Stochastic Geometry-based Throughput Analysis of User-Specific Power-Level-Constrained GF-NOMA

Takeshi Hirai, *Member, IEEE*, Yuta Ueda, Naoki Wakamiya, *Member, IEEE*

***Abstract*—This paper proposes a stochastic geometry-based analytical framework for the throughput of the grant-free power-domain non-orthogonal multiple access (GF-NOMA) with user-specific constraints of selectable power levels and analyzes the achievable throughput. Our analytical framework uses stochastic geometry to reflect selectable power levels constrained by the maximum transmission power and channel of each user to an inhomogeneous offered load per level. This key idea enables our framework to analyze the throughput bounded by the geographical user distribution and derive a suitable selection strategy of power levels under the constraint more accurately than the existing models. Our analytical results showed that our framework analyzed the throughput with only an analysis error of 0.1% compared with the Monte Carlo simulations, although the existing model overestimated 58% higher throughput. By using the proposed analytical model, our results presented decreasing the achievable throughput with increasing the coverage range. This paper also proposes a heuristic method based on our proposed analytical model to derive a suitable selection strategy of power levels. Our results highlight that the derived selection strategy on our analytical framework achieved 20% higher throughput than the baseline strategy, where each user randomly selects a power level under the power level constraint.**

***Index Terms*—mMTC, grant-free NOMA, SIC, power collisions, throughput analysis, stochastic geometry.**

I. Introduction

5G/6G uplink networks require realizing the massive machine-type communications (mMTC) [1], [2]. Such an mMTC applications include smart factories and autonomous vehicles [2]. In such a typical mMTC, a massive number of users, such as sensors, connect a base station (BS) in an uplink. This use case has the following three characteristics. First, mMTC requires high throughput to accommodate such a massive number of users. Second, mMTC needs a grant-free (GF) access protocol (e.g., ALOHA) [3], i.e., a decentralized protocol without scheduling spectrum resources by the connected BS, to mitigate signaling overheads caused by such a massive number of users. Third, each user limits the maximum transmission power based on its battery and hardware limitations. In summary, realizing mMTC requires achieving high throughput in GF access under the maximum transmission power of each user.

This paper was produced by Takeshi Hirai, Yuta Ueda, and Naoki Wakamiya. Takeshi Hirai and Naoki Wakamiya are in the Graduate School of Information Science and Technology, Osaka University, 1-5 Yamadaoka, Suita, Osaka, 565-0871 Japan (email: hirai.takeshi.ist@osaka-u.ac.jp, y-ueda@ist.osaka-u.ac.jp, wakamiya.naoki.ist@osaka-u.ac.jp).

Manuscript received April 19, 2021; revised August 16, 2021.

In mMTC, a technology to increase the throughput has been the power-domain non-orthogonal multiple access (NOMA) with *channel inversion* in GF access [4]–[23], called GF-NOMA. In GF-NOMA, each user selects a pre-designed power resource, called a power level, from the set of its selectable power levels. The user calculates its transmission power based on its channel and the target-received power associated with its selected level before transmitting its packet to its connected BS. The BS uses the successive interference cancellation (SIC) technique to receive multiple packets. GF-NOMA designs each power level to satisfy a required signal-to-interference-and-noise ratio (SINR) under no *power collisions* where multiple packets arrive at a level; namely, power collisions happen, depending on the average packet arrival rate or offered load per level, defined as the per-level offered load. Such a power collision may cause the following three types of packet errors over power levels due to the non-orthogonality: collision, lower-power-level, and higher-power-level errors [15].

The throughput of GF-NOMA complicatedly depends on the selectable power levels constrained by the property of each user, defined as user-specific power level constraints based on channel inversion. Typically, such a user-specific property is the geographical property, like the location of each user , due to its channel, including small-scale fading and path-loss gains. A faraway user from its connected BS experiences a poorer channel and thus needs its larger transmission power, calculated by channel inversion, to use a power level. This characteristic causes such a user to limit its selectable power levels more strictly under its maximum transmission power. Given a homogeneous user distribution within the coverage, more users contain lower power levels in their selectable ones. In such situations, the existing selection strategy of power levels, like a uniform selection strategy, may cause a more inhomogeneous per-level offered load; namely, a lower power level undergoes a larger offered load. Unfortunately, such a per-level offered load yields power collisions more frequently, and as a result, the achievable throughput becomes lower.

The throughput of GF-NOMA has been actively analyzed in related works [4]–[23]. Some works [4]–[15] analyzed the throughput under the common property over users; in other words, all the users can select all the power levels. In particular, the authors in [8]–[15] discussed the throughput in GF-NOMA with a random selection strategy in the general number of power levels. Other works [16]–[21] focused on the selection strategy reflecting the property of each user, like its location and channel. For example, the authors in [16],

[17] proposed to divide a geographic space into concentric circles associated with power levels; the authors in [18]–[20] evaluated the throughput of the channel-dependent GF-NOMA schemes where each user decides its power level based on its channel. Another work [22] modeled a power level constraint reflecting the maximum transmission power in a channel-dependent GF-NOMA scheme with two levels. Another work [23] supported a power level constraint reflecting the maximum transmission power and path-loss gains.

However, these related works [4]–[23] have despised reflecting the property of each user to its power level constraint in the analytical model, and the throughput of the user-specific power-level-constrained GF-NOMA may have been overestimated. The related works [4]–[15], assuming a homogeneous property between all the users, reflected no properties to power level constraints. The related works [16]–[21], considering the selection strategy using the property of each user, also reflected no properties to power level constraints. The related work [22], considering a user-specific power level constraint, supported GF-NOMA with only two levels and treated the power level constraint as a given parameter; in other words, the power level constraint discussed in [22] was irrelevant to the properties of users, like geographical properties. The other work [23], considering a user-specific power level constraint, supported only path-loss gains without small-scale fading gains to discuss the power level constraint; additionally, this model [23] treated only collision errors and lower-power-level errors rather than higher-power-level errors. Note that this limitation was impacted to analyze the throughput based on the previous work [15]. Additionally, these limitations of the existing models [4]–[23] have derived unsuitable selection strategies for the achievable throughput in the user-specific power-level-constrained GF-NOMA.

To overcome the limitations, this paper proposes an analytical framework for the throughput of GF-NOMA with user-specific power level constraints, defined as the user-specific power-level-constrained GF-NOMA, and analyzes the achievable throughput. Our proposed analytical framework uses stochastic geometry to express the power level constraint of each user caused by its geographical property and maximum transmission power. This key idea enables to formulate an inhomogeneous per-level offered load, i.e., as an input of the throughput, reflecting the power level constraints. Also, this paper derives a suitable selection strategy on the proposed analytical framework to analyze the achievable throughput. To this end, this paper proposes a PSO-based heuristic method combined with a deterministic algorithm minimizing the peak offered load at each level under the constraint. The selection strategy derived from our heuristic method mitigates the impacts of power collisions per level while using selectable power levels more effectively than the typically existing strategies. In summary, this paper has the following contributions and findings;

- This paper proposes a stochastic geometry-based analytical framework for the throughput of the user-specific power-level-constrained GF-NOMA by reflecting the power level constraints to the per-level offered load in Sections IV-A and IV-B based on the baseline system described in Section III.
- This paper emphasizes that the proposed analytical model approximated the throughput more accurately than the existing models without power level constraints yielded by the channel of each user, even in strict power level constraints, in Section V.
- This paper shows that using a larger coverage range causes the user-specific power-level-constrained GF-NOMA to provide lower throughput due to poor channels experienced by users in Section V.
- This paper proposes a PSO-based heuristic method to approximately optimize the selection strategy of selectable power levels in the user-specific power-level-constrained GF-NOMA in Sections VI-A and VI-B.
- This paper highlights that the suitable selection strategy by the proposed model mixes using higher power levels actively and mitigating the impacts of power collisions at lower power levels to use as much power domain as possible. Also, this derived strategy outperformed the baseline strategy in the user-specific power-level-constrained GF-NOMA in Section VII.

II. Related Works

This section highlights the novelty of our work compared with the related works [4]–[23], summarized in Table I. The related works are divided into three groups to support user-specific properties and reflect the properties to power level constraints. The following subsections briefly describe the difference between our work and these related works.

The related works in the first group [4]–[15] supported a homogeneous property between users in GF-NOMA; namely, all the users are selectable for any level in a stochastic manner. References [4]–[7] analyzed the throughput of GF-NOMA with two levels. In particular, the two works [6], [7] optimized the selection strategy of power levels by the game theory. The authors in [8], [9] analyzed the throughput of GF-NOMA with any number of levels. The authors in [8], [9] formulated collision errors due to power collisions, and the authors in [10], [11] contained lower-power-level errors in addition to collision errors for an average packet arrival rate. References [12]–[15] formulated all the patterns of packet errors. In particular, the works [14], [15] proposed an accurate analytical model reflecting all the packet error patterns caused by GF-NOMA in a per-level offered load. However, these works have reflected no properties specified by users in the throughput analysis. This drawback has lost the impacts of inhomogeneous per-level offered load caused by power level constraints on the throughput.

The related works in the second group [16]–[21] focused on the property of each user in the selection strategy in GF-

TABLE I
Related Works of GF-NOMA

User property	Power level constraint	Related works
Homogeneous		[4]–[15]
Inhomogeneous	Homogeneous	[16]–[21]
	Partially user-specific	[22], [23]
	User-specific	Only our work

NOMA. The authors in [16], [17] proposed the selection method layering the coverage geographically. This method allows each user to select the power level associated with the circular layer covering its location. The works [18]–[20] discussed the channel-dependent selection strategy where each user selects a power level in the predefined thresholds of its channel gain. The work [20] proposed the game theory to optimize the thresholds for GF-NOMA with two power levels. Also, Reference [21] proposed a self-organized selection method to realize such a channel-dependent strategy. However, these works supported no inhomogeneous power level constraints between users, although selection strategies depend on the properties of users, like their locations and channels. This drawback has unrevealed the throughput of the user-specific power-level-constrained GF-NOMA.

The related work in the third group [22], [23] partially supported the user-specific power level constraint. The work [22] focused on the user-specific power level constraint in a channel-dependent selection strategy in GF-NOMA with two levels. The work evaluated the throughput under the given number of users selecting each of the two levels; then, the work assumed that the users selecting their lower power level did not select the higher one. The work [23] supported only path-loss gains to discuss the power level constraint and treated only collision errors and lower-power-level errors rather than higher-power-level errors. However, these works have two drawbacks. The first one is to unsuitably reflect the property of each user to its power level constraint; the given number of users in the work [22] was not calculated from the geographical property of each user and its maximum transmission power; the work [23] overlooked small-scale fading gains different from each user to simplify derivations. The second one is to unsuitably model higher-power-level errors, which are strongly dominant to the throughput in GF-NOMA as shown in our previous work [15], in three or more power levels. These drawbacks have overlooked the characteristics of the user-specific power-level-constrained GF-NOMA, and also, these works [22], [23] have limited to analyze the achievable throughput using the power domain at maximum.

Unlike these works, this paper models user-specific power level constraints by stochastic geometry. Our analytical framework reflects the geographical property, including small-scale fading gains, and maximum transmission power of each user to an inhomogeneous per-level offered load in GF-NOMA. Some works focused on [24]–[27] discussed stochastic geometry approaches for GF-NOMA, but these system models did not focus on our target NOMA with channel inversion; namely, these works did not reflect the relationship between the maximum transmission power and channels. Note that this paper focused on the power level constraint by channel inversion in GF-NOMA, and thus, other types of GF-NOMA, as shown in [28]–[30], are out of scope. In the next section, we describe the system model for the framework.

