Performance Evaluation of Non-Orthogonal Multiple Access in Visible Light Communication

Citation for published version:

Digital Object Identifier (DOI):
10.1109/TCOMM.2016.2612195

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
IEEE Transactions on Communications

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Performance Evaluation of Non-Orthogonal Multiple Access in Visible Light Communication

Liang Yin, Wasiu O. Popoola, Member, IEEE, Xipeng Wu, Member, IEEE, and Harald Haas, Member, IEEE

Abstract—In this paper, the performance of non-orthogonal multiple access (NOMA) is characterized in a downlink visible light communication (VLC) system for two separate scenarios. For the case of guaranteed quality of service (QoS) provisioning, we derive analytical expressions of the system coverage probability and show the existence of optimal power allocation coefficients on two-user paired NOMA. For the case of opportunistic best-effort service provisioning, we formulate a closed-form expression of the ergodic sum rate, which is applicable for arbitrary power allocation strategies. The probability that NOMA achieves higher individual rates than orthogonal multiple access (OMA) is derived. Also, we give an upper bound of the sum rate gain of NOMA over OMA in high signal-to-noise ratio (SNR) regimes. Both simulation and analytical results prove that the performance gain of NOMA over OMA can be further enlarged by pairing users with distinctive channel conditions. We also find out that the choice of light emitting diodes (LEDs) has a significant impact on the system performance. For the case of guaranteed QoS provisioning, LEDs with larger semi-angles have better performance; while for the case of opportunistic best-effort service provisioning, LEDs with 35° semi-angle give nearly optimal performance.

Index Terms—Non-orthogonal multiple access (NOMA), visible light communication (VLC), coverage probability, ergodic sum rate, order statistics.

I. INTRODUCTION

Driven by the ever-increasing penetration of smartphones, tablets and data-hungry applications such as video streaming and cloud computing, wireless data traffic is expected to increase by over a factor of 100, from 3 exabytes in 2010 to over 500 exabytes by 2020 [1]. Therefore, future wireless networks with ever more network capacity and the concept of “Internet of Things” (IoT) consisting billions of miscellaneous devices are of great research interest. Along with many other emerging 5G technologies such as network densification, mmWave and massive multiple-input multiple-output (MIMO), visible light communication (VLC) [2] has also attracted great attention in both academia and industry for supporting next-generation high-speed wireless communication systems. It has been experimentally demonstrated that 100 Gb/s data rate over 5 m free-space links can be achieved in VLC systems with the use of laser diodes (LDs) [3]. Since the visible light spectrum is unlicensed and currently unused for wireless communications, this available spectrum in the order of terahertz (THz) poses a great opportunity for low-cost broadband communication that could effectively alleviate the spectrum congestion currently evident in radio frequency (RF) systems. As visible light does not penetrate through walls, VLC by nature exhibits high-level data security and a high degree of resource reuse: signals sharing the same frequency block but in adjacent rooms do not interfere with each other. Moreover, the absence of electromagnetic interference to existing RF systems makes VLC particularly promising in electromagnetic sensitive areas such as aircraft cabins, hospitals and intrinsically safe environments.

A. Related Work and Motivation

Over the past four cellular generations, shrinking the cell size (or densification) has proven to be an effective method to increase the network capacity. Cell densification can bring numerous benefits, apart from reducing the path loss, the most important being the reuse of partial or whole spectrum resources across a geographic area. Analogous to the femtocell in cellular systems, the concept of optical attocell was first proposed in [4]. A typical room can be composed of multiple optical attocells, each having a radius of around 3 m. In each optical attocell, light emitting diodes (LEDs), which are routinely deployed as luminaries, can be used as the optical transmitter, and photodiodes (PDs) can be used as the optical receiver. In such a system, intensity modulation and direct detection (IM/DD) is generally employed, where the message signal is modulated onto the intensity of the light, and therefore it must be real-valued and non-negative. Also, since the wavelength of visible light is hundreds of nanometers and the detection area of a typical PD is millions of square wavelengths, this spatial diversity essentially prevents the “multipath fading” effect in VLC.

As a wireless broadband technology, it is essential that VLC can efficiently support multiple users with simultaneous network access. Traditional multiple access (MA) techniques include frequency division multiple access (FDMA), time division multiple access (TDMA) and code division multiple access (CDMA). However, these techniques cannot provide sufficient resource reuse. A transmission scheme using discrete-sequence optical code division multiple access (OCDMA) was proposed in [5], where optical orthogonal codes (OOC) are used to encode the data in the time domain by turning the LEDs on and off. As a default extension from orthogonal frequency division multiplexing (OFDM), orthogonal frequency division multiple access (OFDMA) enables resources to be reused at the subcarrier level, and it has been implemented.
in the long-term evolution (LTE) downlink. For the long-term evolution advanced (LTE-A) uplink, a modified form of OFDMA named single carrier frequency division multiple access (SC-FDMA) [6] is used due to it lower peak-to-average power ratio (PAPR). In [7], self-organizing interference management for the OFDMA-based VLC system using busy burst (BB) signaling was evaluated in an aircraft cabin. Compared with static resource partitioning, both fairness and spectral efficiency were improved using the BB approach. By exploiting different color bands in visible light, a color-clustered MA scheme was proposed in [8], where multiple users are allocated into separate color clusters: red, green and blue (RGB). In [9], CDMA is combined with color-shift keying (CSK), where data is transmitted imperceptibly through the variation in the color emitted by RGB LEDs, to allow MA.

To further enhance system capacity and provide enhanced user experience, especially for users at the cell edge, non-orthogonal multiple access (NOMA) has recently been proposed as a promising solution for 5G wireless networks [10], [11]. Different from orthogonal multiple access (OMA) concepts, where users are allocated exclusive time-frequency (TF) resources, NOMA uses power domain multiplexing to realize MA. Specifically, signals for multiple users are superposed in the power domain so that all of the users can share the same TF resources during signal transmission. After the signal is broadcast through the channel, successive interference cancellation (SIC) is carried out at the receiver side for signal detection. As a result, channel state information (CSI) is required. The concept of NOMA can be related to the special case of superposition coding (SC) developed for broadcast channels (BC) [12]. According to the principle of NOMA, users with poor channel conditions are allocated more signal power to decode their own messages while users with good channel conditions are allocated less power but employ SIC to decode and remove the messages intended for other users with poorer channel conditions before decoding their own messages.

The implementation of NOMA in a coordinated two-point RF system has recently been studied in [13], in which the Alamouti code is employed to improve the data rate of the cell-edge user without degrading the performance of the cell-center user. Assuming a fixed power allocation (FPA) strategy, the performance of NOMA has been investigated and analytically characterized for a cellular downlink scenario in terms of the ergodic sum rate and user outage probability [14]. The problem of user pairing on the performance NOMA in RF systems has recently been studied in [15]. From a fairness standpoint, the power allocation problem was investigated in [16] for both CSI and average CSI at the transmitter side. The optimization problems were solved through proposed low-complexity polynomial algorithms. In [17], the application of multiuser beamforming (BF) has been adapted to NOMA with the aim of maximizing system sum capacity. The design of uplink NOMA was proposed in [18], where a low-complexity proportional fair (PF) based scheduling algorithm was developed. In the same work, fractional frequency reuse (FFR) was applied to further enhance the performance of cell-edge users. A cooperative transmission scheme for NOMA has been proposed in [19], where users with better channel conditions are used as relays to improve the reception reliability for users with poorer channel conditions.

