

# Reliability Analysis of Large Intelligent Surfaces (LISs): Rate Distribution and Outage Probability

Minchae Jung, *Member, IEEE*, Walid Saad, *Fellow, IEEE*,  
Youngrok Jang, Gyuyeol Kong, and Sooyong Choi, *Member, IEEE*

**Abstract**—Large intelligent surfaces (LISs) have been recently proposed as an effective wireless communication solution that can leverage antenna arrays deployed on the entirety of man-made structures such as walls. An LIS can provide space-intensive and reliable communication, enabling the desired wireless channel to exhibit a perfect line-of-sight. However, the outage probability of LIS, which is an important performance metric to evaluate the system reliability, remains uncharacterized. In this paper, the distribution of uplink sum-rate is asymptotically analyzed for an LIS system. Given the derived asymptotic distribution, the outage probability is derived for the considered LIS system. Simulation results show that the results of the proposed asymptotic analyses are in close agreement to the exact mutual information in the presence of a large number of antennas and devices.

## I. INTRODUCTION

THE demand for wireless connectivity has been growing exponentially in recent years, mainly driven by the variety of upcoming Internet of Things (IoT) applications, such as sensors, vehicles, and drones [1], [2]. To support this demand for wireless connectivity for IoT services without additional radio resources, the concept of large intelligent surfaces (LISs) has been newly proposed to exploit the fact that man-made structures, such as buildings and walls, can be made electromagnetically active and used for wireless transmission [3]–[7]. An LIS system enables the desired channels to become line-of-sight (LOS) channels, resulting in more reliable and space-intensive communications compared to conventional massive multiple-input multiple-output (MIMO) systems.<sup>1</sup> Effectively using LISs requires addressing many challenges such as obtaining their achievable rate, evaluating the system reliability, and analyzing the channel hardening effect.

Prior art [3]–[7] has studied some of these challenges. In particular, the work in [3] studied uplink data rate of the matched filter (MF) and derived the performance of the optimal receiver in an LIS system. In [4] and [5], the authors analyzed the uplink data rates to evaluate LIS performance while considering, respectively, hardware impairments and channel estimation errors. Moreover, the work in [5] verified the occurrence of the channel hardening effect theoretically. Meanwhile, the authors in [6] and [7] proposed the use of LIS as a reconfigurable scatterer that reflects transmitted

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

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

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 CNS-1836802 and OAC-1638283.

<sup>1</sup>For an overview on how LIS differs from massive MIMO, see [3]–[7].

signals and developed LIS phase shifter that maximize energy efficiency and received signal strength, respectively. However, these recent works [3]–[7] have not investigated the distribution of the actual data rate which is necessary to analyze the outage probability of an LIS. Since an accurate estimation of the outage probability enables us to operate an LIS network reliably, it is necessary to analyze this outage probability in an LIS system serving a large number of devices.

The main contribution of this paper is to derive an exact closed-form expression for the outage probability and verify the reliability of an LIS system in terms of the outage probability.<sup>2</sup> Given an LIS area serving multiple devices, we analyze the asymptotic distribution of the sum-rate, relying on the Lyapunov central limit theorem (CLT) in the presence of a large number of antennas and devices. This approximation allows for accurate estimations of the outage probability without the need for a large number of simulations, and it enables us to verify the reliability of an LIS system. Simulation results show that an LIS can provide reliable communication regardless of signal-to-noise-ratio (SNR) given that the outage probability keeps constant for varying SNR when the LIS is equipped with a large number of antennas and devices.

## II. SYSTEM MODEL

Consider an uplink LIS system in which a large number of IoT devices are connected to the LIS's large array and a two-dimensional LIS is deployed on the horizontal plane. The LIS consists of  $K$  LIS units, each of which occupies a subarea of the entire LIS and has a square shape with limited area of  $2L \times 2L$  serving a single-antenna device. We assume that the LIS has its own signal processing unit for estimating channel and detecting any data signal, as in [3]–[5]. A large number of antennas,  $M$ , are deployed on the surface of the LIS unit in rectangular lattice form with  $\Delta L$  spacing, centered on the  $(x, y)$  coordinates of the corresponding device. Given that the location of device  $k$  is  $(x_k, y_k, z_k)$ , antenna  $m$  of LIS unit  $k$  is placed at  $(x_{km}^{\text{LIS}}, y_{km}^{\text{LIS}}, 0)$  where  $x_{km}^{\text{LIS}} \in [x_k - L, x_k + L]$  and  $y_{km}^{\text{LIS}} \in [y_k - L, y_k + L]$ . The LIS units may overlap depending on the location of their corresponding devices and this results in severe performance degradation. To prevent this problem, we assume that the devices with partially overlapping LIS units are allocated on orthogonal resources, resulting in an LIS composed of  $K$  non-overlapping LIS units. Moreover, we assume that each device controls its uplink transmit power toward the center of its LIS unit according to a target SNR.