III. System Model of Baseline GF-NOMA

This section describes the baseline GF-NOMA system before modeling the user-specific power-level-constrained GF-NOMA. This baseline system is also based on the typical GF-NOMA. This section also describes the analytical throughput model taken from the related work [15] in the baseline GF-NOMA system. This analytical model is a throughput function with an input of the per-level offered load reflecting the impacts of power level constraints, derived in the next section.

A. Baseline GF-NOMA System

This section describes the system model of a typical GF-NOMA system, i.e., the baseline system, as shown in [15], with a BS and its covered users. This model is used in smart factories [2] All the parameters throughout this paper are summarized in Table II. This system has a BS covering the circle range, defined as Ω, with a radius , denoted as R. Each covered user with its packet k, i.e., the active user, is distributed according to an HPPP, denoted as Φ, with an intensity, denoted as Λ , as the first step. Here, this user is located at $\boldsymbol{\omega}_k \in \Phi$. Let us consider a typical channel model for mMTC referred to as the related work [26]. Its channel, denoted as h_k, is formulated as follows for the mathematical tractability:

$$h_k = \frac{\upsilon_k}{\sqrt{1 + \|\boldsymbol{\omega}_k\|^{\nu}}}, \tag{1}$$

where υ_k is an independent and identically distributed Rayleigh fading channel coefficient, i.e., $\upsilon_k \sim \mathcal{CN}(0,1)$, as a small scale-fading coefficient, ν is the path-loss exponent Here, r_k is the distance between itself and the BS. The power of a channel h, i.e., $|h|^2$, follows the following cumulative density function (CDF) based on the Gaussian–Chebyshev quadrature , as shown in the work [26]:

$$\begin{aligned} F_{|h|^2}(z) &= \frac{2}{R^2}\int_0^R \left[1 - e^{-(1+r^{\nu})z}\right] r\, dr \\ &\approx \frac{1}{2}\sum_{m=1}^{M} \frac{\pi}{M}\sqrt{1-\alpha_m^2}\,(1+\alpha_m)\left(1 - e^{-\delta_m z}\right), \end{aligned} \tag{2}$$

where α_m and δ_m are written as follows:

$$\alpha_m = \cos\left(\frac{2m-1}{2M}\pi\right), \quad \delta_m = 1 + \left(\frac{R}{2} + \frac{R}{2}\alpha_m\right)^{\nu}, \tag{3}$$

where M is a parameter to control the trade-off between complexity and accuracy.

Each active user in Φ immediately transmits its packet k at the time-frequency resource, considering its maximum transmission power, denoted as $p_k^{(+)}$. Generally, each user has its hardware limitations, and thus, the limitations decide $p_k^{(+)}$. Here, $p_k^{(+)}$ follows the probability density function (PDF), denoted as $f_p(\cdot)$. Each time-frequency resource has L available power levels depending on the maximum transmission power in all the users, denoted as $p^{(+)}$, as follows:

$$p^{(+)} = \max_{\boldsymbol{\omega}_k \in \Phi} p_k^{(+)}. \tag{4}$$

A level $\ell \in \mathcal{L} = [1, L]$ is associated with a target-received power value at the BS, denoted as P_ℓ, as follows:

$$P_\ell = \Gamma(\Gamma+1)^{L-\ell} N_0, \tag{5}$$

TABLE II
NOTATIONS THROUGHOUT THIS PAPER

Notation	Definition
$\boldsymbol{\omega}_k$	Location of a user with a packet k
r_k	Distance between a user with k and the connected BS
h_k	Channel coefficient experienced by a user with k
$p^{(+)}$	Maximum transmission power in all the users
$p_k^{(+)}$	Maximum transmission power of k
L	Maximum number of power levels depending on $p^{(+)}$
$\ell_k^{(+)}$	Selectable and highest power level of k, depending on $p_k^{(+)}$
$\mathcal{L}$, $\mathcal{L}_k$	Set of selectable power levels and the set of k
P_ℓ	Target-received power value of a level ℓ
p_k	Transmission power of k
Γ	Required SINR
η	Additive white Gaussian noise (AWGN) with the power density N_0
λ, λ_ℓ	Offered load (or average packet arrival rate) and offered load at ℓ
$\boldsymbol{\lambda}$, $\boldsymbol{\lambda}(\cdot)$	Per-level offered load with a element of λ_ℓ and $\sum_\ell \lambda_\ell = 1$ and a function of per-level offered load
K_ℓ	Number of arrived packets at ℓ
$\boldsymbol{\lambda}'$	Per-level offered load where k selects its selectable and highest power level, i.e., $\ell^{(+)}$, whose element has $\lambda'_{\ell^{(+)}}$, and $\sum_{\ell\in\mathcal{L}} \lambda'_\ell = \lambda'$
Ω	Space with a circle range R for distributed users
Φ	Homogeneous PPP (HPPP) for users with an intensity Λ
$\boldsymbol{\Psi}$	Selection probability matrix of power levels
$\Psi_{\ell^{(+)}\ell}$	Probability where a user experiencing the maximum level, i.e., $\ell^{(+)}$ selects a power level ℓ

where N_0 is the spectrum density of AWGN, denoted as η; note that a smaller index in (5) shows a larger target received power value as follows:

$$P_1 > P_2 > \cdots > P_L. \tag{6}$$

To transmit the packet, the user obeys a strategy, discussed in Section IV, to select a power level ℓ from its selectable power levels, denoted as $\mathcal{L}_k$, and then sets its transmission power, denoted as p_k, by using *channel inversion* as follows:

$$p_k = \frac{P_\ell}{|\tilde{h}_k|^2}. \tag{7}$$

where $\tilde{h}_k$ is an estimated h_k. Its available transmission power is constrained by $p_k^{(+)}$; namely, $p_k \le p_k^{(+)}$, and thus, $\mathcal{L}_k$ is constrained by $p_k^{(+)}$, discussed in Section IV.

The BS receives the superposed signal from active users and uses SIC with *capture effects* to decode this signal. The superposed signal after the ith SIC iteration, $i \ge 0$, is formulated as follows:

$$y^{(i)} = \sum_{k\in\mathcal{K}^{(i)}} \sqrt{p_k} h_k s_k + \eta, \tag{8}$$

where η is AWGN with N_0, and $\mathcal{K}^{(i)}$ is the set of remaining packets after the ith SIC iteration. Given a packet k provides the greatest-received power in $\mathcal{K}^{(i)}$ with the BS, the SINR at the ith SIC iteration, denoted as $\gamma_k^{(i)}$, is presented as follows:

$$\gamma_k^{(i)} = \frac{p_k |h_k|^2}{\sum_{\hat{k}\in\mathcal{K}^{(i)}\setminus\{k\}} p_{\hat{k}} |h_{\hat{k}}|^2 + N_0}. \tag{9}$$

Assumed that the BS perfectly estimates all the channels, i.e., $\tilde{h}_k \approx h_k$, as the first step to analyze the key characteristics of GF-NOMA, the SINR is transformed as follows:

$$\gamma_k^{(i)} = \frac{P_{\ell_k}}{\sum_{\hat{k}\in\mathcal{K}^{(i)}\setminus\{k\}} P_{\ell_{\hat{k}}} + N_0}, \tag{10}$$

where ℓ_k is a power level occupied by k. All the packets require the same data rate, defined as c, and thus, the required SINR, i.e., Γ, is presented as follows:

$$\log_2(1+\gamma_k{}^{(i)}) \ge c \Longleftrightarrow \gamma_k{}^{(i)} \ge 2^c - 1 = \Gamma. \tag{11}$$

After the perfect interference cancellation, the remaining superposed signal is presented as follows:

$$\begin{aligned} y^{(i)} &= y^{(i-1)} - \sqrt{p_k}\tilde{h}_k s_k \\ &\approx \sum_{\hat{k}\in\mathcal{K}^{(i)}\setminus\{k\}} \sqrt{p_{\hat{k}}} h_{\hat{k}} s_{\hat{k}} + \eta. \end{aligned} \tag{12}$$

The approximation is based on assuming the perfect SIC, which assumes the perfect channel estimation of h_k, i.e., $\tilde{h}_k \approx h_k$, as the first step, as shown in many related works [15].

B. Throughput and Packet Errors with Per-Level Offered Load

This paper uses the analytical throughput model derived in our previous work [15]. This model is developed for a typical GF-NOMA. This paper focuses on the number of received packets as the normalized throughput, simply called throughput, defined as $T(\boldsymbol{\lambda})$ where $\boldsymbol{\lambda}$ is a given per-level offered load. The throughput of GF-NOMA is simply modeled with the number of packet arrivals at all the power levels involved by a given per-level offered load, as shown in the related work [15]. Let us consider the number of packet arrivals at each level, denoted as K_ℓ. Then, K_ℓ follows a Poisson distribution with an average packet arrival rate (i.e., an offered load) at ℓ, denoted as λ_ℓ. Let us define $q(n;\lambda_\ell)$ as a probability mass function for $K_\ell = n$, following a Poisson distribution, and then, K_ℓ obeys the following expression:

$$\Pr(K_\ell = n) = q(n;\lambda_\ell) = \frac{e^{-\lambda_\ell}\lambda_\ell^n}{n!}. \tag{13}$$

Then, a per-level offered load is defined as follows.

$$\boldsymbol{\lambda} = \begin{pmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_L \end{pmatrix}, \tag{14}$$

where the offered load over all power levels, i.e., λ, is written as $\lambda = \sum_{\ell\in\mathcal{L}} \lambda_\ell$.

Given a per-level offered load, the related work [15] formulated the analytical model for the throughput. This analytical model has the following lower and upper-bounded throughput equations:

$$T^-(\boldsymbol{\lambda}) \le T(\boldsymbol{\lambda}) \le T^+(\boldsymbol{\lambda}), \tag{15}$$

where $T^\pm(\boldsymbol{\lambda})$ is presented as follows:

$$\begin{aligned} T^\pm(\boldsymbol{\lambda}) &= \sum_{\ell=1}^{L} T_\ell^\pm(\boldsymbol{\lambda}) \\ &= \sum_{\ell=1}^{L} \left(1 - \Pr(\mathfrak{H}_\ell^\mp | K_\ell = 1)\right) q(1;\lambda_\ell) \left(1 - \Pr(\mathfrak{L}_\ell | K_\ell = 1)\right). \end{aligned} \tag{16}$$

$$\Pr(\mathfrak{H}_\ell^- | K_\ell = 1) = \sum_{\ell^{(c)}=\ell+1}^{L} \left(\left(\prod_{\hat{\ell}=\ell+1}^{\ell^{(c)}-1} q\left(1; \lambda_{\hat{\ell}}\right) \right) \left(1 - q\left(0; \lambda_{\ell^{(c)}}\right) - q\left(1; \lambda_{\ell^{(c)}}\right)\right) \right). \tag{18}$$

$$\Pr(\mathfrak{H}_\ell^+ | K_\ell = 1) = \left(1 - \prod_{\hat{\ell}=\ell+1}^{L} \sum_{K_{\hat{\ell}}=0}^{\lfloor \Gamma \rfloor} q\left(K_{\hat{\ell}}; \lambda_{\hat{\ell}}\right)\right) + \sum_{\ell^{(c)}=\ell+1}^{L} \left(\left(\prod_{\hat{\ell}=\ell+1}^{\ell^{(c)}-1} q\left(1; \lambda_{\hat{\ell}}\right) \right) \left(\sum_{K_{\hat{\ell}}=2}^{\lfloor \Gamma \rfloor} q\left(K_{\hat{\ell}}; \lambda_{\hat{\ell}}\right) \right) \left(\prod_{\tilde{\ell}=\ell^{(c)}+1}^{L} \sum_{K_{\tilde{\ell}}=0}^{\lfloor \Gamma \rfloor} q\left(K_{\tilde{\ell}}; \lambda_{\tilde{\ell}}\right) \right) \right). \tag{19}$$

where $\mathfrak{L}_\ell$ is the event set of lower-power-level errors at ℓ, and $\mathfrak{H}_\ell^{\mp}$ is the approximated event set of higher-power-level errors at ℓ. The event probability of lower-power-level errors is calculated as follows:

$$\Pr(\mathfrak{L}_\ell | K_\ell = 1) = 1 - \prod_{\hat{\ell}=1}^{\ell-1} \left(q\left(0; \lambda_{\hat{\ell}}\right) + q\left(1; \lambda_{\hat{\ell}}\right) \right). \tag{17}$$

The upper and lower-bounded event probability of higher-power-level errors is written as (18) and (19).

