

# Performance Analysis of Large Intelligent Surfaces (LISs): Asymptotic Data Rate and Channel Hardening Effects

Minchae Jung, *Member, IEEE*, Walid Saad, *Fellow, IEEE*,

Youngrok Jang, Gyuyeol Kong, *Member, IEEE*, and

Sooyong Choi, *Member, IEEE*

## Abstract

The concept of a large intelligent surface (LIS) has recently emerged as a promising wireless communication paradigm that can exploit the entire surface of man-made structures for transmitting and receiving information. An LIS is expected to go beyond massive multiple-input multiple-output (MIMO) system, insofar as the desired channel can be modeled as a perfect line-of-sight. To understand the fundamental performance benefits, it is imperative to analyze its achievable data rate, under practical LIS environments and limitations. In this paper, an asymptotic analysis of the uplink data rate in an LIS-based large antenna-array system is presented. In particular, the asymptotic LIS rate is derived in a practical wireless environment where the estimated channel on LIS is subject to estimation errors, interference channels are spatially correlated Rician fading channels, and the LIS experiences hardware impairments. Moreover, the occurrence of the channel hardening effect is analyzed and the performance bound is asymptotically derived for the considered LIS system. The analytical asymptotic results are then shown to be in close agreement with the exact mutual information as the number of antennas and devices increase without bounds. Moreover, the derived ergodic rates show that hardware impairments, noise, and interference from estimation errors and the non-line-of-sight path become negligible as the number of antennas increases. Simulation results show that an LIS can achieve a performance that is

A preliminary version of this work was submitted to EuCNC 2020 [1].

M. Jung and W. Saad are with Wireless@VT, Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA 24061 USA (e-mail: hosaly@vt.edu; walids@vt.edu).

Y. Jang, G. Kong, and S. Choi are with School of Electrical Electronic Engineering, Yonsei University, Seoul 03722, Korea (e-mail: dynamics@yonsei.ac.kr; gykong@yonsei.ac.kr; csyong@yonsei.ac.kr).

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2016R1A6A3A11936259) and by the U.S. National Science Foundation under Grants IIS-1633363 and OAC-1638283.

comparable to conventional massive MIMO with improved reliability and a significantly reduced area for antenna deployment.

### Index Terms

large intelligent surface (LIS), large system analysis, channel estimation, ergodic rate, channel hardening effect.

## I. INTRODUCTION

Future man-made structures, such as buildings, roads, and walls, are expected to be electromagnetically active [2]–[5]. As such, these structures can be leveraged to provide wireless connectivity to emerging services, such as Internet of Things (IoT) applications [6]–[11], via the emerging concept of a large intelligent surface (LIS) [2]–[4]. If properly operated and deployed, LISs are expected to provide wireless connectivity to a plethora of IoT devices, such as sensors, vehicles, and surveillance cameras, through man-made structures. The LIS concept can be essentially viewed as a scaled-up version of conventional massive multiple-input and multiple-output (MIMO) systems. However, an LIS exhibits several key differences from massive MIMO systems. First, unlike conventional massive MIMO systems where transmission and reception are carried out via a base station (BS), an LIS can transmit and receive signals through all surfaces of man-made structures. This allows users in close proximity to communicate with an LIS and their transmission power levels can be set to values that are lower than those resulting from massive MIMO. This results in battery savings at the device and reduced interference levels in an LIS. Hence, higher data rates can be achieved because of the reduced interference levels, compared to massive MIMO systems. Second, LISs will be densely located in both indoor and outdoor spaces, making it possible to perform near-field communications through a line-of-sight (LOS) path [2]–[4]. Since an LOS path is highly correlated to the channel components between antennas, antenna spacing of greater than half a wavelength is meaningless in order to obtain full diversity gain. Consequently, an LIS can enable dense antenna arrays<sup>1</sup>. Finally, an LIS enables simpler channel estimation and feedback, compared to conventional massive MIMO systems that typically require channel state information (CSI) for hundreds of antennas. Since an extensive

<sup>1</sup>Indeed, distortions in radiation patterns can occur due to mutual coupling. However, we envision an LIS that can correct the distortions in the radiation patterns of any array with any antenna spacing, as in [12] and [13]. More practically, we consider hardware impairments which include residual coupling loss after decoupling of antenna-RF chains as in [14].

overhead for CSI acquisition resulting from pilot training and CSI feedback can be caused by the massive number of antennas, this overhead can seriously degrade the performance of massive MIMO systems [15], [16]. However, the desired channel of an LIS-based large antenna-array system is highly correlated with the LOS path, facilitating accuracy and simplicity in terms of channel estimation and feedback.

For these reasons, the use of an LIS has recently attracted attention in the wireless literature [2]–[5]. These recent works focus on addressing a number of LIS challenges that include performance analysis, estimating user location, user assignment, and power allocation. For instance, in [2], the authors derived the data rates of the optimal receiver and the matched filter (MF) in the uplink of an LIS-based system. Meanwhile, in [3], the authors obtained the Fisher-information and Cramer–Rao lower bound for user positions using the uplink signal for an LIS. In [4], the authors proposed optimal user assignments to select LIS units that maximize the sum rate and minimum individual rate. The authors in [5] proposed the use of LIS as a relay station for a massive MIMO system and developed a power allocation scheme to maximize energy efficiency. However, these previous studies have not considered practical LIS environments and their limitations, such as imperfect channel estimations and a user-specific channel model. For instance, both [2] and [3] assumed an LIS with an infinite surface area and considered that a single infinite LIS performs the MF over all devices. Moreover, [2]–[5] assumed perfect channel estimations for an LIS. Finally, all of the interference channels in [2]–[4] were assumed as following a LOS path, and [5] considered independent Rayleigh fading both in desired and interference channels. Given that LISs are densely located and devices are reasonably close to their target LISs, desired channels can be modeled as a LOS path whereas interference channels must be modeled depending on the distances between interfering devices and the target LISs. Therefore, the interference channel can be composed of a deterministic LOS path and spatially correlated non-line-of-sight (NLOS) path describing a device-specific spatially correlated multipath environment.

The main contribution of this paper is a rigorous asymptotic analysis of the uplink rate of an LIS-based large antenna-array system that considers a practical LIS environment and its limitations. In this regard, we assume that each device uses a desired surface that maps to a limited area of the entire LIS that we refer to as an LIS unit. Further, the MF procedure across the surface is assumed to be performed under *realistic channel estimation errors* and

*hardware impairments*, such as analog imperfectness, quantization errors, and residual coupling loss, are also considered. The interference channels are modeled as spatially correlated Rician fading channels, composed of a deterministic LOS path and stochastic NLOS path according to the distance between the interfering device and the target LIS unit. We then analyze the uplink ergodic rate of each device in presence of a large number of antennas and devices, and derive an asymptotic ergodic rate of LIS. This approximation allows the estimation of the uplink ergodic rate accurately without the need for extensive simulations, and then it enables to obtain optimal operating parameters such as an optimal size of an LIS unit. The asymptotic variance of the uplink rate is also derived in order to verify the occurrence of *channel hardening effect* theoretically, that is a particularly important phenomenon in large antenna-array systems such as massive MIMO and LIS [17]. Given that the channel hardening determines several practical implications such as system reliability, latency, and scheduling diversity, we analyze the occurrence of the channel hardening effect in an LIS-based system and compare it to a massive MIMO system. On the basis of the asymptotic ergodic rate and variance, the performance bound of an LIS-based system is obtained by using a scaling law for the uplink signal-to-interference-plus-noise ratio (SINR). We then show a particular operating characteristic of LIS whereby noise, estimation errors, hardware impairments, and NLOS interference become negligible compared to LOS interference from other devices. Our simulations show that LIS can be a promising technology beyond massive MIMO given that LIS can provide a comparable rate to massive MIMO, with improved reliability and a significantly reduced area for antenna deployment.

The rest of this paper is organized as follows. Section II presents the LIS-based system model. Section III describes the asymptotic analysis of the uplink data rate and Section IV describes the channel hardening effect and performance bound. Simulation results are provided in Section V to support and verify the analyses, and Section VI concludes the paper.

*Notations:* Throughout this paper, boldface upper- and lower-case symbols represent matrices and vectors respectively, and  $\mathbf{I}_M$  denotes a size- $M$  identity matrix. The conjugate, transpose, and Hermitian transpose operators are denoted by  $(\cdot)^*$ ,  $(\cdot)^T$ , and  $(\cdot)^H$ , respectively. The norm of a vector  $\mathbf{a}$  is denoted by  $|\mathbf{a}|$ .  $E[\cdot]$ ,  $\text{Var}[\cdot]$ , and  $\text{Cov}[\cdot]$  denote expectation, variance, and covariance operators, respectively.  $\mu_X = E[X]$  and  $\sigma_X^2 = \text{Var}[X]$  denote the mean and variance of a random variable  $X$ , respectively.  $\mathcal{O}(\cdot)$ ,  $\otimes$ , and  $\circ$  denote the big O notation, the Kronecker product, and the Hadamard product, respectively. The operators  $\text{Re}(\cdot)$  take the real part.  $\mathcal{CN}(m, \sigma^2)$  and

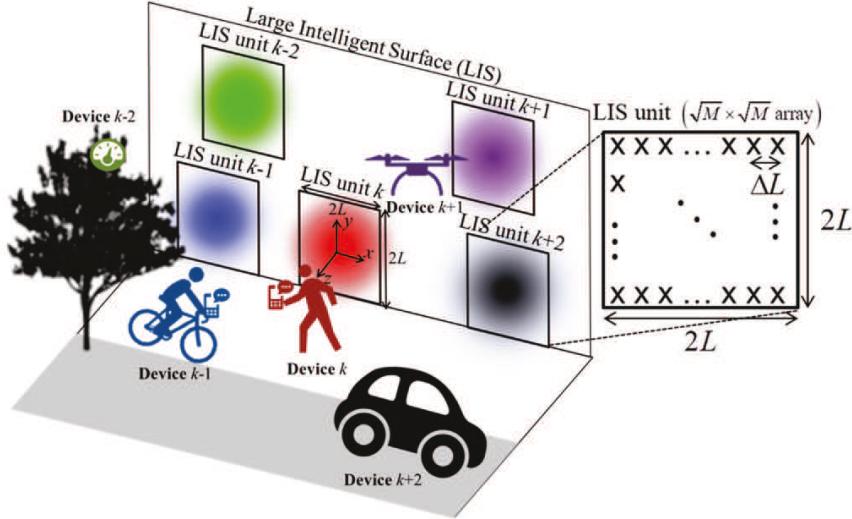


Fig. 1. Illustrative system model of the considered uplink LIS having multiple LIS units and serving  $K$  users.

$\chi_k^2$  denote a complex Gaussian distribution with mean  $m$  and variance  $\sigma^2$ , and a chi-square distribution with  $k$  degrees of freedom, respectively.

## II. SYSTEM MODEL

We consider an uplink LIS-based large antenna-array system that serves  $K$  single-antenna devices, as shown in Fig. 1. The LIS is located in a two-dimensional space along the  $xy$ -plane at  $z = 0$  in Cartesian coordinates. We define the notion of an *LIS unit* which corresponds to a subarea of the entire LIS and has a square shape with an area limited to  $2L \times 2L$  centered on the  $(x, y)$  coordinates of the corresponding device. Each LIS unit has a large number of antennas,  $M$ , distributed on its surface with antenna spacing of  $\Delta L$  in a rectangular lattice form. We assume that each LIS unit has its own signal processing unit for estimating the channel and detecting any data signal, as in [2]–[4], [18], and [19]. Each user device communicates with its corresponding LIS unit and controls the transmission power toward the center of the LIS unit according to target signal-to-noise-ratio (SNR), in order to avoid the near-far problem. We define the location of device  $k$  as  $(x_k, y_k, z_k)$ . Then, antenna  $m$  of LIS unit  $k$  will be located at  $(x_{km}^{\text{LIS}}, y_{km}^{\text{LIS}}, 0)$  where  $x_{km}^{\text{LIS}} \in [x_k - L, x_k + L]$  and  $y_{km}^{\text{LIS}} \in [y_k - L, y_k + L]$ . In contrast to the works in [2] and [3], that consider an infinite  $L$ , we consider a finite  $L$  which is a more practical assumption.

Depending on the location of the device, LIS units may overlap which, in turn, can seriously degrade the performance. To overcome this problem, we assume that the LIS allocates orthogonal resources among devices with similar  $(x, y)$  coordinates using an appropriate resource allocation and scheduling scheme. Therefore, we assume that each device communicates with a non-overlapping LIS unit.

