Beam Domain Optical Wireless Massive MIMO Communications with Transmit Lens

Chen Sun, Student Member, IEEE, Xiqi Gao, Fellow, IEEE,
Zhi Ding, Fellow, IEEE, and Xiang-Gen Xia, Fellow, IEEE

Abstract

We present beam domain optical wireless massive multiple-input multiple-output communications, where base station with massive transmitters communicates with a number of user terminals through transmit lens simultaneously. We focus on LED transmitters, and provide an optical channel model, where the light emitted from one LED passes through the transmit lens and generates a narrow beam. Based on this channel model, we analyze the performance of maximum ratio transmission and regularized zero-forcing linear precoding schemes, and propose concave-convex-procedure based transmit covariance matrix design to maximize the sum-rate. Then, we design the transmit covariance matrix when the number of transmitters goes to infinity, and show that beam division multiple access (BDMA) transmission achieves the asymptotically optimal performance for sum-rate maximization. Compared with the case without transmit lens, BDMA transmission can increase the sum-rate proportionally to $2K$ ($K$ is the number of users) and $K$ under total power and per transmitter power constraints, respectively. Motivated by the asymptotic results, for the non-asymptotic case, we prove the orthogonality conditions of the optimal power allocation in the beam domain and propose beam allocation algorithms satisfying the conditions. Numerical results confirm the significantly improved performances of our proposed beam domain optical massive MIMO communication approaches.

Index Terms

Optical massive MIMO communication, transmit lens, Beam division multiple access (BDMA).

C. Sun and X. Q. Gao are with the National Mobile Communications Research Laboratory, Southeast University, Nanjing, 210096, China. X. Q. Gao is the correspondence author (email: {sunchen, xqgao}@seu.edu.cn).

Z. Ding is with the Department of Electrical and Computer Engineering, University of California, Davis, CA 95616, USA (email: zding@ucdavis.edu).

X.-G. Xia is with the Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19716 USA (e-mail: xxia@ee.udel.edu).
I. INTRODUCTION

Optical wireless communication systems rely on optical radiations to transmit information with wavelengths ranging from infrared to ultraviolet including the visible light spectrum [1]. Base station (BS) commonly employs light-emitting diodes (LEDs) as optical transmitters to convert the electrical signals to optical signals. Recently, laser diodes (LDs) are considered as potential sources for optical communication due to high modulation bandwidth, efficiency, and beam convergence [2]. User terminals (UTs) employ photodetectors like photodiodes as optical receivers to convert the optical power into electrical current. Moreover, BS and UTs can also employ optical transceiver ports to transmit and receive optical signals. Optical communications can significantly relieve the crowded radio frequency (RF) spectrum, provide high speed data transmission [3], and achieve simple and low-cost optical modulation and demodulation through intensity modulation and direct detection (IM/DD) [4]. Thus, optical wireless communication has attracted increasing attention from both academia and industry [5]–[7].

To achieve high data rate in optical communications, multiple separate LED/LD arrays are usually utilized in the BS to provide higher data rate by means of spatial multiplexing. As a result, multiple-input multiple-output (MIMO) technique is a natural progression for optical communication systems [7], [8]. As the nature of optical downlink communication is broadcast network, multiple UTs should be well supported. BSs simultaneously transmit signals to all UTs, resulting in the so-called multi-user interference, consequently degrades the performance. Thus, multi-user MIMO (MU-MIMO) has been studied and several precoding schemes have been proposed, which are different from conventional RF systems since only real-valued non-negative signals can be transmitted [9]–[14]. In [9], the performances of zero forcing and dirty paper coding schemes were compared. An optimal linear precoding transmitter was derived based on the minimum mean-squared error (MMSE) criterion in [10], while block diagonalization precoding algorithm was investigated in [11]. Some works investigated precoding design based on sum-rate maximization [12] and max-min fairness [13]. In [14], a linear precoding method was proposed to minimize the LED power consumption subject to capacity constraint. Furthermore, recently proposed massive MIMO with tens or hundreds of antennas in RF systems [15] was applied into optical communication systems [16] to increase spectrum efficiency.
Since optical communication systems employ intensity modulation and direct detection and line-of-sight (LOS) scenario is mostly considered, highly correlated channels limit the system performance [8]. Imaging receiver is employed to separate signals from different directions, which was originally proposed for infrared communications [17]. In [8], the authors investigated non-imaging and imaging MIMO optical communications, and indicated that the imaging receiver can potentially offer higher spatial diversity. Due to poor imaging quality leading to interference from different LED arrays, some works proposed an imaging receiver with a fisheye lens [18] or a hemispherical lens [19] to provide high-spatial diversity for MIMO signals. Many works study the receive lens to separate the signals from different LED arrays. However, there is no works investigating transmit lens at the BS. Without transmit lens, each transmitter, such as LED, is omni-directional. Since the distance between LED array and UT is much larger than LED array size, the channels between different transmitters and UT are highly correlated [20]. Thus, one LED array transmits only one data stream [8]. To support multiple users, multiple separate LED arrays are required. The number of LED arrays dominates the number of served UTs.

Motivated by the receive lens to separate light from different LED arrays, we employ transmit lens to refract the lights from different transmitters towards different directions, providing high spatial resolution. In this paper, we study beam domain optical wireless massive MIMO communications, where BS equipped with a large number of transmitters serves multiple UTs simultaneously through transmit lens. The fundamental principle of transmit lens is to provide variable refraction angles so as to achieve angle-dependent energy focusing property. Specifically, lights from different transmitters are sufficiently separated by the transmit lens. We focus on LED transmitters as an example, and a similar transmission scheme can be applied to LD transmitters or optical fiber ports connected to optical transceivers. We first provide the optical channel model with transmit lens. The light emitted from one LED passing through the lens will converge to a spot area and generate a beam. As the number of transmitters tends to infinity, the channel vectors of different UTs become asymptotically orthogonal. Based on this channel model, we investigate linear transmit design including maximum ratio transmission (MRT) and regularized zero-forcing (RZF). Moreover, we consider transmit covariance matrix design for sum-rate maximization under total and per LED power constraints. Under both power constraints, utilizing concave-convex-procedure (CCCP), we provide an iteratively procedure to
converge to some stationary point for the transmit covariance matrix design problem. Then, we analyze the transmit covariance matrix design when the number of transmitters goes to infinity, and provide simple and intuitive designs. The asymptotically optimal transmission is that different transmitters transmit independent signals to different UTs, and that the transmit beams for different UTs should be non-overlapping, which indicates that beam division multiple access (BDMA) transmission can achieve the asymptotically optimal performance. Compared with the conventional transmission without transmit lens, BDMA transmission increases the sum-rate proportionally to $2K$ ($K$ is the number of UTs) and $K$ under total power and per LED power constraints, respectively. Moreover, we consider beam domain power allocation for non-asymptotic case. We prove the orthogonality of the optimal power allocation, and provide beam allocation algorithms. Numerical results illustrate the significant performance gains of our proposed optical massive MIMO communication schemes with transmit lens.

We adopt the following notations throughout the manuscript: Upper (lower) bold-face letters denote matrices (column vectors); $I$ denotes the identity matrix, while $1$ denotes an all-one matrix and $0$ denotes zero matrix; numeral subscripts of matrices and vectors, if needed, represent their sizes. Also, matrix superscript $(\cdot)^T$ denotes matrix transpose. We use $\text{tr}(\cdot)$ and $\text{det}(\cdot)$ to represent matrix trace and determinant operations, respectively. The inequality $A \succeq 0$ denotes a positive semi-definite Hermitian matrix $A$. We use $[A]_{mn}$ to denote the $(m, n)$-th element of matrix $A$.

II. SYSTEM MODEL

A. Optical MIMO System with Transmit Lens

We consider an optical massive MIMO system consisting of a BS, equipped with $N^2$ transmitters and a transmit lens, and $K$ UTs, each of which has one photodetector employed as a receiver. The transmitters can be LEDs, LDs, or optical fiber ports connected to optical transceivers. In this paper, we focus on LED transmitters. With a transmit lens at the BS, the light emitted from different LEDs passing through the lens is refracted to different directions, as shown in Fig. [I]. The LED array is located at the focal plane of the lens, and the focal length of the lens is $f$. Here, we consider a square LED array, the size of which is fixed as $d \times d$. The LEDs are symmetric with respect to x-axis and y-axis and there are $N$ LEDs along the x-axis or y-axis. Different LEDs are obstructed by some lightproof material (which is not shown in the figure).
Suppose that each LED can be treated as a circular emitter with radius of \( r = d/(2N) \), and each LED becomes smaller as \( N \) increases. The center position of the \((i, j)\)th LED is denoted as \((x_{ij}, y_{ij})\),

\[
(x_{ij}, y_{ij}) = \left(-\frac{d}{2} + (2i - 1)r, -\frac{d}{2} + (2j - 1)r\right), \quad i, j = 1, 2, \ldots, N. \tag{1}
\]

BS employs the LED array to transmit signals to UTs. Denote \( x_k \in \mathbb{R}^{N^2 \times 1} \) as the signal intended to the \( k \)th UT, and the received signal at the \( k \)th UT can be written as

\[
y_k = h_k^T x + z_k \\
= h_k^T x_k + h_k^T \left( \sum_{k' \neq k} x_{k'} \right) + z_k, \tag{2}
\]

where \( x = \sum_k x_k \) is the summation of all the intended signals, \( h_k^T \in \mathbb{R}^{1 \times N^2} \) is the channel vector from all LEDs to the \( k \)th UT, and \( z_k \) is the receiver noise which can be modeled as the sum of ambient-induced shot noise and thermal noise. It is generally accepted that \( z_k \) is real valued additive white Gaussian noise with zero mean and variance \( \sigma^2 \). Here, without loss of generality, we assume a unit noise variance (i.e., \( \sigma^2 = 1 \)).
B. Physical Channel Model

![Diagram of transmit lens based beam generation](image)

Fig. 2: Schematic diagram of the transmit lens based beam generation.

