DOI: 10.1002/itl2.48

Received: 15 February 2018 | Revised: 26 March 2018 | Accepted: 30 March 2018

DOI: 10.1002/itl2.48

LETTER

WILEY

Accurate log-normal approximation to the signal-to-interference ratio in massive multiple-input multiple-output systems

Michelle S. P. Facina | **Gustavo Fraidenraich**

Communications Department (DECOM), University of Campinas, Campinas, Brazil

Correspondence
Michelle S. P. Facina, DECOM, University of Campinas, Campinas, SP, Brazil.
Email: michelle.facina@gmail.com

Funding Information
São Paulo Research Foundation (FAPESP), 2016/16181-2.

In this paper, a very tight approximation is derived for the signal-to-interference ratio of a multicell massive multiple-input multiple-output system with a finite number of base station (BS) antennas. The approximation is derived considering that each term in the SIR is log-normal distributed. To this end, the first and second moments of the logarithm of each variable are used. In addition, an exact expression is derived for the cumulative distribution function for the net capacity. In order to corroborate our derivations, simulations using the Monte Carlo method were carried out, and it was observed that the proposed analytical results tightly match the numerical simulations. The asymptotic result is also obtained for the case in which the number of BS antennas tends to infinity ($M \to \infty$), considering both uniform and nonuniform spatial user distributions.

KEYWORDS
log-normal shadowing, massive MIMO, signal-to-interference ratio

1 | INTRODUCTION

Massive multiple-input and multiple-output (MIMO) emerges as a promising technology to meet the stringent requirements of future fifth-generation systems. This technology represents a drastic change in the infrastructure of cellular networks, since its communication system is composed of base stations (BS) with hundreds of antennas and can simultaneously serve dozens of user terminals, each having a single antenna. The BS is responsible for sending independent data streams to multiple user terminals in the same time-frequency resource.[1] Furthermore, each user terminal is ideally assigned an orthogonal pilot sequence in the uplink channel during the training stage. However, the maximum number of such sequences is limited by the duration of the coherence interval. So, the available amount of orthogonal pilot sequences, in a multicell system, is finite and can result in pilot contamination in two different situations.

In the first, when the number of pilot sequences is superior or equal to the number of user terminals, orthogonality is assumed between pilot sequences of the same cell. However, the frequency is reused and employed in a regular pattern to ensure that the intercell interference remains below a harmful level. On the other hand, when the number of pilot sequences is inferior to the number of user terminals, pilot sequences are reused within the same cell to reduce the training overhead.[2] In this way, nonorthogonal pilot sequences need to be employed, which is the major source of pilot contamination, known as intracell interference.[1]

The influence of all the interference is normally modeled in the signal-to-interference ratio (SIR) parameter. As our main contribution, this work presents an accurate log-normal approximation to the cumulative distribution function (CDF) of SIR in two scenarios: when the number of antennas at BS is finite and when this number is large enough and can be considered infinite. Moreover, an approximation for the capacity has also been presented, in which the slow fading is a combination of path loss and log-normal shadowing. Practical case studies are provided to illustrate the very good match between simulation results and our analytical approximation. The validity of our approximation is extended not only to uniform user distribution but also to users distributed nonuniformly. To the best of our knowledge, no similar results have been found in the literature.

The remaining of this paper is organized as follows: section 2 presents the system model, while section 3 proposes closed-form expressions for SIR and system capacity. Section 4 shows some examples of a common scheme in the literature, and final remarks are given in section 5.

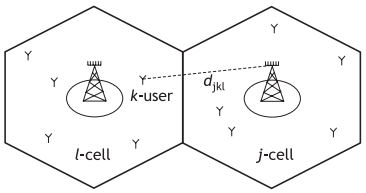

FIGURE 1 System model with hexagonal cells and the distance d_{jkl} between k-user of l-cell and the BS in j-cell

2 | SYSTEM MODEL

Based on Marzetta's work,[3] our model, as shown in Figure 1, consists of a hexagonal cellular geometry. The BSs, located at the center of the cell, hold a determined number of antennas, M. For operation simplicity, the orthogonal frequency division multiplexing (OFDM)/time division duplex (TDD) is used. In this way, it is assumed that the channel is flat and exhibits a reciprocity behavior, reducing the overhead required for the acquisition of channel state information (CSI) by means of uplink training signals. The OFDM symbol interval is denoted by T_s, the subcarrier spacing by Δf, and the useful symbol duration by $T_u = 1/\Delta f$.