Also, the related work [15] formulated the expected number of packet errors caused by each type. The expected number of collision errors $N_{\mathfrak{C}}$, denoted as $\mathbb{E}[N_{\mathfrak{C}}]$, is presented as follows:

$$\mathbb{E}[N_{\mathfrak{C}}] = \sum_{\ell=1}^{L} (\lambda_\ell - \lambda_\ell e^{-\lambda_\ell}). \tag{20}$$

The expected number of lower-power-level errors $N_{\mathfrak{L}}$, denoted as $\mathbb{E}[N_{\mathfrak{L}}]$, is written by using (17) as follows:

$$\mathbb{E}[N_{\mathfrak{L}}] = \sum_{\ell=1}^{L} \Pr(\mathfrak{L}_\ell | K_\ell = 1) q\left(1; \lambda_\ell\right). \tag{21}$$

The upper and lower-bounded expected number of higher-power-level errors $N_{\mathfrak{H}^\pm}$, denoted as $\mathbb{E}[N_{\mathfrak{H}^\pm}]$, is written by using (18) and (19) as follows:

$$\mathbb{E}[N_{\mathfrak{H}^\pm}] = \sum_{\ell=1}^{L} \Pr(\mathfrak{H}_\ell^{\pm} \mid K_\ell = 1) q\left(1; \lambda_\ell\right). \tag{22}$$

IV. Stochastic Geometry-based Analytical Framework for User-Specific Power-Level-Constrained GF-NOMA

This section describes the proposed analytical model for the throughput in the user-specific power-level-constrained GF-NOMA in the following steps on the baseline system model described in Section III. To this end, our framework reflects the user-specific power-level-constraint to the per-level offered load by stochastic geometry. First, we model the two key parameters in this GF-NOMA: a selection strategy of power levels and a power level constraint. Next, by using stochastic geometry, we formulate a per-level offered load based on these design parameters as an input of the throughput. This section qualitatively discusses the characteristics of the proposed analytical model and the throughput.

A. Modeling User-Specific Power-Level-Constrained Selection Strategy

Each user with a packet k has its power level constraint, reformulated as $\mathcal{L}_k = [\ell_k^{(+)}, L]$, to select its power level. The number of power levels, i.e., L, is constrained by the maximum transmission power $p^{(+)}$ in all the distributed users formulated in (4), to reflect the power domain at maximum, unlike [23]; in other words, the target-received power value with the highest power level in L, i.e., P_1, must be smaller than $p^{(+)}$ under no attenuation based on (5) as follows:

$$L = \max\{\hat{L} \in \mathbb{N} \mid \Gamma(\Gamma+1)^{\hat{L}-1} N_0 \leq p^{(+)}\}. \tag{23}$$

$\ell_k^{(+)}$ is constrained by its channel h_k and its maximum transmission power $p_k^{(+)}$ based on (7) as follows:

$$\ell_k^{(+)} = \min\{\ell \in \mathcal{L} \mid P_\ell \leq p_k^{(+)} |h_k|^2\}. \tag{24}$$

Such a user-specific power level constraint is dominant to the selection strategy of power levels. Note that this paper focuses on a stochastic strategy, which is typical and flexible for application, as shown in [15]. To reflect this constraint to the selection strategy, we assume that a user with $\ell_k^{(+)} = \ell^{(+)}$ follows the following probabilistic strategy, denoted as $\mathbf{\Psi}(\cdot)$:

$$\mathbf{\Psi}\left(\ell^{(+)}\right) = \begin{pmatrix} \psi_{\ell^{(+)}1} & \psi_{\ell^{(+)}2} & \cdots & \psi_{\ell^{(+)}L} \end{pmatrix}, \tag{25}$$

where each element in the above vector denotes a conditional probability where a user with $\ell_k^{(+)} = \ell^{(+)}$ selects a power level $\ell_k = \ell$; namely, this probability is formulated as follows:

$$\psi_{\ell^{(+)}\ell} = \Pr\left(\ell_k = \ell \mid \ell_k^{(+)} = \ell^{(+)}\right), \tag{26}$$

where $\psi_{\ell^{(+)}\ell} = 0$ at $\ell^{(+)} > \ell$ by the power level constraint. The summation of these probabilities is presented as follows:

$$\sum_{\ell \in \mathcal{L}} \psi_{\ell^{(+)}\ell} = 1. \tag{27}$$

The matrix of these probabilities in $\ell^{(+)} \in \mathcal{L}$, denoted as $\mathbf{\Psi}$, is presented as follows:

$$\mathbf{\Psi} = \begin{pmatrix} \psi_{11} & \psi_{12} & \cdots & \psi_{1L} \\ \psi_{21} & \psi_{22} & \cdots & \psi_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{L1} & \psi_{L2} & \cdots & \psi_{LL} \end{pmatrix}, \tag{28}$$

and also, the vector of probabilities to select ℓ is formulated as follows:

$$\mathbf{\Psi}_\ell = \begin{pmatrix} \psi_{1\ell} & \psi_{2\ell} & \cdots & \psi_{L\ell} \end{pmatrix}^\top. \tag{29}$$

$$\mathbb{E}_{\Phi}\left[\sum_{\{k|\boldsymbol{\omega}_k\in\Phi\}}\Pr\left(\ell_k=\ell\right)\right]=\mathbb{E}_{\Phi}\left[\sum_{\{k|\boldsymbol{\omega}_k\in\Phi\}}\sum_{\ell^{(+)}\in\mathcal{L}}\Pr\left(\ell_k=\ell\mid\ell_k^{(+)}=\ell^{(+)}\right)\Pr\left(\ell_k^{(+)}=\ell^{(+)}\right)\right]$$
$$=\mathbb{E}_{\Phi}\left[\sum_{\ell^{(+)}\in\mathcal{L}}\psi_{\ell^{(+)}\ell}\sum_{\{k|\boldsymbol{\omega}_k\in\Phi\}}\Pr\left(\ell_k^{(+)}=\ell^{(+)}\right)\right]=\sum_{\ell^{(+)}\in\mathcal{L}}\psi_{\ell^{(+)}\ell}\underbrace{\mathbb{E}_{\Phi}\left[\sum_{\{k|\boldsymbol{\omega}_k\in\Phi\}}\Pr\left(\ell_k^{(+)}=\ell^{(+)}\right)\right]}_{(a)}. \tag{31}$$

$$\mathbb{E}_{\Phi}\left[\sum_{\{k|\boldsymbol{\omega}_k\in\Phi\}}\Pr\left(P_{\ell^{(+)}}\le p_k^{(+)}|h_k|^2<P_{\ell^{(+)}-1}\right)\right]$$
$$=\mathbb{E}_{\Phi}\left[\sum_{\{k|\boldsymbol{\omega}_k\in\Phi\}}\int_0^{p^{(+)}}\Pr\left(P_{\ell^{(+)}}\le p|h_k|^2<P_{\ell^{(+)}-1}\mid p_k^{(+)}=p\right)f_p(p)\,dp\right]$$
$$=\mathbb{E}_{\Phi}\left[\sum_{\{k|\boldsymbol{\omega}_k\in\Phi\}}\int_0^{p^{(+)}}\left(\Pr\left(|h_k|^2\le\frac{P_{\ell^{(+)}-1}}{p}\right)-\Pr\left(|h_k|^2\le\frac{P_{\ell^{(+)}}}{p}\right)\right)f_p(p)\,dp\right]$$
$$\overset{(b)}{=}\Lambda\int_{\boldsymbol{\omega}\in\Omega}\int_0^{p^{(+)}}\left(\Pr\left(\frac{|\upsilon_k|^2}{1+\|\boldsymbol{\omega}\|^{\nu}}\le\frac{P_{\ell^{(+)}-1}}{p}\right)-\Pr\left(\frac{|\upsilon_k|^2}{1+\|\boldsymbol{\omega}\|^{\nu}}\le\frac{P_{\ell^{(+)}}}{p}\right)\right)f_p(p)\,dp\,d\boldsymbol{\omega}$$
$$\overset{(c)}{=}\lambda'\int_0^{p^{(+)}}\frac{2}{R^2}\int_0^{R}\left(\left(1-\exp\left(-\frac{P_{\ell^{(+)}-1}}{p}(1+r^{\nu})\right)\right)-\left(1-\exp\left(-\frac{P_{\ell^{(+)}}}{p}(1+r^{\nu})\right)\right)\right)r\,dr f_p(p)\,dp. \tag{33}$$

B. Stochastic Geometry-based Throughput Analytical Model

The throughput of the target GF-NOMA depends on a per-level offered load reflecting the impacts of the user-specific power level constraints. To this end, each element of a per-level offered load, i.e., λ_ℓ, is characterized by the selection strategy and the expected number of users k experiencing $\ell_k^{(+)}=\ell^{(+)}$, defined as $\lambda'_{\ell^{(+)}}$. Each element of this per-level offered load is formulated as follows.

$$\lambda_\ell=\mathbb{E}\left[\sum_{\{k|\boldsymbol{\omega}_k\in\Phi\}}\mathbb{1}\left(\ell_k=\ell\right)\right]=\mathbb{E}_{\Phi}\left[\sum_{\{k|\boldsymbol{\omega}_k\in\Phi\}}\Pr\left(\ell_k=\ell\right)\right]. \tag{30}$$

This equation is transformed by focusing on user-specific power level constraints and the selection strategy under a power level constraint as (31). Here, the equation of (a) in (31) is $\lambda'_{\ell^{(+)}}$ formulated as follows:

$$\lambda'_{\ell^{(+)}}=\mathbb{E}_{\Phi}\left[\sum_{\{k|\boldsymbol{\omega}_k\in\Phi\}}\Pr\left(\ell_k^{(+)}=\ell^{(+)}\right)\right]$$
$$=\mathbb{E}_{\Phi}\left[\sum_{\{k|\boldsymbol{\omega}_k\in\Phi\}}\Pr\left(P_{\ell^{(+)}}\le p_k^{(+)}|h_k|^2<P_{\ell^{(+)}-1}\right)\right]. \tag{32}$$

Then, this equation is transformed by a conditional probability of $p_k^{(+)}$ by $f_p(\cdot)$ in (33). The transformation of (b) in (33) is based on the Campbell's theorem, and the transformation of (c) is based on $\Lambda_{\ell^{(+)}}=\frac{\lambda'_{\ell^{(+)}}}{\pi R^2}$ and polar coordinate transformation. From (33) and the CDF of $|h_k|^2$ in (2), $\lambda'_{\ell^{(+)}}$ is finally transformed as follows:

$$\lambda'_{\ell^{(+)}}=\lambda'\int_0^{p^{(+)}}\left(F_g\left(\frac{P_{\ell^{(+)}-1}}{p}\right)-F_g\left(\frac{P_{\ell^{(+)}}}{p}\right)\right)f_p(p)\,dp. \tag{34}$$