### A. Wireless Channel Model

Note that the entire LIS environment is active during wireless communication and the signal from the NLOS path

<sup>2</sup>Note that those results cannot be obtained directly from [5].

can be negligible compared to the LOS signal, as proved in [5]. Hence, we consider the LIS channel  $\mathbf{h}_{kk} \in \mathbb{C}^M$  between device  $k$  and LIS unit  $k$  as a LOS path defined as  $\mathbf{h}_{kk} = [\beta_{k1}^L h_{kk1}, \dots, \beta_{kM}^L h_{kkM}]^T$ , where  $\beta_{km}^L = \alpha_{kkm}^L l_{kkm}^L$  and  $h_{kkm} = \exp(-j2\pi d_{kkm}/\lambda)$  denote a LOS channel gain and state, respectively, between device  $k$  and antenna  $m$  of LIS unit  $k$ .  $\alpha_{kkm}^L = \sqrt{z_k/d_{kkm}}$  and  $l_{kkm}^L = 1/\sqrt{4\pi d_{kkm}^2}$  represent, respectively, the antenna gain and path loss attenuation, where  $d_{kkm}$  and  $\lambda$  denote the distance from device  $k$  to antenna  $m$  of LIS unit  $k$  and the wavelength of a signal, respectively.

Given a Rician factor  $\kappa_{jk}$ , we consider the interference channel  $\mathbf{h}_{jk} \in \mathbb{C}^M$  between device  $j$  and LIS unit  $k$  as a Rician fading channel, given by  $\mathbf{h}_{jk} = \sqrt{\frac{\kappa_{jk}}{\kappa_{jk}+1}} \mathbf{h}_{jk}^L + \sqrt{\frac{1}{\kappa_{jk}+1}} \mathbf{h}_{jk}^{\text{NL}}$ , where  $\mathbf{h}_{jk}^L = [\beta_{j1}^L h_{jk1}, \dots, \beta_{jM}^L h_{jkM}]^T$  and  $\mathbf{h}_{jk}^{\text{NL}} = \mathbf{R}_{jk}^{1/2} \mathbf{g}_{jk}$  denote the deterministic LOS and the correlated NLOS component, respectively. Given  $P$  dominant paths among all NLOS paths, we define  $\mathbf{R}_{jk} \in \mathbb{C}^{M \times P}$  and  $\mathbf{g}_{jk} = [g_{jk1}, \dots, g_{jkP}]^T \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_P)$  as the deterministic correlation matrix and an independent fading channel between device  $j$  and LIS unit  $k$ , respectively. Since a two-dimensional LIS is deployed on the  $xy$ -plane, we can model it as a uniform planar array [8]. Then, the correlation matrix can be defined as  $\mathbf{R}_{jk}^{1/2} = \mathbf{l}_{jk}^{\text{NL}} \mathbf{D}_{jk}$ , where  $\mathbf{l}_{jk}^{\text{NL}} = \text{diag}(l_{jk1}^{\text{NL}}, \dots, l_{jkM}^{\text{NL}})$  is a diagonal matrix that includes the path loss attenuation  $l_{jk}^{\text{NL}} = d_{jk}^{-\beta_{\text{PL}}/2}$  with a path loss exponent  $\beta_{\text{PL}}$  and  $\mathbf{D}_{jk} = [\alpha_{jk1}^{\text{NL}} \mathbf{d}(\phi_{jk1}^v, \phi_{jk1}^h), \dots, \alpha_{jkP}^{\text{NL}} \mathbf{d}(\phi_{jkP}^v, \phi_{jkP}^h)]$ . The term  $\mathbf{d}(\phi_{jkP}^v, \phi_{jkP}^h)$  represents the NLOS path  $p$  at given angles  $(\phi_{jkP}^v, \phi_{jkP}^h)$ , defined as  $\mathbf{d}(\phi_{jkP}^v, \phi_{jkP}^h) = \frac{1}{\sqrt{M}} \mathbf{d}_v(\phi_{jkP}^v) \otimes \mathbf{d}_h(\phi_{jkP}^h)$ , where  $\otimes$  is the Kronecker product and