Recently, the application of NOMA has also been studied in VLC systems. Due to the fundamental differences between the VLC channel and the RF channel, applying NOMA to VLC requires careful reconsiderations. In [20], a Gain Ratio Power Allocation (GRPA) strategy was proposed for multiuser VLC systems, and its performance was analyzed using a random walk mobility model to simulate the movements of indoor users. It was shown in [20] that the system sum rate can be further enhanced by adaptively tuning the semi-angle of LEDs and the field of view (FOV) of PDs. In [21], the performance of NOMA was compared with OFDMA in a VLC system with illumination constraints, and its superior performance was shown. In this paper, we extend these two studies by providing a theoretical framework for analyzing the performance of NOMA in a VLC system, and characterizing its performance gain over OMA. Moving on, the new framework is used to devise user pairing strategies to improve system throughput. The results confirm the superiority of NOMA over conventional OMA techniques, and this work suggests possible directions to further improve NOMA systems.

### B. Contributions and Outcomes

As NOMA uses power domain multiplexing for MA in the same TF slot, it is spectrum-efficient for the case of densely deployed user devices, for example, indoor VLC. In this paper, we study the performance of NOMA in a downlink VLC system. The main contributions and outcomes of this paper are summarized as follows.

**Characterization of Distribution Functions of the Optical Channel Gain:** Due to the deterministic nature of the VLC channel, we derive the distribution functions of the channel gain for uniformly distributed users. The result is given in a simple and closed form.

**Performance Evaluation of NOMA:** Two case studies are evaluated. For the first case, assuming all of the users have a target quality of service (QoS) requirement, we derive the outage probability for each user. For the second case, where opportunistic best-effort service is provided for different users, we derive a closed-form expression for the achievable ergodic sum rate. For both cases, the effect of choosing LEDs with different semi-angles on the system performance is evaluated. The effect of shadowing on the system performance is also evaluated.

**Comparison between NOMA and OMA:** For arbitrary power allocation strategies, we derive a closed-form expression of the ergodic sum rate gain of NOMA over OMA. For two-user paired NOMA, the probability that NOMA can achieve higher individual rates than OMA is given. The impact of user pairing on the system performance has also been investigated. We find out that the achievable sum rate gain of NOMA over OMA is upper-bounded in high signal-to-noise ratio (SNR) regimes. An expression of this upper bound is derived and given in a closed form.
Impact of NOMA on LED lighting: The illumination quality of LED lighting is analyzed in a VLC system with the application of NOMA.

Key Findings: Unlike OMA, NOMA can achieve a higher sum rate for a larger number of users. On selecting the optimal LEDs, the following observations are made: for the case of guaranteed QoS provisioning, LEDs with a larger semi-angle achieves higher coverage probability; for the case of opportunistic best-effort service provisioning, using the LED with 35° semi-angle gives nearly the highest sum rate. For two-user paired NOMA, pairing users with more distinctive channel conditions can achieve better performance. Also, we find out that there exists an upper bound of the sum rate gain of NOMA over OMA, which proves to be irrelevant to the power allocation coefficients. Provided that the modulated message signals are zero-mean, LED lighting quality is not affected.

C. Paper Organization

The rest of this paper is organized as follows. The system model for the VLC downlink with the implementation of NOMA is introduced in Section II. The performance of NOMA is evaluated in Section III for two different scenarios. In Section IV, the impact of user pairing on the performance gain of NOMA over OMA is studied. Simulation results are presented and discussed in Section V. The impact on NOMA on LED lighting is studied in Section VI. Finally, conclusions are drawn in Section VII.

II. SYSTEM MODEL

As shown in Fig. 1, we consider a downlink transmission model, where the LED is located on the ceiling and $K$ users are uniformly distributed within a circular area. This model is an extension of the point-to-point VLC system to accommodate multiple users. In the following analysis, we assume that the users to be served are static or quasi-static, so that their CSI is not outdated until the next channel estimation. The maximum cell radius is denoted by $r_c$, and the vertical distance from the LED to the receiving plane is denoted by $L$. In a polar coordinate system, the location of the $k$-th user can be represented by $(r_k, \theta_k)$, where $r_k$ represents its horizontal separation from the LED and $\theta_k$ represents its polar angle from the reference axis. Although a complete VLC channel consists not only the line of sight (LOS) link but also the diffuse components caused by light reflections from interior surfaces, previous work has reported that in a typical indoor environment the strongest diffuse component is at least 7 dB (electrical) lower than the weakest LOS component [22]. For these reasons, only the LOS link is considered in the theoretical analysis. For completeness, simulation results based on a complete VLC channel considering the wideband nature of VLC and the shadowing effect are also presented. The LED is assumed to follow a Lambertian radiation pattern whose order is given by $m = -1/\log_2(\cos(\Phi_{1/2}))$, where $\Phi_{1/2}$ denotes the semi-angle of the LED. The PD at each user is assumed to be facing vertically upwards and its FOV is denoted by $\Psi_{fov}$. In the LOS link, the Euclidean distance between the LED and the $k$-th user is denoted by $d_k$. At the $k$-th user, the angle of irradiance and the angle of incidence are denoted by $\phi_k$ and $\psi_k$, respectively.

Without loss of generality, assume that all of the users are ordered based on their channel qualities:

$$h_1 \leq \cdots \leq h_k \leq \cdots \leq h_K,$$

where $h_k$ denotes the direct current (DC) channel gain of the LOS link between the LED and the $k$-th user, given by [23]:

$$h_k = \frac{(m+1)AR_{ip} \cos^m(\phi_k)T(\psi_k)g(\psi_k)\cos(\psi_k)}{2\pi d_k^2},$$

where $A$ denotes the detection area of the PD; $R_{ip}$ denotes the responsivity of the PD; $T(\psi_k)$ represents the gain of the optical filter used at the receiver; and $g(\psi_k)$ represents the gain of the optical concentrator, given by [23]:

$$g(\psi_k) = \begin{cases} \frac{n^2}{\sin^2(\Psi_{fov})}, & 0 \leq \psi_k \leq \Psi_{fov} \\ 0, & \psi_k > \Psi_{fov} \end{cases},$$

where $n$ is the reflective index of the optical concentrator used at the receiver front-end, and it is defined as the ratio of the speed of light in vacuum and the phase velocity of light in the optical material. For visible light the typical values for $n$ are between 1 and 2.

The principle of NOMA is illustrated in Fig. 2, where the bipolar message signals for different users are superposed in the power domain, and a DC bias is added before signal transmission. After going through the optical channel and
with the addition of noise, SIC is carried out at the receiver side (except the 1-st user). According to this principle, the superposed signal to be transmitted at the LED is given by:

\[ x = \sum_{i=1}^{K} a_i \sqrt{P_{\text{elec}} s_i} + I_{\text{DC}}, \]  

(4)

where \( P_{\text{elec}} \) represents the total electrical power of all the message signals; \( I_{\text{DC}} \) is the DC bias added to the LED to ensure the positive instantaneous intensity; \( s_i \) is the modulated message signal intended for the \( i \)-th user; and \( a_i \) represents the power allocation coefficient for the \( i \)-th user. The message signal for each user is assumed to be zero mean with unity variance. Due to the total power constraint, power allocation coefficients should satisfy:

\[ \sum_{i=1}^{K} a_i^2 = 1. \]  

(5)

The optical transmission power of the LED can be calculated as:

\[ P_{\text{opt}} = \alpha \mathbb{E}[x] = \alpha I_{\text{DC}}, \]  

(6)

where \( \alpha \) is the efficiency of the LED and \( \mathbb{E}[] \) denotes the expectation of a random variable. Without loss of generality, it is assumed that \( \alpha = 1 \). In NOMA, users with poorer channel qualities are allocated more signal power. This implies that \( a_1 \geq \cdots \geq a_K \). After removing the DC term, the received signal at the \( k \)-th user is given by:

\[ y_k = \sqrt{P_{\text{elec}} h_k} \left( \sum_{j=1}^{k-1} a_j s_j + a_k s_k + \sum_{i=k+1}^{K} a_i s_i \right) + z_k, \]  