In this paper, we consider the only LOS propagation path [5], [16]. The light emitted from one LED passing through the lens converges to a spot area, which is called a beam, as shown in Figure 2. The refraction light from the center of the LED is called the center light. According to the geometrical optics [21], [22], the angle between the center light and the z-axis $\phi_{ij}$ can be expressed as

$$\phi_{ij} = \arctan\left(\frac{\sqrt{x_{ij}^2 + y_{ij}^2}}{f}\right).$$

(3)

For one beam generated by one LED, the half width viewing angle is denoted as $\theta_{ij}$, which is the refraction of the light from the edge of the LED. As the LED size is much smaller than the focal length, the angles between refraction lights from the edge of the LED and the center light are approximately the same, and the angle $\theta_{ij}$ can be calculated as

$$\theta_{ij} = \arctan\left(r + \sqrt{x_{ij}^2 + y_{ij}^2}/f\right) - \arctan\left(\sqrt{x_{ij}^2 + y_{ij}^2}/f\right).$$

(4)

Assuming that each beam generated by one LED has Lambertian distribution patterns [8], [23], [24], the luminous intensity $I$ is a cosine function of the viewing angle,

$$I_{ij}(\phi) = \begin{cases} \frac{m_{ij}+1}{2\pi} \cos^{m_{ij}}(\phi) \cos(\phi_{ij}), & \phi < 2\theta_{ij} \\ 0, & \phi \geq 2\theta_{ij}, \end{cases}$$

(5)
where $\phi$ is the viewing angle, and the parameter $m_{ij}$ is calculated as

$$m_{ij} = \frac{-\log 2}{\log(\cos \theta_{ij})},$$

where $\theta_{ij}$ is the transmitter semiangle (at half power), and $2\theta_{ij}$ denotes the field-of-view (FOV) of a transmitter. If a receiver is not in the FOV of a receiver, $I_{ij} = 0$.

Then, the channel gain between transmitter (LED) $(i, j)$ and UT $k$, $h_{k,ij}$ of the channel vector $h_k$ can be presented as

$$h_{k,ij} = \begin{cases} \frac{A_R}{d_k^2} I_{ij}(\phi_{k,ij}) T(\phi_{k,ij}) \cos(\psi_k), & |\psi_k| \leq \psi_C \\ 0, & |\psi_k| > \psi_C \end{cases}$$

(7)

where $A_R$ is the physical area of the receiver, $d_k$ is the distance between the center of the transmit lens and photodetector center of UT $k$, $\phi_{k,ij}$ denotes the angle between the $k$th UT and the center light emitted from LED $(i, j)$, $\psi_k$ denotes the angle of incidence at UT $k$, $\psi_C$ presents the width of the FOV, and $T(\phi_{k,ij})$ reflects the effect of the emission angle dependent energy focusing by the transmit lens. For simplicity, the transmit optical lens gain can be modeled as constant, e.g., $T(\phi_{k,ij}) = T$.

Let $g_k$ be

$$g_k = \begin{cases} \frac{A_R}{c_k} \cos(\psi_k), & |\psi_k| \leq \psi_C \\ 0, & |\psi_k| > \psi_C \end{cases},$$

(8)

and the channel coefficient $h_{k,ij}$ can be expressed as

$$h_{k,ij} = I_{ij}(\phi_{k,ij}) g_k T.$$

(9)

Thus, the channel vector $h_k$ is given by

$$h_k = g_k T \left[ I_{11}(\phi_{k,11}) \quad I_{12}(\phi_{k,12}) \quad \cdots \quad I_{NN}(\phi_{k,NN}) \right]^T.$$

(10)

Consider the conventional channel model without transmit lens. As the distance between BS and UT is much larger than LED size, the distance of different LEDs to UT and the emission angles are almost the same. The DC gain between BS and UT $k$, $\tilde{h}_k$ can be expressed as

$$\tilde{h}_k = I(\theta_k) g_k \mathbf{1}_{N^2 \times 1},$$

(11)

where the luminous intensity $I(\phi)$ is modeled as

$$I(\phi) = \frac{\tilde{m} + 1}{2\pi} \cos^{\tilde{m}}(\phi),$$

(12)
and \( \tilde{m} \) is given by
\[
\tilde{m} = \frac{-\log 2}{\log(\cos(\theta_0))},
\]
where \( \theta_0 \) is the half width viewing angle of a LED.

C. Asymptotic Properties

Next, we will analyze the properties of the channel model with transmit lens, when the number of transmit LEDs grows to infinity. Denote \( a_{ij} = \sqrt{x_{ij}^2 + y_{ij}^2} / f \), and from (4), we have
\[
\cos(\theta_{ij}) = \cos(\arctan(d/(2fN) + a_{ij}) - \arctan(a_{ij})) = \left(1 + \left(\frac{d/(2fN)}{1 + a_{ij}(d/(2fN) + a_{ij})}\right)^2\right)^{-\frac{1}{2}}.
\]
Then, according to (6), the order of Lambertian emission \( m_{ij} \) can be expressed as
\[
m_{ij} = \frac{2 \log 2}{\log \left(1 + \left(\frac{d/(2fN)}{1 + a_{ij}(d/(2fN) + a_{ij})}\right)^2\right)}.
\]
As the number of transmit LEDs \( N \) grows without limit, we have
\[
\lim_{N \to \infty} \frac{m_{ij}}{N^2} = \lim_{N \to \infty} \frac{(2 \log 2)N^{-2}}{\log \left(1 + \left(\frac{d/(2fN)}{1 + a_{ij}(d/(2fN) + a_{ij})}\right)^2\right)} = \frac{(2 \log 2)(a_{ij}^2 + 1)^2}{(d/2f)^2},
\]
which means that \( m_{ij} \) tends to infinity with the same order of \( N^2 \). Then, for \( \phi_{k,ik,jk} = 0 \), which means that the \( k \)th UT is illuminated by the center light of LED \( (i_k, j_k) \), we have
\[
\lim_{N \to \infty} \frac{I_{i_kj_k}(\phi_{k,ik,jk})}{N^2} = \lim_{N \to \infty} \frac{(m_{ik,jk} + 1) \cos^{m_{ik,jk}}(\phi_{k,ik,jk}) \cos(\phi_{ij})}{2\pi N^2} = \frac{(2 \log 2)(a_{ij}^2 + 1)^2 \cos(\phi_{ij})}{(d/2f)^2} \Delta c_{ij}.
\]
Consider that any two UTs \( (k_1, k_2, k_1 \neq k_2) \) are in different positions. As \( N \) goes to infinity, at most one of the emission angles, \( \phi_{k_1,ij} \) or \( \phi_{k_2,ij} \), tends to 0. (If both tend to 0, these two UTs are illuminated by the center light of LED \( (i, j) \), and overlapped.) Without loss of generality,
assume $\phi_{k2,ij}$ does not tend to 0, Employing L’Hôpital’s rule, the luminous intensities for the two UTs have the following asymptotic result,

$$\lim_{N \to \infty} N^2 I_{ij}(\phi_{k1,ij}) I_{ij}(\phi_{k2,ij}) = \lim_{N \to \infty} N^2 \frac{m_{ij} + 1}{2\pi} \cos^{m_{ij}}(\phi_{k1,ij}) \frac{m_{ij} + 1}{2\pi} \cos^{m_{ij}}(\phi_{k2,ij}) \cos^2(\phi_{ij})$$

$$\leq \lim_{N \to \infty} N^2 \frac{(m_{ij} + 1)^2}{(2\pi)^2} \cos^{m_{ij}}(\phi_{k2,ij}) \cos^2(\phi_{ij})$$

$$= 0. \quad (18)$$

Then, we have

$$\lim_{N \to \infty} h_{k1}^T h_{k2} = \lim_{N \to \infty} g_{k1} g_{k2} T^2 \left( \sum_i \sum_j I_{ij}(\phi_{k1,ij}) I_{ij}(\phi_{k2,ij}) \right)$$

$$\leq g_{k1} g_{k2} T^2 \lim_{N \to \infty} N^2 \{ \max I_{ij}(\phi_{k1,ij}) I_{ij}(\phi_{k2,ij}) \} = 0. \quad (19)$$

Moreover, $h_{k1}$ and $h_{k2}$ are non-negative. Therefore, $\lim_{N \to \infty} h_{k1}^T h_{k2} = 0$, and the channel vectors $h_{k1}$ and $h_{k2}$ become asymptotically orthogonal. Combining all user channel vectors, we construct the multi-user channel matrix $H \in \mathbb{C}^{K \times N^2}$ as

$$H = \begin{bmatrix} h_1 & h_2 & \cdots & h_K \end{bmatrix}^T. \quad (20)$$

As $N$ grows to infinity, the rows of matrix $H$ become orthogonal, and thus, the rank of channel matrix $H$ is $K$. In addition, we can construct the channel matrix $\tilde{H}$ without transmit lens as

$$\tilde{H} = \begin{bmatrix} \tilde{h}_1 & \tilde{h}_2 & \cdots & \tilde{h}_K \end{bmatrix}^T = \begin{bmatrix} \tilde{I}(\phi_1) g_1 & \tilde{I}(\phi_2) g_2 & \cdots & \tilde{I}(\phi_K) g_K \end{bmatrix}^T 1_{1 \times N^2}, \quad (21)$$

whose rank is 1, i.e., $\text{rank}(\tilde{H}) = 1$.

Remark 1: For the channel model with transmit lens, the multi-user channel matrix $H$ has full row rank, and thus, BS has the potential to serve $K$ UTs simultaneously. BS without transmit lens cannot support simultaneous multi-user communications. To serve multiple users, in the literature [9]–[13], multiple LED arrays are considered.