The throughput is formulated by this derived per-level offered load. From (31), λ_ℓ is reformulated as follows:

$$\lambda_\ell=\sum_{\ell^{(+)}\in\mathcal{L}}\psi_{\ell^{(+)}\ell}\,\lambda'_{\ell^{(+)}}. \tag{35}$$

To summarize the equation, we introduce $\boldsymbol{\lambda}'$ as the following per-level offered load:

$$\boldsymbol{\lambda}'=\begin{pmatrix}\lambda'_1 & \lambda'_2 & \cdots & \lambda'_{\ell^{(+)}} & \cdots & \lambda'_L\end{pmatrix}. \tag{36}$$

Based on (28) and (36), a per-level offered load defined from (35) is formulated as the following function:

$$\boldsymbol{\lambda}(\boldsymbol{\Psi})=\boldsymbol{\lambda}'\,\boldsymbol{\Psi}, \tag{37}$$

where $\lambda=\lambda'$. Finally, the throughput is calculated in (16) as follows:

$$\mathcal{T}^{\pm}(\boldsymbol{\Psi})=T^{\pm}(\boldsymbol{\lambda}(\boldsymbol{\Psi})). \tag{38}$$

Also, the expectations of three types of packet errors are calculated by substituting the derived $\boldsymbol{\lambda}(\boldsymbol{\Psi})$ to (20)–(22). Note that the upper and lower-approximated event probabilities of higher-power-level errors in this paper do not provide the lower and upper-bounded probabilities for the exact probabilities due to approximating the CDF of (2). As a result, $\mathcal{T}^{\pm}$ and $\mathbb{E}[N_{\mathfrak{H}^{\pm}}]$ do not necessarily guarantee the upper and lower-bounded ones, unlike (16), discussed in Section IV-C. Then, in this paper, $\mathcal{T}^{\pm}$ is called the upper and lower-approximated throughput.

C. Analysis Characteristics

1) Analysis Accuracy: Let us discuss the validation of our analytical framework based on the analysis procedure. Our framework has three approximations. First, the analytical expressions approximate modeling the higher-power-level errors [15], and thus, the analysis error increases as the event

TABLE III
ANALYSIS PARAMETERS

Parameters	Values	Parameters	Values
Γ	$4 \approx 6$ dB	N_0	-110 dBm
ν	3	M	50
$p^{(+)}$	20 dBm	R	100, 500, and 1000 m

probability of this error pattern increases. Increasing higher-power-level errors involves increasing analysis errors above the exact throughput. Thus, increasing R or decreasing $p_k^{(+)}$ causes increasing analysis errors, discussed in Section IV-C2. Second, our framework approximates the number of arrived packets as the Poisson distribution model. This approximation is shown to be accurate enough to express the throughput in the analytical expression in the context of the mMTC in related works [15]. Third, our framework to derive the per-level offered load approximates the CDF of the channel in (2). Increasing M involves decreasing the analysis error. Note that this analysis error prevents the proposed model from approximating the lower and upper-bounded throughput.

2) Analytical Throughput: Let us discuss the throughput of the user-specific power-level-constrained GF-NOMA qualitatively. At first, we discuss the relationship between the throughput and power level constraints based on the three types of packet errors due to power collisions in GF-NOMA. A more strict constraint causes lower power levels to experience a larger offered load and thus causes power collisions to happen more frequently. Due to such more frequent power collisions, a lower power level is expected to experience collision errors and higher-power-level errors more frequently; these characteristics involve less frequent lower-power-level errors. Increasing collision errors is expected to depend on the throughput more significantly than decreasing lower-power-level errors, and as a result, the throughput becomes lower.

Next, we discuss the relationship between factors causing the power level constraints and the throughput in GF-NOMA. The power level constraint of each user k, i.e., $\ell_k^{(+)}$ mainly depends on the geographical property and maximum transmission power. The geographical property depends on the coverage range, i.e., R. A smaller R provides a more strict power level constraint for k. Thus, decreasing R involves decreasing the throughput. Also, using smaller maximum transmission power, $p_k^{(+)}$, undergoes a more strict power level constraint for k. Thus, decreasing $p_k^{(+)}$ involves decreasing the throughput, as well as the trend for R.

V. ANALYTICAL RESULTS

In this section, we validated the analytical model and analyzed the fundamental characteristics of the throughput of the user-specific power-level-constrained GF-NOMA.

A. Parameters

The parameters used in the analysis are summarized in Table III. ν was set to 3 based on the work [26], and M was set to 50 larger than the work [26] to provide more accurate results as possible. This paper presented the throughput for the coverage range, i.e., R, which is one of the key parameters in

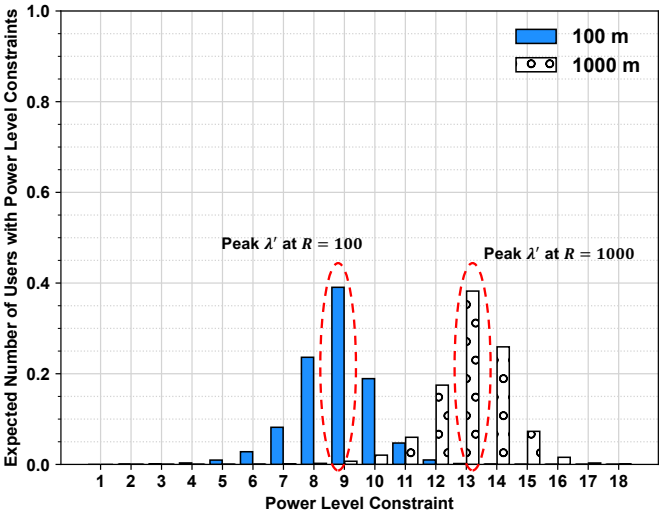


Fig. 1. The expected number of users k where $\ell_k^{(+)} = \ell^{(+)}$ or users with power level constraints at $R = 100, 1000$ at $\lambda = 1$.

the target GF-NOMA, as discussed in Section IV-C. Note that the maximum transmission power also shows similar trends to R, as discussed in Section IV-C. As the first step, the maximum transmission power was common over all the users; then, $p_k^{(+)} = p^{(+)} = 20$ dBm as shown in the related work [25], and namely, this probability mass function (PMF) is written as follows:

$$f_p\left(p_k^{(+)}\right) = \begin{cases} 1 & p_k^{(+)} = 20 \text{ dBm} \\ 0 & \text{otherwise}, \end{cases} \tag{39}$$

where the maximum transmission power provided $L = 18$. As typical scenarios, we validated the analytical model in the coverage ranges of 100, 500, and 1000 m; in particular, based on the above parameters, offered loads where each user selects the highest power level , i.e., $\boldsymbol{\lambda}'$, at $\lambda' = 1$ at $R = 100, 1000$ are shown in Fig. 1. The horizontal axis is the power level constraint, i.e., $\ell^{(+)}$, and the vertical axis is the offered load. Fig. 1 shows that each range had the peak λ', and each user k had a large enough $\mathcal{L}_k$ to select at least a power level. Also, the peak λ' was shown at $\ell^{(+)} = 9$ at $R = 100$ and $\ell^{(+)} = 13$ at $R = 1000$.

Next, we explain the selection strategy of power levels. As a typical strategy, we used the uniform selection strategy, denoted as $\boldsymbol{\Psi}^{(u)}$, to validate the proposed analytical model and analyze the fundamental throughput. $\boldsymbol{\Psi}^{(u)}$ is written as follows:

$$\boldsymbol{\Psi}^{(u)} = \begin{pmatrix} \frac{1}{L} & \frac{1}{L} & \cdots & \frac{1}{L} \\ 0 & \frac{1}{L-1} & \cdots & \frac{1}{L-1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}. \tag{40}$$

To validate the proposed analytical framework, we compared the throughput obtained by our analytical model and Monte Carlo simulations. In the simulations, arrived packets were distributed according to a binomial point process with λ in 1000 uniformly distributed users in Ω to validate the approximation of the channel model. Its Rayleigh fading gain followed a time-variant exponential distribution with the intensity of 1

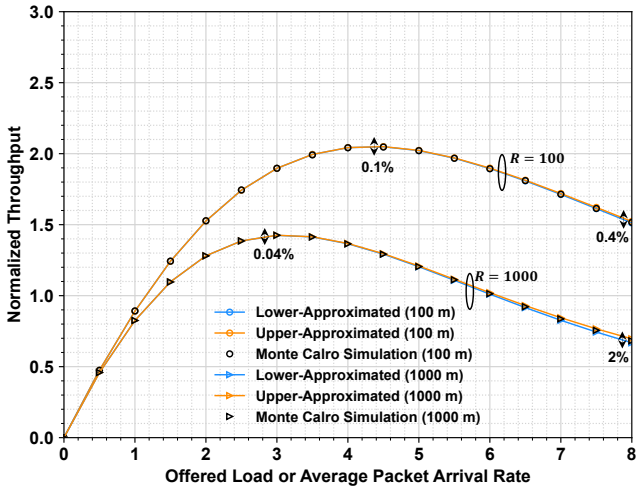


Fig. 2. The normalized throughput of the uniform selection strategy in the user-specific power-level-constrained GF-NOMA at $R = 100, 1000$ with increasing the offered load, i.e., λ.

as the setting of (2), which is a typical model, to validate the approximation of the CDF in (2). Each simulation result was averaged over 10^6 samples. To confirm the impacts of the power level constraint, we evaluated the throughput of the channel-free power-level-constrained GF-NOMA as the typically existing model [15], which expresses the general number of power levels unlike [22]. This channel-free power-level-constrained GF-NOMA allows all the users to uniformly select a level from $\mathcal{L}$ independently from their channels. The matrix of the selection probability, denoted as $\mathbf{\Psi}^{(u+)}$, is written as follows:

$$\mathbf{\Psi}^{(u+)} = \frac{1}{L}\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}. \tag{41}$$

B. Analysis Accuracy

This section shows the analytical throughput in the proposed method with the uniform selection strategy, i.e., $\mathbf{\Psi}^{(u)}$, and $\mathbf{\Psi}^{(u+)}$ to discuss the fundamental characteristics. Fig. 2 shows the normalized throughput of the uniform selection strategy in the user-specific power-level-constrained GF-NOMA with increasing the offered load, i.e., λ, at the coverage range, i.e., R, of 100 and 1000 m. The horizontal axis is λ, and the vertical axis is the normalized throughput. Also, Fig. 3 shows the three types of packet errors due to power collisions, formulated in (20), (21), and (22), in $R = 100, 1000$ with increasing λ. The horizontal axis is λ, and the vertical axis is the expectation of the number of packet errors. Fig. 3 (a) and (b) show the results at $R = 100, 1000$, respectively.