The scheme consists of a tessellation of noncooperative hexagonal cells. The N BSs are deterministically distributed in a circular region. It is assumed that the BSs employ frequency reuse Δ and that the BSs in different bands do not interfere with each other. There are, in total, K user terminals within each of the L active cells reusing the same frequency bands and pilot sequences. The system performance is studied under the identically and independently distributed (iid) fading assumption, including log-normal shadow fading, and geometric attenuation. The user terminals are randomly distributed in a hexagonal cell, whose radius is r_c, except in a disk of radius 100 m centered at the BS. Interfering nodes, which affect a particular cell, are separated into tiers that reuse the same frequency band and are within eight cell-diameters of that cell, as adopted by Marzetta.[3] In addition, perfect synchronism is assumed between the signals received from different cells, which is the worst case in terms of pilot contamination.

As in Filho et al[4], for each subcarrier the vector between the j-th BS and the k-th user at the l-th cell is denoted by $\mathbf{g}_{jkl} = \sqrt{\beta_{jkl}}\underline{g}_{jkl}$, in which β_{jkl} refers to the long-term fading coefficient, comprising path loss and log-normal shadowing. $\underline{g}_{jkl}$ is the short-term fading channel vector that follows a normal distribution with zero mean and unitary variance. In its turn, the orthogonal pilot sequences set is represented by $\Psi = [\Psi_1 \Psi_2 \ldots \Psi_\tau] \in \mathbb{C}^{\tau \times \tau}$, in which τ is the number of available sequences, $\Psi^H \Psi = I_\tau$, and ^{H} is Hermitian operator. Considering Ψ_k the assigned sequence for the k-th user, the received signal at the j-th user BS during the training stage is

$$\mathbf{Y}_j^p = \sqrt{\rho_p} \sum_{l=1}^{L} \sum_{k=1}^{\tau} \mathbf{g}_{jkl} \Psi_k^H + \mathbf{N}_j^p, \tag{1}$$

in which ρ_p is the uplink pilot transmit power, and $\mathbf{N}_j^p \in \mathbb{C}^{M \times \tau}$ is the additive white Gaussian noise (AWGN) matrix with iid elements following a complex normal distribution with zero mean and variance σ_n^2. The j-th user BS estimates the k-th user CSI by correlating $\mathbf{Y}_j^p$ with Ψ_k. By acquiring such estimates, the BS is able to perform linear detection in uplink employing the maximal ratio combining (MRC) scheme in frequency domain. So, the j-th user BS receives the following signal

$$y_j = \sqrt{\rho_u} \sum_{l=1}^{L} \sum_{k=1}^{K} \mathbf{g}_{jkl} x_{kl} + n_j, \tag{2}$$

in which ρ_u is the uplink data transmit power, x_{kl} is the data symbol from the k-th user of the l-th user cell, and n_j is the $M \times 1$ AWGN sample vector.

3 | PROPOSED APPROXIMATION

Next, this work presents the assumptions and steps to derive our proposed approximation for the distribution of SIR, considering a finite and infinite number of antennas M. Without loss of generality, only the uplink SIR is considered, but the downlink case is very similar.

3.1 | Finite number of BS antennas

From the estimates available at the j-th user BS, and the received signal during the uplink data transmission stage (Equation 2), the uplink SIR, employing MRC, can be given by,[4, Eq. 14]

$$\gamma_{jk}^{u} = \frac{\zeta_{jkj}^2}{\sum\limits_{l=1, l \neq j}^{L} \zeta_{jkl}^2 + \frac{\alpha_{jk}^2}{M} \left(\sum\limits_{l=1}^{L} \sum\limits_{i=1}^{K} \zeta_{jil} + \frac{\sigma_n^2}{\gamma \rho_u} \right)}, \tag{3}$$

in which the index jkl refers to a quantity related to the kth terminal in the lth cell and the BS in the jth cell. $\zeta_{jkj} = \frac{\beta_{jkj}}{d_{jkj}^{\varepsilon}}$, in which the variable β_{jkl} represents the shadow fading coefficient, modeled as a log-normal random variable with mean μ and variance σ^2 in the logarithmic scale, and d_{jkl} is the distance between the user terminal and the correspondent BS. The variable ε is the decay exponent, typically $\varepsilon \geq 2$. The AWGN follows a complex normal distribution with zero mean and variance σ_n^2, $\alpha_{jk}^2 = \sum_{l=1}^{L} \zeta_{jkl} + \frac{\sigma_n^2}{\rho^p}$, and $\rho_u = 1$ as it is assumed that all users have unit transmit power. In addition, γ accounts for the transmit power loss due to cyclic prefix (CP).