III. LINEAR Precoding Based Transmission

As the transmit signal may take on negative values, a DC bias should be added to guarantee a non-negative input signal to the LEDs. Assume that the LEDs transmit optical signals with a large DC bias corresponding to the working point of the LEDs. In this way, the bipolar electrical signals are transmitted via the unipolar optical intensity.
A. MRT/RZF Linear Precoding

To communicate with multiple UTs simultaneously, BS transmits the summation of all UTs’ signals, and the signal of the \(k\)th UT \(x_k\) is obtained from symbols through a precoding vector \(w_k \in \mathbb{C}^{N^2 \times 1}\), i.e.,

\[
x_k = w_k s_k,
\]

where \(s_k\) is the message-bearing independent and identically distributed (i.i.d.) symbols with unit variance. Here, we consider two different linear precoding strategies \(w_k\) of practical interest, namely MRT \(w_k^{\text{MRT}}\) and RZF \(w_k^{\text{RZF}}\), which we define, respectively, as

\[
w_k^{\text{MRT}} = \sqrt{\beta^{\text{MRT}}} h_k, \quad w_k^{\text{RZF}} = \sqrt{\beta^{\text{RZF}}} \left( H^T H + \alpha I_{N^2} \right)^{-1} h_k,
\]

where \(\alpha > 0\) is a regularization parameter, \(\beta^{\text{MRT}}\) and \(\beta^{\text{RZF}}\) normalize the total transmit power to \(\sum_k E\{x_k^T x_k\} = P\), i.e.,

\[
\beta^{\text{MRT}} = \frac{P}{\sum_k h_k^T h_k}, \quad \beta^{\text{RZF}} = \frac{P}{\text{tr} \left( H^T H (H^T H + \alpha I)^{-2} \right)}.
\]

Using a standard bound based on the worst-case uncorrelated additive noise \[26\], it yields the achievable sum-rate for linear precoding,

\[
R = \frac{1}{2} \sum_k \log \left( 1 + \frac{(h_k^T w_k)^2}{1 + \sum_{k' \neq k} (h_k^T w_{k'})^2} \right) + \mathcal{O}\left(N^{-4}\right).
\]

Next, we consider \(N\) growing infinitely large and keep total transmit power \(P\) constant. There exists \(\phi_{k,ik,jk} = 0\) and other \(\phi_{k,ij}\) is larger than \(2\theta_{ij}\). From [5] and (17),

\[
\lim_{N \to \infty} \frac{h_k^T h_k}{N^4} = \lim_{N \to \infty} \left( g_k^2 T_2 \left( I_{gk,ij} (\phi_{k,ik,jk}) \right)^2 + g_k^2 T_2 \sum_{\phi_{k,ij} \geq 2\theta_{ij}} \left( I_{ij} (\phi_{k,ij}) \right)^2 \right) = \left( g_k T c_{ik,jk} \right)^2.
\]

Then, we can derive the asymptotic sum-rate for MRT as

\[
\lim_{N \to \infty} \left( R^{\text{MRT}} - \bar{R}^{\text{MRT}} \right) = 0,
\]

where \(\bar{R}^{\text{MRT}}\) is given by

\[
\bar{R}^{\text{MRT}} = \frac{1}{2} \sum_k \log \left( P \frac{N^4 (g_k T c_{ik,jk})^4}{\sum_{k'} (g_{k'} T c_{ik,jk})^2} \right).
\]

From (28), the simplest MRT strategy can vanish the inter-user interference and the asymptotic sum-rate increases with \(N\), which is similar with the asymptotic result in RF massive MIMO system [15].
Similarly, for RZF precoding, we have
\[
\lim_{N \to \infty} \left( R^{\text{RZF}} - \bar{R}^{\text{RZF}} \right) = 0. 
\] (29)
where
\[
\bar{R}^{\text{RZF}} = \frac{1}{2} \sum_k \log \left( \frac{P N^4}{\sum_{k'} (g_{k'} c_{k',j_k})^{-2}} \right). 
\] (30)

**Remark 2:** From the above analysis, the precodings of MRT and RZF can vanish inter-user interference with an infinite number of transmit LEDs. However, the sum-rate of MRT and RZF are not necessarily identical, due to the power normalization factors.

**B. Linear Precoding for Sum-Rate Maximization**

For the asymptotic case, the simplest MRT precoding can vanish the inter-user interference. For large but a limited number of LEDs, RZF can provide better performance than MRT. In this subsection, we directly consider the transmit covariance matrix design maximizing the sum-rate.

Let \( Q_k = \mathbb{E}\{x_k x_k^T\} \) be the covariance matrices of transmitted signals, and \( Q = \sum_k Q_k \) be the covariance matrix of the sum of the transmitted signals. Consider that receivers have knowledge of its CSI, as well as the aggregate interference-plus-noise covariance matrix. When BS employs linear precoding and UTs treat the aggregate interference-plus-noise as Gaussian noise with the same covariance matrix [27], the achievable sum-rate is given by [28], [29]
\[
R_{\text{sum}} = \frac{1}{2} \sum_k \left( \log \left( 1 + h_k^T Q_k h_k \right) - \log \left( 1 + h_k^T \left( \sum_{k' \neq k} Q_{k'} \right) h_k \right) \right). 
\] (31)

Our main objective is to design the transmitted covariance matrices \( Q_k \) maximizing the sum-rate \( R_{\text{sum}} \). A typical power constraint is total power constraint, which can be expressed as \( \sum_k \text{tr}(Q_k) \leq P \), where \( P \) is the total power. Define
\[
f(Q_1, Q_2, \ldots, Q_K) = \frac{1}{2} \sum_k \log \left( I + h_k^T \left( \sum_{k'} Q_{k'} \right) h_k \right),
\] (32)
\[
g(Q_1, Q_2, \ldots, Q_K) = \frac{1}{2} \sum_k \log \left( I + h_k^T \left( \sum_{k' \neq k} Q_{k'} \right) h_k \right). 
\] (33)

The optimization problem under total power constraint can be formulated as
\[
\max_{Q_k} f(Q_1, Q_2, \ldots, Q_K) - g(Q_1, Q_2, \ldots, Q_K)
\]
s.t. \( \sum_k \text{tr}(Q_k) \leq P, \quad Q_k \succeq 0. \) (34)
Due to the concavity of $\log(\cdot)$ function, the sum-rate $R_{\text{sum}}$ is a difference of concave functions (d.c.). To solve problem (34), we utilize the concave-convex-procedure (CCCP), which is an iterative procedure solving a sequence of convex programs. The idea of CCCP program is to linearize the concave part around a solution obtained in the current iteration. Employing the CCCP method, the iterative procedure is expressed as

$$
\left[ Q_1^{(i+1)}, Q_2^{(i+1)}, \ldots, Q_K^{(i+1)} \right] = \arg \max_{Q_k} f(Q_1, Q_2, \ldots, Q_K)
- \sum_k \text{tr} \left( \left( \frac{\partial}{\partial Q_k} g(Q_1^{(i)}, Q_2^{(i)}, \ldots, Q_K^{(i)}) \right)^T Q_k \right)
\text{s.t. } \sum_k \text{tr} (Q_k) \leq P, \quad Q_k \succeq 0.
$$

(35)

To understand the properties of this convex optimization problem better, we present some basic properties of the generated sequences by (35).

**Theorem 1:** Let $\left\{ Q_1^{(i)}, Q_2^{(i)}, \ldots, Q_K^{(i)} \right\}_{i=0}^{\infty}$ be any sequences generated by (35). Then, all limit points of $\left\{ Q_1^{(i)}, Q_2^{(i)}, \ldots, Q_K^{(i)} \right\}_{i=0}^{\infty}$ are stationary points of the d.c. program in (34). In addition, $\lim_{i \to \infty} \left( f(Q_1^{(i)}, Q_2^{(i)}, \ldots, Q_K^{(i)}) - g(Q_1^{(i)}, Q_2^{(i)}, \ldots, Q_K^{(i)}) \right) = f(Q_1^{(s)}, Q_2^{(s)}, \ldots, Q_K^{(s)}) - g(Q_1^{(s)}, Q_2^{(s)}, \ldots, Q_K^{(s)})$, where $\left\{ Q_1^{(s)}, Q_2^{(s)}, \ldots, Q_K^{(s)} \right\}$ is some stationary point of problem (34).

**Proof:** See Appendix A.

**Remark 3:** Theorem 1 establishes that iterative procedure (35) can find a stationary point of d.c. program (34). Although we can obtain a sub-optimal solution by iteratively solving the convex program (35), due to large number of LEDs and large dimension of $Q_k$, the computational complexity is also demanding.