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

$$\mathbf{d}_h(\phi_{jkP}^h) = \left[ 1, e^{j \frac{2\pi \Delta L}{\lambda} \phi_{jkP}^h}, \dots, e^{j \frac{2\pi \Delta L}{\lambda} (\sqrt{M}-1) \phi_{jkP}^h} \right]^T.$$

Here,  $\phi_{jkP}^v = \sin \theta_{jkP}^v$  and  $\phi_{jkP}^h = \sin \theta_{jkP}^h \cos \theta_{jkP}^h$  when the elevation and azimuth angles between device  $j$  and LIS unit  $k$  at path  $p$  are  $\theta_{jkP}^v$  and  $\theta_{jkP}^h$ , respectively. Further,  $\alpha_{jkP}^{\text{NL}} = \sqrt{\cos \theta_{jkP}^v \cos \theta_{jkP}^h}$  denotes the antenna gain at path  $p$  with  $\theta_{jkP} \in [-\frac{\pi}{2}, \frac{\pi}{2}]$  and  $\theta_{jkP} \in \{\theta_{jkP}^v, \theta_{jkP}^h\}$ .

### B. Uplink Data Rate

The received signal from all devices at LIS unit  $k$  is obtained as  $\mathbf{y}_k = \sqrt{\rho_k} \mathbf{h}_{kk} x_k + \sum_{j \neq k}^K \sqrt{\rho_j} \mathbf{h}_{jk} x_j + \mathbf{n}_k$ , where  $x_k$  and  $x_j$  are uplink signals of device  $k$  and  $j$ , respectively,  $\rho_k$  and  $\rho_j$  are their transmit SNRs, and  $\mathbf{n}_k \in \mathbb{C}^M \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$  is noise vector. Given a linear receiver  $\mathbf{f}_k^H$  at LIS unit  $k$ , we have

$$\mathbf{f}_k^H \mathbf{y}_k = \sqrt{\rho_k} \mathbf{f}_k^H \mathbf{h}_{kk} x_k + \sum_{j \neq k}^K \sqrt{\rho_j} \mathbf{f}_k^H \mathbf{h}_{jk} x_j + \mathbf{f}_k^H \mathbf{n}_k. \quad (1)$$

Given an MF receiver with imperfect channel estimation results from a least square estimator, we have  $\mathbf{f}_k = \mathbf{h}_{kk} + \sqrt{\tau_k^2/(1-\tau_k^2)} \mathbf{e}_k$  where  $\mathbf{e}_k$  is the error vector uncorrelated with  $\mathbf{n}_k$  [5]. From (1), the received signal-to-interference-plus-noise ratio (SINR) at LIS unit  $k$  will be  $\gamma_k = \rho_k S_k (1-\tau_k^2)/I_k$ , where  $S_k = |\mathbf{h}_{kk}|^4$  and  $I_k = \rho_k \tau_k^2 X_k + \sum_{j \neq k}^K \rho_j Y_{jk} + Z_k$ . Here,  $X_k$ ,  $Y_{jk}$ , and

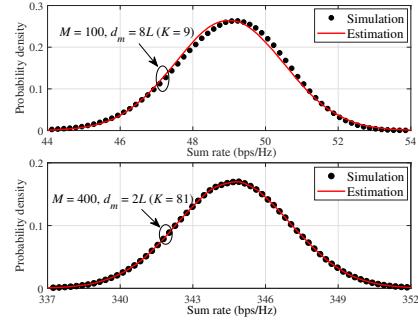


Fig. 1. PDFs of  $R$  when  $(M, d_m)$  is (top)  $(100, 8L)$  and (bottom)  $(400, 2L)$ .

$Z_k$  denote  $|\mathbf{e}_k^H \mathbf{h}_{kk}|^2$ ,  $|\sqrt{1-\tau_k^2} \mathbf{h}_{kk}^H \mathbf{h}_{jk} + \tau_k \mathbf{e}_k^H \mathbf{h}_{jk}|^2$ , and  $|\sqrt{1-\tau_k^2} \mathbf{h}_{kk}^H + \tau_k \mathbf{e}_k^H|^2$ , respectively. From the received SINR at LIS unit  $k$ , the uplink sum-rate is obtained by  $R = \sum_{k=1}^K R_k$ , where  $R_k = \log(1 + \gamma_k)$  is instantaneous rate for device  $k$ . Given this sum-rate, we will analyze the asymptotic distribution of mutual information and derive the closed-form expression for the outage probability to characterize reliability.