(7)

where \( z_k \) denotes the real-valued Gaussian noise with zero mean and variance \( \sigma_z^2 \). A constant noise power spectral density (PSD), denoted by \( N_0 \), is assumed so that \( \sigma_z^2 = N_0 \), where \( B \) is the signal bandwidth. SIC is carried out at the \( k \)-th user to remove the message signal for the other users with poorer channel conditions (the ‘SIC’ term in Eq. (7)). The message signal for the users whose channel gains are stronger than the \( k \)-th user is treated as noise (the ‘interference’ term in Eq. (7)). As a result, after optical to electrical (O/E) conversion, the achievable data rate per bandwidth, i.e., spectral efficiency, for the \( k \)-th user is given by:

\[ R_k = \left\{ \begin{array}{ll}
\frac{1}{2} \log_2 \left( 1 + \frac{(h_k a_k)^2}{\sum_{j=1}^{K} (h_j a_j)^2 + \frac{1}{2}} \right), & k = 1, \cdots, K - 1, \\
\frac{1}{2} \log_2 \left( 1 + \rho (h_k a_k)^2 \right), & k = K \end{array} \right., \]  

(8)

where \( \rho = P_{\text{elec}}/N_0 B \) represents the transmit SNR, and the scaling factor \( 1/2 \) is due to the Hermitian symmetry. Note that Eq. (8) is conditioned on the event that the \( k \)-th user can successfully detect the message for the \( j \)-th user, for \( \forall j \leq k \). Denote \( R_{k \rightarrow j} \) as the rate for the \( k \)-th user to detect the message intended for the \( j \)-th user, and \( \tilde{R}_j \) as the target data rate for successful message detection at the \( j \)-th user. The above condition can be expressed as:

\[ R_{k \rightarrow j} = \left\{ \frac{1}{2} \log_2 \left( 1 + \frac{(h_k a_k)^2}{\sum_{j=1}^{K} (h_j a_j)^2 + \frac{1}{2}} \right) \geq \tilde{R}_j, \quad j \leq k, j \neq K, \right. \]

\[ \left. \frac{1}{2} \log_2 \left( 1 + \rho (h_k a_k)^2 \right) \geq \tilde{R}_j, \quad j = k = K \right\}. \]  

(9)

If Eq. (9) is satisfied, we assume that perfect SIC can be performed in the decoding chain without signal error propagations.

### III. PERFORMANCE EVALUATION

In this section we present analytical results of the performance analysis of NOMA, which lay the foundation for studying the impact of user pairing in the next section.

#### A. Distribution Functions of the Channel Gain

By substituting \( d_k = \sqrt{r_k^2 + L^2} \), \( \cos(\phi_k) = L/\sqrt{r_k^2 + L^2} \), and \( \cos(\psi_k) = L/\sqrt{r_k^2 + L^2} \) in Eq. (2), the LOS channel gain can be expressed as:

\[ h_k = C(m + 1)L^{m+1}/(r_k^2 + L^2)^{m+3/2}, \]  

(10)

where \( C \) is given by:

\[ C = \frac{1}{2\pi} AR_p T(\psi_k) g(\psi_k). \]  

(11)

Define function \( h = u(r) = C(m + 1)L^{m+1}(r^2 + L^2)^{-m/2} \). It is evident that \( h \) is a monotonic decreasing function with respect to \( r \). Therefore the probability density function (PDF) of the unordered channel gain can be calculated using the “change of variable” method as follows:

\[ f_{h_k}(h) = \left| \frac{\partial}{\partial h} u^{-1}(h) \right| \cdot f_{r_k}(u^{-1}(h)), \]  

(12)

where \( u^{-1} \) denotes the inverse function of \( u \), and \( f_{r_k}(r) = 2r/r_k^2 \) is the PDF of variable \( r_k \) following the uniform distribution. Consequently, the PDF of the unordered variable \( h_k^2 \) can be obtained as:

\[ f_{h_k^2}(t) = \frac{1}{r_k^2} \frac{1}{m + 3} \left( C(m + 1)L^{m+1} \right)^{2/m+3} t^{-m/2-1}, \]  

(13)

for \( t \in [\kappa_{\text{min}}, \kappa_{\text{max}}] \), where \( \kappa_{\text{min}} \) and \( \kappa_{\text{max}} \) are given as \( \kappa_{\text{min}} = (C(m + 1)L^{m+1})^2/(r_k^2 + L^2)^{m+3} \) and \( \kappa_{\text{max}} = (C(m + 1)L^{m+1})^2/L^{2(m+3)} \). Integrating Eq. (13) over \( [\kappa_{\text{min}}, \kappa_{\text{max}}] \), the cumulative distribution function (CDF) of the unordered variable \( h_k^2 \) can therefore be obtained as:

\[ F_{h_k^2}(t) = \frac{1}{r_k^2} \left( C(m + 1)L^{m+1} \right)^{2/m+3} t^{-m/2-1} + \frac{L^2}{r_k^2 + L^2} + 1. \]  

(14)
Using order statistics [24], the PDF of the ordered variable $h_k^2$, denoted by $f_{h_k^2}(t)$, can be readily obtained as:

$$
\begin{align*}
    f_{h_k^2}(t) &= \frac{K!}{(k-1)!(K-k)!} F_{h_k^2}(t)^{k-1} \left(1 - F_{h_k^2}(t)\right)^{K-k} f_{h_k^2}(t) \\
    &= \frac{\Omega}{m + 3(k-1)!(K-k)!} \left(\Omega t - \frac{r_1^2}{r_2^2} + 1\right)^{K-k} (\Omega t - \frac{r_1^2}{r_2^2})^{K-k-1},
\end{align*}
$$

where $\Omega = \frac{1}{\pi} \left(C(m+1) L^{m+1}\right) \frac{2}{m+3}$. Note that Eq. (15) is obtained assuming that the total number of users, $K$, is fixed and known. If $K$ is unknown but assumed to follow a certain point process, for example, the Poisson point process (PPP), then Eq. (15) becomes a conditional PDF on $K$. The final result should be obtained by further averaging the conditional PDF with respect to different realizations of $K$. For non-uniform user distribution, the results can be obtained in a similar way by plugging the specific distribution function into Eq. (12).

### B. Case 1: User Outage Probability for Guaranteed Quality of Service

Consider the case that each user has a target data rate, which is determined by its QoS requirement. In this case, user outage probability is an important metric. Service satisfaction at the $k$-th user requires successful detection of messages not only for this user itself but also for other users with poorer channel conditions. If this constraint is met, the sum rate of the system is simply $\sum_{k=1}^{K} R_k$. Therefore, the sum rate is not of interest in this case. Instead, the analysis is focused on the outage probability at each user. Based on Eq. (9), the outage probability at the $k$-th user can be expressed as:

$$
\begin{align*}
    P_{out}^k &= 1 - P\left[ R_{k\rightarrow j} \geq \tilde{R}_j, \forall j \leq k \right] \\
    &= 1 - P\left[ h_k^2 \geq \varepsilon_j, \forall j \leq k \right],
\end{align*}
$$

where $P[\cdot]$ denotes the probability of an event, and the threshold $\varepsilon_j$ is given by:

$$
    \varepsilon_j = \begin{cases} 
        \frac{\beta_j}{r_2^2}, & j \neq k \\
        \frac{\beta_k}{r_2^2}, & j = k 
    \end{cases}
$$

where $\beta_j = 2\tilde{R}_j - 1$ denotes the required signal-to-noise-plus-interference ratio (SNIR) at the $j$-th user for successful message detection. Note that Eq. (16) is obtained based on the following condition:

$$
    a_j^2 > \frac{\beta_j}{r_2^2} \sum_{i=j+1}^{K} a_i^2.
$$

If power allocation coefficients do not satisfy Eq. (18), user outage probability would always be one. Define a new threshold $\varepsilon_k^* = \min\{\varepsilon_1, \cdots, \varepsilon_k, \kappa_{\min}, \kappa_{\max}\}$. Using order statistics [24], the outage probability of the $k$-th user can be obtained as:

$$
    P_{out}^k = 1 - P\left[ h_k^2 \geq \varepsilon_k^* \right] = \sum_{i=k}^{K} \frac{\Omega}{i!(K-i)!} F_{h_k^2}(\varepsilon_k^*)^i \left(1 - F_{h_k^2}(\varepsilon_k^*)\right)^{K-i}.
$$

System coverage probability is defined as the probability that all of the users in the system can achieve reliable detection, which is given by:

$$
    P_{cov} = \prod_{k=1}^{K} (1 - P_{out}^k),
$$

assuming that the outage event at each user is independent.