In optical communications, the common IM/DD schemes require driving current must be non-negative. Thus, the transmit current of each LED is limited to guarantee the non-negative input signal. This constraint can be expressed as

$$
\sum_k [w_k]_n \leq b,
$$

(36)

where $b$ is the bias current. Employing the result in [12], we have

$$
\frac{(\sum_k [w_k]_n)^2}{K} \leq \sum_k [w_k w_k^T]_{nn} = \sum_k [Q_k]_{nn}.
$$

(37)

Thus, $\sum_k [Q_k]_{nn} \leq b^2/K$ can ensure the constraint in (36). Here, we consider the transmit design problem under per LED power constraint. Define unit vector $e_n = [0, \ldots, 0, 1, 0, \ldots, 0]^T$, where
only the \( n \)th element is 1. With the definition of \( f(Q_1, Q_2, \ldots, Q_K) \) and \( g(Q_1, Q_2, \ldots, Q_K) \), the transmit design with per LED power constraint problem can be expressed as

\[
\max_{Q_k} f(Q_1, Q_2, \ldots, Q_K) - g(Q_1, Q_2, \ldots, Q_K)
\]

s.t. \( Q_k \succeq 0, \quad \text{e}_n^T \left( \sum_k Q_k \right) \text{e}_n \leq p, \quad n = 1, 2, \ldots, N^2. \tag{38}
\]

where \( p = b^2 / K \) is the maximal power per LED. Similar to problem (34), due to the concavity of \( f(Q_1, Q_2, \ldots, Q_K) \) and \( g(Q_1, Q_2, \ldots, Q_K) \) on \( Q_k \), problem (38) is a d.c. program. Utilizing the CCCP method, we can solve the d.c. problem by iteratively solving the following convex problem:

\[
\begin{align*}
&\left[ Q_1^{(i+1)}, Q_2^{(i+1)}, \ldots, Q_K^{(i+1)} \right] = \arg \max_{Q_k} f(Q_1, Q_2, \ldots, Q_K) \\
&\quad - \sum_k \text{tr} \left( \left( \frac{\partial}{\partial Q_k} g(Q_1^{(i)}, Q_2^{(i)}, \ldots, Q_K^{(i)}) \right)^T Q_k \right) \\
&\text{s.t.} \quad Q_k \succeq 0, \quad \text{e}_n^T \left( \sum_k Q_k \right) \text{e}_n \leq p, \quad n = 1, 2, \ldots, N^2. \tag{39}
\end{align*}
\]

For the generalized sequences by iteratively solving problem (39), we have the following result.

**Theorem 2:** Let \( \left\{ Q_1^{(i)}, Q_2^{(i)}, \ldots, Q_K^{(i)} \right\}_{i=0}^\infty \) be any sequences generated by (39). Then, all limit points of \( \left\{ Q_1^{(i)}, Q_2^{(i)}, \ldots, Q_K^{(i)} \right\}_{i=0}^\infty \) are stationary points of the d.c. program in (38). In addition, \( \lim_{i \to \infty} \left( f(Q_1^{(i)}, Q_2^{(i)}, \ldots, Q_K^{(i)}) - g(Q_1^{(i)}, Q_2^{(i)}, \ldots, Q_K^{(i)}) \right) = f(Q_1^\ast, Q_2^\ast, \ldots, Q_K^\ast) - g(Q_1^\ast, Q_2^\ast, \ldots, Q_K^\ast) \) where \( \left\{ Q_1^\ast, Q_2^\ast, \ldots, Q_K^\ast \right\} \) is some stationary point of problem (38).

**Proof:** The proof is similar to that of Theorem 1 and is omitted here.

From the above analysis, we can utilize the CCCP method to obtain candidate optimal solutions under both power constraints. The solutions converge to some stationary point of the original d.c. program. For optical massive MIMO communications, as the number of transmit LED increases, the dimension of transmit covariance matrix \( Q_k \) becomes large, and it needs demanding computation to solve the convex problem. In Section IV and Section V, we will analyze the asymptotic performance and propose BDMA transmission.

### IV. BDMA Transmission under Total Power Constraint

In this section, we consider the transmitted covariance matrix \( Q_k \) design under the total power constraint. When the number of LEDs tends to infinity, we design the transmit covariance matrix...
matrix, and compare the optimal performance with the case without transmit lens. Motivated by the asymptotic result, we consider beam domain transmission for non-asymptotic case.

A. Asymptotic Analysis

Let $R_k = h_k h_k^T$. The achievable sum-rate expression in (31) can be rewritten as

$$R_{\text{sum}} = \frac{1}{2} \sum_k \log \left( 1 + \frac{\text{tr}(R_k Q_k)}{1 + \text{tr}(R_k \sum_{k' \neq k} Q_{k'})} \right). \quad (40)$$

Recalling the limit (26), we can derive the asymptotic result of the sum-rate $R_{\text{sum}}$ as follows.

**Theorem 3:** As $N$ goes to infinity, the achievable sum-rate $R_{\text{sum}}$ tends to $\bar{R}_{\text{sum}}$, i.e.,

$$R_{\text{sum}} - \bar{R}_{\text{sum}} \to 0, \quad N \to \infty, \quad (41)$$

where $\bar{R}_{\text{sum}}$ is given by

$$\bar{R}_{\text{sum}} = \frac{1}{2} \sum_k \log \left( 1 + \frac{N^4 (g_k T c_{i_k j_k})^2 [Q_k]_{n_k n_k}}{1 + N^4 (g_k T c_{i_k j_k})^2 \sum_{k' \neq k} [Q_{k'॥n_{k' n_k}}]} \right), \quad (42)$$

where $n_k = (i_k - 1)N + j_k$, and $(i_k, j_k)$ satisfies $\phi_{k,i_k j_k} = 0$.

**Proof:** See Appendix B.

**Remark 4:** Theorem 3 presents that the achievable sum-rate $R_{\text{sum}}$ tends to asymptotic sum-rate $\bar{R}_{\text{sum}}$ when $N$ goes to infinity. For a large but finite $N$, $\bar{R}_{\text{sum}}$ is an approximation of the sum-rate $R_{\text{sum}}$. Moreover, from (42), the sum-rate $\bar{R}_{\text{sum}}$ only depends on the diagonal elements of $Q_k$. This means that for a large number of LEDs, only diagonal elements of $Q_k$ dominant the sum-rate.

Then, we consider the transmit design maximizing the asymptotic sum-rate $\bar{R}_{\text{sum}}$, which is given by

$$\max_{Q_1, Q_2, \ldots, Q_K} \bar{R}_{\text{sum}} \quad \text{s.t.} \quad \sum_k \text{tr}(Q_k) \leq P, \quad Q_k \succeq 0. \quad (43)$$

As UTs are distributed in the different positions, and thus, different UTs are illuminated by different LEDs, i.e., for $k_1 \neq k_2$, we have $(i_{k_1}, j_{k_1}) \neq (i_{k_2}, j_{k_2})$ and $n_{k_1} \neq n_{k_2}$. Thus, we can have the solution of problem (43) as in the following theorem.
**Theorem 4:** The optimal covariance matrix $Q_k$ is a diagonal matrix, and the diagonal elements are the water-filling solution as

$$[Q_k]_{nn} = \begin{cases} \left(\frac{1}{\nu} - \frac{1}{N^4(g_k T c_{i_k j_k})^2}\right)^+, & n = n_k, \\ 0, & n \neq n_k. \end{cases} \tag{44}$$

where $(x)^+ = \max\{x, 0\}$, $\nu$ is the Lagrange multiplier satisfying the condition

$$\sum_n \left(\frac{1}{\nu} - \frac{1}{N^4(g_k T c_{i_k j_k})^2}\right)^+ = P. \tag{45}$$

In the limit of large $N$, the optimal sum-rate $R_{\text{sum}}^o$ can be expressed as

$$R_{\text{sum}}^o - \frac{1}{2} \sum_k \log \left(1 + N^4(g_k T c_{i_k j_k})^2 [Q_k]_{n_k n_k}\right) = 0, \quad N \to \infty. \tag{46}$$

**Proof:** See Appendix C. \hfill \blacksquare

**Remark 5:** For a large $N$, the solution (44) can maximize the sum-rate $R_{\text{sum}}$, which is asymptotically optimal. The asymptotically optimal transmit covariance matrix $Q_k$ should be a diagonal matrix. This means that beams generated by different LEDs transmit independent signals, which is called beam domain transmission. Moreover, in the beam domain transmission, different beams serve different UTs and beams for different UTs are non-overlapping, which is called BDMA transmission [30]. This result shows that BDMA transmission is asymptotically optimal under total power constraint. Theorem 4 also shows that with a large number of LEDs, the performance of the MU-MISO system is asymptotically equal to the summation performance of $K$ SU-SISO systems, without any inter-user interference.

**B. Comparison with the Case without Transmit Lens**

Recall the channel vector without transmit lens $\tilde{h}_k$ in (11). Let

$$\tilde{R}_k = \tilde{h}_k \tilde{h}_k^T = \left(\tilde{I}(\theta_k) g_k\right)^2 1_{N^2 \times N^2}. \tag{47}$$

Then, the transmit design problem can be expressed as

$$\max_{Q_1, Q_2, \ldots, Q_K} \frac{1}{2} \sum_k \left(\log \left(1 + \text{tr} (\tilde{R}_k Q)\right) - \log \left(1 + \text{tr} (\tilde{R}_k (Q - Q_k))\right)\right)$$

s.t. $\sum_k \text{tr} (Q_k) \leq P, \quad Q_k \succeq 0. \tag{48}$
As $\tilde{R}_k$ for different UTs is $1_{N^2 \times N^2}$ with different coefficients $(I(\theta_k)g_k)^2$, the optimal $Q_k$ has the same structure, and can be expressed as $Q_k = \frac{1}{N^2} P_k 1_{N^2 \times N^2}$. Then, problem (48) can be rewritten as

$$\max_{P_1, P_2, \ldots, P_K} \frac{1}{2} \sum_k \left( \log \left( 1 + N^2 \left( \tilde{I}(\theta_k)g_k \right)^2 \left( \sum_{k'} P_{k'} \right) \right) - \log \left( 1 + N^2 \left( \tilde{I}(\theta_k)g_k \right)^2 \left( \sum_{k'} P_{k'} - P_k \right) \right) \right)$$

subject to $\sum_k P_k \leq P$, $P_k \succeq 0$. (49)

The first term only depends on the summation of $P_k$ and the second term is concave on $P_k$, the optimal solution can be obtained as

$$P_k = \begin{cases} P, & k = \arg \max_{k'} \left( \tilde{I}(\theta_{k'})g_{k'} \right)^2 \\ 0, & k \neq \arg \max_{k'} \left( \tilde{I}(\theta_{k'})g_{k'} \right)^2 \end{cases}.$$ (50)

Thus, the sum-rate of the conventional transmission without transmit lens is given by

$$\tilde{R}_\text{sum} = \frac{1}{2} \log \left( 1 + N^2 \left( \tilde{I}(\theta_k)g_k \right)^2 P \right).$$ (51)

Now we can compare the optimal sum-rate performances of transmission schemes with and without transmit lens for the asymptotic case. As $N$ increases to infinity, we can have the sum-rate ratio as

$$\lim_{N \to \infty} \frac{R_\text{sum}}{\tilde{R}_\text{sum}} = \sum_k \lim_{N \to \infty} \frac{1}{2} \log \left( 1 + N^4 (g_k T c_{i_k,j_k})^2 [Q_k]_{n_k,n_k} \right) = 2K.$$ (52)

From the above analysis, we can find that

1) Our proposed BDMA transmission can support multiple users simultaneously, while the conventional transmission without transmit lens can only serve one user.