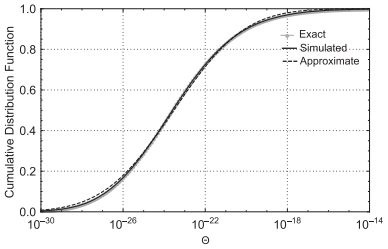

FIGURE 2 Simulated, approximate, and exact cumulative distribution function for Θ ($r = 100$ and $R = 2000$)

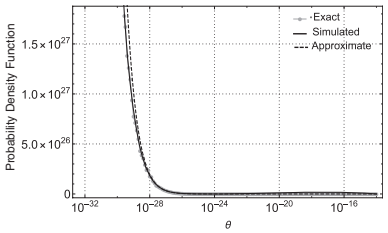

FIGURE 3 Simulated, approximate, and exact probability density function for Θ ($r = 100$ and $R = 2000$)

Given a certain number of users inside the cell, it is assumed that all of them are uniformly distributed from the inner radius r to the outer radius R for a given tier. As the circumference of a circle is proportional to its radius, the probability density function (PDF) $f_{d_{jkl}}(d)$ is also proportional to its radius, that is, $f_{d_{jkl}}(d) = ad$ for some constant a. The condition $\int_r^R f_{d_{jkl}}(d) = 1$ implies $a = \frac{2}{R^2 - r^2}$. Therefore

$$f_{d_{jkl}}(d) = \begin{cases} \frac{2d}{R^2 - r^2}, & \text{if } r \leq \mathrm{d} \leq \mathrm{R} \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

The computation of the exact distribution of Equation 3 is very intricate, and possibly, the final solution would be useless due to its complexity. In order to corroborate this assertion, the exact distribution of the numerator of Equation 3, defined as $\Theta = \beta_{jkj}^2 d_{jkj}^{-2\epsilon}$, is derived. As mentioned before, the random variable β_{jkj} is log-normally distributed, and the distribution of d_{jkj} is given by Equation 4. Finally, the PDF of Θ can be computed as the ratio of two random variables[5] and is given in an exact manner as in Equation 5, where erf$(\cdot)$ is the error function.

Unfortunately, this PDF does not have a closed-form solution for its characteristic function; therefore, the computation of the PDF of the sum, presented in the denominator, would not be possible. In order to circumvent this problem, a proposal is made to approximate Θ by a log-normal distribution X. Through the moment matching method, the mean $\widehat{\mu}_x$ and the variance $\widehat{\sigma}_x^2$ of X are estimated using the equalities $\mathbf{E}[\log X] = \mathbf{E}[\log \Theta]$ and also $\mathbf{V}[\log X] = \mathbf{V}[\log \Theta]$, where $\mathbf{E}[\cdot]$ and $\mathbf{V}[\cdot]$ are the expectation and variance operator, respectively. Therefore, the following can be written

$$f_\Theta(\theta) = \frac{\theta^{-\frac{\epsilon+1}{\epsilon}} e^{\frac{2(\sigma^2+\mu\epsilon)}{\epsilon^2}} \left[\mathrm{erf}\left(\frac{-\frac{1}{2}\epsilon(2\epsilon\log(R)+\log(\theta))+2\sigma^2+\mu\epsilon}{\sqrt{2}\sigma\epsilon} \right) - \mathrm{erf}\left(\frac{-\frac{1}{2}\epsilon(2\epsilon\log(r)+\log(\theta))+2\sigma^2+\mu\epsilon}{\sqrt{2}\sigma\epsilon} \right) \right]}{2\epsilon(r^2 - R^2)} \tag{5}$$

$$\begin{aligned} \widehat{\mu}_x &= \mathbf{E}[\log(\Theta)] = 2\mathbf{E}[\log(\beta)] - 2\epsilon\mathbf{E}[\log(d_{jkl})] = \\ &= 2\mu - \frac{\epsilon\{r^2[1 - 2\log(r)] + R^2[1 - 2\log(R)]\}}{r^2 - R^2}, \end{aligned} \tag{6}$$