### C. Case 2: Ergodic Sum Rate for Opportunistic Best-Effort Service

Consider the case where data rates for different users are opportunistically allocated based on their channel conditions, i.e., $\tilde{R}_j = \tilde{R}_j$. In this scenario it can be readily verified that condition Eq. (9) always holds, and all of the users can be served with zero outage probability but with different data rates. The following theorem gives a closed-form expression of the achievable sum rate.

**Theorem 1.** For arbitrary power allocation strategies, the ergodic sum rate of NOMA with uniformly distributed users is given in Eq. (21), where $w_1(\kappa)$ is defined in Eq. (22), and $w_k(\kappa) = \sum_{i=k}^{K} a_i^2$, $w_{k+1}(\kappa) = \sum_{i=k+1}^{K} a_i^2$ and $2F_1$ denotes the Gauss hypergeometric function. $w_2(\kappa)$ is defined in Eq. (23).

**Proof:** Please refer to the appendix.

Theorem 1 demonstrates that unlike OMA, the capacity of NOMA can be increased with an increase in the total number of users in the system. However, this performance gain is achieved at the cost of the increased computation complexity caused by the SIC process at the receivers.

With the following corollary, the ergodic sum rate gain of NOMA over OFDMA can be obtained.

**Corollary 1.** The ergodic sum rate of an OFDMA-based VLC system with uniformly deployed users is given in Eq. (24), where $w_3(\kappa)$ is defined in Eq. (25).

**Proof:** The ergodic sum rate achieved by OFDMA is calculated as:

$$
    R_{OFDMA} = \sum_{k=1}^{K} \int_{\kappa_{\min}}^{\kappa_{\max}} \frac{1}{2b_k \log_2} \left(1 + \frac{v_k \rho t}{b_k} \right) f_{h_k^2}(t) dt,
$$

where $b_k$ is the fraction of bandwidth occupied by the $k$-th user, and $v_k$ is the fraction of the power allocated to the $k$-th user. The total bandwidth constraint requires that $\sum_{k=1}^{K} b_k = 1$, and the total power constraint requires that $\sum_{k=1}^{K} v_k = 1$. The derivation can follow similar steps as the derivation of $R_{NOMA}$ shown in the appendix.

With Theorem 1 and Corollary 1, the sum rate gain of NOMA over OFDMA can be obtained.
Selecting a subset of users to perform NOMA can effectively reduce the computation complexity of the system. This results in a hybrid MA scheme which consists a combination of NOMA and OMA techniques. As the performance of such a hybrid system is highly dependent on the user selection strategy, in this section we focus on analyzing the effect of user pairing on the system performance. In order to obtain simple but insightful results, the entity of users is divided into groups and each group consists of two users. However, it can be readily extended to the case where an arbitrary number of users are selected and paired to perform NOMA. From a qualitative point of view, the capacity region for NOMA and OFDMA in a two-user scenario is illustrated in Fig. 3. It can be seen that a higher performance gain can be obtained if two users with more distinctive channel conditions are paired to perform NOMA. This finding will be quantitatively validated through theoretical analysis in this section.

Assume that the \(i\)-th user and the \(j\)-th user \((i \leq j)\) in the system are selected to perform NOMA so that \(a_i^2 + a_j^2 = 1\). According to OFDMA, each user is allocated a fraction of orthogonal subcarriers. Therefore the achieved data rate for each user is \(R_k = \frac{1}{2} b_k \log_2 \left(1 + \frac{v_k}{b_k} \rho h_k^2 \right)\), where \(k = \{i, j\}\).
Theorem 2. Assume that the $i$-th and the $j$-th user ($i < j$) are paired together to perform NOMA. A necessary condition for NOMA to achieve higher individual rates than OFDMA ($b_i = b_j$, $v_i = v_j$) is $a_{th1} < a_j < a_{th2}$, where
\[
a_{th1} = \sqrt{-1 + \frac{1 + \rho \kappa_{\text{max}}}{\rho \kappa_{\text{max}}}}, \quad a_{th2} = \sqrt{-1 + \frac{1 + \rho \kappa_{\text{min}}}{\rho \kappa_{\text{min}}}}. \tag{27}
\]

If the above condition is met, the probability that NOMA can achieve higher individual rates than OFDMA is given in Eq. (28), where $\varsigma^* = \min\{\max\{\varsigma, \kappa_{\text{min}}\}, \kappa_{\text{max}}\}$, and $\varsigma = \frac{1 - 2a_j^2}{\rho a_j^3}$.

Proof: Please refer to the appendix.

Theorem 2 demonstrates that given appropriate power allocation coefficients, it is nearly certain for NOMA to outperform OFDMA if two users with highly different channel qualities are paired together.

B. Impact of User Pairing on the Sum Rate

In this subsection, the impact of user pairing on the ergodic sum rate is studied. The following theorem states that the ergodic sum rate gain of NOMA over OFDMA is upper-bounded in high SNR regimes, and this upper-bound remains unchanged for different power allocation strategies.

Theorem 3. Assuming that the $i$-th user and the $j$-th user ($i < j$) are paired to perform NOMA ($a_i^2 + a_j^2 = 1$), as $\rho$ increases, the sum rate gain of NOMA over OFDMA first decreases then increases until it is upper-bounded in high SNR regimes, and this upper bound is given in Eq. (29), where
\[
\varpi_4(\kappa) = -\frac{\kappa^{\frac{p+q+1}{p+q}} (m + 3 + (p + q + 1) \ln(\kappa))}{(p + q + 1)^2}. \tag{30}
\]

Proof: Please refer to the appendix.

Theorem 3 gives the upper bound of the ergodic sum rate gain of NOMA over OFDMA for arbitrary $i$ and $j$ ($i < j$). However, it is of more interest to evaluate how the performance gain varies if $i$ and $j$ change. The following corollary states that the optimum sum rate gain is achieved if two users with the most distinctive channel conditions are paired together to perform NOMA.

Corollary 2. If the $i$-th user and the $j$-th user ($i < j$) are paired to perform NOMA, the sum rate gain of NOMA over OFDMA achieves the maximum by pairing the two users with the most distinctive channel conditions, i.e., $i = 1$ and $j = K$. In high SNR regimes, this maximum gain is upper-bounded by Eq. (31), where
\[
\varpi_5(\kappa) = -\frac{\kappa^{\frac{l+1}{l+1}} (m + 3 + (l + 1) \ln(\kappa))}{(l + 1)^2}. \tag{32}
\]

Proof: Please refer to the appendix.

V. SIMULATION RESULTS

A. Proposed Theoretical Framework

The aim of this section is to substantiate our derived analytical results through Monte Carlo simulations and obtain insights into how the choice of LEDs can affect the system performance. If not otherwise specified, the parameters used for the simulation setup are summarized in Table I.