2) For the asymptotic case ($N \to \infty$), the sum-rate of BDMA transmission is $2K$ times more than that of the conventional transmission without lens.

C. BDMA for Non-Asymptotic Case

Motivated by the asymptotically optimality of BDMA transmission, for non-asymptotic case, we remain focused on the beam domain transmission. Let $Q_k = U_k \Lambda_k U_k^H$, where $U_k$ is the eigenmatrix and $\Lambda_k$ is a diagonal matrix with the corresponding eigenvalues. For beam domain transmission, different LEDs transmit independent signals (i.e., $U_k = I$). Thus, the transmit
covariance matrix $Q_k$ design problem is degraded to a diagonal power allocation matrix $\Lambda_k$ optimization, which can be expressed as

$$
\max_{\Lambda_1, \Lambda_2, \ldots, \Lambda_K} \frac{1}{2} \sum_k \left( \log (1 + \text{tr}(R_k \Lambda)) - \log (1 + \text{tr}(R_k (\Lambda_k - \Lambda))) \right)
$$

s.t. $\text{tr}(\Lambda) \leq P$, $\Lambda_k \succeq 0,$ \hspace{1cm} (53)

where $\Lambda = \sum_k \Lambda_k$. Noting that the first term in the objective function is independent of $\Lambda_k$, we can derive orthogonality conditions of optimal power allocation as follows.

**Theorem 5:** The optimal power allocation for each UT under total power constraint should be non-overlapping (orthogonal) across beams, i.e., the solution of problem (53) satisfies the following conditions:

$$
\Lambda_{k_1} \Lambda_{k_2} = 0, \quad k_1 \neq k_2.
$$

**Proof:** See Appendix D.

**Remark 6:** With a large but limited number of transmit LEDs, if BS transmits independent signals in the beam domain, the optimal power allocation should be orthogonal between UTs. This means that one transmit beam only communicates with one UT, and transmit beams for different UTs should be non-overlapping. Thus, BDMA transmission is optimal for sum-rate maximization in the beam domain, which coincides with the previous results for massive MIMO RF communications [31].

Next, we propose a simple beam allocation algorithm which satisfies the orthogonality conditions (54). Consider equal power allocation for the selected transmit beams. Define $\Lambda_k = \eta B_k$, where $B_k$ is the beam allocation matrix with either 0 or 1 on the diagonal and $\eta$ is an auxiliary variable to satisfy the power constraint. Assume that the maximal number of beams for each UT is $B_m$. The power allocation optimization problem can be degraded to a beam allocation algorithm as

$$
\max_{\eta, B_1, B_2, \ldots, B_K} \frac{1}{2} \sum_k \left( \log \left( 1 + \text{tr} \left( \eta R_k \left( \sum_{k'} B_{k'} \right) \right) \right) - \log \left( 1 + \text{tr}\left( \eta R_k \left( \sum_{k' \neq k} B_{k'} \right) \right) \right) \right)
$$

s.t. $\eta \sum_k \text{tr}(B_k) = P$, $\text{tr}(B_k) \leq B_m$, $B_k(I - B_k) = 0.$ \hspace{1cm} (55)
Then, we propose a beam allocation algorithm, including the following steps:

1) Initialize $i = 1$, $R = 0$, $B_k = 0$.

2) Initialize $j = 1$, and $\Xi[d_1, d_2, \cdots, d_{N^2}]$ is the index of sorted diagonal elements of $R_k$.

3) Set $[B_i]_{d_j,d_j} = 1$ and calculate $\eta$ according to $\eta \sum_k \text{tr}(B_k) = P$ and $R_{\text{sum}}$ according to (55).

4) If $R_{\text{sum}} > R$, set $R = R_{\text{sum}}$, $j = j + 1$, and if $j \leq B_m$, return to Step 3; else, set $[B_i]_{d_j,d_j} = 0$ and recalculate $\eta$.

5) Set $i = i + 1$. If $i \leq K$, return to Step 2; else, stop the algorithm.

For the beam allocation algorithm, the complexity of the algorithm is $O(B_m K)$. Usually, the number of beams for one UT $B_m$ is much smaller than the total number of LEDs $N^2$. Thus, this algorithm has a low computational complexity.

V. BDMA TRANSMISSION UNDER PER LED POWER CONSTRAINT

We have analyzed the optimality of BDMA transmission under total power constraint. In this section, we consider the transmit covariance matrix $Q_k$ design under per LED power constraint.

A. Asymptotic Analysis

From Theorem 3 as $N$ tends to infinity, the achievable sum-rate tends to asymptotic sum-rate $\bar{R}_{\text{sum}}$. Then, we consider the transmit design maximizing the sum-rate $\bar{R}_{\text{sum}}$ under per LED power constraint, which can be expressed as

$$
\max_{Q_1, Q_2, \cdots, Q_K} \bar{R}_{\text{sum}}
$$

s.t. $Q_k \succeq 0$, $e_n^T Q_k e_n \leq p$, $n = 1, 2, \ldots, N^2$. \hfill (56)

Similarly, we can have the following asymptotically optimal transmit covariance matrices.

**Theorem 6:** The optimal transmit covariance matrix $Q_k$ is a diagonal matrix, and the diagonal elements can be expressed as

$$
[Q_k]_{nn} = \begin{cases} 
  p, & n = n_k \\
  0, & n \neq n_k 
\end{cases}
$$

Thus, in the limit of large $N$, the optimal sum-rate $R^o_{\text{sum}}$ can be expressed as

$$
R^o_{\text{sum}} = \frac{1}{2} \sum_k \log \left(1 + N^4 (g_k T c_{ik, jk})^2 p \right) \to 0, \quad N \to \infty.
$$

(58)
Remark 7: With asymptotically large LEDs, BDMA transmission can achieve the optimal performance under per LED power constraint. The asymptotic sum-rate is the summation rate of $K$ SU-SISO systems without inter-user interference.

B. Comparison with the Case without Transmit Lens

To compare the performance of BDMA with the conventional transmission without transmit lens, we first calculate the maximal sum-rate under per LED power constraint. Similar to the transmit design under total power constraint, the optimal $Q_k$ can be written as

$$Q_k = \frac{p_k 1_{N^2 \times N^2}}{\sum_{k'} p_{k'}}^2,$$

and the transmit design problem under per LED power constraint can be expressed as

$$\begin{align*}
\max_{P_1, P_2, \ldots, P_K} & \quad \frac{1}{2} \sum_k \left( \log \left( 1 + N^4 \left( \tilde{I}(\theta_k) g_k \right)^2 \left( \sum_{k'} p_{k'} \right) \right) - \log \left( 1 + N^4 \left( \tilde{I}(\theta_k) g_k \right)^2 \left( \sum_{k'} p_{k'} - p_k \right) \right) \right) \\
\text{s.t.} & \quad \sum_k p_k \leq p, \quad p_k \geq 0.
\end{align*}$$

(59)

As the first term in (59) depends on the total power $\sum_{k'} p_{k'}$ and the second term is concave on $p_k$, we can have the solution as

$$Q_k = \begin{cases} \frac{p 1_{N^2 \times N^2}}{\sum_{k'} p_{k'}}, & k = \arg \max_{k'} \left( \tilde{I}(\theta_{k'}) g_{k'} \right)^2 \\ 0, & k \neq \arg \max_{k'} \left( \tilde{I}(\theta_{k'}) g_{k'} \right)^2. \end{cases}$$

(60)

Thus, the optimal sum-rate of the conventional transmission without transmit lens is given by

$$\tilde{R}_\text{sum} = \frac{1}{2} \log \left( 1 + N^4 \left( \tilde{I}(\theta_k) g_k \right)^2 p \right).$$

(61)

Now, we can compare the performances of transmission schemes with and without transmit lens and have

$$\lim_{N \to \infty} \frac{R^o_k}{\tilde{R}_\text{sum}} = \sum_k \lim_{N \to \infty} \frac{1}{2} \log \left( 1 + N^4 \left( g_k T c_i(j_k) \right)^2 p \right) = K.$$

(62)

When the number of transmit LEDs goes to infinity, the sum-rate of our proposed BDMA transmission is $K$ times more than that of conventional transmission without transmit lens.
C. BDMA for Non-Asymptotic Case

Motivated by the asymptotic result, we consider the beam domain transmission, where each beam transmits independent signals and the transmit design problem is degraded to a power allocation problem, which can be expressed as

\[
\max_{\Lambda_k} \frac{1}{2} \sum_k \left( \log \left( I + \text{tr}(R_k \Lambda) \right) - \log \left( I + \text{tr}(R_k (\Lambda - \Lambda_k)) \right) \right)
\]

s.t. \( \Lambda_k \succeq 0 \), \( e_n^T \Lambda e_n \leq p \), \( n = 1, 2, \ldots, N^2 \). \hspace{1cm} (63)

For the power allocation problem, we can derive the following result.