where $\mathbf{E}[\log(d_{jkl})]$ is evaluated with respect to the distribution of d_{jkl}, as defined in 4. In the same way, the variance $\widehat{\sigma}_x^2$ can be computed as:

$$\begin{aligned} \widehat{\sigma}_x^2 &= \mathbf{V}[\log(\Theta)] = 4\mathbf{V}[\log(\beta)] + 4\epsilon^2\mathbf{V}[\log(d_{jkl})] = \\ &= 4\sigma^2 + \frac{\epsilon^2 \left\{ r^4 - r^2R^2 \left[2 + 4\log^2\left(\frac{R}{r}\right) \right] + R^4 \right\}}{(r^2 - R^2)^2}. \end{aligned} \tag{7}$$

Then, as supposed, the PDF of X is the very well-known log-normal distribution given as[5]

$$f_X(x) = \frac{e^{-\frac{(\log(x) - \widehat{\mu}_x)^2}{2\widehat{\sigma}_x^2}}}{\sqrt{2\pi}x\widehat{\sigma}_x}. \tag{8}$$

The accuracy of the proposed approximation is illustrated in Figures 2 and 3. These figures show the approximate, simulated, and exact CDF and PDF of the random variable Θ, respectively. As can be noted, our approximation is excellent, and this is the key assumption for approximating the distribution of Equation 3 as a log-normal distribution.

It is well known, in the literature, that the ratio of two log-normal variables is also a log-normal variable. In this sense, if a log-normal distribution could approximate the distribution of the denominator of Equation 3, then the entire ratio would also be log-normally distributed.

To this end, every term inside the sum of the denominator is approximated as a log-normal distribution (similar to the X variable), and then, based on the result of Filho et al,[6] the sum of log-normal variables is approximated as a log-normal variable. This can be performed by matching the first two moments of the inverse exact sum with those of the inverse log-normal approximation. In this way, the parameters of the resulting log-normal distribution are determined.

Let's call the first term in the denominator of Equation 3 as S. This random variable can be approximated as the sum of approximate iid log-normal variables, $S = \sum_{l\neq j} \beta_{jkl}^2 d_{jkl}^{-2\epsilon}$. We want to approximate the sum of log-normal variates in S by a single log-normal variable, denoted here as Y. To this end, the moments are matched in the following way: $\mathbf{E}[Y^{-1}] = \mathbf{E}[S^{-1}]$, and also, $\mathbf{E}[Y^{-2}] = \mathbf{E}[S^{-2}]$. As $\mathbf{E}[Y^{-1}] = e^{-(\widehat{\mu}_y - \widehat{\sigma}_y^2/2)}$ and $\mathbf{E}[Y^{-2}] = e^{-2(\widehat{\mu}_y - 2\widehat{\sigma}_y^2/2)}$, with some mathematical manipulations, it is possible to write:

$$\begin{aligned} \widehat{\mu}_y &= 0.5 \ln \mathbf{E}[S^{-2}] - 2 \ln \mathbf{E}[S^{-1}] \\ \widehat{\sigma}_y^2 &= \ln \mathbf{E}[S^{-2}] - 2 \ln \mathbf{E}[S^{-1}], \end{aligned} \tag{9}$$

where $\mathbf{E}[S^n]$ is the n-th moment of S, while $\widehat{\mu}_y$ and $\widehat{\sigma}_y$ are the mean and standard deviation of the approximated log-normal random variable Y, respectively.

Following the same reasoning, the second term in the denominator of Equation 3 presents the terms $V = \frac{\alpha_{jk}^2}{M}$ and $W = \sum_{l=1}^{L} \sum_{k=1}^{K} \frac{\beta_{jkl}}{d_{jkl}^{\epsilon}} + \sigma_n^2$. Both of them are approximated as log-normal variables, and as the product of log-normal distributions is also log-normally distributed, the random variable $Z = VW$ can be approximated as a log-normal variable. So, $\widehat{\mu}_z = \mu_v + \mu_w$ and $\widehat{\sigma}_z^2 = \sigma_v^2 + \sigma_w^2$, where μ_v, μ_w, σ_v^2, and σ_w^2 are means and variances of the variables V and W, respectively. Following the same rationale, the sum $Y + Z$ can be approximated by another log-normal variable, whose parameters are $\widehat{\mu}_{yz}$ and $\widehat{\sigma}_{yz}^2$.