Table I: Simulation Parameters

<table>
<thead>
<tr>
<th>Parameter name, notation</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vertical separation between the LED and PDS, $L$</td>
<td>2.15 m</td>
</tr>
<tr>
<td>Cell radius, $r_c$</td>
<td>3.6 m</td>
</tr>
<tr>
<td>Total number of users, $K$</td>
<td>10</td>
</tr>
<tr>
<td>LED semi-angle, $\Psi_{1/2}$</td>
<td>60°</td>
</tr>
<tr>
<td>Total signal power, $P_{\text{elec}}$</td>
<td>0.25 W</td>
</tr>
<tr>
<td>PD FOV, $\Psi_{\text{fov}}$</td>
<td>60°</td>
</tr>
<tr>
<td>PD responsivity, $R_p$</td>
<td>0.4 A/W</td>
</tr>
<tr>
<td>PD detection area, $A$</td>
<td>1 cm²</td>
</tr>
<tr>
<td>Reflective index, $n$</td>
<td>1.5</td>
</tr>
<tr>
<td>Optical filter gain, $T$</td>
<td>20 MHz</td>
</tr>
<tr>
<td>Signal bandwidth, $B$</td>
<td>20 MHz</td>
</tr>
<tr>
<td>Noise PSD, $N_0$</td>
<td>$10^{-21}$ A²/Hz</td>
</tr>
</tbody>
</table>

Fig. 4 shows the system coverage probability for different power allocation coefficients in a two-user scenario. Along the following discussions on the two-user scenario, the user closer to the LED is referred to as the near user and the user further away from the LED is referred to as the far user, i.e., $d_{\text{near}} < d_{\text{far}}$. It can be seen that the analytical results are consistent with the simulation results, and there exists an optimum set of power allocation coefficients for achieving the maximum coverage probability. For a low target data rate, the system coverage probability is nearly 100%, given that the power allocation coefficients are optimally chosen. As the target data rate increases, the achievable maximum coverage probability decreases. Also, it can be seen that a larger coverage probability can be achieved by pairing two users with more distinctive channel conditions. An interesting finding is that when the target data rate for both users increases, more signal power should be allocated to the far user in order to achieve the optimal coverage probability.

An exhaustive search (ES) method is used to find the optimum power allocation coefficients. Specifically, a lookup table is formed, in which the system coverage probability is saved for each systematic search of the power coefficients. After this, the optimum pair of power coefficients is found by referring to the lookup table and selecting the one that gives the highest coverage probability. Fig. 5 shows the maximum achievable coverage probability as a function of the target data rate is shown. As expected, the maximum coverage probability would decrease as the target data rate increases. Compared with OFDMA, NOMA is shown to be able to provide a larger coverage probability, and this performance gain can be further enlarged by pairing two users with more distinctive channel conditions. For example, when $i = 1$ and $j = 10$, NOMA can provide 2.2 bpcu data rate for both users with 90% coverage probability while OFDMA can only provide 0.7 bpcu data rate for both users with the same coverage probability.

Fig. 6 demonstrates that, in order to achieve the same
\[
\mathbb{P}[R_i > \bar{R}_i, R_j > \bar{R}_j] = \sum_{l=0}^{K} \sum_{i=1}^{j-1} \sum_{j=1}^{K-i} \left[ \frac{(-1)^{K-i-l+p-q+1} \Omega^{i+l+p}}{(j-i+l-q)(p+q+1)! (K-j-l)! p!(i-1-p)! q!(j-1-i-q)! \left( \frac{L^2}{r_e^2} + 1 \right)^{i-1-p} \left( \frac{L^2}{r_e^2} \right)^{K-j-l}} \times \left( \zeta^2 \frac{L^2}{r_e^2} \right)^{- \frac{p+q+1}{m+1} - \frac{j-i+l-q}{m+1} + \frac{j-1-i-q}{m+1}} - \zeta^2 \frac{L^2}{r_e^2} \right],
\]

(28)

\[
\mathbb{E}\left[ R_i + R_j - \bar{R}_i - \bar{R}_j \right] \leq \frac{1}{2} \left( b_i \log_2 \left( \frac{b_i}{\upsilon_i} \right) + b_j \log_2 \left( \frac{b_j}{\upsilon_j} \right) \right) + \frac{\Omega(1 - b_i)}{2 \ln(2)} \sum_{p=0}^{K-i} \sum_{q=0}^{j-1} \left\{ \frac{K! (1)^{p+K-i-q}}{p!(j-1-p)! q!(K-j-q)! \left( \frac{L^2}{r_e^2} + 1 \right)^{j-1-p} \left( \frac{L^2}{r_e^2} \right)^{K-j-q}} \omega_4(\kappa_{\text{min}}) - \omega_4(\kappa_{\text{max}}) \right\} - \frac{\Omega b_j}{2 \ln(2)} \sum_{p=0}^{K-i} \sum_{q=0}^{j-1} \left\{ \frac{K! (1)^{p+K-i-q}}{p!(j-1-p)! q!(K-j-q)! \left( \frac{L^2}{r_e^2} + 1 \right)^{j-1-p} \left( \frac{L^2}{r_e^2} \right)^{K-j-q}} \omega_4(\kappa_{\text{min}}) - \omega_4(\kappa_{\text{max}}) \right\},
\]

(29)

\[
\mathbb{E}\left[ R_i + R_K - \bar{R}_i - \bar{R}_K \right] \leq \frac{1}{2} \left( b_i \log_2 \left( \frac{b_i}{\upsilon_i} \right) + b_K \log_2 \left( \frac{b_K}{\upsilon_K} \right) \right) + \frac{\Omega}{4 \ln(2)} \sum_{l=0}^{K-1} \left\{ \frac{K!}{l!(K-1-l)!} \Omega^l \left( \omega_5(\kappa_{\text{max}}) - \omega_5(\kappa_{\text{min}}) \right) \left( -1 \right)^l \left( \frac{L^2}{r_e^2} + 1 \right)^{K-1-l} - \left( -1 \right)^l \left( \frac{L^2}{r_e^2} \right)^{K-1-l} \right\},
\]

(31)

![Fig. 4. System coverage probability for different power allocation coefficients: (a) \(i = 1\) and \(j = 2\) (users with similar channel conditions); (b) \(i = 1\) and \(j = 10\) (users with distinctive channel conditions).](image)

QoS requirement at both users, using an LED with a larger semi-angle can provide a higher coverage probability for both NOMA and OFDMA techniques.

In Fig. 7, the maximum coverage probability is computed for different transmit SNR values. It can be seen that with the same transmit power, pairing users with more distinctive channel conditions can achieve a higher coverage probability. Also, it shows that more transmit power is required to achieve higher target data rate for both NOMA and OFDMA.

In Fig. 8, the ergodic sum rate is evaluated for different...
numbers of users and LED semi-angles. In this simulation setup, the data rate for each user is opportunistically allocated in a best-effort manner based on their channel conditions. For the FPA strategy, it is assumed that $a_i = \sqrt{2(K + 1 - i)}/K(K + 1)$. For the GRPA strategy, it is assumed that $a_i = \alpha_{GRPA}^{1-1}$, and $\alpha_{GRPA}$ is a constant. Both strategies ensure that a user with poorer channel quality is allocated more signal power and the total power constraint $\sum_{i=1}^{K_i} a_i^2 = 1$ is also satisfied. For OFDMA, each user is allocated an equal number of subcarriers and transmit power, i.e., $b_1 = \cdots = b_K$, and $u_1 = \cdots = u_K$. It can be seen from Fig. 8 that the proposed analytical framework is a general one that can be applied to arbitrary power strategies.