### III. ASYMPTOTIC ANALYSIS OF RELIABILITY VIA OUTAGE

In [5], the mean and variance of the individual rate in an uplink LIS system is asymptotically derived. However, the distribution of the sum-rate must be analyzed in order to obtain the outage probability and evaluate the system reliability. Deriving this distribution is not trivial and cannot be obtained directly from [5], because a sequence of individual rates,  $[R_1, \dots, R_K]$ , is not identically distributed. We begin with a characteristic of LIS systems in which

$$S_k/M^2 - \bar{p}_k/M^2 \xrightarrow{M \rightarrow \infty} 0,$$

$$I_k/M^2 - \bar{\mu}_{I_k}/M^2 \xrightarrow{M, K \rightarrow \infty} 0,$$

where  $\bar{p}_k = \frac{M^2 p_k^2}{16\pi^2 L^4}$ ,  $p_k = \tan^{-1}(L^2/(z_k \sqrt{2L^2 + z_k^2}))$ , and  $\bar{\mu}_{I_k}$  is a deterministic value depending on the correlation matrices and the locations of the devices [5, Lemma 4]. Then, the asymptotic rate,  $\bar{R}_k$ , is obtained by the following lemma.

**Lemma 1.** The asymptotic rate,  $\bar{R}_k$ , can be obtained by the function of a random variable  $I_k$  as  $\bar{R}_k = a_k - b_k I_k$ , where  $a_k = \frac{\rho_k \bar{p}_k (1-\tau_k^2)}{\rho_k \bar{p}_k (1-\tau_k^2) + \bar{\mu}_{I_k}} + \log(1 + \frac{\rho_k \bar{p}_k (1-\tau_k^2)}{\bar{\mu}_{I_k}})$  and  $b_k = \frac{\rho_k \bar{p}_k (1-\tau_k^2) / \bar{\mu}_{I_k}}{\rho_k \bar{p}_k (1-\tau_k^2) + \bar{\mu}_{I_k}}$ .

*Proof:* The detailed proof is presented in Appendix A. ■

Given that  $\bar{p}_k$  and  $\bar{\mu}_{I_k}$  are deterministic values,  $a_k$  and  $b_k$  are also deterministic values. Then, Lemma 1 shows that the distribution of  $\bar{R}_k$  is exclusively determined by that of  $I_k$ , and it allows us to readily obtain the distribution of  $\bar{R} = \sum_k \bar{R}_k$  by analyzing  $I_k$  instead of  $\bar{R}_k$  itself. From Lemma 1, we have  $\bar{R} = \sum_k a_k - b_k I_k$  and the second term  $\sum_k b_k I_k$  determines the distribution of  $\bar{R}$ . By using the Lyapunov CLT [9], we analyze the distribution of  $\sum_k b_k I_k$  and finally obtain the distribution of  $R$ , as follows.

**Theorem 1.** For large  $M$  and  $K$ ,  $R$  approximately follows  $R \sim \mathcal{N}(\bar{\mu}_R, \bar{\sigma}_R^2)$ , where  $\bar{\mu}_R = \sum_k \log(1 + \frac{\rho_k \bar{p}_k (1-\tau_k^2)}{\bar{\mu}_{I_k}})$  and  $\bar{\sigma}_R^2 = \sum_k \frac{\bar{\sigma}_{I_k}^2 \rho_k^2 \bar{p}_k^2 (1-\tau_k^2)^2}{\bar{\mu}_{I_k}^2 (\bar{\mu}_{I_k} + \rho_k \bar{p}_k (1-\tau_k^2))^2}$ .

*Proof:* The detailed proof is presented in Appendix B. ■

Theorem 1 shows that the sum-rate of an LIS system approximately follows a Gaussian distribution for large  $M$  and

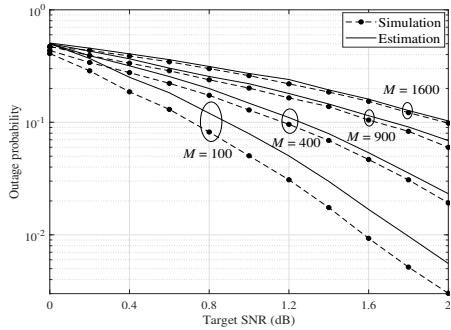


Fig. 2. Outage probabilities as a function of the target SNR when  $d_m = 4L$  ( $K = 25$ ).

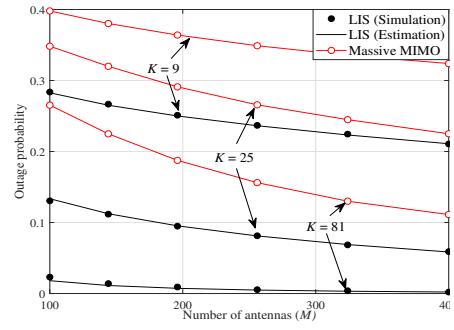


Fig. 3. Outage probabilities as a function of  $M$  when  $\delta = 0.98$ .