As the terms of the numerator and denominator are uncorrelated, the distribution of the ratio of log-normals is also log-normally distributed, so the resultant parameters of the distribution of the SIR are $\widehat{\mu}_{xyz} = \widehat{\mu}_x - \widehat{\mu}_{yz}$ and $\widehat{\sigma}_{xyz}^2 = \widehat{\sigma}_x^2 + \widehat{\sigma}_{yz}^2$. Finally, the CDF of γ_{jk}^{u}, given in Equation 3, can be written as

$$\mathbf{P}[\gamma_{jk}^{\mathrm{u}} \leq \nu] = \frac{1}{2} \mathrm{erfc}\left(\frac{\log \nu - \widehat{\mu}_{xyz}}{\widehat{\sigma}_{xyz}^2 \sqrt{2}} \right), \tag{10}$$

where $\mathrm{erfc}(\cdot)$ is the complementary error function.

According to Marzetta,[3] the net capacity per terminal for uplink, C_{jk}^{u}, in bits/sec/terminal for uplink is given by

$$C_{jk}^{\mathrm{u}} = \frac{B}{\Delta} \left(\frac{T_{\mathrm{slot}} - T_{\mathrm{pilot}}}{T_{\mathrm{slot}}} \right) \left(\frac{T_{\mathrm{u}}}{T_{\mathrm{s}}} \right) \log_2(1 + \gamma_{jk}^{\mathrm{u}}), \tag{11}$$

where B is the total bandwidth in Hz, T_{slot} is the slot length, T_{pilot} is the time to transmit pilot sequences, T_{u} is the useful symbol duration, and T_{s} is the OFDM symbol interval, where the time is measured in seconds. Then, defining a constant $U = \frac{B(T_{\mathrm{slot}} - T_{\mathrm{pilot}})T_u}{\Delta T_{\mathrm{slot}} T_s}$, it is possible to obtain a closed-form expression for the CDF of C_{jk}^{u} as shown in Equation 12.

$$\mathbf{P}[C_{jk}^{\mathrm{u}} \leq c] = \frac{|\widehat{\mu}_{xyz} - \log(2^{c/U} - 1)| \mathrm{erf}\left[\frac{|\widehat{\mu}_{xyz} - \log(2^{c/U} - 1)|}{\sqrt{2}\widehat{\sigma}_{xyz}} \right] |- \widehat{\mu}_{xyz} + \log(2^{c/U} - 1)|}{2|\widehat{\mu}_{xyz} - \log(2^{c/U} - 1)|} \tag{12}$$

3.2 | Infinite number of BS antennas

As shown in Madhusudhanan et al,[7] as the number of BS antennas is very large ($M \to \infty$), the effects of uncorrelated noise and fast fading vanish completely, and there is no interference between data transmissions inside a cell. So, the simplest linear precoders and detectors are proved to be optimal. However, as every terminal is assigned an orthogonal time-frequency pilot sequence reused in other cells according to Δ, the only source of pilot contamination is the intercell interference. The asymptotic uplink SIR can be expressed as

$$\gamma_{jk}^{\mathrm{u}} = \frac{\zeta_{jkj}^2}{\sum_{l=1, l\neq j} \zeta_{jkl}^2}. \tag{13}$$

Note that as Equation 13 is composed in the numerator by Θ and in the denominator by S, it can be approximated by X and Y, respectively, as a ratio of two log-normal random variables whose mean and variance parameters are, respectively, given by $\widehat{\mu}_{xy} = \widehat{\mu}_x - \widehat{\mu}_y$ and $\widehat{\sigma}_{xy}^2 = \widehat{\sigma}_x^2 + \widehat{\sigma}_y^2$. Finally, when $M \to \infty$, the CDFs of γ_{jk}^{u} and C_{jk}^{u} can be obtained from Equations 10 and 12, respectively, assuming $\widehat{\mu}_{xyz} = \widehat{\mu}_{xy}$ and $\widehat{\sigma}_{xyz}^2 = \widehat{\sigma}_{xy}^2$.