### A. Wireless Channel Model

The desired channel  $\mathbf{h}_{kk} \in \mathbb{C}^M$  between device  $k$  and LIS unit  $k$  is assumed to be a LOS path. Then,  $\mathbf{h}_{kk}$  can be given by  $\mathbf{h}_{kk} = [\beta_{kk1}^L h_{kk1}, \dots, \beta_{kkM}^L h_{kkM}]^T$ , where  $\beta_{kkm}^L = \alpha_{kkm}^L l_{kkm}^L$  denotes a LOS channel gain between device  $k$  and antenna  $m$  of LIS unit  $k$ . Here,  $\alpha_{kkm}^L = \sqrt{\cos \theta_{kkm}}$  denotes the antenna gain and  $\theta_{kkm}$  is the azimuth angle-of-arrival between device  $k$  and antenna  $m$  of LIS unit  $k$ . Given that the antennas on each LIS unit are placed at different locations within the  $2L \times 2L$  square-shaped area,  $\theta_{kkm}$  has different values for different  $m$  considering the non-isotropic characteristic of an LIS. Since device  $k$  has a distance of  $z_k$  from the center of its target LIS unit, we obtain  $\cos \theta_{kkm} = z_k / d_{kkm}$ , where  $d_{kkm}$  denotes the distance between device  $k$  and antenna  $m$  of LIS unit  $k$ , given by  $d_{kkm} = \sqrt{(x_k - x_{km}^{\text{LIS}})^2 + (y_k - y_{km}^{\text{LIS}})^2 + z_k^2}$ . Also,  $l_{kkm}^L = 1 / \sqrt{4\pi d_{kkm}^2}$  is the free space path loss attenuation, and  $h_{kkm}$  is the LOS channel state between device  $k$  and antenna  $m$  of LIS unit  $k$ , obtained as  $h_{kkm} = \exp(-j2\pi d_{kkm}/\lambda)$ , where  $\lambda$  denotes the wavelength of a signal [20]. In fact, the desired channel can be generated by Rician fading composed of a deterministic LOS path and spatially correlated NLOS path. However, the signal from the NLOS path becomes negligible compared to the one from the LOS path as  $M$  increases, as will be proved in Section IV. Therefore, the desired channels are modeled as a *perfect LOS path* in the considered LIS as assumed in many prior works [2]–[4]. Further,  $\mathbf{h}_{jk} \in \mathbb{C}^M$  is the interference channel between device  $j$  and LIS unit  $k$ , expressed as a combination of LOS and NLOS:

$$\mathbf{h}_{jk} = \sqrt{\frac{\kappa_{jk}}{\kappa_{jk} + 1}} \mathbf{h}_{jk}^L + \sqrt{\frac{1}{\kappa_{jk} + 1}} \mathbf{h}_{jk}^{\text{NL}}, \quad (1)$$

where  $\kappa_{jk}$  is the Rician factor between device  $j$  and LIS unit  $k$ . Here,  $\mathbf{h}_{jk}^L \in \mathbb{C}^M$  is the deterministic LOS component from device  $j$  to LIS unit  $k$  given by  $\mathbf{h}_{jk}^L = [\beta_{j1}^L h_{jk1}, \dots, \beta_{jM}^L h_{jkM}]^T$ , where  $\beta_{jm}^L = \alpha_{jkm}^L l_{jkm}^L$  and  $h_{jkm}$  are LOS channel gain and state, respectively, between device  $j$  and antenna  $m$  of LIS unit  $k$ .  $\mathbf{h}_{jk}^{\text{NL}}$  is the correlated NLOS component defined as

$\mathbf{h}_{jk}^{\text{NL}} = [h_{jk1}^{\text{NL}}, \dots, h_{jkM}^{\text{NL}}]^T = \mathbf{R}_{jk}^{1/2} \mathbf{g}_{jk}$ , where  $\mathbf{R}_{jk} \in \mathbb{C}^{M \times P}$  is the deterministic correlation matrix from device  $j$  to LIS unit  $k$  and  $\mathbf{g}_{jk} = [g_{jk1}, \dots, g_{jkP}]^T \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_P)$  is an independent fast-fading channel vector. Here,  $P$  represents the number of dominant paths among all NLOS paths and is related to the amount of scattering in the wireless channel environment [21]. Since we consider an LIS located in a two-dimensional space along the  $xy$ -plane, it can be modeled as a uniform planar array (UPA) [22]. Given a UPA model,  $\mathbf{R}_{jk}$  is obtained as  $\mathbf{R}_{jk}^{1/2} = \mathbf{l}_{jk}^{\text{NL}} \mathbf{D}_{jk}$ , where  $\mathbf{D}_{jk} = [\alpha_{jk1}^{\text{NL}} \mathbf{d}(\phi_{jk1}^{\text{v}}, \phi_{jk1}^{\text{h}}), \dots, \alpha_{jkP}^{\text{NL}} \mathbf{d}(\phi_{jkP}^{\text{v}}, \phi_{jkP}^{\text{h}})]$  and  $\mathbf{l}_{jk}^{\text{NL}} = \text{diag}(l_{jk1}^{\text{NL}}, \dots, l_{jkM}^{\text{NL}})$  is a diagonal matrix including the path loss attenuation factors  $l_{jkM}^{\text{NL}} = d_{jkM}^{-\beta_{\text{PL}}/2}$  with a path loss exponent  $\beta_{\text{PL}}$ . Here,  $\mathbf{d}(\phi_{jkp}^{\text{v}}, \phi_{jkp}^{\text{h}}) \in \mathbb{C}^M$  represents NLOS path  $p$  at given angles of  $(\phi_{jkp}^{\text{v}}, \phi_{jkp}^{\text{h}})$ . By using a UPA model,  $\mathbf{d}(\phi_{jkp}^{\text{v}}, \phi_{jkp}^{\text{h}})$  will be given by:

$$\mathbf{d}(\phi_{jkp}^{\text{v}}, \phi_{jkp}^{\text{h}}) = \frac{1}{\sqrt{M}} \mathbf{d}_{\text{v}}(\phi_{jkp}^{\text{v}}) \otimes \mathbf{d}_{\text{h}}(\phi_{jkp}^{\text{h}}), \quad (2)$$

$$\mathbf{d}_{\text{v}}(\phi_{jkp}^{\text{v}}) = \left[ 1, e^{j \frac{2\pi \Delta L}{\lambda} \phi_{jkp}^{\text{v}}}, \dots, e^{j \frac{2\pi \Delta L}{\lambda} (\sqrt{M}-1) \phi_{jkp}^{\text{v}}} \right]^T, \quad (3)$$

$$\mathbf{d}_{\text{h}}(\phi_{jkp}^{\text{h}}) = \left[ 1, e^{j \frac{2\pi \Delta L}{\lambda} \phi_{jkp}^{\text{h}}}, \dots, e^{j \frac{2\pi \Delta L}{\lambda} (\sqrt{M}-1) \phi_{jkp}^{\text{h}}} \right]^T, \quad (4)$$

where  $\phi_{jkp}^{\text{v}} = \sin \theta_{jkp}^{\text{v}}$  and  $\phi_{jkp}^{\text{h}} = \sin \theta_{jkp}^{\text{h}} \cos \theta_{jkp}^{\text{v}}$  when the elevation and azimuth angles of path  $p$  between device  $j$  and LIS unit  $k$  are  $\theta_{jkp}^{\text{v}}$  and  $\theta_{jkp}^{\text{h}}$ , respectively [23].  $\alpha_{jkp}^{\text{NL}}$  indicates the antenna gain of path  $p$  which can be obtained by  $\alpha_{jkp}^{\text{NL}} = \sqrt{\cos \theta_{jkp}^{\text{v}} \cos \theta_{jkp}^{\text{h}}}$  where  $\theta_{jkp} \in [-\frac{\pi}{2}, \frac{\pi}{2}]$  and  $\theta_{jkp} \in \{\theta_{jkp}^{\text{v}}, \theta_{jkp}^{\text{h}}\}$ .

### B. Uplink Data Rate

The instantaneous uplink data rate of device  $k$  will be given by  $R_k = \log(1 + \gamma_k)$ , where  $\gamma_k$  is the instantaneous SINR of device  $k$  received at LIS unit  $k$ . The uplink signal received from all devices at LIS unit  $k$  is expressed as follows:

$$\mathbf{y}_k = \sqrt{\rho_k} \mathbf{h}_{kk} x_k + \sum_{j \neq k}^K \sqrt{\rho_j} \mathbf{h}_{jk} x_j + \mathbf{w}_k + \mathbf{n}_k, \quad (5)$$

where  $x_k$  and  $x_j$  are uplink transmit signals of devices  $k$  and  $j$ , respectively, assumed as independent Gaussian variables with zero means and unit variances. Further,  $\rho_k$  and  $\rho_j$  are the uplink transmit SNRs of devices  $k$  and  $j$ , respectively, and  $\mathbf{n}_k \in \mathbb{C}^M \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$  is the noise vector. Moreover,  $\mathbf{w}_k \in \mathbb{C}^M$  represents the residual noise caused by hardware impairments,

as given by

$$\mathbf{w}_k = \mathbf{c}_k \circ \left( \sqrt{\rho_k} \mathbf{h}_{kk} x_k + \sum_{j \neq k}^K \sqrt{\rho_j} \mathbf{h}_{jk} x_j \right), \quad (6)$$

where  $\mathbf{c}_k = [c_1, \dots, c_M]^T$  represents hardware impairments at LIS unit  $k$  which can be modeled using a Gaussian distribution such that  $\mathbf{c}_k \sim \mathcal{CN}(\mathbf{0}, \delta \mathbf{I}_M)$  [14], [24]. Therefore, in (5),  $\sqrt{\rho_k} \mathbf{h}_{kk} x_k$  is uplink desired signal of device  $k$ ,  $\sum_{j \neq k}^K \sqrt{\rho_j} \mathbf{h}_{jk} x_j$  is the aggregate uplink interference from other devices, and  $\mathbf{w}_k + \mathbf{n}_k$  is the sum of the noise components. We consider a linear receiver  $\mathbf{f}_k^H$  for signal detection. Then, the received signal at LIS unit  $k$  is obtained as

$$\mathbf{f}_k^H \mathbf{y}_k = \sqrt{\rho_k} \mathbf{f}_k^H \mathbf{h}_{kk} x_k + \sum_{j \neq k}^K \sqrt{\rho_j} \mathbf{f}_k^H \mathbf{h}_{jk} x_j + \mathbf{f}_k^H \mathbf{w}_k + \mathbf{f}_k^H \mathbf{n}_k. \quad (7)$$

We consider an MF receiver defined by  $\mathbf{f}_k = \hat{\mathbf{h}}_{kk}$  where  $\hat{\mathbf{h}}_{kk}$  is the estimated channel of  $\mathbf{h}_{kk}$ . In the case of perfect channel estimation,  $\mathbf{f}_k = \mathbf{h}_{kk}$ . Under the imperfect CSI results from an least square estimator, we have  $\mathbf{f}_k = \mathbf{h}_{kk} + \sqrt{\frac{\tau^2}{1-\tau^2}} \mathbf{e}_k$ , where parameter  $\tau_k \in [0, 1]$  represents the imperfectness of  $\hat{\mathbf{h}}_{kk}$ .  $\mathbf{e}_k = [\beta_{k1}^L e_{k1}, \dots, \beta_{kM}^L e_{kM}]^T \in \mathbb{C}^M$  denotes the estimation error vector uncorrelated with  $\mathbf{h}_{kk}$  and  $\mathbf{n}_k$ . The elements of the estimation error vector has independent random variables of  $e_{km} \sim \mathcal{CN}(0, 1)$ . Using (7), we can write the received SINR of device  $k$  at LIS unit  $k$  as follows:

$$\gamma_k = \frac{\rho_k (1 - \tau_k^2) |\mathbf{h}_{kk}|^4}{\rho_k \tau_k^2 |\mathbf{e}_k^H \mathbf{h}_{kk}|^2 + \sum_{j \neq k}^K \rho_j \left| \sqrt{1 - \tau_k^2} \mathbf{h}_{kk}^H \mathbf{h}_{jk} + \tau_k \mathbf{e}_k^H \mathbf{h}_{jk} \right|^2 + |\mathbf{f}_k^H \tilde{\mathbf{w}}_k|^2 + \left| \sqrt{1 - \tau_k^2} \mathbf{h}_{kk}^H + \tau_k \mathbf{e}_k^H \right|^2}, \quad (8)$$

where  $\tilde{\mathbf{w}}_k = [\tilde{w}_1, \dots, \tilde{w}_M]^T = \mathbf{c}_k \circ \sum_{i=1}^K \sqrt{\rho_i} \mathbf{h}_{ik}$ . Then, (8) can be simplified to

$$\gamma_k = \frac{\rho_k S_k (1 - \tau_k^2)}{I_k}, \quad (9)$$

where

$$S_k = |\mathbf{h}_{kk}|^4, \forall k \quad (10)$$

$$I_k = \rho_k \tau_k^2 X_k + \sum_{j \neq k}^K \rho_j Y_{jk} + Z_k, \forall k \quad (11)$$

and

$$X_k = |\mathbf{e}_k^H \mathbf{h}_{kk}|^2, \forall k \quad (12)$$

$$Y_{jk} = \left| \sqrt{1 - \tau_k^2} \mathbf{h}_{kk}^H \mathbf{h}_{jk} + \tau_k \mathbf{e}_k^H \mathbf{h}_{jk} \right|^2, \forall j, k \quad (13)$$

$$Z_k = Z_k^w + Z_k^n = \left| \sqrt{1 - \tau_k^2} \mathbf{h}_{kk}^H \tilde{\mathbf{w}}_k + \tau_k \mathbf{e}_k^H \tilde{\mathbf{w}}_k \right|^2 + \left| \sqrt{1 - \tau_k^2} \mathbf{h}_{kk}^H + \tau_k \mathbf{e}_k^H \right|^2, \forall k. \quad (14)$$

In fact, the considered LIS system is significantly different from a classical massive MIMO because of a *key difference in the SINR expression*. In the considered LIS system, the desired signal power,  $|\mathbf{h}_{kk}|^4$ , is calculated by the squared sum of the squared LOS channel gains over all antennas, i.e.,  $(\sum_{m=1}^M (\beta_{kkm}^L)^2)^2$ , and this is a deterministic value known at the LIS by measuring the signal strength of the reference signals. However, in a conventional massive MIMO system, this desired signal power,  $|\mathbf{h}_{kk}|^4$ , is not a deterministic value and cannot be known at the BS accurately because of an NLOS fading. Therefore, the BS can detect the desired signal using only the estimated CSI,  $\hat{\mathbf{h}}_{kk}$ , resulting in  $S_k = |\hat{\mathbf{h}}_{kk}|^4$  and  $X_k = |\hat{\mathbf{h}}_{kk}^H \mathbf{e}_k|^2$  as most prior studies on massive MIMO systems have considered (e.g., see [21], [25], and references therein). Given the uplink data rate  $R_k$ , we will analyze the moments of mutual information asymptotically as  $M$  and  $K$  increase without bounds.

### III. ASYMPTOTIC RATE ANALYSIS

We consider an LIS-based large antenna-array system composed of a large number of discrete antennas that are densely distributed on a contiguous LIS and each LIS unit occupies a subarea of the LIS with  $M$  antennas. Given a massive number of IoT devices will be connected via wireless communication systems in the near future [7], we present an asymptotic analysis of the data rate in an LIS-based system as  $M$  and  $K$  increase. In conventional massive MIMO systems, there is a relationship between  $M$  and  $K$  such that  $M/K \geq 1$  and  $M/K$  is constant as in [21] and [26]. In contrast, LIS enables wireless communications without any constraint on the relationship between  $M$  and  $K$ .