Fig. 8 shows that NOMA with FPA can achieve a higher sum rate for a larger number of users while the sum rate for OFDMA remains constant for arbitrary number of users. Also, it is interesting to see that the selection of LEDs with different semi-angles can have a significant effect on the system sum rate. For OFDMA, the optimum choice of LED is that with 50° semi-angle. However, for NOMA with FPA, the optimum choice of LED varies for different number of users. For example, when $K = 2$ the optimum choice of LED is that with 45° semi-angle, while when $K = 10$ the optimum choice is the LED with 35° semi-angle.

In Fig. 9, the probability that NOMA can achieve higher individual rates than OFDMA ($\alpha_1 = 9/10$, $\alpha_2 = 1/10$).

It can be seen that the derived theoretical bound shows good consistency with simulation results. Also, it can be seen from Fig. 10 that, as the transmit SNR increases, the sum rate gain of NOMA over OFDMA first decreases and reaches a minimum.
As $\rho$ continues to increase, the sum rate gain increases until it reaches the upper bound. This trend is consistent with Theorem 3. Also, it shows that a higher sum rate gain can be achieved by pairing users with more distinctive channel conditions.

### B. Multipath Reflections and Shadowing Effect

One of the advantages of VLC is that when the LOS is blocked by opaque objects, signal transmission is still possible via the diffuse links but with a lower data rate. Denote the LOS blockage event as $X$, whose probability mass function can be modeled by the Bernoulli distribution, given by:

$$\Pr[X = \chi] = \begin{cases} p, & \chi = 1 \\ 1-p, & \chi = 0 \end{cases},$$

where $p$ is the probability that the LOS is blocked. When $\chi = 0$, the VLC channel consists both the LOS and diffuse components but the dominant part is the LOS link. Theoretical analysis for this case has been presented in previous sections of this paper. When $\chi = 1$, the LOS is blocked with probability of $p$. In this case, signal is transmitted via the diffuse links. Different from narrowband infrared (IR) wireless communication, VLC uses a wide spectrum, which ranges from 380 nm to 780 nm. This means that the wavelength-dependent properties of the power spectral density (PSD) of the LED and the reflectance of indoor reflectors should be well considered when modeling the VLC channel. It has been reported that the received signal power from multipath reflections and the root mean square (RMS) delay spread in VLC are generally smaller than those in IR systems [25]. In the following, the performance of NOMA is studied in a more realistic scenario where multipath reflections and LOS blockage are considered. Specifically, a room of size $5 \times 5 \times 3 \text{ m}^3$ is considered. The LED is located at the ceiling center while receivers are randomly distributed in the room. Monte Carlo simulations are carried out and the system performance is evaluated over 1000 independent trials. The channel impulse response is simulated using the recursive algorithm reported in [25], which includes both the wavelength-dependent white LED characteristics and spectral reflectance of indoor reflectors. As shown in Fig. 11, the performance of NOMA in a VLC system degrades when the LOS blockage probability increases. However, due to the existence of multipath reflections in indoor VLC, signal transmission is still possible even if the LOS link is totally blocked.

### VI. LED LIGHTING

VLC is built upon indoor luminaries without affecting the main functionality of the LED: lighting. The impact of VLC on light emission quality of white LEDs, such as correlated color temperature (CCT), color rendering index (CRI) and chromaticity has recently been reported in [26]. It is shown that, as long as the average of the LED driving current remains constant, VLC does not affect the emitted light quality. In this section, the illumination quality is analyzed for a VLC system with the application of NOMA.

It is shown in Eq. (6) that the average optical power transmitted by the LED is determined by the added DC bias $I_{DC}$, not the variance of the information-carrying signal $s_i$. This indicates that the functionality of LED lighting is not affected by the application of NOMA in VLC, since the driving current of the LED is typically modulated at a high frequency so that changes in the output optical power is not perceptible to the human eye. In order to calculate the illumination quality, one should not only know the optical power of the LED, but also its PSD curve $\Phi_e(\lambda_v)$. Because human eyes are not equally sensitive to all wavelengths of the visible light, luminous flux $\Phi_v$ is calculated by weighting $\Phi_e(\lambda_v)$ with the luminosity function of the human eye $V(\lambda_v)$ as [27]:

$$\Phi_v = K_v \int_{750\text{nm}}^{380\text{nm}} \Phi_e(\lambda_v) V(\lambda_v) d\lambda_v,$$

where $\lambda_v$ represents the wavelength of the visible light and $K_v = 683 \text{ lm/W}$ is the maximum photopic luminous efficacy.
of the human eye. For a Lambertian LED, the irradiated luminous intensity at angle $\phi_k$ is:

$$I_v(\phi_k) = \frac{m + 1}{2\pi} \Phi_v \cos^m(\phi_k).$$  \hfill (35)

Indoor lighting level can be quantified by parameters such as illuminance $E_v$, which in photometry is defined as the the luminous flux per unit area:

$$E_v = \frac{I_v(\phi_k) \cos(\psi_k)}{d^2_k} = \frac{I_v(0) \cos^m(\phi_k) \cos(\psi_k)}{d^2_k},$$  \hfill (36)

where $I_v(0)$ is the center luminous intensity. Assuming that the center luminous intensity of the LED is 2000 cd, the illuminance as a function of the horizontal distance from the cell center is plotted in Fig. 12. It is shown that the highest illuminance is always at the cell center, and it decreases as the horizontal distance increases. For example, the illuminance is found to be 500.0 lx at the cell center and 27.8 lx at the cell edge, when the detection surface is 2 m below the LED. As the vertical distance $L$ decreases, the illuminance increases significantly at the cell center and decreases slightly at the cell edge.

VII. CONCLUSION

In this paper, a theoretical framework for analyzing the performance of NOMA in a downlink VLC system has been presented, which can be applied to either traditional grid-based network models or stochastic models such as the PPP. In a single attocell deployment, the performance of NOMA has been analytically and numerically evaluated and compared with the traditional OMA technique. We have derived closed-form expressions of the system coverage probability and the ergodic sum rate in scenarios of providing guaranteed QoS and opportunistic best-effort service, respectively. Results show that, compared with OMA, NOMA can offer a high performance gain by utilizing power-domain multiplexing in a multiuser VLC system, and this performance gain can be further enlarged by pairing users with more distinctive channel conditions. Also, we find out that a careful selection of LEDs can have a significant improvement on the system performance. With certain constrains on the modulated signal, LED lighting quality is not affected.

APPENDIX

A. Proof of Theorem 1

The ergodic sum rate of NOMA is given by:

$$R_{\text{sum}} = \sum_{k=1}^{K} \mathbb{E}[R_k] = \sum_{k=1}^{K-1} \int_{\kappa_{\text{min}}}^{\kappa_{\text{max}}} \log_2 \left( 1 + \frac{a_k^2 t}{t \sum_{i=1}^{K} a_i^2 + \frac{1}{\rho}} \right) f_{h_k^2}(t) \, dt$$

$$+ \int_{\kappa_{\text{min}}}^{\kappa_{\text{max}}} \log_2 \left( 1 + \rho a_k^2 t \right) f_{h_k^2}(t) \, dt,$$ \hfill (37)

where $Q_k$ denotes the ergodic data rate for the $k$-th user, $k \in \{1, \cdots, K-1\}$, and $Q_K$ denotes the ergodic data rate for the $K$-th user. Applying binomial expansion, $Q_K$ can be written as:

$$Q_K = \frac{\Omega K}{2(m+3)} \sum_{l=0}^{K-1} \left( \frac{(K-1)!}{(K-1-l)!} \left( \frac{L^2}{p^2} + 1 \right)^{K-1-l} \right)$$

$$\times \left( \int_{\kappa_{\text{min}}}^{\kappa_{\text{max}}} \log_2 \left( 1 + \rho a_K^2 t \right) t^{\frac{l+1}{m+3}} \, dt \right).$$ \hfill (38)

Using geometric series, the integration part of Eq. (38) is derived in Eq. (39), in which $(\cdot)_t$ represents the Pochhammer symbol. By replacing the power series in Eq. (39) with the Gaussian hypergeometric function $\text{2F}_1$, the expression of $Q_K$ is obtained as in Eq. (21). In a similar way, the expression of $Q_k$ can be obtained by applying the quotient rule of the logarithmic function. To this end, the proof of Theorem 1 is completed.