3.2.1 | Nonuniform spatial distributions

In this section, an approximation for Equation 13 considering two types of nonuniform spatial distributions for terminal users is presented. Center-intensive user distribution and edge-intensive user distribution, which are not only analytically simple but can model practical scenarios. The center-intensive user distribution is particularly suitable for urban scenarios with populated buildings, while the edge-intensive user distribution is used to model rural mountainous propagation scenarios.[8]

For center-intensive user distribution, the PDF of the distance can be expressed by

$$f_{d_{jkl}}(d) = \frac{a_c(R - d)^2 + 2d}{R^2 - r^2} b_c, \quad r \leq d \leq R \tag{14}$$

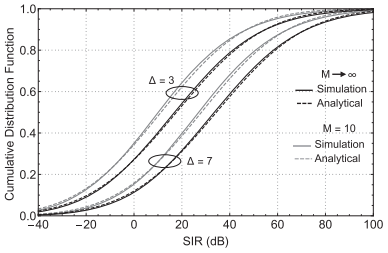

FIGURE 4 Cumulative distribution of SIR (dB) when $M = 10$ (gray curves) and $M \to \infty$ (black curves)

in which the scaling factor a_c controls how strongly the users cluster towards the BS, and $b_c = \frac{3(\mathrm{r}+\mathrm{R})}{a_c(\mathrm{r}^2-2\mathrm{rR}+\mathrm{R}^2)+3(\mathrm{r}+\mathrm{R})}$. In this case, $\widehat{\mu}_x$ and $\widehat{\sigma}_x$ are given, respectively, by Equations 15 and 16. The variable $\widehat{\sigma}_x$ is a function of the parameter A.

$$\begin{aligned}\widehat{\mu}_x = (r-R)\frac{-7rRa_c + r\epsilon(2ra_c+9) + 11R^2a_c + 9R)}{3(r-R)[r^2a_c + r(3-2Ra_c) + R(Ra_c+3)]} + \frac{6\mu(r^2a_c - 2rRa_c + R^2a_c + 3r + 3R)}{3(r-R)[r^2a_c + r(3-2Ra_c) + R(Ra_c+3)]} \\ + \frac{6R^2\epsilon\log(R)(Ra_c+3)}{3(r-R)[r^2a_c + r(3-2Ra_c) + R(Ra_c+3)]} + \frac{6r\epsilon\log(r)(a_cr^2 - 3a_crR + 3a_cR^2 + 3r)}{3(r-R)[r^2a_c + r(3-2Ra_c) + R(Ra_c+3)]}\end{aligned} \tag{15}$$

$$\begin{aligned}\widehat{\sigma}_x &= -\sqrt{A(\mathrm{r}-\mathrm{R})^2 - 36\mathrm{rR}^2\epsilon^2\log\left(\frac{\mathrm{r}}{\mathrm{R}}\right)\left\{(a_c\mathrm{R}+3)(-3a_c\mathrm{rR} + \mathrm{r}(a_c\mathrm{r}+3) + 3a_c\mathrm{R}^2)\log\left(\frac{\mathrm{r}}{\mathrm{R}}\right) + a_c(\mathrm{r}-\mathrm{R})[3a_c\mathrm{rR} + \mathrm{R}(9-5a\mathrm{R}) + \mathrm{r}]\right\}} \\ A &= \epsilon^2\{4a_c^2\mathrm{r}^4 + 2a_c\mathrm{r}^3(21-17a_c\mathrm{R}) + 3\mathrm{r}^2[a_c\mathrm{R}(59a_c\mathrm{R}-26)+27] + 2\mathrm{rR}[a_c\mathrm{R}(123-116a_c\mathrm{R})+81] + \mathrm{R}^2[a_c\mathrm{R}(49a_c\mathrm{R}+366)+81]\} \\ &\quad +36\sigma^2[a_c\mathrm{r}^2 + \mathrm{r}(3-2a_c\mathrm{R}) + \mathrm{R}(a_c\mathrm{R}+3)]^2\end{aligned} \tag{16}$$

On the other hand, for the edge-intensive user distribution, the PDF of the distance is given by

$$f_{d_{jkl}}(d) = \frac{a_ed^2 + 2d}{R^2 - r^2}b_e, \quad r \le d \le R \tag{17}$$

in which a_e controls how strongly the users cluster toward the cell edge and $b_e = \frac{3(\mathrm{r}+\mathrm{R})}{a_e(\mathrm{r}^2+\mathrm{rR}+\mathrm{R}^2)+3(\mathrm{r}+\mathrm{R})}$. In this case, $\widehat{\mu}_x$ and $\widehat{\sigma}_x$ are given by Equation 18, and the variable $\widehat{\sigma}_x$ is a function of the parameter B. From Equations 14 and 17, it can be noted that, for the special case $a_c = a_e = 0$, both distributions reduce to the conventional uniform distribution.