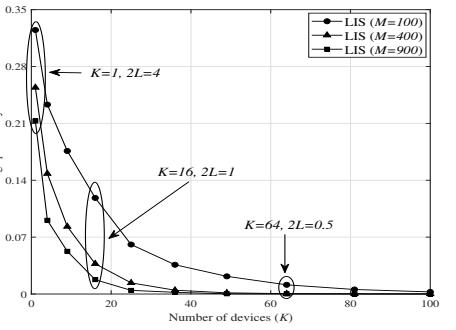


Fig. 4. Outage probabilities as a function of  $K$  when  $\delta = 0.98$  and  $A = 16$ .

$K$ , and its mean and variance can be obtained deterministically. This allows us to evaluate the performance of an LIS system in terms of outage probability, ergodic rate, reliability, and scheduling diversity, without extensive simulations. Then, we can characterize the closed-form expression for the outage probability, defined as the probability when the instantaneous sum-rate falls below a certain threshold value.

**Corollary 1.** For large  $M$  and  $K$ , the outage probability of event  $\{R < R_D\}$  is approximately obtained as follows:

$$P_o = \Pr [R < R_D] = 1 - Q\left(\frac{R_D - \bar{\mu}_R}{\bar{\sigma}_R}\right),$$

where  $R_D$  is a desired rate when the probability of the outage event is  $P_o$ , and  $Q(\cdot)$  is a Gaussian Q-function.

Assume that  $R_D = \delta \cdot \bar{\mu}_R$  for  $0 < \delta \leq 1$ , then we have  $P_o = Q\left((1 - \delta) \frac{\bar{\mu}_R}{\bar{\sigma}_R}\right)$ . As  $M$  increases,  $\bar{\mu}_R$  increases and converges to its bound, and  $\bar{\sigma}_R$  decreases, as proved in [5]. Therefore  $\bar{P}_o$  decreases as  $M$  increases. On the basis of scaling law for  $K$ ,  $\bar{\mu}_R$  and  $\bar{\sigma}_R$  follow  $\mathcal{O}(K \log(1 + \frac{1}{K})) = \mathcal{O}(1)$  and  $\mathcal{O}(\frac{1}{\sqrt{K}})$ , respectively. Then,  $\frac{\bar{\mu}_R}{\bar{\sigma}_R}$  increases with  $\mathcal{O}(\sqrt{K})$  and finally  $\bar{P}_o$  also decreases, as  $K$  increases. Therefore, an LIS can provide *reliable communication* in the presence of a large number of antennas and devices. Moreover, since  $\bar{\mu}_R$  and  $\bar{\sigma}_R$  are deterministic values obtained from the correlation matrices and the locations of the devices, we can estimate the outage probability of LIS without the need for extensive simulations.

#### IV. SIMULATION RESULTS AND ANALYSIS

In this section, all simulations are statistically averaged over a large number of independent runs. For our simulations, we assume that target SNR is 3 dB,  $\lambda = 0.1$  m,  $L = 0.25$  m,  $\tau_k^2 = 0.5$ , and  $\beta_{PL} = 3.7$ , and the probability of LOS path and Rician factor are applied based on the 3GPP model [5]. We consider a scenario in which devices are located on a two-dimensional  $xy$ -plane at  $z = 1$  in parallel to the LIS within the range of  $-2 \leq x \leq 2$  and  $0 \leq y \leq 4$  (in meters) and the distance between the adjacent devices is equally set to  $d_m$ . Therefore, a total of 81, 25, and 9 devices are located in a two-dimensional rectangular lattice form when  $d_m = 2L$ ,  $d_m = 4L$ , and  $d_m = 8L$ , respectively. For a massive MIMO system, we assume a single BS with  $M$  antennas serving  $K$  users based on a uniform linear array, as in [5].

Fig. 1 shows the probability density function (PDF) of the instantaneous rate according to the independent channel realizations. As proved in Theorem 1, the PDF of the sum-rate follows a Gaussian distribution and the accuracy of the Gaussian approximation improves as  $M$  and  $K$  increase. As

can be seen, the two PDFs are almost aligned with each other, and the approximation error is negligible when  $M = 400$  and  $d_m = 2L$ , corroborating the result of Theorem 1.