From (10), we can describe the desired signal power,  $S_k$ , as

$$S_k = |\mathbf{h}_{kk}|^4 = \left( \sum_m^M |\beta_{km}^L h_{kkm}|^2 \right)^2 = \left( \sum_m^M (\beta_{km}^L)^2 \right)^2. \quad (15)$$

$\sum_m^M (\beta_{km}^L)^2$  is the summation of the desired signal power received at LIS unit  $k$ , and this is equivalent to the summation of the power received within the ranges of  $-L \leq x \leq L$  and  $-L \leq y \leq L$  when the signal is transmitted from the location of  $(0, 0, z_k)$  in Cartesian coordinates [2]. Then,  $S_k$  converges as  $S_k - \bar{S}_k^2 \xrightarrow{M \rightarrow \infty} 0$  where we define

$$\begin{aligned} \bar{S}_k &= \frac{1}{4\pi\Delta L^2} \iint_{-L \leq (x,y) \leq L} \frac{z_k}{(x^2 + y^2 + z_k^2)^{\frac{3}{2}}} dx dy \\ &= \frac{1}{2\pi\Delta L^2} \int_{-L \leq y \leq L} \frac{z_k L}{(y^2 + z_k^2) \sqrt{L^2 + y^2 + z_k^2}} dy \\ &= \frac{1}{\pi\Delta L^2} \tan^{-1} \left( \frac{L^2}{z_k \sqrt{2L^2 + z_k^2}} \right) = \frac{p_k}{\pi\Delta L^2}, \forall k \end{aligned} \quad (16)$$

where  $p_k = \tan^{-1} \left( \frac{L^2}{z_k \sqrt{2L^2 + z_k^2}} \right)$ . Then, we have

$$S_k - \bar{p}_k \xrightarrow{M \rightarrow \infty} 0, \quad (17)$$

where  $\bar{p}_k = \frac{p_k^2}{\pi^2 \Delta L^4}$ . (17) shows that  $S_k$  converges to a constant value,  $\bar{p}_k$ , depending on  $z_k$  and  $L$ , and the total captured energy by the LIS can increase as  $M$  increases within the constrained physical area of LIS unit (i.e.,  $\Delta L^2$  decreases). In fact, the work in [27] showed that the total captured energy is limited by the product of the physical area used for deploying antennas and the channel's solid angle, and it remains unchanged as  $M$  increases within the constrained physical area. However, since the authors in [27] assumed perfect NLOS channels and obtained the array response by using the far-field approximation, [27] does not directly apply to our LIS system. In the considered LIS system, the increase of  $M$  indicates an increase of the number of LOS paths and this results in an increase in the number of spatial channels and the channel's solid angle. Then, the total captured energy by an LIS can increase as  $M$  increases within the constrained physical area of the LIS unit. Therefore, the mean and variance of  $R_k$  can be readily derived, as follows.

**Corollary 1.** The mean and variance of  $R_k$  can be respectively approximated as follows:

$$\mu_{R_k} \approx \log(1 + \mu_{\gamma_k}) - \frac{\sigma_{\gamma_k}^2}{2(1 + \mu_{\gamma_k})^2}, \quad (18)$$

$$\sigma_{R_k}^2 \approx \frac{\sigma_{\gamma_k}^2}{(1 + \mu_{\gamma_k})^2} - \frac{\sigma_{\gamma_k}^4}{4(1 + \mu_{\gamma_k})^4}, \quad (19)$$

where  $\mu_{\gamma_k}$  and  $\sigma_{\gamma_k}^2$  represent the mean and variance of the uplink SINR, respectively, which are likewise approximated as

$$\mu_{\gamma_k} \approx \rho_k S_k (1 - \tau_k^2) \left( \frac{1}{\mu_{I_k}} + \frac{\sigma_{I_k}^2}{\mu_{I_k}^3} \right), \quad (20)$$

$$\sigma_{\gamma_k}^2 \approx \rho_k^2 S_k^2 (1 - \tau_k^2)^2 \left( \frac{\sigma_{I_k}^2}{\mu_{I_k}^4} - \frac{\sigma_{I_k}^4}{\mu_{I_k}^6} \right), \quad (21)$$

*Proof:* From [28], the mean of a function  $f$  for a random variable  $X$  using Taylor expansions can be approximated by

$$E[f(X)] \approx f(\mu_X) + \frac{f''(\mu_X)}{2} \sigma_X^2. \quad (22)$$

The variance of a function  $f$  for a random variable  $X$  can be approximated as

$$\text{Var}[f(X)] \approx (f'(\mu_X))^2 \sigma_X^2 - \frac{(f''(\mu_X))^2 \sigma_X^4}{4}. \quad (23)$$

Since  $S_k$  is constant, (20) and (21) can be obtained when  $f(I_k) = \rho_k S_k (1 - \tau_k^2) / I_k$  from (9).

Similarly, (18) and (19) can be obtained when  $f(\gamma_k) = \log(1 + \gamma_k)$ .  $\blacksquare$

Corollary 1 shows that both the mean and variance of the uplink rate are determined exclusively by a random variable  $I_k$ . Based on the results from Corollary 1, the mean and variance of  $I_k$  will be analyzed asymptotically.

#### A. Asymptotic Analysis of $R_k$

We provide an asymptotic analysis of uplink data rate,  $R_k$ , by following three steps. Given that  $R_k$  exclusively depends on a random variable  $I_k$ , we first analyze the moments of random variables  $X_k$ ,  $Y_{jk}$ , and  $Z_k$  from (11). We then asymptotically obtain the asymptotic moments of  $I_k$  given the covariances between  $X_k$ ,  $Y_{jk}$ , and  $Z_k$ . We finally derive asymptotic moments of  $R_k$  from Corollary 1 using the derived asymptotic moments of  $I_k$ .

In order to obtain the moments of  $I_k$ , we first derive the following lemmas from the asymptotic analyses.

**Lemma 1.** The mean and variance of  $X_k$  follow  $\mu_{X_k} = \sum_m^M (\beta_{km}^L)^4$  and  $\sigma_{X_k}^2 = \left( \sum_m^M (\beta_{km}^L)^4 \right)^2$ , respectively.

*Proof:* The detailed proof is presented in Appendix A.  $\blacksquare$

**Lemma 2.** The mean and variance of  $Y_{jk}$  follow  $\mu_{Y_{jk}} - \bar{\mu}_{Y_{jk}} \xrightarrow[M \rightarrow \infty]{} 0$  and  $\sigma_{Y_{jk}}^2 - \bar{\sigma}_{Y_{jk}}^2 \xrightarrow[M \rightarrow \infty]{} 0$ , respectively, where

$$\bar{\mu}_{Y_{jk}} = s_{jk}^L + s_{jk}^{N1} + s_{jk}^{N2} + |\mu_{jk}^L|^2, \quad (24)$$

$$\bar{\sigma}_{Y_{jk}}^2 = (s_{jk}^L + s_{jk}^{N1} + s_{jk}^{N2})^2 + 2|\mu_{jk}^L|^2 (s_{jk}^L + s_{jk}^{N1} + s_{jk}^{N2}), \quad (25)$$

and

$$\mu_{jk}^L = \sqrt{\frac{\kappa_{jk}(1 - \tau_k^2)}{\kappa_{jk} + 1}} \mathbf{h}_{kk}^H \mathbf{h}_{jk}^L, \quad (26)$$

$$s_{jk}^L = \frac{\kappa_{jk}\tau_k^2}{\kappa_{jk} + 1} \sum_m^M (\beta_{km}^L \beta_{jm}^L)^2, \quad (27)$$

$$s_{jk}^{N1} = \frac{1 - \tau_k^2}{\kappa_{jk} + 1} \sum_p^P |\mathbf{h}_{kk}^H \mathbf{r}_{j kp}|^2, \quad (28)$$

$$s_{jk}^{N2} = \frac{\tau_k^2}{\kappa_{jk} + 1} \sum_{m,p}^{M,P} (\alpha_{j kp}^{\text{NL}} \beta_{km}^L l_{j km}^{\text{NL}})^2 / M. \quad (29)$$

*Proof:* The detailed proof is presented in Appendix B.  $\blacksquare$

**Lemma 3.** The mean and variance of  $Z_k$  follow  $\mu_{Z_k} - \bar{\mu}_{Z_k} \xrightarrow[M \rightarrow \infty]{} 0$  and  $\sigma_{Z_k}^2 - \bar{\sigma}_{Z_k}^2 \xrightarrow[M \rightarrow \infty]{} 0$ , respectively, where  $\bar{\mu}_{Z_k} = \sum_m^M (\beta_{km}^L)^2 + \mu_{Z_k^w}$ ,  $\bar{\sigma}_{Z_k}^2 = \tau_k^2 (2 - \tau_k^2) \sum_m^M (\beta_{km}^L)^4 + \sigma_{Z_k^w}^2$ , and

$$\mu_{Z_k^w} = \sum_m^M \left( \sigma_{z_{km}^{\text{wL}}}^2 + \sigma_{z_{km}^{\text{wR}}}^2 + 2\text{Re}(\omega_{km}^{\text{wLR}}) \right), \quad (30)$$

$$\sigma_{Z_k^w}^2 = \left( \sum_m^M \left( \sigma_{z_{km}^{\text{wL}}}^2 + \sigma_{z_{km}^{\text{wR}}}^2 + 2\text{Re}(\omega_{km}^{\text{wLR}}) \right) \right)^2. \quad (31)$$

The terms  $\sigma_{z_{km}^{\text{wL}}}^2$ ,  $\sigma_{z_{km}^{\text{wR}}}^2$ , and  $\omega_{km}^{\text{wLR}}$  are given in (70), (71), and (76), respectively.

*Proof:* The detailed proof is presented in Appendix C.  $\blacksquare$

$\sum_m^M (\beta_{km}^L)^2$  in Lemma 3 can be asymptotically obtained by  $\sqrt{\bar{p}_k}$  using (17). Similarly,  $\sum_m^M (\beta_{km}^L)^4$  in Lemma 1 and 3 can be obtained by  $\sum_m^M (\beta_{km}^L)^4 - \bar{q}_k \xrightarrow[M \rightarrow \infty]{} 0$ , where  $\bar{q}_k = \frac{q_k}{16\pi^2 \Delta L^2}$  and  $q_k$  is given by

$$\begin{aligned} q_k &= \iint_{-L \leq (x,y) \leq L} \frac{z_k^2}{(x^2 + y^2 + z_k^2)^3} dx dy \\ &= \frac{L^2}{(L^2 + z_k^2)(2L^2 + z_k^2)} + \frac{L(2L^2 + 3z_k^2)}{(z_k^2(L^2 + z_k^2))^{3/2}} \tan^{-1} \left( \frac{L}{\sqrt{L^2 + z_k^2}} \right). \end{aligned} \quad (32)$$

Next, we asymptotically derive the covariances between  $X_k$ ,  $Y_{jk}$ , and  $Z_k$ , and then the asymptotic mean and variance of  $I_k$  are obtained by the following lemma.

**Lemma 4.** The mean and variance of  $I_k$  follow  $\mu_{I_k} - \bar{\mu}_{I_k} \xrightarrow[M \rightarrow \infty]{} 0$  and  $\sigma_{I_k}^2 / M^2 - \bar{\sigma}_{I_k}^2 / M^2 \xrightarrow[M, K \rightarrow \infty]{} 0$ , respectively, where  $\bar{\mu}_{I_k}$  and  $\bar{\sigma}_{I_k}^2$  are given as

$$\bar{\mu}_{I_k} = \rho_k \tau_k^2 \bar{q}_k + \sum_{j \neq k}^K \rho_j \bar{\mu}_{Y_{jk}} + \sqrt{\bar{p}_k} + \mu_{Z_k^w}, \quad (33)$$

$$\bar{\sigma}_{I_k}^2 = \rho_k^2 \tau_k^4 \bar{q}_k^2 + \tau_k^2 (2 - \tau_k^2) \bar{q}_k + \sum_{j \neq k}^K \rho_j^2 \bar{\sigma}_{Y_{jk}}^2 + \sum_{i, j \neq k: i \neq j}^K \rho_i \rho_j \bar{\omega}_{ijk} + \sigma_{Z_k^w}^2, \quad (34)$$

where

$$\bar{\omega}_{ijk} = 2 \operatorname{Re} \left( \mu_{c_{ik}} \mu_{c_{jk}}^* \sum_m^M \mu_{a_{imk}}^* \mu_{a_{jmk}} \right), \quad (35)$$

$$\mu_{c_{tk}} = \sqrt{\frac{\kappa_{tk} (1 - \tau_k^2)}{\kappa_{tk} + 1}} \mathbf{h}_{kk}^H \mathbf{h}_{tk}^L, \quad (36)$$

$$\mu_{a_{tmk}} = \sqrt{\frac{\tau_k^2 \kappa_{tk}}{\kappa_{tk} + 1}} \beta_{km}^L \beta_{tm}^L h_{tkm}. \quad (37)$$

*Proof:* The detailed proof is presented in Appendix D.  $\blacksquare$

Lemma 4 shows that  $\bar{\mu}_{I_k}$  and  $\bar{\sigma}_{I_k}^2$  are deterministic values depending on locations of the devices and the correlation matrices. Therefore, we can approximate the mean and variance of  $R_k$  as follows.