**Theorem 7:** The optimal power allocation for each UT under per LED power constraint should be non-overlapping (orthogonal) across beams, i.e., the solution of problem (63) satisfies the following conditions:

\[
\Lambda_k \Lambda_{k'} = 0, \quad k_1 \neq k_2.
\]

\hspace{1cm} (64)

**Proof:** The proof is similar with that of Theorem 5 and is omitted here. \hfill \blacksquare

**Remark 8:** For beam domain transmission, it is optimal that different LEDs transmit signals to different UTs. Combined with Theorem 5 under both power constraints, BDMA transmission can achieve optimal performance.

Consider beam allocation with equal power on the selected beams. Define \( \Lambda_k = \eta B_k \), and the beam allocation algorithm is similar with the algorithm in Section IV. The difference is the power factor \( \eta \). Here, we set \( \eta = p \) under per LED power constraint.

VI. Simulation Results

In this section, we give some examples to illustrate the performance of our proposed optical massive MIMO communication approaches with transmit lens, in comparison with the conventional transmission without lens. We consider two typical optical massive MIMO communication scenarios. One is for optical communication in a small area, such as a meeting room, where BS equipped with \( 12 \times 12 \) LEDs serves 20 UTs. The room size is \( 5 \times 5 \) m, and the height is \( 3 \) m. UTs are randomly distributed in the room. The other is for a wide area, for example airport lounge or stadium, where BS with \( 80 \times 80 \) LEDs serves 484 UTs. The area size is \( 16 \times 16 \) m, and the height is \( 10 \) m. We consider two user distributions in the wide area: randomly
distributed and uniformly distributed. For uniformly distributed, the \((i, j)\)th UT position is given by
\[(X_i, Y_j) = (-7.37 + 0.67(i - 1), -7.37 + 0.67(j - 1)), \quad i, j = 1, 2, \cdots, 22. \tag{65}\]
We define the transmit signal-to-noise ratio (SNR) as \(\text{SNR} = P/\sigma^2\), and consider the same total transmit power under both power constraint, i.e., \(P = pN^2\).

Fig. 3 illustrates the beam pattern on the receive plane with \(4 \times 4\) and \(8 \times 8\) LEDs. As the number of LEDs grows, the spatial resolution for one beam increases, and the LED array can distinguish more directions to serve more users. Moreover, the channel gains of each beam with \(N = 8\) is larger than that with \(N = 4\).

![Fig. 3](image_url)

Fig. 3: Illumination intensity of the summation of total beams.

Fig. 4 compares the sum-rate in small area scenario. The RZF precoding and the BDMA using beam allocation algorithm (BDMA-BA) can approach the CCCP based optimal transmission (CCCP-OT) performance. The BDMA based on asymptotic design (BDMA-AD) as in (44) has a slight performance loss compared to BDMA-BA. For RZF and MRT under per LED power constraint, we multiply the precoding vectors by a power factor which makes the maximal power on one LED equal to power constraint. Thus, the sum-rate of RZF becomes lower than BDMA-BA. Under both power constraints, the sum-rate of the conventional transmission without lens (CT-w/o lens) is much smaller than that of BDMA-BA.
Fig. 4: Comparison of sum-rate for small area scenario, (a) under total power constraint, (b) under per LED power constraint.

Fig. 5: Comparison of sum-rate for wide area scenario, (a) under total power constraint, (b) under per LED power constraint.

Fig. 5 compares the sum-rate in wide area scenario. For uniform distribution, as UTs are separated enough, there is little interference between UTs. All the transmit schemes have similar performance. For random distribution, BDMA-BA can approach RZF under total power constraint, and achieve the best under per LED power constraint. For uniform distribution under total power constraint, the sum-rate can achieve 2000 bits/channel use, and rate of each UT is
approximate 4 bits/channel use.

Fig. 6: Comparison of sum-rate (a) and sum-rate ratio (b) for different number of LEDs $N$.

Fig. 6(a) compares the sum-rate of asymptotic and non-asymptotic cases as the number of transmit LEDs increases. Here, we consider that $K = 500$ UTs are randomly distributed. Under both power constraints, the BDMA-BA can approach the asymptotic sum-rate as in (46) and (58). MRT scheme has a significant performance loss under per LED power constraint. Recall the theoretic analysis of the sum-rate ratio in Section IV-B and Section V-B. Fig. 6(b) compares the sum-rate ratios as the number of transmit LEDs increases. Compared with the theoretic results, in (52), the ratio for total power constraint is $2K = 1000$, and in the figure, when $N \geq 70$, the ratio is larger than 900. Under per LED power constraint, in (62), the asymptotic ratio is $K = 500$, while in the simulation, the ratio is larger than 400.

VII. CONCLUSION

We have investigated beam domain optical wireless massive MIMO communications with a large number of transmitters and a transmit lens equipped at BS. We focused on LED transmitters in this work, and provided the optical channel model. With transmit lens and a large number of transmit LEDs, the channel vectors for different UTs become asymptotically orthogonal. For this channel model, we analyzed the performance of MRT/RZF linear precoding, and provided a transmit covariance matrix design based on CCCP. As the number of LEDs tends to infinity, we designed the transmit covariance matrix under both total and per LED power constraints.
For both power constraints, the optimal transmit policy is to transmit signals to different UTs by non-overlapping beams. Thus, BDMA transmission can achieve the optimal performance. Compared with the conventional transmission without transmit lens, the sum-rate of BDMA transmission is improved by $2K$ times under total power constraint, and $K$ times under per LED power constraint. In addition, for non-asymptotic case, we provided beam allocation algorithms to allocate non-overlapping beams to different UTs. Simulations illustrate the extremely high spectrum efficiency with massive LEDs and hundreds of UTs.

**Appendix A**

**Proof of Theorem 1**

As \( \log \det (\cdot) \) function is concave, for the \( i \)th and \( (i + 1) \)th iteration results, we have

\[
f(Q_{1}^{(i+1)}, Q_{2}^{(i+1)}, \ldots, Q_{K}^{(i+1)}) - g(Q_{1}^{(i+1)}, Q_{2}^{(i+1)}, \ldots, Q_{K}^{(i+1)}) \\
\geq f(Q_{1}^{(i)}, Q_{2}^{(i)}, \ldots, Q_{K}^{(i)}) - g(Q_{1}^{(i)}, Q_{2}^{(i)}, \ldots, Q_{K}^{(i)}) \\
- \sum_{k} \text{tr} \left( \left( \frac{\partial}{\partial Q_{k}} g(Q_{1}^{(i)}, Q_{2}^{(i)}, \ldots, Q_{K}^{(i)}) \right)^{T} \left( Q_{k}^{(i+1)} - Q_{k}^{(i)} \right) \right) \\
\geq f(Q_{1}^{(i)}, Q_{2}^{(i)}, \ldots, Q_{K}^{(i)}) - g(Q_{1}^{(i)}, Q_{2}^{(i)}, \ldots, Q_{K}^{(i)}) \\
- \sum_{k} \text{tr} \left( \left( \frac{\partial}{\partial Q_{k}} g(Q_{1}^{(i)}, Q_{2}^{(i)}, \ldots, Q_{K}^{(i)}) \right)^{T} \left( Q_{k}^{(i)} - Q_{k}^{(i)} \right) \right) \\
= f(Q_{1}^{(i)}, Q_{2}^{(i)}, \ldots, Q_{K}^{(i)}) - g(Q_{1}^{(i)}, Q_{2}^{(i)}, \ldots, Q_{K}^{(i)}). \tag{66}
\]

The objective function is monotonic and bounded. Moreover, the set of \( Q_{k} \) is closed and bounded. Invoking Theorem 4 in [32], we have

\[
\lim_{i \to \infty} \left( f(Q_{1}^{(i)}, Q_{2}^{(i)}, \ldots, Q_{K}^{(i)}) - g(Q_{1}^{(i)}, Q_{2}^{(i)}, \ldots, Q_{K}^{(i)}) \right) \\
= f(Q_{1}^{(*)}, Q_{2}^{(*)}, \ldots, Q_{K}^{(*)}) - g(Q_{1}^{(*)}, Q_{2}^{(*)}, \ldots, Q_{K}^{(*)}). \tag{67}
\]
where \( \{Q_k^{(s)}\} \) is a generalized fixed point. Then, there exists Lagrange multipliers \( \eta^{(s)} \) and \( \{A_k^{(s)}\}_{k=1}^K \) such that the following KKT conditions hold

\[
\frac{\partial}{\partial Q_k} f(Q_1^{(s)}, Q_2^{(s)}, \ldots, Q_K^{(s)}) - \frac{\partial}{\partial Q_k} g(Q_1^{(s)}, Q_2^{(s)}, \ldots, Q_K^{(s)}) - \eta^{(s)} + A_k^{(s)} = 0
\]

\[
\sum_k \text{tr} \left( Q_k^{(s)} \right) \leq P, \eta^{(s)} \geq 0, \eta^{(s)} \left( \sum_k \text{tr} \left( Q_k^{(s)} \right) - P \right) = 0,
\]

\[
Q_k^{(s)} \geq 0, A_k^{(s)} \geq 0, \text{tr} \left( A_k^{(s)} Q_k^{(s)} \right) = 0,
\]

which is exactly the KKT condition of problem (34), and therefore, \( \{Q_k^{(s)}\} \) is a stationary point of (34). This completes the proof. 

\[\blacksquare\]

**APPENDIX B**

**PROOF OF THEOREM 3**

The \((n_1, n_2)\)th element of \( R_k \) can be expressed as

\[
[R_k]_{n_1 n_2} = (g_k T)^2 I_{i_1 j_1} (\phi_{k,i_1j_1}) I_{i_2 j_2} (\phi_{k,i_2j_2})
\]

(69)

where \( n_1 = (i_1 - 1)N + j_1, n_2 = (i_2 - 1)N + j_2 \). Then, \( \text{tr}(R_k Q_k) \) is calculated as

\[
\text{tr}(R_k Q_k) = \sum_{n_1} \sum_{n_2} [R_k]_{n_1 n_2} [Q_k]_{n_2 n_1}
\]

\[
= \sum_{n_1} \sum_{n_2} (g_k T)^2 I_{i_1 j_1} (\phi_{k,i_1j_1}) I_{i_2 j_2} (\phi_{k,i_2j_2}) [Q_k]_{n_2 n_1}.
\]