Fig. 2 and 3 compare the outage probabilities resulting from the simulations to the estimations from Corollary 1 as a function of the target SNR and  $M$ , respectively. We assume that  $R_D = \bar{\mu}_R$  when the target SNR is 0 dB in Fig. 2 and  $\delta = 0.98$  in Fig. 3. In both figures, the asymptotic results from Corollary 1 become close to the results of our simulations as  $M$  increases, verifying the accuracy of our analyses. As shown in Fig. 2, even when  $M = 100$ , the performance gap between the results of the asymptotic estimation and the simulation is less than 0.2 dB in terms of the target SNR. Moreover, since the noise component of an LIS system becomes negligible as  $M$  increases [5], we can observe that the target SNR gradually lessens its effect on the outage probability as  $M$  increases (e.g., when the target SNR is 2 dB, the outage probabilities for  $M = 100$  and 1600 are  $3 \cdot 10^{-3}$  and  $10^{-1}$ , respectively). Furthermore, Fig. 3 shows that the outage probability decreases and eventually reaches zero as  $M$  and  $K$  increase. In addition, the outage probabilities resulting from LIS are always lower than those resulting from massive MIMO system over the entire range of  $M$ , which shows that an LIS system can be more reliable than massive MIMO.

Fig. 4 shows the impact of the surface-area of each LIS unit on the outage probability with a fixed total allocated area of an LIS (e.g.,  $A = 16$  [m<sup>2</sup>]). Given a fixed  $A$ ,  $K$  increases as  $L$  decreases and, hence,  $P_o$  decreases, as shown in Fig. 4.

#### V. CONCLUSIONS

In this paper, we have analyzed the sum-rate of an uplink LIS system asymptotically, under practical LIS environments. In particular, we have derived the distribution of the sum-rate by considering a practical LIS environment in which the interference can be generated by device-specific Rician fading with spatial correlation in presence of imperfect channel estimation. Moreover, we have studied the outage probability of the sum-rate from the derived asymptotic distribution. We have shown that the asymptotic results can accurately and analytically determine the performance of an LIS, including the distribution of the sum-rate and outage probability, without the need for a large number of simulations. Simulation results have verified the results of the asymptotic analyses. These numerical results have shown that the analytic results were in close agreement with the results arising from extensive

simulations, and an LIS can provide reliable communication in terms of the outage probability.

## APPENDIX A PROOF OF LEMMA 1

We begin with the individual rate for device  $k$  as  $R_k = \log(1 + \frac{\rho_k S_k (1 - \tau_k^2)}{I_k})$ . Then,  $R_k$  can be divided into two terms as  $R_k = R_k^L - R_k^R$ , where  $R_k^L = \log(\rho_k S_k (1 - \tau_k^2) + I_k)$  and  $R_k^R = \log I_k$ . Since  $I_k/M^2 - \bar{\mu}_{I_k}/M^2$  converges to zero as  $M, K \rightarrow \infty$ , we have the following series of equalities using the exponential function definition  $e^x = \lim_{n \rightarrow \infty} (1 + x/n)^n$ :

$$R_k^L = \log\left(1 + \frac{(I_k - \bar{\mu}_{I_k})/M^2}{(\rho_k \bar{\rho}_k (1 - \tau_k^2) + \bar{\mu}_{I_k})/M^2}\right) + c_k,$$

$$\bar{R}_k^L = \frac{I_k - \bar{\mu}_{I_k}}{\rho_k \bar{\rho}_k (1 - \tau_k^2) + \bar{\mu}_{I_k}} + c_k,$$

where  $c_k = \log(\rho_k \bar{\rho}_k (1 - \tau_k^2) + \bar{\mu}_{I_k})$  and  $R_k^L - \bar{R}_k^L \xrightarrow{M, K \rightarrow \infty} 0$ . Similarly, we have  $R_k^R = (I_k - \bar{\mu}_{I_k})/\bar{\mu}_{I_k} + \log \bar{\mu}_{I_k}$  and finally, we can obtain  $\bar{R}_k = \bar{R}_k^L - R_k^R$ , which completes the proof.

## APPENDIX B PROOF OF THEOREM 1

Given the definition of  $I_k = \rho_k \tau_k^2 X_k + \sum_{j \neq k}^K \rho_j Y_{jk} + Z_k$ , we make use of the following asymptotic convergence from [5] pertaining to the distribution of  $I_k$ :  $I_k - \sum_{j \neq k}^K \rho_j Y_{jk} \xrightarrow{M \rightarrow \infty} 0$ . Then,  $\sum_{j \neq k}^K \rho_j Y_{jk}$  approximately determines the distribution of  $I_k$  for large  $M$ . Therefore, we have the following numerically accurate approximation for large  $M$ :  $\sum_k b_k I_k = \sum_k b_k \sum_{j \neq k}^K \rho_j Y_{jk}$ . Given a random variable  $\sum_k b_k \sum_{j \neq k}^K \rho_j Y_{jk}$ , we can observe that  $b_k \sum_{j \neq k}^K \rho_j Y_{jk}$  is a random variable independent across  $k$  since  $\mathbf{g}_{jk}$  and  $\mathbf{e}_k$  are independent random vectors for different  $k$ . Therefore, the following Lyapunov's condition should be satisfied for some  $\delta > 0$  and large  $M$  to prove that  $\sum_k b_k \sum_{j \neq k}^K \rho_j Y_{jk}$  is Gaussian distributed [9]:

