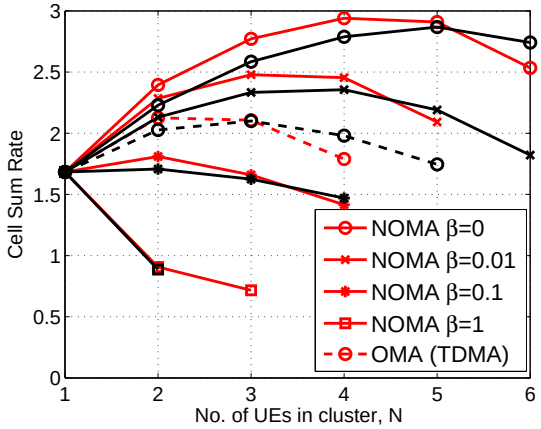


Fig. 6: $\mathcal{R}_{\text{tot}}$ vs. $N$ for Model 1 using OMA and NOMA with $\mathcal{T} = 0.3$ and different $\beta$ values. Black lines are for MSP-based UE ordering, while red lines are for ISINR-based ordering. The curves end at the largest $N$ that can be supported given the TMR constraint and $\beta$.

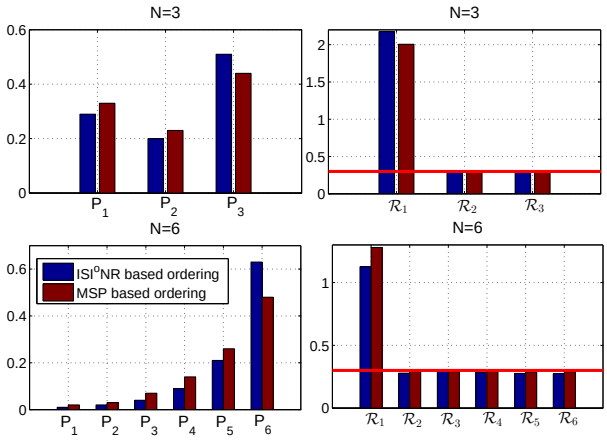


Fig. 7: Comparison of individual NOMA UE power and rates between ISINR- and MSP-based UE ordering for $N = 3$ and $N = 6$ using Model 1 with $\beta = 0$ and $\mathcal{T} = 0.3$.

explained previously, this is because $\text{UE}_N$ for ISINR-based ordering is weaker than its MSP counterpart.

RA for the remaining figures is done according to Algorithms 1 and 2 depending on whether the objective is constrained to a TMR (i.e., $\mathcal{P}1$) or symmetric rate (i.e., $\mathcal{P}2$), respectively. Unless specified otherwise, Model 1 is employed.

Fig. 6 is a plot of the cell sum rate against the number of UEs in a NOMA cluster, $N$, for both MSP-based and ISINR-based UE ordering given a TMR constraint. We have included

$N = 1$ in these plots which has the same $\mathcal{R}_{\text{tot}}$ for all the curves since it only has one UE in a resource block ($\because$ independent of $\beta$) which maximizes its rate ($\because$ independent of $\mathcal{T}$). Given $\mathcal{T}$ and $\beta$, there exists an optimum $N$ that maximizes $\mathcal{R}_{\text{tot}}$. When $\beta$ is large we observe that using NOMA may not necessarily be more beneficial compared to OMA in terms of $\mathcal{R}_{\text{tot}}$. Otherwise, for small $N$, increasing $N$ enhances $\mathcal{R}_{\text{tot}}$ because interference cancellation is efficient in this regime, and more UEs are covered. Also, increasing $N$ results in a stronger $\text{UE}_1$ as it decreases $R_1$ on average in the case of MSP-based ordering and improves $Z_1$ on average for ISI$\not{0}$NR-based ordering; this enhances $\mathcal{R}_1$ given a $P_1$. This in turn enhances $\mathcal{R}_{\text{tot}}$ which in the TMR constraint problem ($\mathcal{P}1$) receives the largest contribution from $\mathcal{R}_1$. However, increasing $N$ beyond the optimum leaves too little power for $\text{UE}_1$ to boost $\mathcal{R}_1$ with. For a given $\mathcal{T}$ and $\beta$, the resources are only sufficient to support a maximum cluster size; after this $N$ (discontinuation of the plots), not all of the UEs are able to achieve $\mathcal{T}$. Increasing $\beta$ results in a decrease in the maximum cluster size that can be supported. Similarly, increasing $\mathcal{T}$ results in a decrease in the maximum cluster size that can be supported [1]; this has not been shown for brevity. NOMA outperforms OMA significantly if $\beta$ is small and can support the same number of UEs or more.