B. Proof of Theorem 2

The probability that NOMA achieves higher individual data rates than OFDMA can be calculated as:

$$P \left[ R_i > \bar{R}_i, R_j > \bar{R}_j \right]$$

$$= P \left[ 1 + \frac{h_i^2 a_i^2}{h_j^2 a_j^2 + \frac{1}{\rho}} > \left( 1 + \frac{v_i}{b_i} \rho h_i^2 \right)^{b_i} \right] \left[ 1 + \rho h_j^2 a_j^2 > \left( 1 + \frac{v_j}{b_j} \rho h_j^2 \right)^{b_j} \right].$$ \hfill (40)
\[
\int \log_2 \left(1 + \rho a_R^2 t \right) t^{-\frac{i+1}{m+3}} dt
\]
\[
= -\frac{m+3}{l+1} \log_2 \left(1 + \rho a_R^2 t \right) t^{-\frac{i+1}{m+3}} + \frac{m+3}{\ln(2)} \frac{1}{l+1} \int \frac{\rho a_R^2 t}{1 + \rho a_R^2 t} t^{-\frac{i+1}{m+3}} dt
\]
\[
= -\frac{m+3}{l+1} \log_2 \left(1 + \rho a_R^2 t \right) t^{-\frac{i+1}{m+3}} + \frac{m+3}{\ln(2)} \frac{1}{l+1} \int \left(1 - \sum_{n=0}^{\infty} \left(-\rho a_R^2 t \right)^n \right) t^{-\frac{i+1}{m+3}} dt
\]
\[
= -\frac{m+3}{l+1} \log_2 \left(1 + \rho a_R^2 t \right) t^{-\frac{i+1}{m+3}} - \frac{1}{\ln(2)} \frac{m+3}{l+1} \int \left(1 + \sum_{n=0}^{\infty} \frac{(-\rho a_R^2 t)^n}{m+3+n} \right) t^{-\frac{i+1}{m+3}} dt
\]
\[
= -\frac{m+3}{l+1} \log_2 \left(1 + \rho a_R^2 t \right) t^{-\frac{i+1}{m+3}} + \frac{1}{\ln(2)} \frac{m+3}{l+1} \int \left(1 + \sum_{n=0}^{\infty} \frac{(-\rho a_R^2 t)^n}{m+3+n} \right) t^{-\frac{i+1}{m+3}} dt
\]
\[
= -\frac{m+3}{l+1} \log_2 \left(1 + \rho a_R^2 t \right) t^{-\frac{i+1}{m+3}} + \frac{1}{\ln(2)} \frac{m+3}{l+1} \int \left(1 + \sum_{n=0}^{\infty} \frac{(1)^n \left(-\frac{1}{m+3} \right)^n}{n!} \right) t^{-\frac{i+1}{m+3}} dt,
\]
(39)

Note that the joint PDF of \( h^2 \) and \( h^2 \) can be obtained as [24]:
\[
f_{h^2, h^2} (u, v) = \omega f_{h^2} (u) f_{h^2} (v) \left[1 - F_{h^2} (v) \right]^{K-j} \left[ F_{h^2} (v) - F_{h^2} (u) \right]^{j-1-i},
\]
(41)
\[
\Xi_2 (v) = \sum_{p=0}^{i-1} \sum_{q=0}^{j-1-i} \frac{1}{p+q+1} \frac{(i-1)!}{(i-1)! (j-1-i)!} \frac{(i-1)!}{(i-1)! (j-1-i)!} \frac{(i-1)!}{(i-1)! (j-1-i)!} \frac{(i-1)!}{(i-1)! (j-1-i)!} \frac{(i-1)!}{(i-1)! (j-1-i)!} \frac{(i-1)!}{(i-1)! (j-1-i)!} \frac{(i-1)!}{(i-1)! (j-1-i)!} \frac{(i-1)!}{(i-1)! (j-1-i)!} \frac{(i-1)!}{(i-1)! (j-1-i)!} \frac{(i-1)!}{(i-1)! (j-1-i)!}
\]
\[
\Xi_2 (v) = \sum_{p=0}^{i-1} \sum_{q=0}^{j-1-i} \frac{1}{p+q+1} \frac{(i-1)!}{(i-1)! (j-1-i)!} \frac{(i-1)!}{(i-1)! (j-1-i)!} \frac{(i-1)!}{(i-1)! (j-1-i)!} \frac{(i-1)!}{(i-1)! (j-1-i)!} \frac{(i-1)!}{(i-1)! (j-1-i)!} \frac{(i-1)!}{(i-1)! (j-1-i)!} \frac{(i-1)!}{(i-1)! (j-1-i)!} \frac{(i-1)!}{(i-1)! (j-1-i)!} \frac{(i-1)!}{(i-1)! (j-1-i)!} \frac{(i-1)!}{(i-1)! (j-1-i)!}
\]
(46)

Combining Eq. (44)–Eq. (46), Theorem 2 is proved.

C. Proof of Theorem 3

The sum rate gain of NOMA over OFDMA can be formulated as:
\[
\mathbb{E} \left[ R_i + R_j - \bar{R}_i - \bar{R}_j \right]
\]
\[
= \frac{1}{2} \int_{\xi_{\min}}^{\xi_{\max}} \left( \log_2 \left(1 + \frac{a_i^2 t}{a_i^2 t + b_i} \right) - b_i \log_2 \left(1 + \frac{v_i}{b_i} \rho \right) \right) f_{h^2} (t) dt
\]
\[
+ \frac{1}{2} \int_{\xi_{\min}}^{\xi_{\max}} \left( \log_2 \left(1 + \rho a_j^2 t \right) - b_j \log_2 \left(1 + \frac{v_j}{b_j} \rho \right) \right) f_{h^2} (t) dt.
\]
(47)

It can be seen from Eq. (47) that \( \mathbb{E} \left[ R_i + R_j - \bar{R}_i - \bar{R}_j \right] \) approximates zero when \( \rho \) is extremely small. According to
Leibniz integral rule, we have

\[
\frac{\partial \mathbb{E} [R_i + R_j - \bar{R}_i - \bar{R}_j]}{\partial \rho} = \frac{1}{2 \ln(2)} \int_{\kappa_{\min}}^{\kappa_{\max}} \left( \frac{t}{1 + \rho t} - \frac{a_j^2 t}{1 + \rho a_j^2 t} - \frac{b_j t}{v_j} + \rho t \right) f'_R(t) dt \\
+ \frac{1}{2 \ln(2)} \int_{\kappa_{\min}}^{\kappa_{\max}} \left( \frac{a_j^2 t}{1 + \rho a_j^2 t} - \frac{b_j t}{v_j} + \rho t \right) f'_R(t) dt.
\]

(48)

As \( \rho \) increases, the derivative of \( \mathbb{E} [R_i + R_j - \bar{R}_i - \bar{R}_j] \) first drops below zero and then increases to a positive value. Therefore, the trend of the sum rate gain of NOMA over OFDMA is proved. In high SNR regimes, it is straightforward to show:

\[
\lim_{\rho \to \infty} \mathbb{E} [R_i] = \lim_{\rho \to \infty} \int_{\kappa_{\min}}^{\kappa_{\max}} \frac{1}{2} \log_2 \left( 1 + \frac{a_j^2 t}{a_j^2 + \frac{1}{\rho}} \right) f'_R(t) dt \\
= \frac{1}{2} \log_2 \left( 1 + \frac{a_j^2}{a_j^2} \right) = - \log_2 a_j.
\]