$$\begin{aligned}\widehat{\mu}_x &= \frac{\epsilon[\mathrm{r}^2(2a_e\mathrm{r}+9) - 6\mathrm{r}^2(a_e\mathrm{r}+3)\log(\mathrm{r}) - 2a_e\mathrm{R}^3 + 6\mathrm{R}^2(a_e\mathrm{R}+3)\log(\mathrm{R}) - 9\mathrm{R}^2]}{3a_e\mathrm{r}^3 - 3\mathrm{R}^2(a_e\mathrm{R}+3) + 9\mathrm{r}^2} + 2\mu \\ \widehat{\sigma}_x &= \sqrt{\frac{\epsilon^2 B}{9(a_e\mathrm{r}^3 - \mathrm{R}^2(a_e\mathrm{R}+3) + 3\mathrm{r}^2)^2} + 4\sigma^2} \\ B &= 4a_e^2\mathrm{r}^6 - 8a_e^2\mathrm{r}^3\mathrm{R}^3 + 4a_e^2\mathrm{R}^6 + 42a_e\mathrm{r}^5 - 42a_e\mathrm{r}^3\mathrm{R}^2 - 42a_e\mathrm{r}^2\mathrm{R}^3 - 36\mathrm{r}^2\mathrm{R}^2(a_e\mathrm{r}+3)(a_e\mathrm{R}+3)\log^2(\mathrm{r}) \\ &\quad -36\mathrm{r}^2\mathrm{R}^2(a_e\mathrm{r}+3)(a_e\mathrm{R}+3)\log^2(\mathrm{R}) + 36a_e\mathrm{r}^2\mathrm{R}^2(\mathrm{r}-\mathrm{R})\log(\mathrm{R}) + 36\mathrm{r}^2\mathrm{R}^2\log(\mathrm{r})[a_e(\mathrm{R}-\mathrm{r}) \\ &\quad +2(a_e\mathrm{r}+3)(a_e\mathrm{R}+3)\log(\mathrm{R})] + 42a_e\mathrm{R}^5 + 81\mathrm{r}^4 - 162\mathrm{r}^2\mathrm{R}^2 + 81\mathrm{R}^4\end{aligned} \tag{18}$$

4 | SIMULATION RESULTS

For the comparison between the analytical and simulated results, it is assumed that the cellular area of interest is a tessellation filled by hexagonal cells with a radius of $r_c = 1600$ meters. The BSs are distributed deterministically in the center of each hexagon cell whose whole radius is 100 meters.

The OFDM parameters are identical to long-term evolution (LTE): $T_s = 500/7$ μs, $\Delta f = 15$ kHz, $T_u = 1/\Delta f$, $\frac{T_{\mathrm{slot}} - T_{\mathrm{pilot}}}{T_{\mathrm{slot}}} = 3/7$, and $B = 20$ MHz. The shadow fading is modeled as a random variable, β_{jkl}, that follows a log-normal distribution with $\mu = 0$ and $\sigma = 8$ dB. The frequency reuse factor varies as $\Delta = 3, 7$, while the decay exponent is $\epsilon = 3.8$. There is no power control, and the power of BSs and the user terminals are unitary, in addition to γ.

Figure 4 shows the CDFs of SIR, in which the analytical result is represented by dashed black lines, while the solid ones refer to the simulated data. The circles indicate the curves for $\Delta = 3, 7$. As can be seen, our approximation perfectly matches the simulated results for all the range of SIR.

In order to investigate the influence of the number of antennas M, Figure 4 also shows the CDF for $\sigma_n = 0$, $K = 5$, and $M = 10$ in gray color and for $M \to \infty$ in black color. Using $\Delta = 7$, for an arbitrary value of 0.6 for the CDF, a gap of 6 dB between the SIRs can be observed. The CDF was also calculated for $M = 10^6$, and in this case, the gap between the finite M and the asymptotic case, although small, still does not vanish. So, it has been observed that the convergence between Equations 3 and 13 is slow. In this sense, our derived expression can be used to assess how many antennas are necessary to achieve a target performance. In addition, as expected, as Δ increases, the number of interferers decreases, and the SIR becomes larger, so the curve of $\Delta = 7$ is positioned to the right in this plot, while the curve of $\Delta = 3$ is on the left side.