**Theorem 1.** The mean and variance of  $R_k$  follow  $\mu_{R_k} - \bar{\mu}_{R_k} \xrightarrow[M \rightarrow \infty]{} 0$  and  $\sigma_{R_k}^2 - \bar{\sigma}_{R_k}^2 \xrightarrow[M, K \rightarrow \infty]{} 0$ , where  $\bar{\mu}_{R_k}$  and  $\bar{\sigma}_{R_k}^2$  are

$$\bar{\mu}_{R_k} = \log(1 + \bar{\mu}_{\gamma_k}) - \frac{\bar{\sigma}_{\gamma_k}^2}{2(1 + \bar{\mu}_{\gamma_k})^2}, \quad (38)$$

$$\bar{\sigma}_{R_k}^2 = \frac{\bar{\sigma}_{\gamma_k}^2}{(1 + \bar{\mu}_{\gamma_k})^2} - \frac{\bar{\sigma}_{\gamma_k}^4}{4(1 + \bar{\mu}_{\gamma_k})^4}. \quad (39)$$

*Proof:* By respectively replacing  $\mu_{I_k}$  and  $\sigma_{I_k}^2$  in (20) and (21) with  $\bar{\mu}_{I_k}$  and  $\bar{\sigma}_{I_k}^2$  from Lemma 4, we obtain the mean and variance of the asymptotic uplink SINR. Let us define the mean and variance of the asymptotic uplink SINR of device  $k$  as  $\bar{\mu}_{\gamma_k}$  and  $\bar{\sigma}_{\gamma_k}^2$ , respectively, then  $\bar{\mu}_{\gamma_k}$  and

$\bar{\sigma}_{\gamma_k}^2$  are also deterministic values, as follows:

$$\bar{\mu}_{\gamma_k} = \rho_k \bar{p}_k (1 - \tau_k^2) \left( \frac{1}{\bar{\mu}_{I_k}} + \frac{\bar{\sigma}_{I_k}^2}{\bar{\mu}_{I_k}^3} \right), \quad (40)$$

$$\bar{\sigma}_{\gamma_k}^2 = \rho_k^2 \bar{p}_k^2 (1 - \tau_k^2)^2 \left( \frac{\bar{\sigma}_{I_k}^2}{\bar{\mu}_{I_k}^4} - \frac{\bar{\sigma}_{I_k}^4}{\bar{\mu}_{I_k}^6} \right). \quad (41)$$

We can obtain  $\bar{\mu}_{R_k}$  and  $\bar{\sigma}_{R_k}^2$  by respectively replacing  $\mu_{\gamma_k}$  and  $\sigma_{\gamma_k}^2$  in (18) and (19) with  $\bar{\mu}_{\gamma_k}$  and  $\bar{\sigma}_{\gamma_k}^2$ , which completes the proof.  $\blacksquare$

We refer to  $\bar{\mu}_{R_k}$  and  $\bar{\sigma}_{R_k}^2$  as the asymptotic mean and variance of  $R_k$ , respectively. Given deterministic values of  $\bar{\mu}_{\gamma_k}$  and  $\bar{\sigma}_{\gamma_k}^2$ , Theorem 1 shows that the mean and variance of uplink data rate in an LIS-based large antenna-array system can be obtained based on deterministic values such as the locations of the devices and the correlation matrices. Then, we can evaluate the performance of an LIS-based system in terms of ergodic rate, reliability, and scheduling diversity, without extensive simulations. In particular, we can easily estimate the ergodic rate from (38), and verify system reliability and the scheduling diversity gain from (39). Furthermore, the results from Theorem 1 will be in close agreement with the moments of mutual information resulting from an actual LIS-based system as the number of devices and antennas increase.

#### IV. CHANNEL HARDENING EFFECT AND PERFORMANCE BOUND

The channel hardening effect is an important feature in large antenna-array systems whereby the variance of mutual information shrinks as the number of antennas grows [17]. Since a wireless system's reliability and scheduling diversity depend on the fluctuations of the mutual information, it is important to estimate the fluctuations that can be expected in a given large antenna-array systems such as an LIS. Given the importance of the channel hardening effect, next, we verify its occurrence in an LIS-based large antenna-array system and we then derive the performance bound of the ergodic rate.

Since  $M = (2L/\Delta L)^2$ , (9) is given by using (17) as

$$\bar{\gamma}_k = \frac{\rho_k (1 - \tau_k^2) \frac{p_k^2}{L^4 \pi^2}}{16 \bar{I}_k / M^2}, \quad (42)$$

where  $\bar{\gamma}_k$  denotes the asymptotic value of  $\gamma_k$  and  $\bar{I}_k$  denotes a random variable with a mean and variance of  $\bar{\mu}_{I_k}$  and  $\bar{\sigma}_{I_k}^2$  from Lemma 4, respectively. With  $M = (2L/\Delta L)^2$ ,  $\bar{\mu}_{I_k}$  and  $\bar{\sigma}_{I_k}^2$  are

represented by

$$\bar{\mu}_{I_k} = \frac{M\rho_k\tau_k^2q_k}{64\pi^2L^2} + \frac{Mp_k}{4\pi L^2} + \sum_{j \neq k}^K \rho_j \bar{\mu}_{Y_{jk}} + \mu_{Z_k^w}, \quad (43)$$

$$\bar{\sigma}_{I_k}^2 = \frac{M^2\rho_k^2\tau_k^4q_k^2}{4096\pi^4L^4} + \frac{M\tau_k^2q_k(2-\tau_k^2)}{64\pi^2L^2} + \sum_{j \neq k}^K \rho_j^2 \bar{\sigma}_{Y_{jk}}^2 + \sum_{i,j \neq k:i \neq j}^K \rho_i \rho_j \bar{\omega}_{ijk} + \sigma_{Z_k^w}^2, \quad (44)$$

We can observe in (42) that the mean and variance of  $\bar{\gamma}_k$  are determined by  $\bar{\mu}_{I_k}/M^2$  and  $\bar{\sigma}_{I_k}^2/M^4$ , respectively. Lemma 5 is used to determine the scaling laws of  $\bar{\mu}_{I_k}/M^2$  and  $\bar{\sigma}_{I_k}^2/M^4$  according to  $M$ .

**Lemma 5.** According to the scaling laws for  $M$ , the mean and variance of  $\bar{I}_k/M^2$  follow  $\bar{\mu}_{I_k}/M^2 - \hat{\mu}_{I_k} \xrightarrow[M \rightarrow \infty]{} 0$  and  $\bar{\sigma}_{I_k}^2/M^4 \xrightarrow[M \rightarrow \infty]{} 0$ , respectively, where  $\hat{\mu}_{I_k} = \sum_{j \neq k}^K \frac{\rho_j \kappa_{jk}(1-\tau_k^2)}{M^2(1+\kappa_{jk})} |\mathbf{h}_{kk}^H \mathbf{h}_{jk}^L|^2$ .

*Proof:* From (44),  $\bar{\sigma}_{I_k}^2/M^4$  is obtained as follows:

$$\frac{\bar{\sigma}_{I_k}^2}{M^4} = \frac{\rho_k^2\tau_k^4q_k^2}{4096\pi^4L^4M^2} + \frac{\tau_k^2q_k(2-\tau_k^2)}{64\pi^2L^2M^3} + \frac{\sigma_{Z_k^w}^2}{M^4} + \frac{\sum_{j \neq k}^K \rho_j^2 \bar{\sigma}_{Y_{jk}}^2 + \sum_{i,j \neq k:i \neq j}^K \rho_i \rho_j \bar{\omega}_{ijk}}{M^4}. \quad (45)$$

From the scaling laws for  $M$ , as discussed in [25],  $\frac{\rho_k^2\tau_k^4q_k^2}{4096\pi^4L^4M^2}$  and  $\frac{\tau_k^2q_k(2-\tau_k^2)}{64\pi^2L^2M^3}$  in (45) converge to zero as  $M$  goes to infinity.  $\sigma_{Z_k^w}^2$  in (45) is calculated by the squared sum of  $M$  elements from (31), and then  $\sigma_{Z_k^w}^2$  increases with  $\mathcal{O}(M^2)$  as  $M$  increases based on the scaling law. Hence,  $\sigma_{Z_k^w}^2/M^4$  converges to zero as  $M \rightarrow \infty$ . Also,  $\bar{\sigma}_{Y_{jk}}^2/M^4$  in (45) is given by using (25) as follows:

$$\frac{\bar{\sigma}_{Y_{jk}}^2}{M^4} = \left( \frac{s_{jk}^L + s_{jk}^{N1} + s_{jk}^{N2}}{M^2} \right)^2 + \frac{2|\mu_{jk}^L|^2 (s_{jk}^L + s_{jk}^{N1} + s_{jk}^{N2})}{M^4}. \quad (46)$$

In order to verify the scaling law of (46), we determine the scaling laws of  $s_{jk}^L$ ,  $s_{jk}^{N1}$ ,  $s_{jk}^{N2}$ , and  $|\mu_{jk}^L|^2$  according to  $M$ . From (27),  $s_{jk}^L$  is calculated by the sum of  $(\beta_{km}^L \beta_{jm}^L)^2$  over all  $m$  where  $m = 1, \dots, M$ . Consequently,  $s_{jk}^L$  increases with  $\mathcal{O}(M)$  as  $M$  increases. From (28),  $s_{jk}^{N1}$  is calculated by the sum of  $|\mathbf{h}_{kk}^H \mathbf{r}_{jkp}|^2$  over all  $p$  where  $p = 1, \dots, P$ . Given that  $\mathbf{r}_{jkp}$  is a correlation vector normalized by  $\sqrt{M}$  from (2) and  $\mathbf{h}_{kk}^H \mathbf{r}_{jkp}$  is calculated by the sum of  $M$  elements,  $s_{jk}^{N1}$  increases with  $\mathcal{O}(M)$  as  $M$  increases. From (29),  $s_{jk}^{N2}$  is calculated by the sum of  $(\alpha_{jkp}^{NL} \beta_{km}^L l_{jkm}^{NL})^2/M$  for all  $m$  and  $p$ . Thus,  $s_{jk}^{N2}$  follows  $\mathcal{O}(1)$  as  $M$  increases. From (26),  $|\mu_{jk}^L|^2$  is obtained from  $\mathbf{h}_{kk}^H \mathbf{h}_{jk}^L$  which is calculated by the sum of  $M$  elements. Therefore  $|\mu_{jk}^L|^2$  increases with  $\mathcal{O}(M^2)$  as  $M$  increases. Hence, (46) goes to zero as  $M \rightarrow \infty$  and we have

$\sum_{j \neq k}^K \rho_j^2 \bar{\sigma}_{Y_{jk}}^2 / M^4 \xrightarrow[M \rightarrow \infty]{} 0^2$ . Similarly,  $\bar{\omega}_{ij}$  from (35) increases with  $\mathcal{O}(M^3)$  as  $M$  increases. Then,  $\sum_{i,j \neq k: i \neq j}^K \rho_i \rho_j \bar{\omega}_{ijk} / M^4 \xrightarrow[M \rightarrow \infty]{} 0$  and (45) eventually converges to zero as  $M \rightarrow \infty$ .

From (43),  $\bar{\mu}_{I_k} / M^2$  is obtained by

$$\frac{\bar{\mu}_{I_k}}{M^2} = \frac{\rho_k \tau_k^2 q_k}{64\pi^2 L^2 M} + \frac{p_k}{4\pi L^2 M} + \frac{\sigma_{Z_k^w}^2}{M^2} + \sum_{j \neq k}^K \frac{\rho_j \bar{\mu}_{Y_{jk}}}{M^2}. \quad (47)$$

From the scaling laws for  $M$ ,  $\frac{\rho_k \tau_k^2 q_k}{64\pi^2 L^2 M}$  and  $\frac{p_k}{4\pi L^2 M}$  in (47) converge to zero as  $M \rightarrow \infty$ .  $\mu_{Z_k^w}^2$  in (47) is calculated by the sum of  $M$  elements from (30), and then  $\mu_{Z_k^w}$  increases with  $\mathcal{O}(M)$ , finally  $\mu_{Z_k^w}^2 / M^2$  converges to zero as  $M \rightarrow \infty$ . Also,  $\bar{\mu}_{Y_{jk}} / M^2$  in (47) is presented by using (24) as  $\frac{\bar{\mu}_{Y_{jk}}}{M^2} = \frac{s_{jk}^L + s_{jk}^{N1} + s_{jk}^{N2} + |\mu_{jk}^L|^2}{M^2}$ . From the scaling laws of  $s_{jk}^L$ ,  $s_{jk}^{N1}$ ,  $s_{jk}^{N2}$ , and  $|\mu_{jk}^L|^2$  according to  $M$ , we have  $\bar{\mu}_{Y_{jk}} / M^2 - |\mu_{jk}^L|^2 / M^2 \xrightarrow[M \rightarrow \infty]{} 0$ . Therefore, (47) is asymptotically obtained as  $\frac{\bar{\mu}_{I_k}}{M^2} - \sum_{j \neq k}^K \frac{\rho_j |\mu_{jk}^L|^2}{M^2} \xrightarrow[M \rightarrow \infty]{} 0$ , which completes the proof.  $\blacksquare$

Lemma 5 shows that  $\bar{I}_k / M^2$  does not converge to a random variable, but rather to a constant without any variance as  $M$  increases. Therefore, we can prove the following result related to the occurrence of the channel hardening effect and the performance bound of the uplink data rate.

**Theorem 2.** The asymptotic variance of  $R_k$  goes to zero as  $M \rightarrow \infty$ , and an asymptotic mean of  $R_k$  is given by  $\bar{\mu}_{R_k} - \hat{\mu}_{R_k} \xrightarrow[M \rightarrow \infty]{} 0$ , where  $\hat{\mu}_{R_k} = \log \left( 1 + \frac{p_k^2 \rho_k (1 - \tau_k^2)}{16L^4 \pi^2 \hat{\mu}_{I_k}} \right)$  is an asymptotic bound of the uplink data rate.

*Proof:* From Lemma 5,  $\bar{I}_k / M^2$  converges on a constant value of  $\hat{\mu}_{I_k}$  as  $M \rightarrow \infty$ . Then,  $\bar{\gamma}_k$  from (42) also converges on a constant value without any variance as  $M \rightarrow \infty$ , as given by  $\bar{\gamma}_k - \frac{p_k^2 \rho_k (1 - \tau_k^2)}{16L^4 \pi^2 \hat{\mu}_{I_k}} \xrightarrow[M \rightarrow \infty]{} 0$ . Therefore, the asymptotic uplink data rate,  $\bar{R}_k$ , converges in distribution to a constant value:  $\bar{R}_k - \log \left( 1 + \frac{p_k^2 \rho_k (1 - \tau_k^2)}{16L^4 \pi^2 \hat{\mu}_{I_k}} \right) \xrightarrow[M \rightarrow \infty]{} 0$ . Finally, the asymptotic mean and variance of  $R_k$  converge, respectively, as  $\bar{\mu}_{R_k} - \log \left( 1 + \frac{p_k^2 \rho_k (1 - \tau_k^2)}{16L^4 \pi^2 \hat{\mu}_{I_k}} \right) \xrightarrow[M \rightarrow \infty]{} 0$  and  $\bar{\sigma}_{R_k}^2 \xrightarrow[M, K \rightarrow \infty]{} 0$ .  $\blacksquare$

Theorem 2 shows that the channel fading of an LIS-based large antenna-array system behaves as a static channel and its impact on the uplink data rate becomes negligible as  $M$  increases. This shows that an LIS-based large antenna-array system is subject to the *channel hardening effect* resulting in several practical implications. First, an LIS-based system lacks scheduling diversity given that the fluctuations of the mutual information are small. Further, an LIS offers

<sup>2</sup> If the number of devices that dominantly transmits interference signals to a target LIS unit is as large as  $M$ , it will not converge to 0. However, there are actually far fewer such devices than  $M$  because the minimum distance of the  $(x, y)$  coordinates between adjacent devices is considered under the assumption that the LIS units do not overlap. Therefore, we assume that this value always converges to 0.