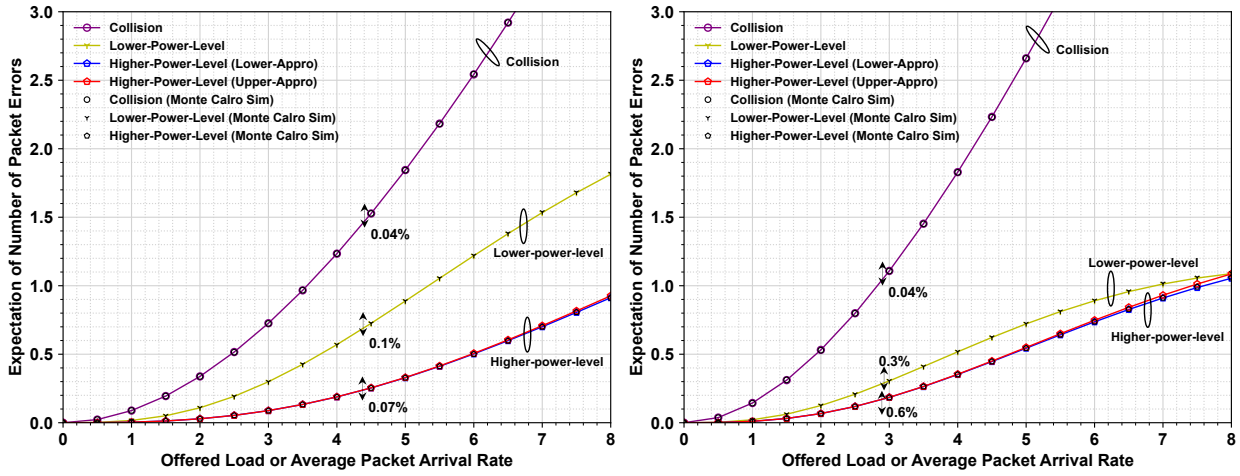
Fig. 2 emphasizes that the proposed analytical model showed enough accurate approximations to follow the throughput trends of the user-specific power-level-constrained GF-NOMA. At $R = 100$, the proposed lower-approximated model, i.e., $\mathcal{T}^-$, had the analysis error of only 0.1% compared with the simulation result at the offered load providing the peak throughput, i.e., $\lambda = 4.5$; then, the proposed upper-approximated one, i.e., $\mathcal{T}^+$, provided the analysis error of only 0.15%, compared with the simulation result. At $R = 1000$, the proposed models also followed the throughput trends obtained by the simulation results; the analysis error was at most 0.04% at $\lambda = 3$, providing the peak throughput at $R = 1000$. As discussed in IV-C, at the same λ, the analysis error increased as R increased due to approximating higher-power-level errors; specifically, the analysis errors were at most 0.4% and 2% at $\lambda = 8$ at $R = 100$ and $R = 1000$, respectively. Fig. 3 shows that our analytical model expressed the expected number of packet errors in each pattern; in particular, at $R = 100$ and $R = 1000$, the expected number of higher-power-level errors was analyzed by the lower-approximated one with only analysis error of 0.07% and 0.6%, compared with the simulation results at $\lambda = 4.5$ and $\lambda = 3$ around the peak throughput, respectively. These results suggested that the lower-approximated model expressed better characteristics in the target GF-NOMA.

C. Throughput and Packet Error Analysis

This section discusses the analytical results of the throughput and packet errors in the user-specific power-level-constrained GF-NOMA. Fig. 4 shows the throughput with increasing the offered load at $R = 100, 500, 1000$ and the channel-free one, i.e., $\mathbf{\Psi}^{(u+)}$, which is the existing model with all the packet error patterns [15]. The axes are the same as Fig. 2. Also, Fig. 5 shows the expected number of packet errors with increasing the offered load. The axes are the same as Fig. 3. Fig. 6 shows the offered load at each level in the uniform selection strategy at $\lambda = 4$ at $R = 100, 500, 1000$. The horizontal axis is the power level, and the vertical axis is the offered load.

Fig. 4 highlights the impacts of the power level constraints on the throughput of the user-specific power-level-constrained GF-NOMA. Fig. 4 shows that increasing the coverage range, i.e., R, monotonically decreased the throughput due to suppressing the number of selectable power levels, as discussed in Section IV-C2. The user-specific power-level-constrained GF-NOMA provided 30% lower peak throughput at $R = 1000$ than $R = 100$; also, this GF-NOMA provided 14% higher peak throughput at $R = 500$ than $R = 1000$. To analyze more deeply, we compared the packet errors and offered loads at $R = 100$ with $R = 1000$. Fig. 5 shows that using $R = 100$ involved 33% fewer collision errors and 46% fewer higher-power-level errors, calculating lower-approximated one, but 10% more lower-power-level errors than using $R = 1000$ at $\lambda = 4$, as discussed in Section IV-C. Fig. 6 highlighted that a lower power level experienced a larger offered load, compared with $R = 100$ and 1000 due to more severe constraints, as discussed in Section IV-C; in particular, the 14th to 18th highest power levels experienced larger offered loads at $R = 100$ than $R = 1000$.

Also, from Fig. 4, the proposed model expressed the throughput reflected by the geographical property more remarkably than the existing model [15]. Our proposed model at $R = 100$ provided 28% lower peak throughput than the existing model, i.e., $\mathbf{\Psi}^{(u+)}$, due to the constraint, as discussed in Section IV-C; in other words, the existing model overestimated 33% higher throughput than the simulation result at



(a) $R = 100$ (b) $R = 1000$

Fig. 3. The three types of packet errors due to power collisions (collision, lower-power-level, and higher-power-level errors) with increasing λ at $R = 100$ (a) and $R = 1000$ (b).

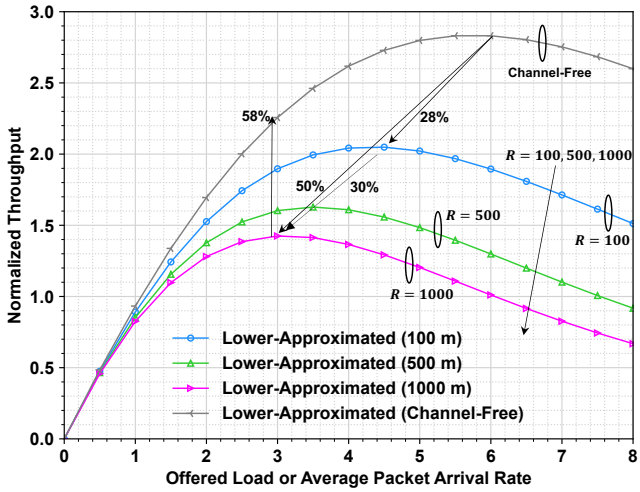


Fig. 4. The lower-approximated throughput of the uniform selection strategy in the user-specific power-level-constrained GF-NOMA and channel-free power-level-constrained GF-NOMA at $R = 100, 500, 1000$ with increasing the offered load, i.e., λ.

$\lambda = 4.5$ at $R = 100$. Also, this GF-NOMA at $R = 1000$ provided 50% lower peak throughput than $\mathbf{\Psi}^{(u+)}$; namely, the existing model overestimated 58% higher throughput than the simulation result at $\lambda = 3$ at $R = 1000$.

VI. Maximizing Throughput of User-Specific Power-Level-Constrained GF-NOMA

This section formulates the throughput maximization problem for the selection strategy in the user-specific power-level-constrained GF-NOMA by using the proposed analytical model in Section IV. To this end, we propose a PSO-based heuristic method to solve this problem based on our analytical model.

A. Throughput Maximization Problem for Selection Strategy

This section describes the throughput maximization problem for the user-specific power-level-constrained GF-NOMA. We use $\mathcal{T}^-$ to approximate the throughput, which is the most accurate analytical model in $\mathcal{T}^{\pm}$ in (16) based on Section V. Given the properties of users, this problem is formulated as follows:

$$\max_{\mathbf{\Psi}} \quad \mathcal{T}^-(\mathbf{\Psi}) \tag{42a}$$

$$\text{s.t.} \quad \psi_{\ell^{(+)}\ell} \geq 0, \qquad {}^\forall\ell,\ {}^\forall\ell^{(+)}, \tag{42b}$$

$$\psi_{\ell^{(+)}\ell} = 0, \qquad {}^\forall\ell^{(+)},\ {}^\forall\ell \leq \ell^{(+)}, \tag{42c}$$

$$\sum_{\ell} \psi_{\ell^{(+)}\ell} = 1, \quad {}^\forall\ell^{(+)}. \tag{42d}$$

This constrained problem is transformed into the following problem with the objective function adding a penalty term to satisfy (42b):

$$\max_{\mathbf{\Psi}} \quad \Upsilon(\mathbf{\Psi}) \tag{43a}$$

$$\text{s.t.} \quad (42c), (42d), \tag{43b}$$

where $\Upsilon(\mathbf{\Psi})$ is the following function:

$$\Upsilon(\mathbf{\Psi}) = \mathcal{T}^-(\mathbf{\Psi}) - w_e \sum_{\hat{\ell}\in\mathcal{L}} \sum_{\hat{\ell}\leq\ell} \max(-\psi_{\hat{\ell}\ell}, 0) - w_e \sum_{\ell\in\mathcal{L}} \max(\lambda_\ell(\mathbf{\Psi}) - 1, -\lambda_\ell(\mathbf{\Psi}), 0), \tag{44}$$

where w_e is the weight of the penalty term.

B. PSO-based Heuristic Method for Selection Strategy

This section describes a PSO-based heuristic method to solve the aforementioned problem in our analytical framework. This problem is too complex to be solved directly. To this end, we aim to approximately solve this problem by using PSO, a basic meta-heuristic technique adopted in [15], combined with a deterministic algorithm, defined as *rumbling algorithm*, inspired by [21]. Our heuristic method first discovers a sub-optimal and near-uniform per-level offered load, denoted as $\tilde{\mathbf{\Psi}}^*$, by the rumbling algorithm. Then, this method adopts $\tilde{\mathbf{\Psi}}^*$ obtained by the rumbling algorithm as one of the initial solutions in PSO and discovers a more near-optimal $\mathbf{\Psi}$, denoted as $\mathbf{\Psi}^*$, than $\tilde{\mathbf{\Psi}}^*$, referred to [15]. The

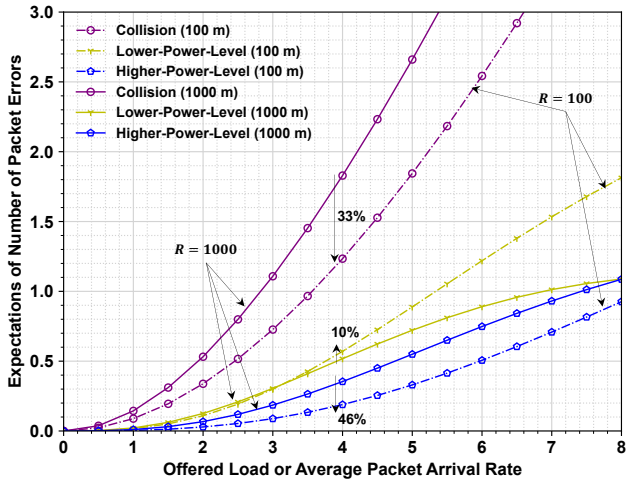


Fig. 5. The expected number of packet errors due to power collisions with increasing the offered load, i.e., λ at $R = 100, 1000$.

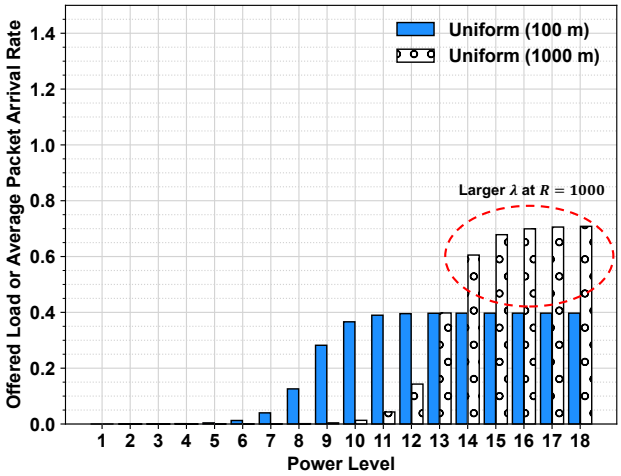


Fig. 6. The offered load at each level, i.e., λ_ℓ, at $\lambda = 4$ at $R = 100, 1000$.

derived strategy, i.e., $\mathbf{\Psi}^*$, is expected to balance the highest-power-level selection strategy at higher power levels and an inhomogeneous selection strategy at lower power levels.