In Fig. 6 we observe that for a given $\beta$ ISI$\not{0}$NR-based UE ordering outperforms MSP-based ordering when $N$ is not larger than a certain value. After this, MSP performs better. In Fig. 7, we compare the individual NOMA UE powers and rates when $\beta = 0$ and $\mathcal{T} = 0.3$ for $N = 6$ (where MSP-based ordering outperforms) with $N = 3$ (where ISI$\not{0}$NR-based ordering outperforms). Unlike the other UEs, we observe that $\text{UE}_N$ requires more power to achieve $\mathcal{T}$ in the ISI$\not{0}$NR-based ordering case than MSP-based ordering. This can be attributed to: 1) $\text{UE}_N$ in the ISI$\not{0}$NR-based case is worse than $\text{UE}_N$ in the MSP-based case causing it to require more power, 2) $\text{UE}_N$ is unable to cancel any intraference; in ISI$\not{0}$NR-based ordering $\text{UE}_N$ may not be the farthest UE, the impact of intraference may therefore be higher than its MSP-based counterpart where $\text{UE}_N$ is guaranteed to be farthest. The other non-strongest UEs in ISI$\not{0}$NR-based ordering require lower powers to achieve TMR than their MSP counterparts implying they are on average stronger. For smaller $N$ ($N = 3$ in Fig. 7) when ISI$\not{0}$NR-based ordering is employed, despite the increased power requirement by $\text{UE}_N$, there is still enough power left for $\text{UE}_1$ to maximize its rate with so that $\mathcal{R}_{\text{tot}}$ exceeds the MSP case. Although $\text{UE}_1$ in the ISI$\not{0}$NR-based case is stronger, when $N$ is larger, the $P_1$ left is too little and $\mathcal{R}_{\text{tot}}$ is lower than its MSP-based counterpart. The figures highlight that when $N$ is large, using the simpler MSP-based ordering scheme results in better performance than the complex ISI$\not{0}$NR-based ordering.

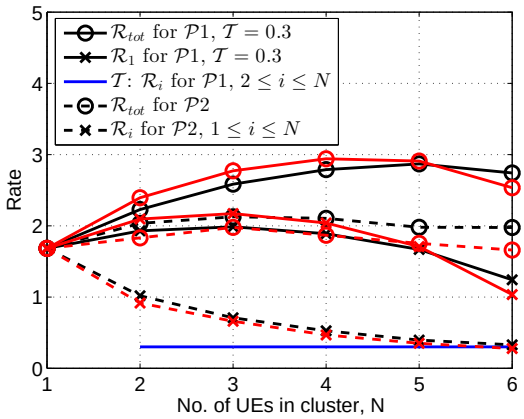


Fig. 8: Individual UE and cell sum rate vs. $N$ using Model 1 with $\beta = 0$ for: $\mathcal{P}1$ with $\mathcal{T} = 0.3$ and $\mathcal{P}2$. Black lines are for MSP-based UE ordering, while red lines are for ISI̊NR-based. The blue line is the TMR achieved by $\text{UE}_2, \ldots, \text{UE}_N$ in $\mathcal{P}1$.

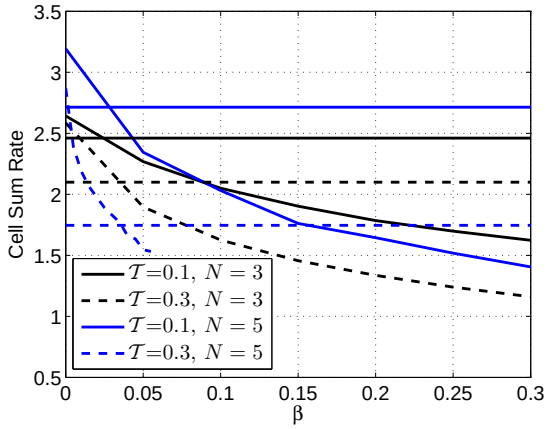


Fig. 9: $\mathcal{R}_{\text{tot}}$ vs. $\beta$ for Model 1 with different values of $N$ and $\mathcal{T}$ using MSP-based ordering. Curves represent NOMA and horizontal lines (independent of $\beta$) represent OMA.