TABLE I  
SIMULATION PARAMETERS

Parameter	Value
Carrier frequency	3 GHz
Uplink target SNR	3 dB
Hardware impairments ( $\delta$ )	1
Channel imperfectness ( $\tau_k^2$ )	0.5
Length of LIS unit ( $2L$ )	0.5 m
Rician factor ( $\kappa$ [dB]) [29]	$13 - 0.03d$ [m]
LOS path loss model [20]	$11 + 20\log_{10}d$ [m]
NLOS path loss model [21]	$37\log_{10}d$ [m] ( $\beta_{PL} = 3.7$ )

an improved reliability insofar as it has a nearly deterministic data rate. Also, an LIS provides a low latency of having a deterministic data rate. Furthermore, Theorem 2 shows that the ergodic rate of an LIS converges to the asymptotic bound  $\hat{\mu}_{R_k}$  as  $M$  increases. We can observe that  $\hat{\mu}_{R_k}$  is a function of  $\hat{\mu}_{I_k}$  which depends exclusively on  $|\mathbf{h}_{kk}^H \mathbf{h}_{jk}^L|^2$ . Therefore, the asymptotic bound is only affected by the interference signals through the LOS path from other devices. *Hardware impairments, noise, and interference from estimation errors and the NLOS path become negligible* compared to LOS interference as  $M$  increases. If all of the interference is generated from the NLOS path, the asymptotic bound goes to infinity as  $M$  increases. Moreover, Lemma 5 and Theorem 2 show that the approximation gap resulting from the Taylor expansions in Corollary 1 goes to zero as  $M \rightarrow \infty$ . As  $M$  increases,  $\bar{\mu}_{I_k}$  and  $\bar{\sigma}_{I_k}^2$  follow, respectively,  $\mathcal{O}(M^2)$  and  $\mathcal{O}(M^3)$  as proved in Lemma 5, and thus,  $\bar{\sigma}_{I_k}^2 / \bar{\mu}_{I_k}^2$  follows  $\mathcal{O}(1/M)$  and eventually converges to zero. Hence, since the terms of a higher order than the second degree of the Taylor expansion become negligible compared to the first and second-order terms, the approximation gap resulting from Taylor expansions in (20) and (21) goes to zero as  $M \rightarrow \infty$ . Similarly,  $\bar{\sigma}_{\gamma_k}^2$  goes to zero as proved in Theorem 2 and the gap in (18) and (19) eventually goes to zero, as  $M \rightarrow \infty$ .

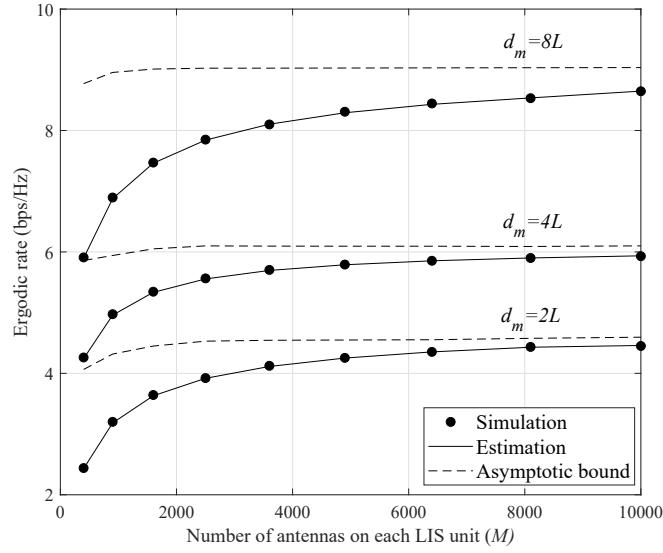


Fig. 2. Ergodic rates of an LIS-based system with LOS interference as a function of the number of antennas on the LIS unit.

## V. SIMULATION RESULTS AND ANALYSES

In this section, simulation results for the uplink rate in an LIS-based large antenna-array system are presented under a practical-sized environment with finite  $M$  and  $K$ . Further, the asymptotic analyses are compared with the numerical results obtained from Monte Carlo simulations (all simulations are statistically averaged over a large number of independent runs). The simulation parameters are provided in Table I and we do not consider shadowing given that the desired channel of LIS can be modeled as a perfect LOS. In our results, the labels “Estimation” and “Asymptotic bound” refer to the results obtained from Theorems 1 and 2, respectively, while the label “Simulation” captures a practical, simulated deployment of the considered LIS system.

In Figs. 2 and 3, Theorems 1 and 2 are verified in the following scenario. The devices are located at  $z = 1$  in parallel with the LIS on a two-dimensional plane. The devices are located in the ranges of  $-10 \leq x \leq 10$  and  $-10 \leq y \leq 10$  (in meters). The distance between the adjacent devices is set equally to  $d_m$  and the target device is located at  $(0, 0, 1)$ . Therefore, a total of 1681, 441, and 121 devices are located in a two-dimensional rectangular lattice form when  $d_m = 2L$ ,  $d_m = 4L$ , and  $d_m = 8L$ , respectively.

Figs. 2 and 3 compare the ergodic rates resulting from the simulations to the estimations from

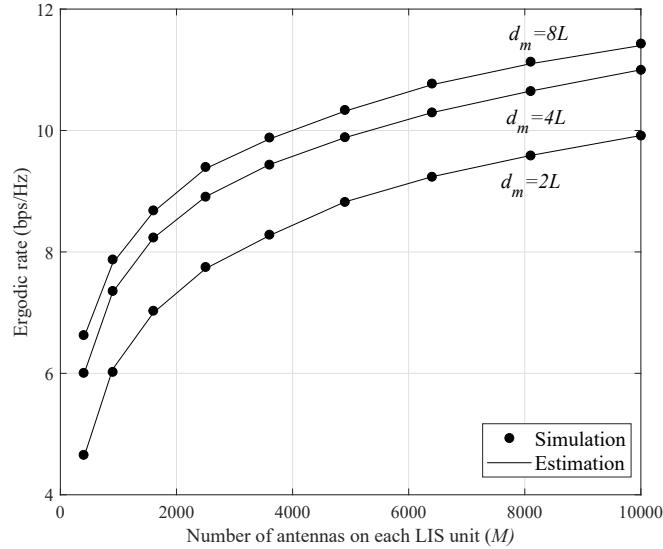


Fig. 3. Ergodic rates of an LIS-based system with NLOS interference as a function of the number of antennas on the LIS unit.

Theorem 1 as  $M$  increases. In Fig. 2, we assume that every interference signal from the other devices is generated entirely from the LOS path and in Fig. 3, it is generated entirely from the NLOS path. As shown in Figs. 2 and 3, the asymptotic mean values derived from Theorem 1 are close to the results of our simulations over the entire range of  $M$ . We can also observe from Fig. 2 that the ergodic rate converges to the asymptotic bound obtained from Theorem 2 as  $M$  increases. However, in Fig. 3, the ergodic rate goes to infinity as  $M$  increases without bound. As proved in Theorem 2, we can see that only the interference stemming from a LOS path affects the ergodic rate of an LIS-based system.

In Figs. 4–8, we consider that the devices are uniformly distributed within a three-dimensional space. In particular, these figures are generated for a scenario in which we randomly and uniformly deploy the devices in a  $4 \text{ m} \times 4 \text{ m} \times 2 \text{ m}$  space. Based on the 3GPP model in [29], the existence of a LOS path depends on the distance from the transmitter and receiver. The probability of LOS is then given as follows:

$$P_{jk}^{\text{LOS}} = \begin{cases} (d_C - d_{jk}) / d_C, & 0 < d_{jk} < d_C, \\ 0, & d_{jk} \geq d_C, \end{cases} \quad (48)$$

where  $d_{jk}$  is the distance in meters between device  $j$  and the center of LIS unit  $k$ , and  $d_C$  denotes

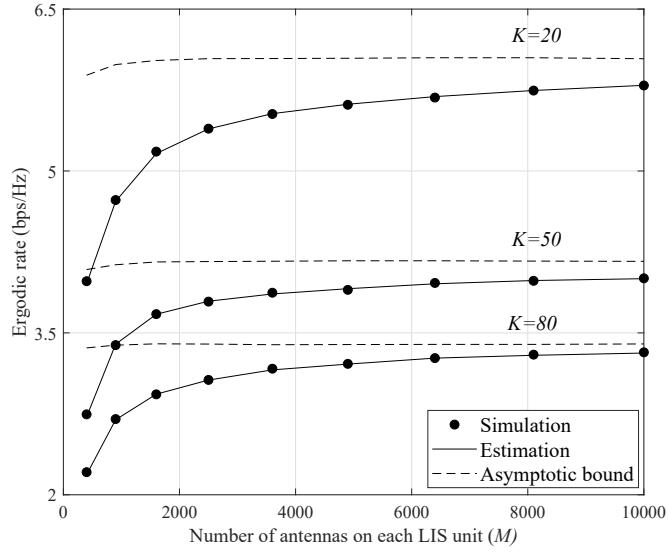


Fig. 4. Ergodic rates of an LIS-based system with randomly located devices as a function of the number of antennas on the LIS unit.

a cutoff point, which is typically set to 300 m in a cellular environment [29]. Since the antenna of a BS in a cellular system is located at a high altitude,  $d_C$  takes a large value such as 300 m. However, in an LIS environment, a relatively smaller  $d_C$  value is more reasonable. In Figs. 4–8,  $d_C = 10$  m is assumed to jointly consider the LOS and NLOS path simultaneously. If a LOS path occurs, the Rician factor,  $\kappa_{jk}$ , is calculated according to  $d_{jk}$  as per in Table I.

Figs. 4 and 5 show the ergodic rate and variance of the rate, respectively, as a function of  $M$ . The asymptotic mean and variance derived from Theorem 1 are close to the results of our simulations and the accuracy improves as  $M$  increases. In Fig. 4, the asymptotic means closely approximate the results of the simulations regardless of  $K$ , whereas the asymptotic variances approach the results of the simulations as  $K$  increases as shown in Fig. 5. Based on Theorem 1, the gap between the actual mean and asymptotic mean approaches zero as  $M$  increase, while the gap between the actual variance and asymptotic variance approaches zero as both  $M$  and  $K$  increase. Fig. 4 also shows that the ergodic rate gradually converges to the asymptotic bound obtained from Theorem 2. Given that the interference power increases as  $K$  increases, the ergodic rate gradually decreases as  $K$  increases.

Fig. 5 shows the channel hardening effect whereby the rate variance gradually converges

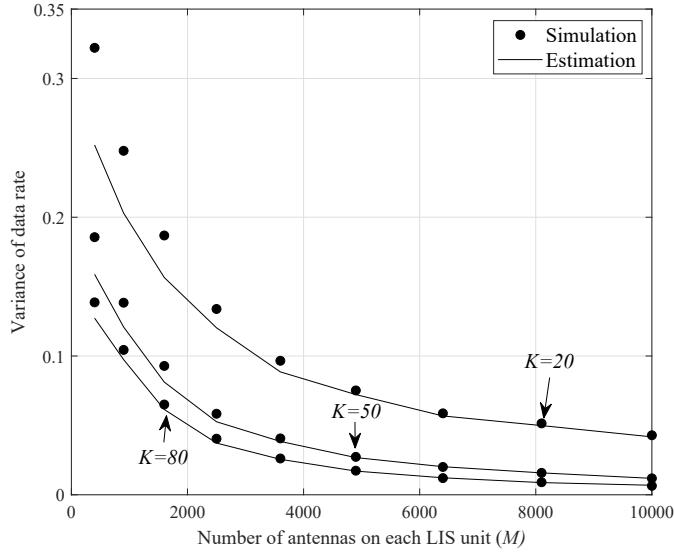


Fig. 5. Variances of uplink rates of an LIS-based system with randomly located devices as a function of the number of antennas on the LIS unit.

to zero as  $M$  increases. Moreover, the rate variance gradually decreases as  $K$  increases for a fixed  $M$ . From the scaling laws of  $\bar{\mu}_{I_k}$  and  $\bar{\sigma}_{I_k}^2$  in Lemma 4,  $\bar{\mu}_{I_k}$  and  $\bar{\sigma}_{I_k}^2$  follow  $\mathcal{O}(K)$  and  $\mathcal{O}(K^2)$ , respectively. Then,  $\bar{\sigma}_{R_k}^2$  follows  $\mathcal{O}(1/K^2)$  from (19)–(21). Therefore,  $\bar{\sigma}_{R_k}^2$  decreases as  $K$  increases.

In Figs. 6 and 7, we compare the performances of an LIS-based large antenna-array system and a massive MIMO system. We consider a multi-user massive MIMO system in which an MF is used for uplink signal detection. Since massive MIMO systems typically operate via far-field communications, we assume that every wireless signal is from an NLOS path and the distance from a device to all BS antennas is taken as equal [21]. For a massive MIMO system with a uniform linear array (ULA), the wireless channels from the NLOS path can be modeled using (1) with  $\kappa_{jk} = 0$  and

$$\mathbf{d}(\theta_{jkp}^h) = \frac{1}{\sqrt{M}} \left[ 1, e^{j\frac{2\pi\nu}{\lambda} \sin \theta_{jkp}^h}, \dots, e^{j\frac{2\pi\nu}{\lambda} (M-1) \sin \theta_{jkp}^h} \right]^T, \quad (49)$$

which is then applied to (2). Here,  $\nu$  is the antenna spacing of a massive MIMO system assuming  $\nu = \lambda/2$ . We also assume a single BS with  $M$  antennas and  $P = M/2$ , as in [21], and the same device distribution is considered as in the case of the LIS. For a fair comparison, we assume

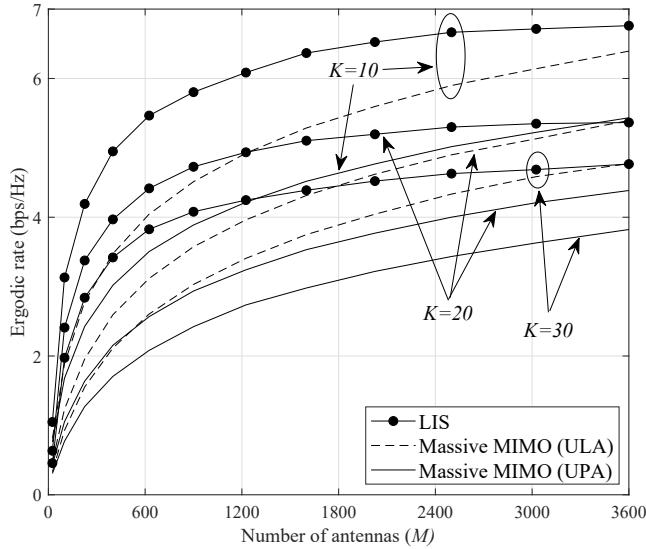


Fig. 6. Performance comparison between the ergodic rates of an LIS-based system and a massive MIMO system as a function of the number of antennas.

that the antenna gain is always equal to 1 in both cases (i.e., with massive MIMO and the LIS).