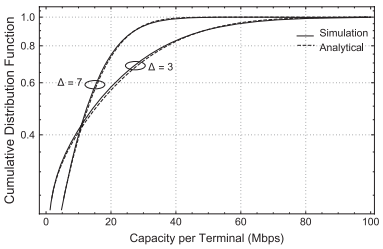

FIGURE 5 Cumulative distribution function of the net uplink capacity per terminal (Mbps)

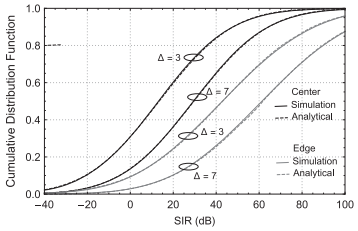

FIGURE 6 Cumulative distribution function of SIR (dB), considering center- and edge-intensive user distributions

Figure 5 shows the analytical (Equation 12) and simulated CDFs of the net uplink capacity per terminal represented by dashed and solid lines, respectively. It is observed that, for larger Δ, the capacity increases when the asymptotic SIR is low. In this region, the gains due to SIR compensate the loss by the frequency reuse, associated with the reduction in the bandwidth used by each cell. On the other side, when the SIR is high, a higher frequency reuse factor causes a net decrease in the system performance. As before, in all the cases, both simulated and analytical curves are almost indistinguishable.

Finally, considering infinite M and nonuniform user distribution, Figure 6 shows the CDFs of SIR, in which the analytical result is represented by dashed lines, while the solid ones refer to the simulated data. The circles indicate the curves for $\Delta = 3, 7$. The black and gray colors refer to the curves obtained from the center and edge user distributions, respectively. As it can be seen, our approximation again presents an excellent agreement.

5 | CONCLUSIONS

In this work, the performance of a massive MIMO system is evaluated, modeling the shadowing as a log-normal distribution. We have presented an approximated closed-form expression for the PDF and CDF of the SIR for finite and infinite number of antennas. Using the proposed approximation, the CDF for the net capacity is also derived, which allows the analysis of the system performance. All approximations are based on the moment matching method. The results have been validated by numerical simulations and have shown an excellent agreement, considering both uniform and nonuniform user distributions.

ACKNOWLEDGMENT

The authors thank the Grant #2016/16181-2, São Paulo Research Foundation (FAPESP) for support in this research.

ORCID

Michelle S. P. Facina http://orcid.org/0000-0002-0785-9606
Gustavo Fraidenraich http://orcid.org/0000-0002-0517-1496

REFERENCES

1. Larsson EG, Edfors O, Tufvesson F, Marzetta TL. Massive MIMO for next generation wireless systems. *IEEE Commun Mag*. 2014;*52*(2):186-195.
2. Bai T, Heath RW. Analyzing uplink SINR and rate in massive MIMO systems using stochastic geometry. *IEEE Trans Commun*. 2016;*64*(11):4592-4606.
3. Marzetta TL. Noncooperative cellular wireless with unlimited numbers of base station antennas. *IEEE Trans Wirel Commun*. 2010;*9*(11):3590-3600.
4. Filho JCM, Panazio C, Abrao T. Uplink performance of single-carrier receiver in massive MIMO with pilot contamination. *IEEE Access*. 2017;*5*:8669-8681.
5. Papoulis A, Pillai SU. *Probability, Random Variables, and Stochastic Processes*. 4th ed. New York: McGraw-Hill Higher Education; 2002.
6. Filho JCSS, Cardieri P, Yacoub MD. Simple accurate lognormal approximation to lognormal sums. *Electr. Lett.* 2005;*41*(18):1016-1017.
7. Madhusudhanan P, Li X, Liu Y, Brown TX. Stochastic geometric modeling and interference analysis for massive MIMO systems. 11th International Symposium and Workshops on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks, Tsukuba Science City, 2013, pp. 15–22.
8. Kong C, Zhong C, Zhang Z. Performance of ZF precoder in downlink massive MIMO with non-uniform user distribution. *J Commun Netw*. 2016;*18*(5):688-698.

How to cite this article: Facina MSP, Fraidenraich G. Accurate log-normal approximation to the signal-to-interference ratio in massive multiple-input multiple-output systems. *Internet Technology Letters* 2018;1:e48. https://doi.org/10.1002/itl2.48