Fig. 8 is a plot of both the cell sum rate and individual UE rates against the number of UEs in a NOMA cluster for MSP and ISI̊NR-based UE ordering. We compare the maximum symmetric rate objective in $\mathcal{P}2$ (dashed lines) with the TMR constrained objective where $\mathcal{T} = 0.3$ in $\mathcal{P}1$ (solid lines). The curves for $\mathcal{P}1$ end at the largest $N$ that can be supported given the TMR constraint and $\beta$. The symmetric rate objective does not have a limit on the largest $N$ that can be supported but for comparison with $\mathcal{P}1$, we plot them up to the same value of $N$. The TMR constraint of $\mathcal{T} = 0.3$ always outperforms the symmetric rate objective in terms of $\mathcal{R}_{\text{tot}}$ for the values of $N$ that it can support; this is in accordance with what is anticipated from the rate region in Fig. 3. Additionally, for $\mathcal{P}2$, MSP-based ordering always has a superior $\mathcal{R}_{\text{tot}}$ compared to its ISI̊NR-based counterpart, which is in line with our conclusions regarding Fig. 3.

In Fig. 8 the symmetric rate objective of $\mathcal{P}2$, like the case of $\mathcal{P}1$, has an optimum $N$ that maximizes $\mathcal{R}_{\text{tot}}$. For the problem in $\mathcal{P}2$, the largest symmetric rate is limited by the weakest UE, $\text{UE}_N$, which requires the largest power. As $N$ grows, the weakest UE becomes worse and the total power needs to be shared among a larger number of UEs. This causes the individual UE's (symmetric) rate to decrease with $N$ as shown. However, increasing $N$ at first enhances $\mathcal{R}_{\text{tot}}$ because SIC is efficient in this regime and more UEs are covered, so the gains from the larger number of UEs are more significant. For larger $N$, the individual UE rate becomes too small. Consequently $\mathcal{R}_{\text{tot}}$ starts decreasing after the optimum $N$. Additionally, as long as $N$ is not too

large, the individual UEs perform better for $\mathcal{P}2$ compared to $\text{UE}_2, \ldots, \text{UE}_N$ in $\mathcal{P}1$ which achieve $\mathcal{T}$. Also, $\mathcal{R}_1$ in $\mathcal{P}1$ outperforms the individual UE rates in $\mathcal{P}2$ as anticipated. More interestingly, $\mathcal{R}_1$ has an optimum $N$ for which it is maximized. This is due to the improving strength of $\text{UE}_1$ as $N$ grows, followed by a decrease in $\mathcal{R}_1$ because of lower available $P_1$ when $N$ is too large.

Fig. 9 plots the cell sum rate against $\beta$ for different $N$ and $\mathcal{T}$ using MSP-based ordering. Since OMA does not use SIC, the corresponding $\mathcal{R}_{\text{tot}}$ plots are horizontal lines independent of $\beta$. The figure shows the existence of a maximum $\beta$ value until which a NOMA system with a particular $N$ and $\mathcal{T}$ is able to outperform the corresponding OMA system. This highlights that there is a critical minimum level of SIC required for NOMA to outperform OMA. We also observe that the decrease in $\mathcal{R}_{\text{tot}}$ as a function of $\beta$ is steeper for larger $N$ and $\mathcal{T}$ highlighting their increased susceptibility to RI.

The results highlight the importance and impact of choosing network parameters such as $N$ and the UE ordering technique depending on the network objective and $\beta$. As an example, if complete user fairness is required, i.e., the objective is $\mathcal{P}2$, MSP-based ordering would result in higher $\mathcal{R}_{\text{tot}}$, while $N$ would be chosen according to $\beta$. Similarly, if the network requires a certain TMR, the objective is $\mathcal{P}1$. To enhance $\mathcal{R}_{\text{tot}}$, ISINR-based ordering may be chosen if $\mathcal{T}$ is not too high with a smaller $N$; otherwise MSP-based ordering would be a better option. The value of $N$ would also depend on $\beta$.