Fig. 6 compares the ergodic rates of an LIS-based large antenna-array system and a massive MIMO system as  $M$  increases. This figure shows that the ergodic rates resulting from LIS are higher than those resulting from massive MIMO in the range of practical-sized  $M$ , since the desired signal power of the LIS channel (i.e., LOS channel) is higher than that of the massive MIMO channel (i.e., NLOS channel). The performance gap decreases as  $M$  increases because the interference signal from the NLOS path becomes negligible, and eventually the massive MIMO system becomes an interference-free environment. When  $K = 30$ , an LIS shows about 2-fold increase in the ergodic rate compared to massive MIMO with ULA at  $M = 100$ , but two systems achieve a nearly equal ergodic rate at  $M = 3600$ . However, the increase of  $M$  indicates an increase in the physical area for deploying the massive antennas, whereas the physical area of the LIS remains constant at  $2L \times 2L$ . For example, the ergodic rates resulting from the LIS and the massive MIMO systems are almost equal when  $K = 30$  and  $M = 3600$ . The total physical length of the massive MIMO antennas is equal to 180 m under the assumption of a ULA with  $\lambda/2$ -spacing. Even if we consider a two-dimensional antenna deployment, a  $60 \times 60$  antenna-array occupies an area of roughly 9 m<sup>2</sup>. However, the LIS unit only occupies an area

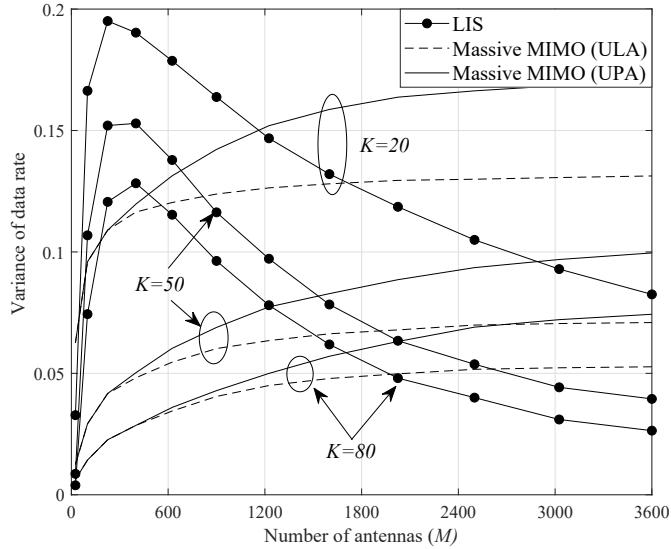


Fig. 7. Performance comparison between variances of uplink rates resulting from an LIS-based system and a massive MIMO system as a function of the number of antennas.

of  $0.25 \text{ m}^2$ . An interference-free MIMO environment is practically impossible given that the size of its array would have to be tremendous. This clearly shows the advantages of an LIS for space-intensive wireless communication.

Fig. 7 compares the variances of uplink rates resulting from an LIS-based large antenna-array system and a massive MIMO system as  $M$  increases. In Fig. 7, we plot the rate variance of LIS using the estimated value obtained from Theorem 1. We can observe that the rate variance of massive MIMO increases as  $M$  increases and then eventually converge to constant value exemplifying the so-called reduced channel hardening effect [17]. However, the rate variance of LIS converges to zero as  $M$  increases due to the channel hardening effect. Therefore, an LIS has improved reliability having a deterministic rate and it results in lower latency compared to a massive MIMO system.

Fig. 8 shows the ergodic rates resulting from an LIS-based large antenna-array system as a function of  $L$  when  $M = 100$ . In the LIS, the maximum SINR is achieved at the central antenna of the LIS unit and the SINR gradually decreases as the antenna moves from the center to the edge. Thus, the ergodic rate increases as  $L$  increases when  $L$  is small and decreases when  $L$  exceeds some threshold point. As shown in Fig. 8, the maximum ergodic rates can be achieved

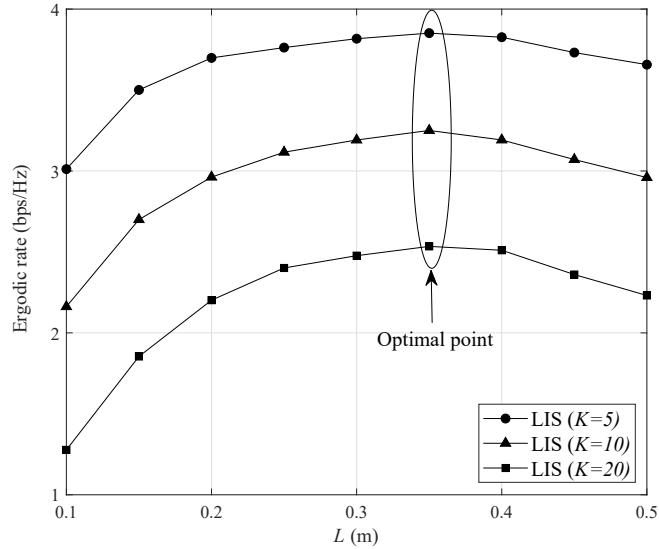


Fig. 8. Ergodic rates of an LIS-based system as a function of  $L$  when  $M = 100$ .

through optimal  $L$  values. Furthermore, an optimal  $L$  can be obtained numerically using the asymptotic analysis from Theorem 1 as given by  $L = 0.35$  m in Fig. 8.

## VI. CONCLUSIONS

In this paper, we have asymptotically analyzed the uplink data rate of an LIS-based large antenna-array system. We have derived the asymptotic moments of mutual information by considering a practical LIS environment in which a large LIS can be divided into smaller LIS units, each of which having a limited area. We have studied the uplink rate in presence of limitations such as hardware impairments, imperfect channel estimation, and interference that is generated by device-specific spatially correlated Rician fading. We have shown that our analyses can accurately determine the performance of an LIS analytically, without the need for extensive simulations. Furthermore, we have demonstrated that a channel hardening effect will occur in an LIS-based system. We have also derived the asymptotic bound of the uplink data rate and shown that hardware impairments, noise, and interference from channel estimation errors and the NLOS path become negligible as  $M$  increases. The simulation results have shown that the results of our asymptotic analyses agree with those resulting from extensive simulations, and the ergodic rate and the variance of rate respectively converge to the derived asymptotic bound and zero as

$M$  and  $K$  increase. Moreover, we have observed that an LIS enables reliable and space-intensive communication, which renders it a promising technology beyond massive MIMO systems. We expect that our asymptotic analyses will be invaluable to predict the theoretical performance of an LIS-based large antenna-array system when conducting system-level simulations and developing prototypes.

## APPENDIX A

### PROOF OF LEMMA 1

Given the definition of  $X_k = |\mathbf{e}_k^H \mathbf{h}_{kk}|^2$  from (12), we have

$$X_k = \left| \sum_m^M (\beta_{km}^L)^2 e_{km}^* h_{kkm} \right|^2. \quad (50)$$

Let us define  $\tilde{X}_{km} = (\beta_{km}^L)^2 e_{km}^* h_{kkm} \forall k, m$ . Then,  $\tilde{X}_{km} \sim \mathcal{CN}(0, (\beta_{km}^L)^4)$  and  $X_k$  can be described as follows:

$$X_k \sim \frac{1}{2} \sum_m^M (\beta_{km}^L)^4 \chi_2^2, \quad (51)$$

where  $\chi_k^2$  denotes the chi-square distribution with  $k$  degrees of freedom, which completes the proof.

## APPENDIX B

### PROOF OF LEMMA 2

Given the definition of  $Y_{jk} = \left| \sqrt{1 - \tau_k^2} \mathbf{h}_{kk}^H \mathbf{h}_{jk} + \tau_k \mathbf{e}_k^H \mathbf{h}_{jk} \right|^2$  from (13), we have

$$Y_{jk} = |\Sigma_{Y_{jk}}|^2 = |Y_{jk}^L + Y_{jk}^{N1} + Y_{jk}^{N2}|^2, \quad (52)$$

where we define  $\Sigma_{Y_{jk}} = Y_{jk}^L + Y_{jk}^{N1} + Y_{jk}^{N2} \forall j, k$  and

$$Y_{jk}^L = \sqrt{\frac{\kappa_{jk}}{\kappa_{jk} + 1}} \left( \sqrt{1 - \tau_k^2} \mathbf{h}_{kk}^H \mathbf{h}_{jk}^L + \tau_k \sum_m^M \beta_{km}^L \beta_{jm}^L e_{km}^* h_{jk} \right), \forall j, k \quad (53)$$

$$Y_{jk}^{N1} = \sqrt{\frac{1}{\kappa_{jk} + 1}} \left( \sqrt{1 - \tau_k^2} \mathbf{h}_{kk}^H \mathbf{R}_{jk}^{1/2} \mathbf{g}_{jk} \right) = \sqrt{\frac{1 - \tau_k^2}{\kappa_{jk} + 1}} \sum_p^P \mathbf{h}_{kk}^H \mathbf{r}_{jkp} g_{jk}, \forall j, k \quad (54)$$

$$Y_{jk}^{N2} = \sqrt{\frac{1}{\kappa_{jk} + 1}} \left( \tau_k \mathbf{e}_k^H \mathbf{R}_{jk}^{1/2} \mathbf{g}_{jk} \right) = \sqrt{\frac{\tau_k^2}{\kappa_{jk} + 1}} \sum_{m,p}^{M,P} \beta_{km}^L r_{jkmp} e_{km}^* g_{jk}, \forall j, k \quad (55)$$

where  $\mathbf{R}_{jk}^{1/2} = [\mathbf{r}_{jk1}, \dots, \mathbf{r}_{jkP}]$  and  $\mathbf{r}_{jkp} = [r_{jk1p}, \dots, r_{jkMp}]^T \forall j, k, p$ . We first prove that  $Y_{jk}^L$  and  $Y_{jk}^{N1}$  follow independent complex Gaussian distributions. We then asymptotically obtain the distribution of  $Y_{jk}^{N2}$  by using the law of large numbers for a large  $M$  and the Lyapunov central limit theorem (CLT). Finally, we asymptotically derive the mean and variance of  $Y_{jk}$ . Given that  $e_{km}$ ,  $g_{jkp}$ , and  $e_{km}g_{jkp}$  are independent of each other,  $Y_{jk}^L$ ,  $Y_{jk}^{N1}$ , and  $Y_{jk}^{N2}$  are independent random variables. Further, since  $e_{km}$  and  $g_{jkp}$  are standard complex Gaussian random variables respectively independent across  $m$  and  $p$ , we have  $Y_{jk}^L \sim \mathcal{CN}(\mu_{jk}^L, s_{jk}^L)$  and  $Y_{jk}^{N1} \sim \mathcal{CN}(0, s_{jk}^{N1})$ , where

$$\mu_{jk}^L = \sqrt{\frac{\kappa_{jk}(1 - \tau_k^2)}{\kappa_{jk} + 1}} \mathbf{h}_{kk}^H \mathbf{h}_{jk}^L, \quad (56)$$

$$s_{jk}^L = \frac{\kappa_{jk}\tau_k^2}{\kappa_{jk} + 1} \sum_m^M (\beta_{km}^L \beta_{jm}^L)^2, \quad (57)$$

$$s_{jk}^{N1} = \frac{1 - \tau_k^2}{\kappa_{jk} + 1} \sum_p^P |\mathbf{h}_{kk}^H \mathbf{r}_{jkp}|^2. \quad (58)$$

Here,  $\mathbf{h}_{kk}^H \mathbf{h}_{jk}^L$  in (56),  $\sum_m^M (\beta_{km}^L \beta_{jm}^L)^2$  in (57), and  $\sum_p^P |\mathbf{h}_{kk}^H \mathbf{r}_{jkp}|^2$  in (58) are deterministic values depending on the locations of the devices and the correlation matrices.

In order to obtain a random variable  $Y_{jk}^{N2}$ , we use the law of large numbers to approximate (55) for a large  $M$ .  $\sum_m^M \beta_{km}^L r_{jkmp} e_{km}^* g_{jkp}$  in (55) is thus expressed as follows:

$$\sum_m^M \beta_{km}^L r_{jkmp} e_{km}^* g_{jkp} = \alpha_{jkp}^{\text{NL}} \sum_m^M \beta_{km}^L l_{jkm}^{\text{NL}} d_{jkmp} e_{km}^* g_{jkp}, \quad (59)$$

where  $d_{jkmp}$  is element  $m$  of  $\mathbf{d}(\phi_{jkp}^v, \phi_{jkp}^h)$ . We define  $\tilde{Y}_{jkp} = \sum_m^M \beta_{km}^L l_{jkm}^{\text{NL}} d_{jkmp} y_{jkmp} \forall j, k, p$  where  $y_{jkmp} = e_{km}^* g_{jkp}$ . Then, the random variable  $\tilde{Y}_{jkp}$  follows Corollary 2.