(49)

The date rate for the \( j \)-th user participated in NOMA can be divided into two parts:

\[
\mathbb{E} [R_j] = \int_{\kappa_{\min}}^{\kappa_{\max}} b_j \log_2 \left( 1 + \rho a_j^2 t f'_R(t) \right) dt \\
+ \int_{\kappa_{\min}}^{\kappa_{\max}} b_j \log_2 \left( 1 + \rho a_j^2 t f'_R(t) \right) dt.
\]

(50)

In high SNR regimes the difference between \( N_2 \) and \( \mathbb{E} [\bar{R}_j] \) can be calculated as:

\[
\lim_{\rho \to \infty} \left( N_2 - \mathbb{E} [\bar{R}_j] \right) = \lim_{\rho \to \infty} \frac{b_j}{2} \int_{\kappa_{\min}}^{\kappa_{\max}} \log_2 \left( 1 + \frac{\rho a_j^2 t}{1 + \frac{1}{\rho^2} \rho t} f'_R(t) \right) dt \\
= b_j \log_2 a_j + \frac{1}{2} b_j \log_2 \left( \frac{b_j}{v_j} \right).
\]

(51)

Applying integration by parts, \( N_1 \) in high SNR regimes can be calculated as:

\[
\lim_{\rho \to \infty} N_1 = b_j \log_2 a_j + \lim_{\rho \to \infty} \frac{b_j}{2} \log_2 \left( 1 + \rho \kappa_{\min} \right) \\
+ \frac{b_j}{2 \ln(2)} \int_{\kappa_{\min}}^{\kappa_{\max}} \frac{1}{t} \left( 1 - F'_R(t) \right) dt,
\]

(52)

where \( F'_R(t) \) represents the CDF of the order variable \( h^2_i \), and the integration in Eq. (52) can be obtained as:

\[
\int_{\kappa_{\min}}^{\kappa_{\max}} \frac{1}{t} \left( 1 - F'_R(t) \right) dt \\
= \sum_{p=0}^{j-1} \sum_{q=0}^{K-j} \frac{K!(-1)^{p+K-j-q} \Omega^{p+q+1}}{p! (j-1-p)! q! (K-j-q)!} \left( \frac{L^2}{\rho^2} + 1 \right)^{j-1-p} \\
\times \left( \frac{L^2}{\rho^2} \right)^{K-j-q} \left( \varpi^4(\kappa_{\max}) - \varpi^4(\kappa_{\min}) \right).
\]

(53)

In a similar way, \( \bar{R}_i \) in high SNR regimes can be calculated as:

\[
\lim_{\rho \to \infty} \mathbb{E} [\bar{R}_i] \\
= \lim_{\rho \to \infty} \frac{b_i}{2} \log_2 \left( 1 + \frac{\nu_i}{b_i} \rho \kappa_{\min} \right) - \frac{b_i}{2 \ln(2)} \ln(\kappa_{\min}) \\
+ \frac{\Omega b_i}{2 \ln(2)} \sum_{p=0}^{i-1} \sum_{q=0}^{K-i} \frac{K!(-1)^{p+K-i-q} \Omega^{p+q}}{p! (i-1-p)! q! (K-i-q)!} \\
\times \left( \frac{L^2}{\rho^2} + 1 \right)^{i-1-p} \left( \frac{L^2}{\rho^2} \right)^{K-i-q} \\
\times \left( \varpi^4(\kappa_{\max}) - \varpi^4(\kappa_{\min}) \right).
\]

(54)

Combining Eq. (49)–Eq. (54) Theorem 3 is proved.

\( \square \)

D. Proof of Corollary 2

The proof of is divided into two parts. First, we prove that the maximum sum rate gain is achieved when \( i = 1 \) and \( j = K \). Second, we prove the expression of the maximum sum rate gain in Eq. (31).

For the first part, it is equivalent to prove the following:

\[
\lim_{\rho \to \infty} \mathbb{E} [R_{i+1} + R_j - \bar{R}_{i+1} - \bar{R}_j] \leq \lim_{\rho \to \infty} \mathbb{E} [R_i + R_j - \bar{R}_i - \bar{R}_j],
\]

(55)

for \( \forall i < K \), and

\[
\lim_{\rho \to \infty} \mathbb{E} [R_i + R_{j+1} - \bar{R}_i - \bar{R}_{j+1}] > \lim_{\rho \to \infty} \mathbb{E} [R_i + R_j - \bar{R}_i - \bar{R}_j],
\]

(56)

for \( \forall j < K \). The expression in Eq. (55) is equivalent to

\[
\lim_{\rho \to \infty} \mathbb{E} [R_{i+1} + R_i - \bar{R}_{i+1} + \bar{R}_i] < 0.
\]

(58)

From Eq. (49), it can be shown that \( \lim_{\rho \to \infty} \mathbb{E} [R_{i+1}] = \lim_{\rho \to \infty} \mathbb{E} [R_i] \). Therefore, in order to prove Eq. (55), we need to prove:

\[
\lim_{\rho \to \infty} \mathbb{E} [R_{i+1} - R_i - \bar{R}_{i+1} + \bar{R}_i] = 0
\Rightarrow \lim_{\rho \to \infty} \mathbb{E} [R_i - \bar{R}_{i+1}] < 0
\Rightarrow \int \frac{1}{t} \left( 1 - F'_R(t) \right) dt < \int \frac{1}{\kappa_{\min}} \left( 1 - F'_R(t) \right) dt
\Rightarrow F'_R(t) > F'_R(t).
\]

(59)

As \( F'_R(t) \) represents the CDF of the \( i \)-th largest variable \( h^2_i \), it is obvious that Eq. (59) is true. To this end Eq. (55) is proved, and the proof of Eq. (56) can be conducted in a similar way.

For the second part, the expression of Eq. (31) can be obtained by setting \( i = 1 \) and \( j = K \) in Eq. (29). To this end, Corollary 2 is proved.

\( \square \)
Harald Haas currently holds the Chair of Mobile Communications at the University of Edinburgh, and is co-founder and Chief Scientific Officer of pureLiFi Ltd as well as the Director of the LiFi Research and Development Center at the University of Edinburgh. His main research interests are in optical wireless communications, hybrid optical wireless and RF communications, spatial modulation, and interference coordination in wireless networks. He first introduced and coined spatial modulation and LiFi. LiFi was listed among the 50 best inventions in TIME Magazine 2011. Prof. Haas was an invited speaker at TED Global 2011, and his talk, “Wireless Data from Every Light Bulb” has been watched online more than 2.4 million times. He gave a second TED Global lecture in 2015 on the use of solar cells as LiFi data detectors and energy harvesters. This has been viewed online more than 1.6 million times. Professor Haas holds 31 patents and has more than 30 pending patent applications. He has published 300 conference and journal papers including a paper in Science. He co-authors a book entitled: “Principles of LED Light Communications Towards Networked Li-Fi” published with Cambridge University Press in 2015. Prof. Haas is editor of IEEE Transactions on Communications and IEEE Journal of Lightwave Technologies. He was co-recipient of recent best paper awards at the IEEE Vehicular Technology Conference (VTC-Fall) in Las Vegas in 2013, VTC-Spring in Glasgow in 2015, and ICC 2016. He was co-recipient of the EURASIP Best Paper Award for the Journal on Wireless Communications and Networking in 2015, and co-recipient of the Jack Neubauer Memorial Award of the IEEE Vehicular Technology Society. In 2012, he was the recipient of the prestigious Established Career Fellowship from the EPSRC (Engineering and Physical Sciences Research Council) within Information and Communications Technology in the UK. In 2014, he was selected by EPSRC as one of ten RISE (Recognising Inspirational Scientists and Engineers) Leaders in the UK.