1) Rumbling Algorithm: At first, this section discusses $\tilde{\mathbf{\Psi}}^*$ found by this algorithm under the power level constraints. As a basic approach, this algorithm aims to minimize the frequency of power collisions and thus minimize the offered load at each level. To this end, $\tilde{\mathbf{\Psi}}^*$ uses as many levels as possible and decreases the peak offered load at each level as follows:

- First, $\tilde{\mathbf{\Psi}}^*$ aims to use the power domain at maximum to minimize the potential peak offered load at each level. Each user selects its selectable and highest power level. The selection strategy, called the highest-power-level selection strategy, is written as follows:

$$\mathbf{\Psi}^{(h)} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}. \tag{45}$$

- Second, $\tilde{\mathbf{\Psi}}^*$ aims to offload a larger offered load at ℓ, i.e., λ_ℓ, than other levels to lower power levels experiencing smaller offered loads than ℓ, supporting the power level constraint to minimize the peak offered load at each level. This offloading aims for a homogeneous per-level offered load, achieving a near-optimal solution [15] to reduce causing power collisions rather than the impacts of power collisions.

Based on the above discussions, the proposed algorithm initializes $\mathbf{\Psi}$ to $\mathbf{\Psi}^{(h)}$ and offloads the peak offered load under the power level constraint. By this initialization, to satisfy the power level constraint, this algorithm only transfers a selection probability of ℓ to $\hat{\ell} > \ell$. The detailed algorithm is written in Algorithm 1. This algorithm repeatedly uniformizes the offered load at λ_ℓ and $\lambda_{\ell+1}$ by operating $\psi_{\ell^{(+)}\ell}$ at $\ell > \ell^{(+)}$ in the lines 10–14 until satisfying $\lambda_\ell - \lambda_{\ell+1} \leq \epsilon$ at all the levels in the lines 5–18, judged by a flag variable, denoted as θ in the line 19; here, ϵ is a constant value small enough to avoid the loss of digits. Note that this algorithm is similar to the algorithm in the related work [23], but our algorithm is more general in terms of discussing the offered load rather than the number of users in the related work [23]; specifically, our algorithm can support heterogeneous traffic demand per user.

Algorithm 1 Rumbling Algorithm

1: **Input**: L, $\mathbf{\Psi}^{(h)}$, $\boldsymbol{\lambda}'$, ϵ
2: **Output**: $\tilde{\mathbf{\Psi}}^*$
3: **Initialize**: $\theta = 0$, $\mathbf{\Psi} = \mathbf{\Psi}^{(h)}$
4: **while** *True* **do**
5: **for all** $\ell \leftarrow 1, \ldots, L-1$ **do**
6: $\boldsymbol{\lambda} = \boldsymbol{\lambda}(\mathbf{\Psi})$
7: **if** $\lambda_\ell - \lambda_{\ell+1} > \epsilon$ **then**
8: $\zeta \leftarrow \frac{\lambda_\ell + \lambda_{\ell+1}}{2}$
9: $\theta = 0$
10: **for all** $i \leftarrow \ell, \ldots, 1$ **do**
11: $\mu \leftarrow \min(\zeta, \lambda'_i \psi_{i\ell})$
12: $\zeta \leftarrow \zeta - \mu$
13: $\psi_{i(\ell+1)} \leftarrow \psi_{i(\ell+1)} + \frac{\mu}{\lambda'_i}, \quad \psi_{i\ell} \leftarrow \psi_{i\ell} - \frac{\mu}{\lambda'_i}$
14: **end for**
15: **else**
16: $\theta \leftarrow \theta + 1$
17: **end if**
18: **end for**
19: **if** $\theta > L$ **then**
20: **break**
21: **end if**
22: **end while**

2) PSO: Our PSO-based heuristic method aims for discovering a more suitable $\mathbf{\Psi}$, i.e., $\mathbf{\Psi}^*$, to achieve the optimal $\boldsymbol{\lambda}$ than $\tilde{\mathbf{\Psi}}^*$, shown in [15]. The optimal $\boldsymbol{\lambda}$ reduces the impacts of packet errors based on the level causing power collisions; specifically, the highest power level experiences a larger offered load than the second-highest one, and otherwise, a lower power level undergoes a larger offered load. The optimal $\boldsymbol{\lambda}$ achieves slightly higher throughput than a homogeneous one based on [15]. Thus, the derived $\mathbf{\Psi}$ is expected to show similar throughput to $\tilde{\mathbf{\Psi}}^*$, a near-homogeneous $\boldsymbol{\lambda}$ under power level constraints.

To discover $\boldsymbol{\Psi}^*$, the proposed PSO-based framework initially contains a particle indicating $\tilde{\boldsymbol{\Psi}}^*$ in the set of particles, denoted as $\mathfrak{Z}$. Each particle $z \in \mathfrak{Z}$ on a position at a time step t, denoted as $\boldsymbol{x}_z^{(t)}$, explores a position experiencing a higher fitness value than the past positions of itself and other particles. Its position is updated by using its velocity, denoted as $\boldsymbol{v}_z^{(t)}$, as follows:

$$\boldsymbol{x}_z^{(t+1)} = \boldsymbol{x}_z^{(t)} + \boldsymbol{v}_z^{(t+1)}, \tag{46}$$

where this velocity is defined as follows:

$$\begin{aligned}\boldsymbol{v}_z^{(t+1)} = w_m \boldsymbol{v}_z^{(t)} + \boldsymbol{w}_p^{(t)} \odot \left(\boldsymbol{x}_{z,p}^{(t)} - \boldsymbol{x}_z^{(t)}\right) \\ + \boldsymbol{w}_g^{(t)} \odot \left(\boldsymbol{x}_g^{(t)} - \boldsymbol{x}_z^{(t)}\right),\end{aligned} \tag{47}$$

where w_m is a constant value, and $\boldsymbol{w}_p^{(t)} = [0, w_p]^{L\times L}$ and $\boldsymbol{w}_g^{(t)} = [0, w_g]^{L\times L}$ are matrices with random values from zero to constant values, defined w_p and w_g, respectively. Also, $\boldsymbol{x}_{z,p}^{(t)}$ is the personally best position in a particle z until t, and $\boldsymbol{x}_g^{(t)}$ is the globally best position in all the particles until t. These calculations are repeated until the number of time steps.

To independently calculate each element in $\boldsymbol{\Psi}$ while satisfying the condition of (42d), our framework designs these positions of particles as ratios. Let us define that a function, denoted as $\boldsymbol{\rho}(\boldsymbol{x})$, transforms a ratio matrix $\boldsymbol{x} = (x_{ij})_{L\times L}$ to a probability matrix $\boldsymbol{\Psi}$ as follows:

$$\begin{aligned}\boldsymbol{\Psi} &= \boldsymbol{\rho}(\boldsymbol{x}) \\ &= \begin{pmatrix} \frac{x_{11}}{\sum_{\hat{\ell}=1}^{L} x_{1\hat{\ell}}} & \frac{x_{12}}{\sum_{\hat{\ell}=1}^{L} x_{1\hat{\ell}}} & \cdots & \frac{x_{1L}}{\sum_{\hat{\ell}=1}^{L} x_{1\hat{\ell}}} \\ 0 & \frac{x_{22}}{\sum_{\hat{\ell}=1}^{L} x_{2\hat{\ell}}} & \cdots & \frac{x_{2L}}{\sum_{\hat{\ell}=1}^{L} x_{2\hat{\ell}}} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{x_{\ell(+)_L}}{\sum_{\hat{\ell}=1}^{L} x_{\ell(+)\hat{\ell}}} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.\end{aligned} \tag{48}$$

By using $\boldsymbol{\rho}(\cdot)$, $\boldsymbol{x}_{z,p}^{(t)}$ and $\boldsymbol{x}_g^{(t)}$ are calculated as follows:

$$\begin{aligned}\boldsymbol{x}_{z,p}^{(t)} &= \underset{\{\boldsymbol{x}_z^{(\hat{t})} | \hat{t} \le t\}}{\operatorname{argmax}} \ \Upsilon\left(\boldsymbol{\rho}\left(\boldsymbol{x}_z^{(\hat{t})}\right)\right), \\ \boldsymbol{x}_g^{(t)} &= \underset{\{\boldsymbol{x}_{z,p}^{(t)} | z \in \mathfrak{Z}\}}{\operatorname{argmax}} \ \Upsilon\left(\boldsymbol{\rho}\left(\boldsymbol{x}_{z,p}^{(t)}\right)\right).\end{aligned} \tag{49}$$

Also, the initial position and velocity of each particle are upper triangular matrices where each element is randomly selected to satisfy the condition of (42c) as follows:

$$\boldsymbol{x}_z^{(0)} \in [0, w_o]^{L\times L}, \quad \boldsymbol{v}_z^{(0)} \in [0, w_v]^{L\times L}, \tag{50}$$

where w_o and w_v are upper bounds of the initial position and velocity of each particle, and $\boldsymbol{x}_0^{(0)} = \tilde{\boldsymbol{\Psi}}^*$ to improve the throughput obtained by $\tilde{\boldsymbol{\Psi}}^*$; namely, given the objective function is smaller than the throughput, the algorithm outputs $\boldsymbol{\Psi}^*$.

C. Characteristics of Derived Selection Strategy

First, we qualitatively discuss the characteristics of the derived strategy by the proposed heuristic method discussed in Section VI-B. Lower-power-level errors provide a bigger impact on the throughput than higher-power-level errors based on [15]. To mitigate this impact while suppressing the occurrence of collision errors, $\boldsymbol{\Psi}^*$ transfers offered loads at higher power levels except for levels experiencing small enough offered loads to lower power levels in $\tilde{\boldsymbol{\Psi}}^*$; namely, a lower power level experiences a larger offered load. This transfer increases as the offered load, i.e., λ, increases, but above a specific λ, this transfer decreases to satisfy the penalty term in (44), and thus, $\boldsymbol{\Psi}^*$ becomes more similar to $\tilde{\boldsymbol{\Psi}}^*$.

Next, we qualitatively compare the derived strategy with the two typical and existing selection strategies. First, we compare $\boldsymbol{\Psi}^*$ and the uniform selection strategy, i.e., $\boldsymbol{\Psi}^{(u)}$. In this uniform one, each user offloads packets to the overall of its selectable power levels, although its selectable and highest power level occupies a small enough offered load to suppress the occurrence of power collisions. This characteristic causes a lower power level to experience a larger offered load monotonically. $\boldsymbol{\Psi}^*$ uses the overall power domain more efficiently and thus is expected to cause fewer collision errors and higher-power-level errors than the uniform one, providing slightly more lower-power-level errors. Thus, the derived strategy is also expected to provide higher throughput than the uniform one.

Also, we compare $\boldsymbol{\Psi}^*$ and a channel-dependent selection strategy extended from [16]–[23] to adopt the power level constraint, defined as a channel-dependent and uniform selection strategy. This strategy allows each user k to select a power level ℓ as follows:

$$\ell = \max\left(\beta_k, \hat{\ell}\right), \tag{51}$$

where $\beta_k \in \mathcal{L}_k$ is a uniformly random value to offload k over selectable power levels, and $\hat{\ell}$ is formulated as follows:

$$\hat{\ell} = \begin{cases} 1 & \tau_1 \le |h_k|^2 \\ \hat{\ell} & \tau_{\hat{\ell}} \le |h_k|^2 < \tau_{\hat{\ell}-1}, \quad {}^\forall\hat{\ell} > 1. \end{cases} \tag{52}$$

Each channel threshold τ_ℓ is calculated from the CDF of channels in (2) by the *golden section search*. Formulating the selection probability matrix of this strategy in the power-level-constrained GF-NOMA is out of the scope of this paper, and thus, this paper uses Monte Carlo simulations to show the throughput of this strategy. A better relationship between τ_ℓ and power level constraints enables this strategy to use the power domain more efficiently; then, this strategy is expected to become a more similar strategy to $\tilde{\boldsymbol{\Psi}}^*$ and also show similar throughput.