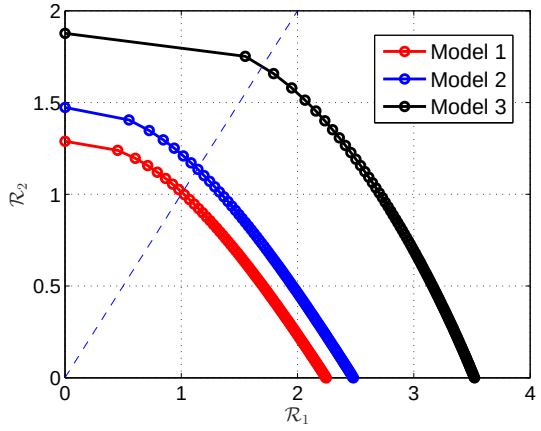


Fig. 10: Rate region for NOMA with MSP-based UE ordering, $\beta = 0$ and $N = 2$ for Models 1, 2, and 3.

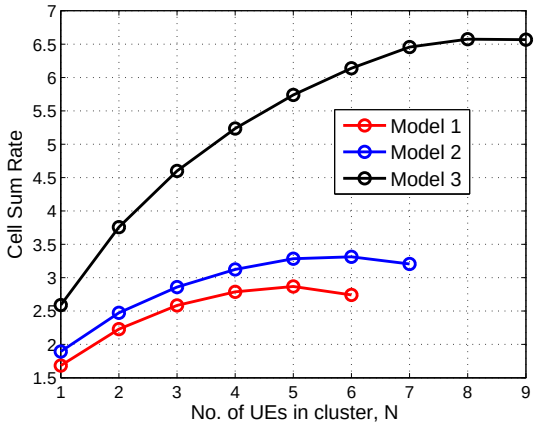


Fig. 11: $\mathcal{R}_{\text{tot}}$ vs. $N$ for $\mathcal{T} = 0.3$ using NOMA and $\beta = 0$ for MSP-based ordering. The curves end at the largest $N$ that can be supported given $\mathcal{T}$ and $\beta$.

Fig. 10 is a plot of the rate regions for the $N = 2$ case using MSP-based UE ordering and with $\beta = 0$ for Models 1, 2, and 3. We observe that the optimum performance (rate region boundary)

of Model 2 can be viewed as an upper bound on that of Model 1. This can be explained by the fact that Model 2 selects UEs that are located farther from the closest cell edge than the UEs in Model 1 resulting in better average interference conditions and consequently, performance. Similarly Model 3, which selects UEs that are located even farther from the closest cell edge than both Models 1 and 2 and thereby have the best average interference conditions, upper bounds the other two models in terms of performance.

Fig. 11 is a plot of the cell sum rate with increasing $N$ for all three models with MSP-based ordering, $\beta = 0$ and $\mathcal{T} = 0.3$. In general, the smaller the sector angle $\phi$, the more clustered NOMA UEs are towards the center of the cell and the better their performance is on average. Accordingly, we observe that Model 2 outperforms Model 1 for each value of $N$, and that Model 3 outperforms both Models 1 and 2. This highlights how a superior UE clustering technique that selects UEs with good interference conditions is able to significantly improve the performance. In particular, for the given $\beta$ and $\mathcal{T}$ at its optimum $N$, Model 2 outperforms the optimum of Model 1 by 15.5%, and Model 3 outperforms Models 1 and 2 by 129% and 98.4%, respectively. Additionally, we observe that a more superior clustering technique is able to support a larger maximum cluster size given a TMR constraint; for $\mathcal{T} = 0.3$, Models 2 and 3 are able to support a largest $N$ of 7 and 9, respectively, compared to the largest $N$ of 6 in the case of Model 1.

It ought to be highlighted that although Model 3 shows a significant improvement in performance, its main purpose is to serve as an upper bound. In a practical setting, a model with a sector angle such as $\phi = \pi/2$ (i.e., "Model 2.5"), as shown in Fig. 1, may provide a very good trade off between having enough UEs available for NOMA and the interference conditions.