(70)

As the number of LED increases, there exists LED \((i_k, j_k)\) satisfying \( \phi_{k,i_kj_k} = 0 \). Thus, we have

\[
\lim_{N \to \infty} \left( R_{\text{sum}} - \bar{R}_{\text{sum}} \right)
\]

\[
= \frac{1}{2} \sum_k \lim_{N \to \infty} \left( \log \left( 1 + \frac{\text{tr}(R_k Q_k)}{1 + \text{tr} \left( R_k \sum_{k' \neq k} Q_{k'} \right)} \right) - \log \left( 1 + \frac{N^4 (g_k T c_{i_kj_k})^2 [Q_k]_{n_k n_k}}{1 + N^4 (g_k T c_{i_kj_k})^2 \sum_{k' \neq k} [Q_{k'}]_{n_k n_k}} \right) \right)
\]

\[
= \frac{1}{2} \sum_k \lim_{N \to \infty} \log \left( \frac{1 + \text{tr} \left( R_k \sum_{k'} Q_{k'} \right)}{1 + N^4 (g_k T c_{i_kj_k})^2 \sum_{k'} [Q_{k'}]_{n_k n_k}} \frac{1 + N^4 (g_k T c_{i_kj_k})^2 \sum_{k' \neq k} [Q_{k'}]_{n_k n_k}}{1 + \text{tr} \left( R_k \sum_{k' \neq k} Q_{k'} \right)} \right).
\]

(71)
Thus, we have

\[
\lim_{N \to \infty} \frac{1 + \text{tr} (R_k \sum_{k'} Q_{k'})}{1 + N^4 (g_k T c_{i,j,k})^2 \sum_{k'} [Q_{k'}]_{n_k n_k}} = 1 + \lim_{N \to \infty} \text{tr} \left( R_k \sum_{k'} Q_{k'} \right)
\]

As \(\phi_{k,i,j,k} = 0\) and other \(\phi_{k,i,j} \geq 2\theta_{ij}\),

\[
\lim_{N \to \infty} (g_k T)^2 I_{i,j,i} (\phi_{k,i,j,i}) I_{i,j,j} (\phi_{k,i,j,j}) [Q_{k'}]_{n_2 n_1} = 0.
\] (73)

Thus, the limit (72) can be expressed as

\[
\lim_{N \to \infty} \frac{1 + \text{tr} (R_k \sum_{k'} Q_{k'})}{1 + N^4 (g_k T c_{i,j,k})^2 \sum_{k'} [Q_{k'}]_{n_k n_k}} = 1 + \lim_{N \to \infty} (g_k T I_{i,k,j,k}(0))^2 \sum_{k'} [Q_{k'}]_{n_k_n_k} = 1.
\] (74)

For the case of \(\sum_{k'} [Q_{k'}]_{n_k n_k} \neq 0\), we have

\[
\lim_{N \to \infty} \frac{1 + \text{tr} (R_k \sum_{k'} Q_{k'})}{1 + N^4 (g_k T c_{i,j,k})^2 \sum_{k'} [Q_{k'}]_{n_k n_k}} = \lim_{N \to \infty} \frac{1}{N^4} + \frac{1}{N^4} \text{tr} (R_k \sum_{k'} Q_{k'}).
\] (75)

The limit of \(\frac{1}{N^4} \text{tr} (R_k \sum_{k'} Q_{k'})\) exists, and can be calculated as

\[
\lim_{N \to \infty} \frac{1}{N^4} \text{tr} \left( R_k \sum_{k'} Q_{k'} \right) = \sum_{k'} \lim_{N \to \infty} \frac{1}{N^4} \sum_{n_1} \sum_{n_2} (g_k T)^2 I_{i,j,i} (\phi_{k,i,j,i}) I_{i,j,j} (\phi_{k,i,j,j}) [Q_{k'}]_{n_2 n_1}
\]

\[
= \sum_{k'} \lim_{N \to \infty} \frac{(I_{i,k,j,k}(0))^2}{N^4} (g_k T)^2 [Q_{k'}]_{n_k n_k}
\]

\[
= \sum_{k'} (g_k T c_{i,j,k})^2 [Q_{k'}]_{n_k n_k}
\] (76)

Then, the limit (75) is

\[
\lim_{N \to \infty} \frac{1 + \text{tr} (R_k \sum_{k'} Q_{k'})}{1 + N^4 (g_k T c_{i,j,k})^2 \sum_{k'} [Q_{k'}]_{n_k n_k}} = \sum_{k'} (g_k T c_{i,j,k})^2 [Q_{k'}]_{n_k n_k} = 1.
\] (77)

Thus, we have

\[
\lim_{N \to \infty} \frac{1 + \text{tr} (R_k \sum_{k'} Q_{k'})}{1 + N^4 (g_k T c_{i,j,k})^2 \sum_{k'} [Q_{k'}]_{n_k n_k}} = 1,
\] (78)
and similarly, the limit of the second term in log operation in (71) can be calculated as

\[
\lim_{N \to \infty} \frac{1 + N^4(g_k T c_{i_k, j_k})^2 \sum_{k' \neq k} [Q_{k'}]_{n_k n_k}}{1 + \text{tr} \left( R_k \sum_{k' \neq k} Q_{k'} \right)} = 1.
\] (79)

As log is a continuous function and \(\log(1) = 0\), the limit of \(R_{\text{sum}} - \overline{R}_{\text{sum}}\) is

\[
\lim_{N \to \infty} (R_{\text{sum}} - \overline{R}_{\text{sum}}) = 0.
\] (80)

This completes the proof. ■

APPENDIX C

PROOF OF THEOREM 4

As UTs are in different positions, different UTs receive signals from different LEDs, i.e., \(n_k \neq n_{k'}\), for \(k \neq k'\). From the asymptotic sum-rate \(\overline{R}_{\text{sum}}\) in (42), the optimal transmit covariance matrix \(Q_k\) should satisfy

\[
[Q_k]_{n_k' n_k} = 0, \quad n_k' \neq n_k.
\] (81)

Moreover, as \(Q_k\) is a positive-semidefinite matrix, \(Q_k\) must be a diagonal matrix. Under this condition, the problem (43) is reduced to

\[
\max_{Q_1, Q_2, \ldots, Q_K} \frac{1}{2} \sum_k \log \left( 1 + N^4(g_k T c_{i_k, j_k})^2 [Q_k]_{n_k n_k} \right)
\]

s.t.

\[
\sum_k [Q_k]_{n_k n_k} \leq P,
\]

\[
[Q_k]_{n_k n_k} \geq 0.
\] (82)

For this problem, we can have the water-filling result:

\[
[Q_k]_{n_k n_k} = \left( \frac{1}{\nu} - \frac{1}{N^4(g_k T c_{i_k, j_k})^2} \right)^+, \quad (x)^+ = \max\{x, 0\},
\] (83)

where \((x)^+ = \max\{x, 0\}\), \(\nu\) is the Lagrange multiplier with the condition

\[
\sum_n \left( \frac{1}{\nu} - \frac{1}{N^4(g_k T c_{i_k, j_k})^2} \right)^+ = \overline{P}.
\] (84)

According to Theorem 4, the asymptotically optimal sum-rate can be expressed as

\[
R_{\text{sum}}^\alpha = \frac{1}{2} \sum_k \log \left( 1 + N^4(g_k T c_{i_k, j_k})^2 [Q_k]_{n_k n_k} \right) \to 0.
\] (85)

This completes the proof. ■
Let \( \Lambda_k = \Lambda B_k \), where \( B_k \) is an auxiliary diagonal matrix satisfying \( \sum_k B_k = I \) and \( B_k \succeq 0 \).

As the first term in problem \((53)\) is independent of \( \Lambda_k \), problem \((53)\) is equal to

\[
\max_{\Lambda} \frac{1}{2} \sum_k \log (I + \text{tr}(R_k \Lambda)) - \left\{ \min_{B_1, \cdots, B_K} \frac{1}{2} \sum_k \log (I + \text{tr}(R_k \Lambda (I - B_k))) \right\}
\]

s.t. \( \sum_k B_k = I, \quad B_k \succeq 0, \)
\( \text{tr}(\Lambda) \leq P, \quad \Lambda \succeq 0. \) (86)

Consider the inner optimization problem on \( B_k \), which is given by

\[
\min_{B_1, \cdots, B_K} \frac{1}{2} \sum_k \log (1 + \text{tr}(R_k \Lambda) - \text{tr}(\tilde{R}_k \Lambda B_k))
\]

s.t. \( \sum_k B_k = I, \quad B_k \succeq 0. \) (87)

Define \( \tilde{R}_k = R_k \odot I \), where \( \tilde{R}_k \) is a diagonal matrix, which consists of the diagonal elements of \( R_k \). As \( \Lambda \) and \( B_k \) are diagonal matrices, we have

\[
\text{tr}(R_k \Lambda B_k) = \text{tr}(\tilde{R}_k \Lambda B_k). \quad (88)
\]

Let diagonal matrices \( A_k \) and \( C \) be Langrange multipliers, and we have the cost function as

\[
\mathcal{L} = \frac{1}{2} \sum_k \log \left( 1 + \text{tr}(R_k \Lambda) - \text{tr}(\tilde{R}_k \Lambda B_k) \right) + \text{tr} \left( C \left( \sum_k B_k - I \right) \right) - \text{tr}(A_k B_k). \quad (89)
\]