VII. Throughput Maximization Results

This section derives a suitable selection strategy of the user-specific power-level-constrained GF-NOMA and highlights the derived strategy outperforms other strategies, including baseline and some typical strategies.

TABLE IV
ANALYSIS AND OPTIMIZATION PARAMETERS

Parameters	Values	Parameters	Values
Γ	$4 \approx 6$ dB	N_0	-110 dBm
ν	3	M	50
$p^{(+)}$	20 dBm	R	1000 m
w_e	100	ϵ	10^{-10}
$\|\mathfrak{Z}\|$	1000	Number of time steps	100
w_m	0.8	w_p, w_g	0.3
w_o	1	w_v	0.3

A. Parameters

The parameters used in the evaluations are summarized in Table IV. The wireless parameters were the same as V-A, i.e., a typical setting. We used $R = 1000$ to improve the throughput lowest in the analyzed ranges in Section V. The parameters for our heuristic method, i.e., rumbling algorithm and PSO, were also summarized in Table III. We adopted these values of PSO parameters to discover the solutions efficiently. ϵ was a small enough value to show the difference between offered loads at two neighboring levels.

To discuss the impact of each selection strategy on the calculation process, we evaluated the following strategies; namely, the highest-power-level selection strategy, i.e., $\mathbf{\Psi}^{(h)}$, the selection strategy applied by only the rumbling algorithm, i.e., $\tilde{\mathbf{\Psi}}^*$, defined as rumbling-only, and the selection strategy applied by only PSO, defined as PSO-only. To discuss the contribution of the rumbling algorithm in PSO, we also evaluated the throughput obtained by the PSO using the uniform selection strategy, i.e., $\mathbf{\Psi}^{(u)}$, as one of the initial particles, labeled PSO + Uniform, instead of $\tilde{\mathbf{\Psi}}^*$. Then, we compared the throughput of the calculation results obtained by the derived selection strategy and the other selection strategies discussed in Section VI-C, i.e., $\mathbf{\Psi}^{(u)}$ and the channel-dependent selection strategy.

B. Achievable Throughput by Suitable Selection Strategy

This section describes the achievable throughput of the user-specific power-level-constrained GF-NOMA at $R = 1000$. Fig. 7 highlights the normalized throughput of the user-specific power-level-constrained GF-NOMA with selection strategies with increasing λ at $R = 1000$. The axes are the same as Fig. 2. Fig. 8 shows the expected number of packet errors at each pattern at $R = 1000$. The axes are the same as Fig. 3. Fig. 9 (a)–(c) show the selection probabilities calculated by only the PSO (labeled as PSO-only), only the rumbling algorithm (labeled as Rumbling-only), and the proposed heuristic method (labeled as PSO + Rumbling), respectively. The x-axes show the power level constraint, the y-axes show the selected power level, and the z-axes show the selection probability. Fig. 10 (a)–(c) show the offered load at each level by the highest-power-level selection strategy, $\tilde{\mathbf{\Psi}}^*$, and $\mathbf{\Psi}^*$, respectively, at $R = 1000$ and $\lambda = 3.5$. The horizontal axes show the power level, and the vertical axes show the offered load at each level.

Fig. 7 shows that the derived strategy, i.e., $\mathbf{\Psi}^*$, achieved 20% higher throughput than the uniform selection strategy; $\mathbf{\Psi}^*$

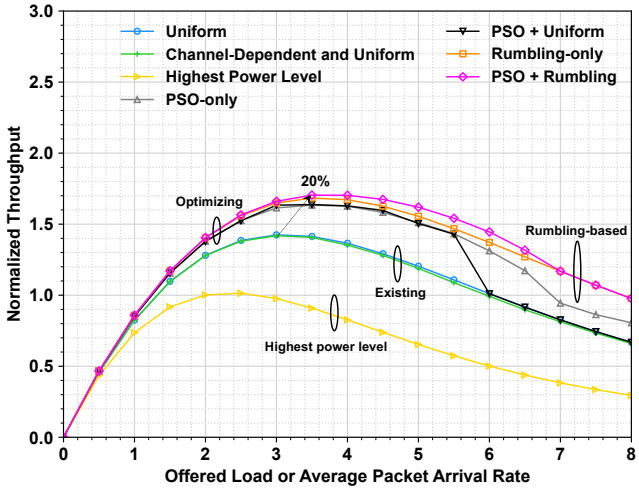


Fig. 7. The normalized throughput of the user-specific power-level-constrained GF-NOMA with each selection strategy at $R = 1000$ with increasing λ.

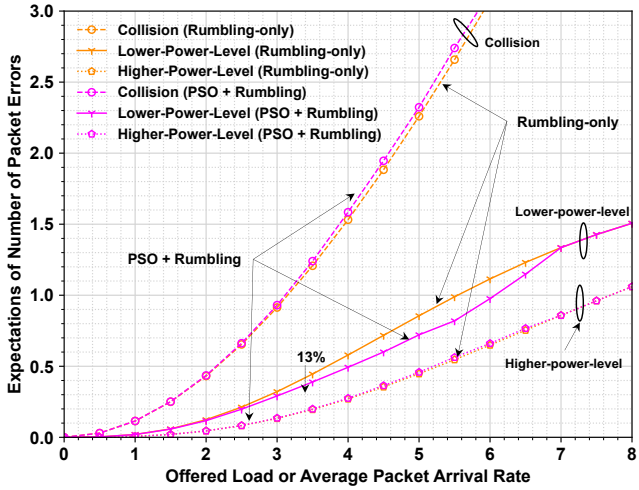


Fig. 8. The expectations of the number of packet errors of the user-specific power-level-constrained GF-NOMA with the derived selection strategy and the uniform one at $R = 1000$ with increasing λ.

provided the peak throughput at $\lambda = 3.5$. Let us discuss each algorithm in the proposed algorithm: the rumbling algorithm and PSO. Then, the rumbling algorithm discovered a near-optimal solution enough for the mixed algorithm; the difference in the throughput between the two algorithms was only 0.18%. The rumbling algorithm and the PSO-only showed 18% and 13% higher peak throughput than the uniform one, respectively. Thus, the rumbling algorithm contributed more significantly than the PSO-only. As discussed in Section VI-C, at $3 \leq \lambda \leq 7$, PSO contributed to increasing the throughput above the rumbling algorithm, but at $\lambda > 7$, PSO showed the same $\mathbf{\Psi}$ as $\mathbf{\Psi}^*$. The PSO-only discovered a similar strategy to PSO with the uniform selection strategy at $\lambda \leq 5$. The proposed algorithm showed 5% higher peak throughput than the PSO-only. Fig. 8 highlights that PSO decreased lower-power-level errors by 13% below $\tilde{\mathbf{\Psi}}^*$. Also, the derived selection strategy by the proposed method achieved 20% higher throughput than the channel-dependent and uniform selection strategy that showed a similar selection strategy to the uniform

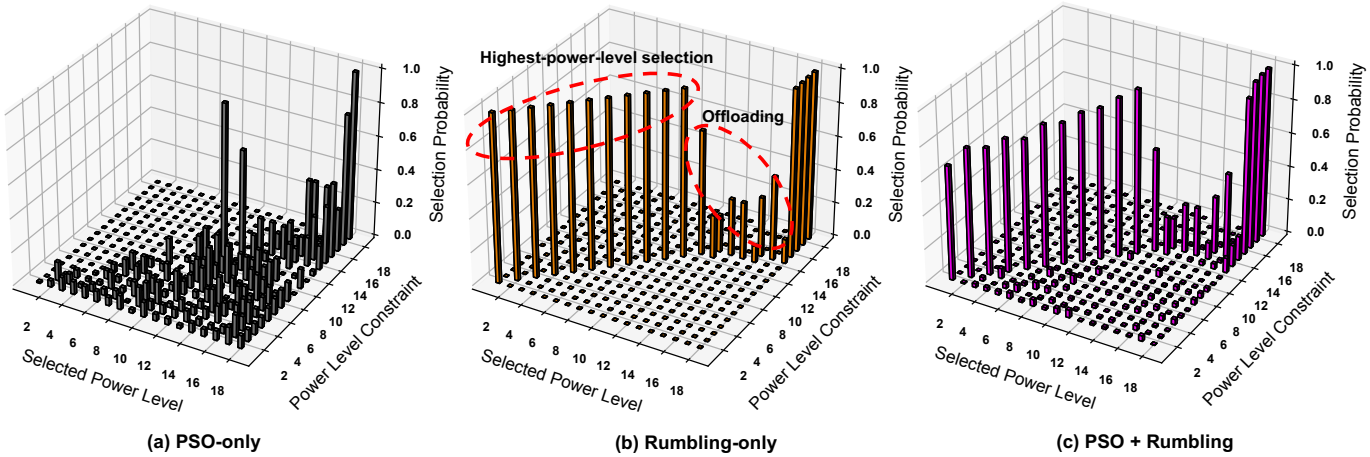


Fig. 9. The selection strategy derived by the PSO-only (a), rumbling algorithm-only (b), and PSO with the rumbling algorithm (c) at $R = 1000$ and $\lambda = 3.5$.

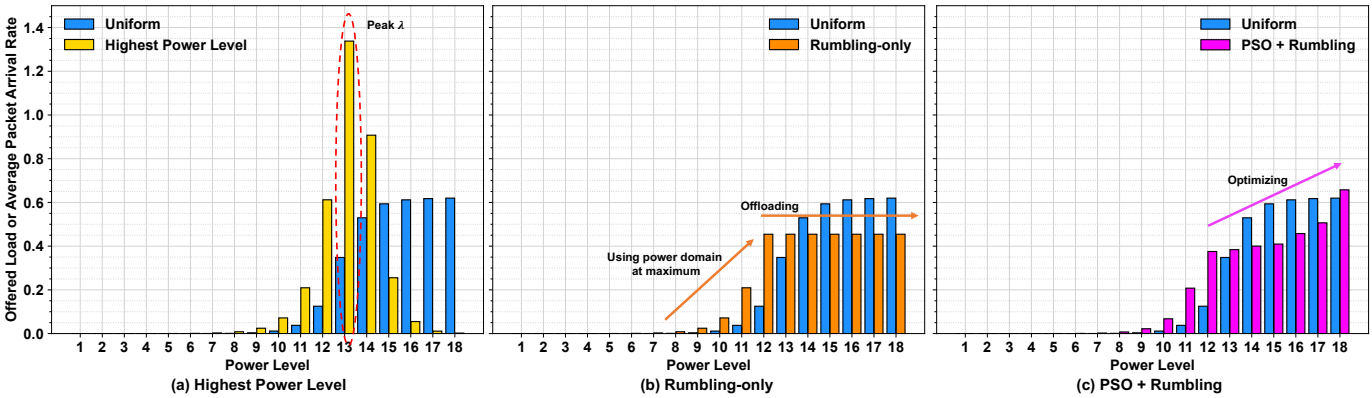


Fig. 10. The offered load at each level by the highest-power-level selection strategy (a), the selection strategy derived by only the rumbling algorithm (b), and the selection strategy derived by PSO with the rumbling algorithm (c) at $R = 1000$ and $\lambda = 3.5$, compared with the uniform selection strategy, respectively.

one.