## VI. Conclusion

In this paper a large cellular network that employs NOMA in the downlink is studied. As NOMA requires ordering of the UEs based on some measure of link quality, two kinds of UE ordering techniques are analyzed and compared: 1) MSP-based ordering, 2) ISINR-based ordering. An SINR analysis that takes into account the SIC chain and RI from imperfections in SIC is developed for each ordering technique. We show that neither ordering technique is consistently superior to the other and highlight scenarios where each outperforms the other. Two non-convex problems of maximizing the cell sum rate $\mathcal{R}_{\text{tot}}$ subject to a constraint are formulated: a TMR constraint $\mathcal{T}$ in $\mathcal{P}1$ and the symmetric rate constraint in $\mathcal{P}2$. Since the optimum solution for RA to solve each problem requires an exhaustive search, two efficient algorithms for general

NOMA cluster size $N$ that give feasible solutions for intercell interference-aware PA and choice of target rate are proposed. Additionally, the existence of an optimum NOMA cluster size that maximizes $\mathcal{R}_{\text{tot}}$ for each problem is shown. It is observed that $\mathcal{P}1$ a higher $\mathcal{R}_{\text{tot}}$; however, $\mathcal{P}2$ guarantees better individual UE performance. Additionally, it is shown that NOMA outperforms OMA provided $\beta$ is below a certain critical value. The results highlight the importance and impact of choosing network parameters such as $N$ and the UE ordering technique, depending on the network objective and $\beta$. Three models to show the impact of UE clustering in NOMA are introduced. The models demonstrate that efficient UE clustering which selects UEs with good interference conditions can improve performance significantly; in fact, with efficient UE clustering the cell sum rate can be doubled.

## REFERENCES


[1] K. S. Ali, H. Elsawy, A. Chaaban, M. Haenggi, and M. S. Alouini, "Analyzing non-orthogonal multiple access (NOMA) in downlink Poisson cellular networks," in *to appear in Proc. of IEEE International Conference on Communications (ICC18)*, 2018.

[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. NJ: John Wiley, 2006.

[3] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2004.

[4] P. Patel and J. Holtzman, "Analysis of a simple successive interference cancellation scheme in a DS/CDMA system," *IEEE J. Selec. Areas Commun.*, vol. 12, no. 5, pp. 796–807, June 1994.

[5] S. Vanka, S. Srinivasa, Z. Gong, P. Vizi, K. Stamatiou, and M. Haenggi, "Superposition coding strategies: Design and experimental evaluation," *IEEE Trans. Wireless Commun.*, vol. 11, no. 7, pp. 2628–2639, July 2012.

[6] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for cellular future radio access," in *Proc. of IEEE 77th Vehicular Technology Conference (VTC Spring 2013)*, June 2013, pp. 1–5.

[7] 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 Proc. Letters*, vol. 21, no. 12, pp. 1501–1505, Dec. 2014.

[8] S. Timotheou and I. Krikidis, "Fairness for non-orthogonal multiple access in 5G systems," *IEEE Signal Proc. Letters*, vol. 22, no. 10, pp. 1647–1651, Oct. 2015.

[9] J. Choi, "Power allocation for max-sum rate and max-min rate proportional fairness in NOMA," *IEEE Comm. Letters*, vol. 20, no. 10, pp. 2055–2058, Oct. 2016.

[10] Z. Ding, M. Peng, and H. V. Poor, "Cooperative non-orthogonal multiple access in 5G systems," *IEEE Comm. Letters*, vol. 19, no. 8, pp. 1462–1465, Aug. 2015.

[11] Y. Liu, Z. Ding, M. Elkashlan, and H. V. Poor, "Cooperative non-orthogonal multiple access with simultaneous wireless information and power transfer," *IEEE J. Select. Areas Commun.*, vol. 34, no. 4, pp. 938–953, Apr. 2016.

[12] Y. Liu, Z. Qin, M. Elkashlan, Y. Gao, and A. Nallanathan, "Non-orthogonal multiple access in massive MIMO aided heterogeneous networks," in *Proc. of IEEE Global Communications Conference (GLOBECOM16)*, Dec. 2016.

[13] H. Tabassum, E. Hossain, and M. J. Hossain, "Modeling and analysis of uplink non-orthogonal multiple access (NOMA) in large-scale cellular networks using Poisson cluster processes," *IEEE Trans. Commun.*, vol. 65, no. 8, pp. 3555–3570, Aug. 2017.