$$\lim_{K \rightarrow \infty} \frac{1}{s_K^{2+\delta}} \sum_k \mathbb{E} \left[ \left| b_k \sum_{j \neq k}^K \rho_j (Y_{jk} - \bar{\mu}_{Y_{jk}}) \right|^{2+\delta} \right] = 0, \quad (2)$$

where  $s_K^2 = \sum_k b_k^2 \bar{\sigma}_{I_k}^2$  and  $\bar{\mu}_{Y_{jk}}$  is asymptotic mean of  $Y_{jk}$ . Given that  $Y_{jk}$  is a random variable correlated across  $j$ , we have  $Y_{jk} = s_{jk} |x_j|^2 + |\mu_{jk}^L|^2 + 2\text{Re}(\sqrt{s_{jk}} \mu_{jk}^L x_j^*)$ , where  $s_{jk} = s_{jk}^L + s_{jk}^{N1} + s_{jk}^{N2}$  and  $x_j$  is a standard complex Gaussian random variable correlated across  $j$ . The terms  $s_{jk}^L, s_{jk}^{N1}, s_{jk}^{N2}$ , and  $|\mu_{jk}^L|^2$  are provided in [5, Lemma 2], and  $s_{jk}$  and  $|\mu_{jk}^L|^2$  are respectively increase with  $\mathcal{O}(M)$  and  $\mathcal{O}(M^2)$  as  $M$  increases. We then calculate  $Y_{jk} - \bar{\mu}_{Y_{jk}}$  in (2) as  $s_{jk}(|x_j|^2 - 1) + 2\text{Re}(\sqrt{s_{jk}} \mu_{jk}^L x_j^*)$ . Given that  $s_{jk}$  and  $\sqrt{s_{jk}} \mu_{jk}^L$  respectively increase with  $\mathcal{O}(M)$  and  $\mathcal{O}(M\sqrt{M})$ , we have  $Y_{jk} - \bar{\mu}_{Y_{jk}} = 2\text{Re}(\sqrt{s_{jk}} \mu_{jk}^L x_j^*)$  for large  $M$ . Then, we have the following inequality when  $\delta = 2$  in (2):

$$\lim_{K \rightarrow \infty} \frac{16 \sum_k |b_k|^4 \mathbb{E} \left[ \left| \sum_{j \neq k}^K \rho_j \text{Re}(\sqrt{s_{jk}} \mu_{jk}^L x_j^*) \right|^4 \right]}{\left( \sum_k b_k^2 \bar{\sigma}_{I_k}^2 \right)^2} \leq \lim_{K \rightarrow \infty} \frac{16 \sum_k |b_k|^4 \mathbb{E} \left[ \left| \sum_{j \neq k}^K \rho_j |x_j| \sqrt{s_{jk}} \mu_{jk}^L \right|^4 \right]}{\left( \sum_k b_k^2 \bar{\sigma}_{I_k}^2 \right)^2}. \quad (3)$$

The expectation term in the numerator of (3) is bounded as

$$\begin{aligned} \mathbb{E} \left[ \left| \sum_{j \neq k}^K \rho_j |x_j| \sqrt{s_{jk}} \mu_{jk}^L \right|^4 \right] &\stackrel{(a)}{\leq} \mathbb{E} \left[ \left( d_k \sum_{j \neq k}^K \rho_j^2 |x_j|^2 \right)^2 \right] \\ &= d_k^2 \left( \text{Var} \left[ \sum_{j \neq k}^K \rho_j^2 |x_j|^2 \right] + \mathbb{E} \left[ \sum_{j \neq k}^K \rho_j^2 |x_j|^2 \right]^2 \right) \\ &\stackrel{(b)}{\leq} d_k^2 \left( \sum_{j \neq k}^K \rho_j^4 + \sum_{i,j \neq k: i \neq j}^K \rho_i^2 \rho_j^2 + \left( \sum_{j \neq k}^K \rho_j^2 \right)^2 \right), \end{aligned} \quad (4)$$