From Figs. 9 and 10, we confirmed the derived selection strategy by the proposed heuristic method at $R = 1000$ at $\lambda = 3.5$, as discussed in Section IV. Fig. 9 (b) shows that the rumbling algorithm had only the selection probability of the highest power level at $\ell^{(+)} \leq 11$, i.e., $\psi_{\ell^{(+)}\ell^{(+)}} = 1$. The selection probability offloaded the offered load at $12 \leq \ell^{(+)} \leq 14$ experiencing larger λ'_ℓ than other levels as shown in Fig. 1. The selection probabilities of the lowest power level were one at $15 \leq \ell^{(+)} \leq 18$. After the rumbling algorithm, PSO adjusted the offered load, as shown in Fig. 9 (c). The proposed heuristic method (i.e., PSO + Rumbling) used the power domain and offloaded the offered load more efficiently than the PSO-only method in Fig. 9 (a). Also, we confirmed the per-level offered load through the optimization process. As shown in Fig. 10 (a), the highest-power-level selection strategy showed packet arrivals along with $\boldsymbol{\lambda}'$ shown in Fig. 1. Fig. 10 (b) shows that the rumbling algorithm uniformized the offered load at $\ell \geq 12$. Based on Fig. 10 (b) and (c), $\boldsymbol{\Psi}^*$ provided a larger offered load in a lower power level, compared with $\tilde{\boldsymbol{\Psi}}^*$.

VIII. Applications of Selection Strategy to Selection Methods

This section discusses the applications of $\boldsymbol{\Psi}$ to selection methods of power levels. Note that the detailed implementations are mainly out of scope in this paper. We can consider the following two applications. The first application is to apply the derived $\boldsymbol{\Psi}$ to a probabilistic selection method. In this application, given the user-specific property (i.e., user distribution, channel distribution, and the maximum transmission power), the system pre-designs the probability matrix, and each user directly uses the probability to select its power level under its power level constraint. The second application is to apply $\tilde{\boldsymbol{\Psi}}^*$ to the self-organized selection method proposed in [21] without the above information except for $p^{(+)}$.

IX. Conclusion

This paper proposed an analytical framework for the throughput of the user-specific power-level-constrained GF-NOMA in mMTC and derived a suitable selection strategy of power levels based on our framework to increase the achievable throughput. Our analytical framework applies stochastic geometry to calculating an inhomogeneous per-level offered load. This key idea reflects the power level constraint of each user, depending on its maximum transmission power and geographical property, including path-loss and small-scale fading gains, to the per-level offered load more accurately than the existing model. Our analysis results highlighted that our analytical framework provided the throughput accurately enough for the exact throughput presented by the Monte

Carlo simulations as a key insight; the difference was at most 0.1% at the offered load with the peak throughput at the coverage radius of 100 and 1000 m, although the existing model overestimated 58% higher throughput at the coverage radius of 1000 m. These analysis results showed that using a larger coverage radius, i.e., a more strict power level constraint, provided lower throughput as another key insight; specifically, the throughput at 1000 m showed 30% lower throughput than at 100 m. Also, we derived the selection strategy based on our proposed model to maximize the throughput; this derived strategy was to unify the selection probabilities at higher power levels and offload the selection probabilities at lower power levels, respectively. Our results showed that the derived selection strategy achieved 20% higher throughput than the uniform selection one, i.e., the baseline, at the coverage radius of 1000 m.

ACKNOWLEDGMENTS

This paper was supported by JSPS KAKENHI under 23K16868, the Support Center for Advanced Telecommunications Technology Research (SCAT), and the MIC/SCOPE (JP235006102).

REFERENCES

[1] P. Popovski, K. F. Trillingsgaard, O. Simeone, and G. Durisi, "5G wireless network slicing for eMBB, URLLC, and mMTC: A communication-theoretic view," *IEEE Access*, vol. 6, pp. 55 765–55 779, 2018.

[2] S. R. Pokhrel, J. DIng, J. Park, O. S. Park, and J. Choi, "Towards enabling critical mmtc: A review of urllc within mmtc," *IEEE Access*, vol. 8, pp. 131 796–131 813, 2020.

[3] H. Grama Srinath, M. Rana, and N. M. Balasubramanya, "Grant-free access for mmtc: A performance analysis based on number of preambles, repetitions, and retransmissions," *IEEE Internet of Things Journal*, vol. 9, no. 16, pp. 15 169–15 183, 2022.

[4] Z. Chen, Y. Liu, S. Khairy, L. X. Cai, Y. Cheng, and R. Zhang, "Optimizing non-orthogonal multiple access in random access networks," in *IEEE VTC2020-Spring*, 2020, pp. 1–5.

[5] J.-B. Seo, B. C. Jung, and H. Jin, "Online backoff control for NOMA-enabled random access procedure for cellular networks," *IEEE Wireless Communications Letters*, vol. 10, no. 6, pp. 1158–1162, 2021.

[6] J. Choi, "Multichannel NOMA-ALOHA game with fading," *IEEE Transactions on Communications*, vol. 66, no. 10, pp. 4997–5007, 2018.

[7] J. Choi and Y. Ko, "On asymmetric game for NOMA-ALOHA under fading," in *IEEE VTC2022-Spring*, 2022, pp. 1–5.

[8] J. Choi, "NOMA-based random access with multichannel ALOHA," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 12, pp. 2736–2743, 2017.

[9] ——, "Layered non-orthogonal random access with SIC and transmit diversity for reliable transmissions," *IEEE Transactions on Communications*, vol. 66, no. 3, pp. 1262–1272, 2017.

[10] Y. Jin and T.-J. Lee, "Throughput analysis of NOMA-ALOHA," *IEEE Transactions on Mobile Computing*, vol. 21, no. 4, pp. 1463–1475, 2020.

[11] W. Yu, C. H. Foh, A. U. Quddus, Y. Liu, and R. Tafazolli, "Throughput analysis and user barring design for uplink NOMA-enabled random access," *IEEE Transactions on Wireless Communications*, vol. 20, no. 10, pp. 6298–6314, 2021.

[12] J. Choi, "On throughput bounds of NOMA-ALOHA," *IEEE Wireless Communications Letters*, vol. 11, no. 1, pp. 165–168, 2021.

[13] J.-B. Seo, B. C. Jung, and H. Jin, "Performance analysis of NOMA random access," *IEEE Communications Letters*, vol. 22, no. 11, pp. 2242–2245, 2018.

[14] T. Hirai, R. Oda, and N. Wakamiya, "Power level design-aware throughput analysis of grant-free power-domain NOMA in mMTC," in *IEEE GLOBECOM 2022*, 2022, pp. 105–110.

[15] ——, "Power-level-design-aware scalable framework for throughput analysis of GF-NOMA in mMTC," *IEEE IoT Journal*, 2024, (accepted).

[16] H. Jiang, Q. Cui, Y. Gu, X. Qin, X. Zhang, and X. Tao, "Distributed layered grant-free non-orthogonal multiple access for massive MTC," in *IEEE PIMRC 2018*, 2018, pp. 1–7.

[17] J.-B. Seo, S. Pack, and H. Jin, "Uplink NOMA random access for UAV-assisted communications," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 8, pp. 8289–8293, 2019.

[18] J.-B. Seo, S. De, and H. Jin, "Real-time transmission control for multichannel NOMA random access systems," *IEEE Internet of Things Journal*, vol. 10, no. 10, pp. 8984–8995, 2023.

[19] J.-B. Seo, H. Jin, and B. C. Jung, "Multichannel uplink NOMA random access: Selection diversity and bistability," *IEEE Communications Letters*, vol. 23, pp. 1515–1519, 2019.

[20] J. Choi and J. B. Seo, "Evolutionary game for hybrid uplink NOMA with truncated channel inversion power control," *IEEE Transactions on Communications*, vol. 67, no. 12, pp. 8655–8665, 2019.

[21] R. Oda, T. Hirai, and N. Wakamiya, "Energy-efficient path-loss-based self-organized power level selection for GF-NOMA in mMTC," in *IEEE ICCE 2024*, 2024, pp. 1–6.

[22] Z. Chen, Y. Liu, L. X. Cai, Y. Cheng, R. Zhang, and M. Han, "Performance study of random access NOMA with truncated channel inversion power control," *IEEE International Conference on Communications 2021*, 2021.

[23] H.-h. Choi, S. Member, K.-m. Kang, H. Lee, and S. Member, "NOMA-based ALOHA protocol for air-to-ground communications with maximum transmit power limits," *IEEE Internet of Things Journal*, pp. 1–11, 2024.

[24] R. Abbas, M. Shirvanimoghaddam, Y. Li, and B. Vucetic, "A novel analytical framework for massive grant-free NOMA," *IEEE Transactions on Communications*, vol. 67, no. 3, pp. 2436–2449, 2019.

[25] J. Liu, G. Wu, X. Zhang, S. Fang, and S. Li, "Modeling, analysis, and optimization of grant-free NOMA in massive MTC via stochastic geometry," *IEEE Internet of Things Journal*, vol. 8, no. 6, pp. 4389–4402, 2021.

[26] X. Bai and X. Gu, "NOMA Assisted Semi-Grant-Free Scheme for Scheduling Multiple Grant-Free Users," *IEEE Systems Journal*, vol. 17, no. 2, pp. 3294–3305, 2023.

[27] H. Lu, X. Xie, Z. Shi, H. Lei, H. Yang, and J. Cai, "Advanced NOMA Assisted Semi-Grant-Free Transmission Schemes for Randomly Distributed Users," *IEEE Transactions on Wireless Communications*, vol. 22, no. 7, pp. 4638–4653, 2023.

[28] Y. Liu, Y. Deng, H. Zhou, M. Elkashlan, and A. Nallanathan, "Deep reinforcement learning-based grant-free NOMA optimization for mURLLC," *IEEE Transactions on Communications*, vol. 71, no. 3, pp. 1475–1490, 2023.

[29] T. Hirai and P. Spasojevic, "Link-level performance evaluations of sparse code multiple access for PC5-based cellular-V2X with heterogeneous channel estimation errors," in *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*, 2021, pp. 1–5.

[30] A. Celik, "Grant-free NOMA: A low-complexity power control through user clustering," *Sensors*, vol. 23, no. 19, 2023.

Takeshi Hirai was born in Toyama, Japan, in 1994. He received a B.S. in engineering in 2016, an M.S. in information science in 2018, and a Ph.D. in informatics in 2021 from Nagoya University, Nagoya, Japan. He has been an assistant professor at Osaka University, Osaka, Japan, since 2021.

His research interests include wireless communications, wireless LAN systems, cellular networks, V2X, UAV communications, grant-free, non-orthogonal multiple access, IoT, QoS control, congestion control, broadcast communications, MAC layer protocol, driver assistance systems, and autonomous driving.

Dr. Hirai is a member of IEEE and IEICE. He received the Research Fellowship for Young Scientists (DC2) from the Japan Society for the Promotion of Science (JSPS), the Kurata Grant from the Hitachi Global Foundation, the IEICE Young Researcher's Award 2019, the IEICE Tokai Section Student Award 2017, a travel grant from the Telecommunications Advancement Foundation, and publication grants from the Research Foundation for the Electrotechnology of Chubu in 2018 and 2020.

Yuta Ueda received the B.S. in engineering at Osaka University, Japan.

His research interests include wireless communications, wireless LAN systems, cellular networks, and non-orthogonal multiple access.

Naoki Wakamiya (Member, IEEE) received the M.E. and Ph.D. degrees from Osaka University, Osaka, Japan, in 1994 and 1996, respectively. He has been a Research Associate with the Graduate School of Engineering Science and the Educational Center for Information Processing, an Assistant Professor with the Graduate School of Engineering Science, and an Associate Professor with the Graduate School of Information Science and Technology, Osaka University, where he became a Professor in 2011.

His research interests include biologically and brain inspired information and communication technology and self-organizing network control.