The Karush-Kuhn-Tucker (KKT) conditions for optimum \( B_k, A_k \) and \( C \) can be written as

\[
\frac{\partial}{\partial B_k} \mathcal{L} = -\frac{1}{2} \left( 1 + \text{tr}(R_k \Lambda) - \text{tr}(\tilde{R}_k \Lambda B_k) \right)^{-1} \tilde{R}_k \Lambda + C - A_k = 0
\]
\[
\sum_k B_k - I = 0, \quad A_k B_k = 0,
\]
\[
B_k \succeq 0, \quad A_k \succeq 0. \quad (90)
\]

Consider the KKT conditions for \( UT_{k_1} \) and \( UT_{k_2} \), and we have

\[
C = A_{k_1} + \frac{1}{2} \left( 1 + \text{tr}(R_{k_1} \Lambda) - \text{tr}(\tilde{R}_{k_1} \Lambda B_{k_1}) \right)^{-1} R_{k_1} \Lambda,
\]
\[
C = A_{k_2} + \frac{1}{2} \left( 1 + \text{tr}(R_{k_2} \Lambda) - \text{tr}(\tilde{R}_{k_2} \Lambda B_{k_2}) \right)^{-1} R_{k_2} \Lambda. \quad (91)
\]
For the $n$th diagonal elements, if $[B_{k_1}]_{nn} \neq 0$ and $[B_{k_2}]_{nn} \neq 0$, there exists $[A_{k_1}]_{nn} = 0$ and $[A_{k_2}]_{nn} = 0$. Thus, we have

$$[C]_{nn} = \frac{1}{2} \left[ \left( 1 + \text{tr}(R_{k_1}\Lambda) - \text{tr}(\tilde{R}_{k_1}AB_{k_1}) \right)^{-1} R_{k_1}\Lambda \right]_{nn}$$

$$= \frac{1}{2} \left[ \left( 1 + \text{tr}(R_{k_2}\Lambda) - \text{tr}(\tilde{R}_{k_2}AB_{k_2}) \right)^{-1} R_{k_2}\Lambda \right]_{nn}.$$  \hfill (92)

Due to the concavity of objective function in (87), the solution of equation (92) is a maximum point. Therefore, the solution of problem (87) should satisfies $[B_{k_1}]_{nn}[B_{k_2}]_{nn} = 0$. This means that for different UTs $k$ and $k'$, we have

$$B_kB_{k'} = 0.$$ \hfill (93)

Therefore we have (54). This completes the proof. \hfill \blacksquare

REFERENCES

[1] H. Elgala, R. Mesleh, and H. Haas, “Indoor optical wireless communication: Potential and state-of-the-art,” IEEE Commun. Mag., vol. 49, no. 9, pp. 56–62, Sept. 2011.

[2] F. Zafar, M. Bakaul, and R. Parthiban, “Laser-diode-based visible light communication: Toward gigabit class communication,” IEEE Commun. Mag., vol. 55, no. 2, pp. 144–151, Feb. 2017.

[3] A. H. Azhar, T. A. Tran, and D. O’Brien, “Demonstration of high-speed data transmission using MIMO-OFDM visible light communications,” in Proc. GLOBECOM Workshops, Jan. 2010, pp. 1052–1056.

[4] Q. Gao, C. Gong, S. Li, and Z. Xu, “DC-informative modulation for visible light communications under lighting constraints,” IEEE Wireless Commun., vol. 22, no. 2, pp. 54–60, Apr. 2015.

[5] T. Fath and H. Haas, “Performance comparison of MIMO techniques for optical wireless communications in indoor environments,” IEEE Trans. Commun., vol. 61, no. 61, pp. 733–742, Dec 2013.

[6] D. Kedar and S. Arnon, “Urban optical wireless communication networks: The main challenges and possible solutions,” IEEE Commun. Mag., vol. 42, no. 5, pp. S2 – S7, June 2004.

[7] A. Burton, H. L. Minh, Z. Ghassemlooy, and E. Bentley, “Experimental demonstration of 50-Mb/s visible light communications using $4 \times 4$ MIMO,” IEEE Photon. Technol. Lett., vol. 26, no. 9, pp. 945–948, Mar. 2014.

[8] L. Zeng, D. O’Brien, H. Minh, G. Faulkner, K. Lee, D. Jung, Y. Oh, and E. T. Won, “High data rate multiple input multiple output (MIMO) optical wireless communications using white LED lighting,” IEEE J. Sel. Areas Commun., vol. 27, no. 9, pp. 1654–1662, Dec. 2009.

[9] Z. Yu, R. J. Baxley, and G. T. Zhou, “Multi-user MISO broadcasting for indoor visible light communication,” in Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, May 2013, pp. 4849–4853.

[10] H. Ma, L. Lampe, and S. Hranilovic, “Robust MMSE linear precoding for visible light communication broadcasting systems,” in Proc. IEEE GLOBECOM Workshops, Dec. 2013, pp. 1081–1086.

[11] Y. Hong, J. Chen, Z. Wang, and C. Yu, “Performance of a precoding MIMO system for decentralized multiuser indoor visible light communications,” IEEE Photonics Journal, vol. 5, no. 4, July 2013.
[12] T. V. Pham, H. L. Minh, Z. Ghassemlooy, T. Hayashi, and A. T. Pham, “Sum-rate maximization of multi-user MIMO visible light communications,” in Proc. IEEE International Conference on Communication Workshop (ICCW), June 2015, pp. 1344–1349.

[13] T. V. Pham and A. T. Pham, “Max-min fairness and sum-rate maximization of MU-VLC local networks,” in Proc. IEEE GLOBECOM Workshops, Dec. 2015, pp. 1–6.

[14] R. Feng, M. Dai, H. Wang, and B. Chen, “Linear precoding for multiuser visible-light communication with field-of-view diversity,” IEEE Photonics Journal, vol. 8, no. 2, pp. 1–8, Mar. 2016.

[15] T. L. Marzetta, “Noncooperative cellular wireless with unlimited numbers of base station antennas,” IEEE Trans. Wireless Commun., vol. 9, no. 5, pp. 3590–3600, Nov. 2010.

[16] X. Gao, L. Dai, Y. Hu, and Y. Zhang, “Low-complexity signal detection for large-scale MIMO in optical wireless communications,” IEEE J. Sel. Areas Commun., vol. 33, no. 9, pp. 1903–1912, July 2015.

[17] G. Yun and M. Kavehrad, “Spot-diffusing and fly-eye receivers for indoor infrared wireless communications,” in Proc. IEEE International Conference on Selected Topics in Wireless Communications, June 1992, pp. 262–265.

[18] T. Chen, L. Liu, B. Tu, Z. Zheng, and W. Hu, “High-spatial-diversity imaging receiver using fisheye lens for indoor MIMO VLCs,” IEEE Photonics Technology Letters, vol. 26, no. 22, pp. 2260–2263, Nov. 2014.

[19] T. Q. Wang, Y. A. Sekercioglu, and J. Armstrong, “Analysis of an optical wireless receiver using a hemispherical lens with application in MIMO visible light communications,” Journal of Lightwave Technology, vol. 31, no. 11, pp. 1744–1754, June 2013.

[20] P. M. Butala, H. Elgala, and T. D. C. Little, “Performance of optical spatial modulation and spatial multiplexing with imaging receiver,” in Wireless Communications and NETWORKING Conference, 2014, pp. 394–399.

[21] B. E. Saleh, M. C. Teich, and B. E. Saleh, Fundamentals of photonics. Wiley New York, 1991, vol. 22.

[22] J. E. Greivenkamp, Field guide to geometrical optics. SPIE Press Bellingham, Washington, 2004, vol. 1.

[23] S. Cool, J. G. Pieters, K. C. Mertens, S. Mora, F. Cointault, J. Dubois, T. Van de Gucht, and J. Vangeyte, “Development of a high irradiance LED configuration for small field of view motion estimation of fertilizer particles,” Sensors, vol. 15, no. 11, pp. 28 627–28 645, 2015.

[24] I. Moreno, M. Avendaño-Alejo, and R. I. Tzonchev, “Designing light-emitting diode arrays for uniform near-field irradiance,” Applied optics, vol. 45, no. 10, pp. 2265–2272, Apr. 2006.

[25] T. Fath and H. Haas, “Performance comparison of MIMO techniques for optical wireless communications in indoor environments,” IEEE Trans. Commun., vol. 61, no. 2, pp. 733–742, Feb. 2013.

[26] B. Hassibi and B. M. Hochwald, “How much training is needed in multiple-antenna wireless links,” IEEE Trans. Inf. Theory, vol. 49, no. 4, pp. 951–963, Apr. 2003.

[27] T. L. Marzetta, “How much training is required for multiuser MIMO?” in Proc. IEEE Asilomar Conf. on Signals, Systems, and Computers (ACSSC), Pacific Grove, CA, Oct. 2006.

[28] S. S. Christensen, R. Agarwal, E. Carvalho, and J. M. Cioffi, “Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design,” IEEE Trans. Wireless Commun., vol. 7, no. 12, pp. 4792–4799, Dec. 2008.

[29] J. Shin and J. Moon, “Weighted sum rate maximizing transceiver design in MIMO interference channel,” in Proc. IEEE GLOBECOM, Dec. 2011, pp. 1–5.

[30] C. Sun, X. Gao, S. Jin, M. Matthaiou, Z. Ding, and C. Xiao, “Beam division multiple access transmission for massive MIMO communications,” IEEE Trans. Commun., vol. 63, no. 6, pp. 2170–2184, June 2015.
[31] C. Sun, X. Gao, and Z. Ding, “BDMA in multicell massive MIMO communications: Power allocation algorithms,” *IEEE Trans. Signal Process.*, vol. 65, no. 11, pp. 2962–2974, 2017.

[32] G. R. Lanckriet and B. K. Sriperumbudur, “On the convergence of the concave-convex procedure,” in *Proc. Advances in neural information processing systems*, 2009, pp. 1759–1767.