where  $d_k = \sum_j s_{jk} |\mu_{jk}^L|^2$ , (a) results from Cauchy–Schwarz inequality, and (b) results from the covariance inequality [10]:

$$\sum_{i,j \neq k: i \neq j}^K \rho_i^2 \rho_j^2 \text{Cov}[|x_i|^2, |x_j|^2] \stackrel{(c)}{\leq} \sum_{i,j \neq k: i \neq j}^K \rho_i^2 \rho_j^2 \sqrt{\text{Var}[|x_i|^2] \text{Var}[|x_j|^2]} = \sum_{i,j \neq k: i \neq j}^K \rho_i^2 \rho_j^2.$$

The covariance inequality (c) always holds because the covariance matrix of a random vector  $\mathbf{x}$  is a real symmetric matrix [5], and it is therefore positive-definite when we define  $\mathbf{x} = [x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_K]^T$ . We then determine the scaling laws of (4) according to  $M$  and  $K$ . Since  $s_{jk} |\mu_{jk}^L|^2$  increases with  $\mathcal{O}(M^3)$ ,  $d_k^2$  increases with  $\mathcal{O}(K^2 M^6)$  and then (4) increases with  $\mathcal{O}(K^4 M^6)$  as  $M, K \rightarrow \infty$ . Further, given that  $\bar{\mu}_{I_k}$  and  $\bar{\sigma}_{I_k}^2$  respectively increase with  $\mathcal{O}(KM^2)$  and  $\mathcal{O}(K^2 M^3)$  as proved in [5],  $b_k$  in (3) decreases with  $\mathcal{O}(\frac{1}{K^2 M^4})$  as  $M, K \rightarrow \infty$ . Then, the numerator and denominator of the upper bound of (3) decrease with  $\mathcal{O}(\frac{1}{K^3 M^{10}})$  and  $\mathcal{O}(\frac{1}{K^2 M^{10}})$ , respectively, as  $M, K \rightarrow \infty$ . Therefore, the upper bound of (3) decreases with  $\mathcal{O}(\frac{1}{K})$  and eventually reaches zero as  $K \rightarrow \infty$ . In conclusion, (2) is satisfied and  $\sum_k b_k I_k$  approximately follows a Gaussian distribution. Given that  $b_k I_k$  is a random variable independent across  $k$ , we can ultimately obtain the distribution of  $\sum_k b_k I_k$  for large  $M$  and  $K$ , as follows:  $\sum_k b_k I_k \sim \mathcal{N}(\sum_k b_k \bar{\mu}_{I_k}, \sum_k b_k^2 \bar{\sigma}_{I_k}^2)$ . For large  $M$  and  $K$ , we have the following numerically accurate approximation:  $R = \sum_k a_k - b_k I_k$ . Therefore,  $R \sim \mathcal{N}(\sum_k a_k - b_k \bar{\mu}_{I_k}, \sum_k b_k^2 \bar{\sigma}_{I_k}^2)$ , which completes the proof.

## REFERENCES

- [1] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *available online: arxiv.org/abs/1902.10265*, Feb. 2019.
- [2] M. Mozaffari, A. Kasgari, W. Saad, M. Bennis, and M. Debbah, "Beyond 5G with UAVs: Foundations of a 3D wireless cellular network," *IEEE Trans. Wireless Commun.*, vol. 18, no. 1, pp. 357–372, Jan 2019.
- [3] 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.
- [4] —, "Capacity degradation with modeling hardware impairment in large intelligent surface," *available online: arxiv.org/abs/1810.09672*, Oct. 2018.
- [5] M. Jung, W. Saad, Y. Jang, G. Kong, and S. Choi, "Performance analysis of large intelligent surfaces (LISs): Asymptotic data rate and channel hardening effects," *available online: arxiv.org/abs/1810.05667*, Oct. 2018.
- [6] C. Huang, G. Alexandropoulos, A. Zaponne, M. Debbah, and C. Yuen, "Energy efficient multi-user MISO communication using low resolution large intelligent surfaces," in *IEEE GLOBECOM*, 2018.
- [7] C. Huang, G. Alexandropoulos, C. Yuen, and M. Debbah, "Indoor signal focusing with deep learning designed reconfigurable intelligent surfaces," *available online: arxiv.org/abs/1905.07726*, May 2019.
- [8] 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.
- [9] R. Ash and C. Doléans-Dade, *Probability and Measure Theory*. Academic Press, 2000.
- [10] N. Mukhopadhyay, *Probability and Statistical Inference*. CRC Press, 2000.