**Corollary 2.** On the basis of the Lyapunov CLT, a random variable  $\tilde{Y}_{jkp}$  asymptotically follows a complex Gaussian distribution:

$$\sqrt{\frac{M}{\sum_m^M (\beta_{km}^L l_{jkm}^{\text{NL}})^2}} \tilde{Y}_{jkp} \xrightarrow[M \rightarrow \infty]{\text{d}} \mathcal{CN}(0, 1), \quad (60)$$

where “ $\xrightarrow[M \rightarrow \infty]{\text{d}}$ ” denotes the convergence in distribution.

*Proof:* In order to follow the Lyapunov CLT,  $y_{jkmp}$  should be independent random variable across  $m$  and the following Lyapunov's condition should be satisfied for some  $\delta > 0$  [30]:

$$\lim_{M \rightarrow \infty} \frac{1}{s_M^{2+\delta}} \sum_m^M \mathbb{E} \left[ |\beta_{km}^L l_{jkm}^{\text{NL}} d_{jkmp} y_{jkmp} - \mu_m|^{2+\delta} \right] = 0, \quad (61)$$

where  $s_M^2 = \sum_m^M \sigma_m^2$ , and  $\mu_m$  and  $\sigma_m^2$  are the mean and variance of the random variable  $\beta_{km}^L l_{jkm}^{\text{NL}} d_{jkmp} y_{jkmp}$ , respectively. Here,  $y_{jkmp} = e_{km}^* g_{jkp}$  is a random variable product of two independent random variables. Since  $e_{km}$  is an independent random variable across  $m$  and independent with  $g_{jkp}$ ,  $\{y_{jk1p}, \dots, y_{jkMp}\}$  is a sequence of independent random variables, each with zero mean and unit variance. Then, we can obtain  $\mu_m = 0$  and  $\sigma_m^2 = (\beta_{km}^L l_{jkm}^{\text{NL}})^2 |d_{jkmp}|^2 = (\beta_{km}^L l_{jkm}^{\text{NL}})^2 / M \forall m$ . We consider  $\delta = 2$ , such that (61) is obtained as follows:

$$\lim_{M \rightarrow \infty} \frac{\sum_m^M (\beta_{km}^L l_{jkm}^{\text{NL}})^4 \mathbb{E} [|y_{jkmp}|^4]}{\left( \sum_m^M (\beta_{km}^L l_{jkm}^{\text{NL}})^2 \right)^2} \stackrel{(a)}{\underset{M \rightarrow \infty}{\rightarrow}} \frac{4 \sum_m^M (\beta_{km}^L l_{jkm}^{\text{NL}})^4}{\left( \sum_m^M (\beta_{km}^L l_{jkm}^{\text{NL}})^2 \right)^2} = \lim_{M \rightarrow \infty} \frac{4 \bar{\alpha}_{jk}}{M \tilde{\alpha}_{jk}^2}, \quad (62)$$

where we define  $\bar{\alpha}_{jk} = \sum_m^M (\beta_{km}^L l_{jkm}^{\text{NL}})^4 / M$  and  $\tilde{\alpha}_{jk} = \sum_m^M (\beta_{km}^L l_{jkm}^{\text{NL}})^2 / M \forall j, k$ . (a) results from  $\mathbb{E} [|y_{jkmp}|^4] = 4 \forall m$  since  $y_{jkmp}$  is product of two independent random variables that follow an identical standard complex Gaussian distribution. Given that  $0 < \beta_{km}^L \leq 1$  and  $0 < l_{jkm}^{\text{NL}} \leq 1$ , we have  $0 < \bar{\alpha}_{jk} \leq 1$  and  $0 < \tilde{\alpha}_{jk} \leq 1$ . Therefore, (62) goes to zero, which completes the proof. ■

Based on Corollary 2, we have  $\frac{1}{\sqrt{s_{jk}^{\text{N2}}}} Y_{jk}^{\text{N2}} \xrightarrow[M \rightarrow \infty]{\text{d}} \mathcal{CN}(0, 1)$ , where

$$s_{jk}^{\text{N2}} = \frac{\tau_k^2}{\kappa_{jk} + 1} \sum_{m,p}^{M,P} (\alpha_{jkp}^{\text{NL}} \beta_{km}^L l_{jkm}^{\text{NL}})^2 / M. \quad (63)$$

Here,  $\alpha_{jkp}^{\text{NL}} \beta_{km}^L l_{jkm}^{\text{NL}}$  is a deterministic value depending on the locations of the devices and the correlation matrices. Given that  $Y_{jk}^{\text{L}}$ ,  $Y_{jk}^{\text{N1}}$ , and  $Y_{jk}^{\text{N2}}$  are independent of each other, we have  $\frac{1}{\sqrt{s_{jk}^{\text{L}} + s_{jk}^{\text{N1}} + s_{jk}^{\text{N2}}}} (\Sigma_{Y_{jk}} - \mu_{jk}^{\text{L}}) \xrightarrow[M \rightarrow \infty]{\text{d}} \mathcal{CN}(0, 1)$ . For a random variable  $Y = \sum_i |X_i|^2$ , where  $X_i$  is a complex Gaussian random variable independent across  $i$  such as  $X_i \sim \mathcal{CN}(m_i, \sigma_i^2)$ , the mean and variance of  $Y$  are respectively obtained by  $\mu_Y = \sum_i (\sigma_i^2 + |m_i|^2)$  and  $\sigma_Y^2 = \sum_i (\sigma_i^4 + 2|m_i|^2 \sigma_i^2)$ . Then,  $\bar{\mu}_{Y_{jk}}$  and  $\bar{\sigma}_{Y_{jk}}^2$  are ultimately obtained by

$$\bar{\mu}_{Y_{jk}} = s_{jk}^{\text{L}} + s_{jk}^{\text{N1}} + s_{jk}^{\text{N2}} + |\mu_{jk}^{\text{L}}|^2, \quad (64)$$

$$\bar{\sigma}_{Y_{jk}}^2 = (s_{jk}^{\text{L}} + s_{jk}^{\text{N1}} + s_{jk}^{\text{N2}})^2 + 2|\mu_{jk}^{\text{L}}|^2 (s_{jk}^{\text{L}} + s_{jk}^{\text{N1}} + s_{jk}^{\text{N2}}). \quad (65)$$

## APPENDIX C

### PROOF OF LEMMA 3

Given the definition of  $Z_k$  from (14), we have

$$Z_k^n = \left| \sqrt{1 - \tau_k^2} \mathbf{h}_{kk}^H + \tau_k \mathbf{e}_k^H \right|^2 = \sum_m^M |z_{km}^n|^2, \quad (66)$$

$$Z_k^w = \left| \sqrt{1 - \tau_k^2} \mathbf{h}_{kk}^H \tilde{\mathbf{w}}_k + \tau_k \mathbf{e}_k^H \tilde{\mathbf{w}}_k \right|^2 = \left| \sum_m^M z_{km}^w \right|^2, \quad (67)$$

where we define  $z_{km}^n = \sqrt{1 - \tau_k^2} \beta_{km}^L h_{kkm}^* + \tau_k \beta_{km}^L e_{km}^*$  and  $z_{km}^w = \sqrt{1 - \tau_k^2} \beta_{km}^L h_{kkm}^* \tilde{w}_m + \tau_k \beta_{km}^L e_{km}^* \tilde{w}_m = \tilde{w}_m z_{km}^n$ . Then,  $z_{km}^n$  follows a complex Gaussian distribution:

$$z_{km}^n \sim \mathcal{CN} \left( \sqrt{1 - \tau_k^2} \beta_{km}^L h_{kkm}^*, \tau_k^2 (\beta_{km}^L)^2 \right). \quad (68)$$

Then, the mean and variance of  $Z_k^n$  are  $\mu_{Z_k^n} = \sum_m^M (\beta_{km}^L)^2$  and  $\sigma_{Z_k^n}^2 = \tau_k^2 (2 - \tau_k^2) \sum_m^M (\beta_{km}^L)^4$ , respectively. Similarly, as proved in Corollary 2,  $\sum_m^M z_{km}^w$  follows a complex Gaussian distribution based on the Lyapunov CLT as  $M \rightarrow \infty$ . From (6),  $z_{km}^w$  can be represented as follows:

$$z_{km}^w = c_m z_{km}^n (\sqrt{\rho_k} \beta_{km}^L h_{kkm} + h_{km}^R), \quad (69)$$

where  $h_{km}^R = \sum_{j \neq k}^K \frac{\sqrt{\rho_j} (\sqrt{\kappa_{jk}} \beta_{jm}^L h_{jkm} + h_{jkm}^{\text{NL}})}{\sqrt{\kappa_{jk} + 1}}$ . We define  $z_{km}^{wL} = \sqrt{\rho_k} \beta_{km}^L h_{kkm} c_m z_{km}^n$  and  $z_{km}^{wR} = c_m z_{km}^n h_{km}^R$ . Since  $c_m$  is a zero-mean complex Gaussian random variable independent with  $z_{km}^n$ , the mean and variance of  $z_{km}^{wL}$  are obtained from [31], respectively, as  $\mu_{z_{km}^{wL}} = 0$  and

$$\sigma_{z_{km}^{wL}}^2 = \rho_k (\beta_{km}^L)^2 \text{Var} [c_m] \left( \left| \mathbb{E} [z_{km}^n] \right|^2 + \text{Var} [z_{km}^n] \right) = \delta \rho_k (\beta_{km}^L)^4. \quad (70)$$

Similarly, the mean and variance of  $z_{km}^{wR}$  can be obtained, respectively, as  $\mu_{z_{km}^{wR}} = 0$  and

$$\sigma_{z_{km}^{wR}}^2 = \delta \left( \left| \mathbb{E} [z_{km}^n h_{km}^R] \right|^2 + \text{Var} [z_{km}^n h_{km}^R] \right). \quad (71)$$

Since  $z_{km}^n$  and  $h_{km}^R$  are independent, we have

$$\left| \mathbb{E} [z_{km}^n h_{km}^R] \right|^2 = (1 - \tau_k^2) \left| \sum_{j \neq k}^K \sqrt{\frac{\rho_j \kappa_{jk}}{\kappa_{jk} + 1}} \beta_{km}^L \beta_{jm}^L h_{jkm}^* h_{jkm} \right|^2. \quad (72)$$

Also,  $\text{Var} [z_{km}^n h_{km}^R]$  in (71) is obtained in a manner similar to (70), as follows:

$$\text{Var} [z_{km}^n h_{km}^R] = (\beta_{km}^L)^2 \left( \sigma_{h_{km}^R}^2 + \tau_k^2 |\mu_{h_{km}^R}|^2 \right), \quad (73)$$

where

$$\mu_{h_{km}^R} = \sum_{j \neq k}^K \sqrt{\frac{\rho_j \kappa_{jk}}{\kappa_{jk} + 1}} \beta_{jm}^L h_{jkm}, \quad (74)$$

$$\sigma_{h_{km}^R}^2 = \sum_{j \neq k}^K \frac{\rho_j}{\kappa_{jk} + 1} \sum_p^P |r_{jkm}^p|^2. \quad (75)$$

(72) and (73) respectively show that  $|\mathbb{E}[z_{km}^n h_{km}^R]|^2$  and  $\text{Var}[z_{km}^n h_{km}^R]$  can be calculated based on the deterministic values such as the locations of the devices and the correlation matrices. Since  $z_{km}^{\text{wL}}$  and  $z_{km}^{\text{wR}}$  include a common random variable  $c_m z_{km}^n$ , they are dependent of each other with the following covariance:

$$\omega_{km}^{\text{wLR}} = \mathbb{E}[z_{km}^{\text{wL}} (z_{km}^{\text{wR}})^*] = \delta \sqrt{\rho_k} (\beta_{km}^L)^3 \sum_{j \neq k}^K \sqrt{\frac{\rho_j \kappa_{jk}}{\kappa_{jk} + 1}} \beta_{jm}^L h_{kkm} h_{jkm}^*. \quad (76)$$

Note that  $z_{km}^{\text{wL}}$  and  $z_{km}^{\text{wR}}$  are independent random variables across  $m$ . Therefore, we can finally obtain the mean and variance of  $Z_k^w$  respectively as follows:

$$\mu_{Z_k^w} = \sum_m^M \left( \sigma_{z_{km}^{\text{wL}}}^2 + \sigma_{z_{km}^{\text{wR}}}^2 + 2\text{Re}(\omega_{km}^{\text{wLR}}) \right), \quad (77)$$

$$\sigma_{Z_k^w}^2 = \left( \sum_m^M \left( \sigma_{z_{km}^{\text{wL}}}^2 + \sigma_{z_{km}^{\text{wR}}}^2 + 2\text{Re}(\omega_{km}^{\text{wLR}}) \right) \right)^2, \quad (78)$$

which completes the proof.