[14] Y. Liu, Z. Ding, M. Elkashlan, and J. Yuan, "Nonorthogonal multiple access in large-scale underlay cognitive radio networks," *IEEE Trans. Vehicular Tech.*, vol. 65, no. 12, pp. 10 152–10 157, Dec. 2016.

[15] J. Zhu, J. Wang, Y. Huang, S. He, X. You, and L. Yang, "On optimal power allocation for downlink non-orthogonal multiple access systems," *IEEE J. Selec. Areas Commun.*, vol. 35, no. 12, pp. 2744–2757, Dec. 2017.

[16] C. L. Wang, J. Y. Chen, and Y. J. Chen, "Power allocation for a downlink non-orthogonal multiple access system," *IEEE Wireless Comm. Letters*, vol. 5, no. 5, pp. 532–535, Oct. 2016.

[17] Y. Sun, D. W. K. Ng, Z. Ding, and R. Schober, "Optimal joint power and subcarrier allocation for full-duplex multicarrier non-orthogonal multiple access systems," *IEEE Trans. Commun.*, vol. 65, no. 3, pp. 1077–1091, Mar. 2017.

[18] K. S. Ali, H. Elsawy, A. Chaaban, and M. S. Alouini, "Non-orthogonal multiple access for large-scale 5G networks: Interference aware design," *IEEE Access*, vol. 5, pp. 21 204–21 216, 2017.

[19] B. Blaszczyszyn, M. Haenggi, P. Keeler, and S. Mukherjee, *Stochastic Geometry Analysis of Cellular Networks*. Cambridge University Press, 2018.

[20] J. Andrews, F. Baccelli, and R. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 3122–3134, Nov. 2011.

[21] H. ElSawy, A. Sultan-Salem, M. S. Alouini, and M. Z. Win, "Modeling and analysis of cellular networks using stochastic geometry: A tutorial," *IEEE Commun. Surveys and Tutorials*, vol. 19, no. 1, pp. 167–203, 2017.

[22] W. Lu and M. D. Renzo, "Stochastic geometry modeling of cellular networks: Analysis, simulation and experimental validation," *CoRR*, vol. abs/1506.03857, 2015. [Online]. Available: http://arxiv.org/abs/1506.03857

[23] Z. Zhang, H. Sun, and R. Q. Hu, "Downlink and uplink non-orthogonal multiple access in a dense wireless network," *IEEE J. Selec. Areas Commun.*, vol. 35, no. 12, pp. 2771–2784, Dec. 2017.

[24] Z. Zhang, H. Sun, R. Q. Hu, and Y. Qian, "Stochastic geometry based performance study on 5G non-orthogonal multiple access scheme," in *Proc. of IEEE Global Communications Conference (GLOBECOM16)*, Dec. 2016, pp. 1–6.

[25] A. H. Sakr and E. Hossain, "Location-aware cross-tier coordinated multipoint transmission in two-tier cellular networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 11, pp. 6311–6325, Nov. 2014.

[26] A. H. Sakr, H. ElSawy, and E. Hossain, "Location-aware coordinated multipoint transmission in OFDMA networks," in *Proc. of IEEE International Conference on Communications (ICC14)*, June 2014, pp. 5166–5171.

[27] V. Olsbo, "On the correlation between the volumes of the typical Poisson Voronoi cell and the typical Stienen sphere," *Advances in Applied Probability*, vol. 39, no. 4, pp. 883–892, 2007.

[28] G. Geraci, M. Wildemeersch, and T. Q. S. Quek, "Energy efficiency of distributed signal processing in wireless networks: A cross-layer analysis," *IEEE Trans. Signal Proc.*, vol. 64, no. 4, pp. 1034–1047, Feb. 2016.

[29] M. Wildemeersch, T. Q. S. Quek, M. Kountouris, A. Rabbachin, and C. H. Slump, "Successive interference cancellation in heterogeneous networks," *IEEE Trans. Commun.*, vol. 62, no. 12, pp. 4440–4453, Dec. 2014.

[30] H. A. David, *Order statistics*. NJ: John Wiley, 1970.

[31] J. Choi, "NOMA: Principles and recent results," in *International Symposium on Wireless Communication Systems (ISWCS17)*, Aug. 2017, pp. 349–354.