## APPENDIX D

### PROOF OF LEMMA 4

Given the definition of  $I_k = \rho_k \tau_k^2 X_k + \sum_{j \neq k}^K \rho_j Y_{jk} + Z_k$  from (11), we have

$$I_k = \rho_k \tau_k^2 |\mathbf{e}_k^H \mathbf{h}_{kk}|^2 + \sum_{j \neq k}^K \rho_j \left| \sqrt{1 - \tau_k^2} \mathbf{h}_{kk}^H \mathbf{h}_{jk} + \tau_k \mathbf{e}_k^H \mathbf{h}_{jk} \right|^2 + \left| \sqrt{1 - \tau_k^2} \mathbf{h}_{kk}^H + \tau_k \mathbf{e}_k^H \right|^2. \quad (79)$$

Based on Lemmas 1–3,  $\bar{\mu}_{I_k}$  can be obtained as  $\bar{\mu}_{I_k} = \rho_k \tau_k^2 \bar{q}_k + \sum_{j \neq k}^K \rho_j \bar{\mu}_{Y_{jk}} + \sqrt{p_k}$ . We can observe from (79) that  $X_k$ ,  $\sum_j \rho_j Y_{jk}$ , and  $Z_k$  are function of a common random variable vector  $\mathbf{e}_k$ . Similarly,  $Y_{ik}$  and  $Y_{jk} \forall i \neq j$  also include a common random variable vector  $\mathbf{e}_k$ . As  $K$  increases, the sum of covariances between  $Y_{ik}$  and  $Y_{jk} \forall i \neq j$  becomes dominant and the covariances between  $X_k$ ,  $\sum_j \rho_j Y_{jk}$ , and  $Z_k$  becomes negligible. Then, the variance of  $I_k$  can be asymptotically obtained based on Lemmas 1–3 as  $\sigma_{I_k}^2 / M^2 - \bar{\sigma}_{I_k}^2 / M^2 \xrightarrow[M, K \rightarrow \infty]{} 0$ , where

$$\bar{\sigma}_{I_k}^2 = \rho_k^2 \tau_k^4 \bar{q}_k^2 + \tau_k^2 (2 - \tau_k^2) \bar{q}_k + \sum_{j \neq k}^K \rho_j^2 \bar{\sigma}_{Y_{jk}}^2 + \sum_{i,j \neq k: i \neq j}^K \rho_i \rho_j \bar{\omega}_{ijk}. \quad (80)$$

$\bar{\omega}_{ijk}$  denotes the asymptotic value of  $\omega_{ijk} = \text{Cov}[Y_{ik}, Y_{jk}] = \mu_{Y_{ik}Y_{jk}} - \mu_{Y_{ik}}\mu_{Y_{jk}}$   $\forall i \neq j$ , where  $\mu_{Y_{ik}Y_{jk}} = \text{E}[Y_{ik}Y_{jk}]$ ,  $\mu_{Y_{ik}} = \text{E}[Y_{ik}]$ , and  $\mu_{Y_{jk}} = \text{E}[Y_{jk}]$ . For convenience, we use the following notations.

$$c_{tk} = \sqrt{1 - \tau_k^2} \mathbf{h}_{kk}^H \mathbf{h}_{tk} = \sqrt{\frac{1 - \tau_k^2}{\kappa_{tk} + 1}} \left( \sqrt{\kappa_{tk}} \mathbf{h}_{kk}^H \mathbf{h}_{tk}^L + \sum_p^P \mathbf{h}_{kk}^H \mathbf{r}_{tkp} g_{tkp} \right), \quad (81)$$

$$a_{tmk} = \sqrt{\frac{\tau_k^2}{\kappa_{tk} + 1}} \beta_{km}^L \left( \sqrt{\kappa_{tk}} \beta_{tm}^L h_{tkm} + \sum_p^P r_{tkmp} g_{tkp} \right), \quad (82)$$

for  $t \in \{i, j\}$  and  $\forall m, k$ . With  $a_{tmk}$ , we have  $\tau_k \mathbf{e}_k^H \mathbf{h}_{tk} = \sum_m^M a_{tmk} e_{km}^*$ . Here,  $c_{tk}$  and  $a_{tmk}$  are complex Gaussian random variables independent with  $e_{km}$  such as  $c_{tk} \sim \mathcal{CN}(\mu_{c_{tk}}, \sigma_{c_{tk}}^2)$  and  $a_{tmk} \sim \mathcal{CN}(\mu_{a_{tmk}}, \sigma_{a_{tmk}}^2)$ , where  $\mu_{c_{tk}} = \sqrt{\frac{\kappa_{tk}(1 - \tau_k^2)}{\kappa_{tk} + 1}} \mathbf{h}_{kk}^H \mathbf{h}_{tk}^L$ ,  $\sigma_{c_{tk}}^2 = \frac{1 - \tau_k^2}{\kappa_{tk} + 1} \sum_p^P |\mathbf{h}_{kk}^H \mathbf{r}_{tkp}|^2$ ,  $\mu_{a_{tmk}} = \sqrt{\frac{\tau_k^2 \kappa_{tk}}{\kappa_{tk} + 1}} \beta_{km}^L \beta_{tm}^L h_{tkm}$ , and  $\sigma_{a_{tmk}}^2 = \frac{\tau_k^2 |\beta_{km}^L|^2}{\kappa_{tk} + 1} \sum_p^P |r_{tkmp}|^2$ .

By using the above notations, we have

$$\begin{aligned} Y_{tk} &= \left| \sqrt{1 - \tau_k^2} \mathbf{h}_{kk}^H \mathbf{h}_{tk} + \tau_k \mathbf{e}_k^H \mathbf{h}_{tk} \right|^2 = \left| c_{tk} + \sum_m^M a_{tmk} e_{km}^* \right|^2 \\ &= |c_{tk}|^2 + 2\text{Re} \left( c_{tk} \sum_m^M a_{tmk}^* e_{km} \right) + \sum_m^M a_{tmk} e_{km}^* \sum_m^M a_{tmk}^* e_{km}. \end{aligned} \quad (83)$$

With a standard complex Gaussian distribution for  $e_{km}$ , we have  $\text{E}[e_{km}] = 0$ ,  $\text{E}[e_{km}^2] = 0$ ,  $\text{E}[e_{km}^3] = 0$ , and  $\text{E}[|e_{km}|^2] = 1$ . Then,  $\mu_{Y_{tk}}$  and  $\mu_{Y_{ik}Y_{jk}}$  are obtained as  $\mu_{Y_{tk}} = \mu_{C_{tk}} + \sum_m^M \mu_{A_{tmk}}$  and

$$\begin{aligned} \mu_{Y_{ik}Y_{jk}} &= \text{E} \left[ |c_{ik}|^2 |c_{jk}|^2 + |c_{ik}|^2 \sum_m^M |a_{jm} e_{km}|^2 |e_{km}|^2 \right. \\ &\quad \left. + |c_{jk}|^2 \sum_m^M |a_{im} e_{km}|^2 |e_{km}|^2 + 2\text{Re} \left( c_{ik} c_{jk}^* \sum_m^M a_{im}^* a_{jm} |e_{km}|^2 \right) + \xi \right] \\ &= \mu_{C_{ik}} \mu_{C_{jk}} + \mu_{C_{ik}} \sum_m^M \mu_{A_{jm}} + \mu_{C_{jk}} \sum_m^M \mu_{A_{im}} + 2\text{Re} \left( \mu_{c_{ik}} \mu_{c_{jk}}^* \sum_m^M \mu_{a_{im}}^* \mu_{a_{jm}} \right) + \mu_\xi, \end{aligned} \quad (84)$$

where we define  $\mu_{C_{tk}} = \text{E}[|c_{tk}|^2]$  and  $\mu_{A_{tmk}} = \text{E}[|a_{tmk}|^2]$  for  $t \in \{i, j\}$  and  $\forall m, k$ . Then, we have  $\mu_{C_{tk}} = \sigma_{c_{tk}}^2 + |\mu_{c_{tk}}|^2$  and  $\mu_{A_{tmk}} = \sigma_{a_{tmk}}^2 + |\mu_{a_{tmk}}|^2$ . Further,  $\xi$  and  $\mu_\xi$  are given by

$$\xi = \sum_m^M a_{im} e_{km}^* \sum_m^M a_{im}^* e_{km} \sum_m^M a_{jm} e_{km}^* \sum_m^M a_{jm}^* e_{km}, \quad (85)$$

$$\mu_\xi = \text{E} \left[ \sum_m^M |a_{im}|^2 |e_{km}|^2 \sum_m^M |a_{jm}|^2 |e_{km}|^2 + \sum_m^M a_{im} a_{jm}^* |e_{km}|^2 \sum_{n \neq m}^M a_{in}^* a_{jn} |e_{kn}|^2 \right]. \quad (86)$$

Then,  $\omega_{ijk}$  is obtained by

$$\omega_{ijk} = 2\text{Re} \left( \mu_{c_{ik}} \mu_{c_{jk}}^* \sum_m^M \mu_{a_{imk}}^* \mu_{a_{jmk}} \right) + \mu_\xi - \sum_m^M \mu_{A_{imk}} \sum_m^M \mu_{A_{jmk}}. \quad (87)$$

Here,  $\mu_{c_{tk}}$  is obtained from  $\mathbf{h}_{kk}^H \mathbf{h}_{tk}^L$  which is calculated by the sum of  $M$  elements. Then, the first term of  $\omega_{ijk}$  in (87) increases with  $\mathcal{O}(M^3)$  as  $M \rightarrow \infty$ . Similarly, the second and the last term of  $\omega_{ijk}$  increase with  $\mathcal{O}(M^2)$  as  $M \rightarrow \infty$ . According to the scaling laws for  $M$ , we have  $\omega_{ijk}/M^2 - \bar{\omega}_{ijk}/M^2 \xrightarrow[M \rightarrow \infty]{} 0$ , where

$$\bar{\omega}_{ijk} = 2\text{Re} \left( \mu_{c_{ik}} \mu_{c_{jk}}^* \sum_m^M \mu_{a_{imk}}^* \mu_{a_{jmk}} \right). \quad (88)$$

We can observe from (88) that  $\bar{\omega}_{ijk}$  can be obtained by a deterministic value and the LOS component of the interference channel exclusively produces  $\bar{\omega}_{ijk}$ .

## REFERENCES

- [1] M. Jung, W. Saad, Y. Jang, G. Kong, and S. Choi, “Uplink data rate in large intelligent surfaces: Asymptotic analysis under channel estimation errors,” in *European Conference on Networks and Communications (EuCNC)*, Dubrovnik, Croatia, Jun. 2020.
- [2] S. Hu, F. Rusek, and O. Edfors, “Beyond massive MIMO: The potential of data transmission with large intelligent surfaces,” *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2746–2758, May 2018.
- [3] ———, “Beyond massive MIMO: The potential of positioning with large intelligent surfaces,” *IEEE Trans. Signal Process.*, vol. 66, no. 7, pp. 1761–1774, Apr. 2018.
- [4] S. Hu, K. Chitti, F. Rusek, and O. Edfors, “User assignment with distributed large intelligent surface (LIS) systems,” in *IEEE PIMRC*, Bologna, Italy, Sep. 2018.
- [5] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, “Reconfigurable intelligent surfaces for energy efficiency in wireless communication,” *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4157–4170, Aug 2019.
- [6] W. Saad, M. Bennis, and M. Chen, “A vision of 6G wireless systems: Applications, trends, technologies, and open research problems,” *IEEE Network*, to appear, 2019.
- [7] Ericsson White Paper, “More than 50 billion connected devices,” Ericsson, Tech. Rep. 284 23-3149 Uen, Feb. 2011.
- [8] Z. Dawy, W. Saad, A. Ghosh, J. Andrews, and E. Yaacoub, “Towards massive machine type cellular communications,” *IEEE Commun. Mag.*, vol. 24, no. 1, pp. 120–128, Feb. 2017.
- [9] T. Park, N. Abuzainab, and W. Saad, “Learning how to communicate in the Internet of Things: Finite resources and heterogeneity,” *IEEE Access*, vol. 4, pp. 7063–7073, Nov. 2016.
- [10] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, “Unmanned aerial vehicle with underlaid device-to-device communications: Performance and tradeoffs,” *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, pp. 3949–3963, Jun. 2016.
- [11] T. Zeng, O. Semiari, W. Saad, and M. Bennis, “Joint communication and control for wireless autonomous vehicular platoon systems,” available online: [arxiv.org/abs/1804.05290](https://arxiv.org/abs/1804.05290), Apr. 2018.

- [12] R. Janaswamy, D. Gupta, and D. Schaubert, "Adaptive correction to array coefficients through dithering and near-field sensing," *IEEE Trans. Antennas Propag.*, vol. 58, no. 11, pp. 3558–3567, Nov. 2010.
- [13] P. Chakravorty and D. Mandal, "Radiation pattern correction in mutually coupled antenna arrays using parametric assimilation technique," *IEEE Trans. Antennas Propag.*, vol. 64, no. 9, pp. 4092–4095, Sep. 2016.
- [14] C. Studer, M. Wenk, and A. Burg, "MIMO transmission with residual transmit-RF impairments," in *ITG WSA*, Feb 2010.
- [15] X. Rao and V. Lau, "Distributed compressive CSIT estimation and feedback for FDD multi-user massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3261–3271, Jun. 2014.
- [16] S. Noh, M. Zoltowski, Y. Sung, and D. J. Love, "Pilot beam pattern design for channel estimation in massive MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 787–801, Oct. 2014.
- [17] B. M. Hochwald, T. L. Marzetta, and V. Tarokh, "Multiple-antenna channel hardening and its implications for rate feedback and scheduling," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1893–1909, Sep. 2004.
- [18] E. J. Black, "Holographic beam forming and MIMO," *Pivotal Commware, Technical Report, available online: pivotalcommware.com/wp-content/uploads/2017/12/ Holographic-Beamforming-WP-v.6C-FINAL.pdf*, Dec. 2017.
- [19] W. Tang, X. Li, J. Y. Dai, S. Jin, Y. Zeng, Q. Cheng, and T. J. Cui, "Wireless communications with programmable metasurface: Transceiver design and experimental results," *available online: arxiv.org/abs/1811.08119*, Nov. 2018.
- [20] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge Univ. Press, 2005.
- [21] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.
- [22] J. Song, J. Choi, and D. J. Love, "Common codebook millimeter wave beam design: Designing beams for both sounding and communication with uniform planar arrays," *IEEE Trans. Commun.*, vol. 65, no. 4, pp. 1859–1872, Apr. 2017.
- [23] Y. Han, S. Jin, X. Li, Y. Huang, L. Jiang, and G. Wang, "Design of double codebook based on 3D dual-polarized channel for multiuser MIMO system," *EURASIP J. Adv. Signal Process.*, vol. 2014, no. 1, pp. 1–13, Jul. 2014.
- [24] S. Hu, F. Rusek, and O. Edfors, "Capacity degradation with modeling hardware impairment in large intelligent surface," *available online: arxiv.org/abs/1810.09672*, Oct. 2018.
- [25] T. Kim, K. Min, M. Jung, and S. Choi, "Scaling laws of optimal training lengths for TDD massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7128–7142, Aug. 2018.
- [26] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan 2013.
- [27] A. S. Y. Poon, R. W. Brodersen, and D. N. C. Tse, "Degrees of freedom in multiple-antenna channels: A signal space approach," *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 523–536, Feb 2005.
- [28] H. Benaroya, S. Han, and M. Nagurka, *Probability Models in Engineering and Science*. Boca Raton, CRC Press, 2005.
- [29] 3rd Generation Partnership Project, "Technical Specification Group Radio Access Network; Spatial channel model for Multiple Input Multiple Output (MIMO) simulations," TR 25.996, V14.0.0, Mar. 2017.
- [30] R. Ash and C. Doléans-Dade, *Probability and Measure Theory*. Academic Press, 2000.
- [31] L. A. Goodman, "On the exact variance of products," *Journal of the American Statistical Association*, vol. 55, no. 292, pp. 708–713, 1960.