How Accurate is Your Gaussian/Gamma Approximation?

Peter J. Smith, *Fellow, IEEE*, Rajitha Senanayake, *Member, IEEE*, and Pawel A. Dmochowski, *Senior Member, IEEE*

***Abstract*—Gaussian and Gamma approximations appear very frequently in communications theory applications. However, in most of the applications, numerical methods are used to illustrate the accuracy of such approximations as there are no simple general purpose analytical tools available to provide a more rigorous assessment of the error. We present a simple and accurate tool to measure the accuracy of Gaussian and Gamma approximations. Specifically, we consider the first order correction terms of the approximations and derive their upper bounds. Our tool provides an accurate approximation of the error term based on the third and the fourth moments of the random variable. To demonstrate the usefulness of our analysis we consider two topical applications based on mmWave channel modeling and distributed massive multiple-input-multiple-output systems. In both cases, the proposed technique provides valuable insights into the accuracy of the commonly used approximations.**

***Index Terms*—Gaussian distribution, Gamma distribution.**

I. Introduction

In theoretical analysis of communication systems we often come across random variables for which the exact distributions are very challenging to compute and do not result in closed form expressions. This motivates us to approximate the distributions of these complex random variables with simpler and well-known distributions, such as Gaussian or gamma distributions [1], [2]. While this approach leads to simplifications in performance characterization and network design, a thorough analysis of the accuracy of such approximations is very important yet often done only via numerical methods.

In this letter, we take a more rigorous approach and develop a theoretical tool to measure the accuracy of approximating a given distribution by either a Gaussian or gamma distribution. While our methodology is applied to the Gaussian and gamma distributions, which are commonly encountered in communications theory, the analysis can be extended to other distributions such as the beta distribution.

An unknown distribution can be approximated by a Gaussian or gamma distribution, by matching the first and second moments in what is known as *moment matching*. Consider a random variable, V, for which the distribution is unknown and the expected value and the variance of V are given by $\nu_1 = \mathrm{E}[V]$ and $\nu_2 = \mathrm{Var}[V]$, respectively. A Gaussian random variable, W, with the same first and second moments has a probability density function (PDF) given by $\phi_W(x) = \frac{1}{\sqrt{2\pi\nu_2}} e^{\frac{(x-\nu_1)^2}{2\nu_2}}$. If preliminary tests suggest that the Gaussian distribution is a good approximation to the distribution of V, we can approximate the PDF of V by $\phi_W(x)$ based on a second moment match. Similarly, the distribution of a gamma random variable, W, with the same first and second moments has a PDF given by $f_W(x) = x^{m-1}\frac{e^{-x/\theta}}{\theta^k \Gamma(k)}$, where $m = \frac{\nu_1^2}{\nu_2}$ and $\theta = \frac{\nu_2}{\nu_1}$. If preliminary tests suggest that the gamma distribution is a good approximation to the distribution of V, we can approximate the distribution of V by $f_W(x)$ based on second moment matching. These approximations can then be strengthened by providing measures of their accuracy.

In this letter, we present a theoretical analysis to measure the accuracy of Gaussian and gamma approximations. Based on the series expansion methodology in [3, eq. 12. 51], for the Gaussian approximation we use the Edgeworth series expansion [4] and, for the gamma approximation we use the Laguerre series expansion, to derive easy-to-evaluate approximations of the correction terms. Our contributions also extend to two interesting applications in communications theory related to mmWave communications and distributed massive MIMO. Based on a rigorous theoretical analysis we illustrate how the developed tools can provide interesting insights that are difficult to gain through numerical methods.

II. Accuracy of the Gaussian Approximation

Let V denote the original random variable for which the distribution is unknown or difficult to analyse. Based on V, we define the standardized random variable X as

$$X = (V - \nu_1)/\sqrt{\nu_2}, \tag{1}$$

where ν_1 and ν_2 denote the expected value and the variance of V, respectively. If preliminary results suggest a Gaussian distribution is a good approximation for X, based on the Edgeworth series expansion in [3, eq. 12.59'] we can write the cumulative distribution function (CDF) of X as

$$F_X(x) = \Phi_X(x) - \left[\frac{\mu_3}{6}(x^2-1) - \frac{(\mu_4-3)}{24}(x^3-3x) - \frac{\mu_3^2}{72}(x^5-10x^3+15x)\right]\phi_X(x) + \dots, \tag{2}$$

where $\Phi_X(x) = \frac{1}{2}[1+\mathrm{erf}(\frac{x}{\sqrt{2}})]$ and $\phi_X(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ represent the CDF and the PDF of the standard normal distribution, respectively, and, μ_3 and μ_4 represent the third and the fourth

Manuscript received March 7, 2018; revised March 31, 2018; accepted April 8, 2018. Date of publication April 12, 2018; date of current version October 11, 2018. This work was supported by the Australian Research Council through the Discovery Early Career Researcher under Award DE180100501. The associate editor coordinating the review of this paper and approving it for publication was D. So. *(Corresponding author: Rajitha Senanayake.)*

P. J. Smith is with the School of Mathematics and Statistics, Victoria University of Wellington, Wellington 6140, New Zealand (e-mail: peter.smith@sms.vuw.ac.nz).

R. Senanayake is with the Department of Electrical and Electronic Engineering, University of Melbourne, Melbourne, VIC 3010, Australia (e-mail: rajitha.senanayake@unimelb.edu.au).

P. A. Dmochowski is with the School of Engineering and Computer Science, Victoria University of Wellington, Wellington 6140, New Zealand (e-mail: pawel.dmochowski@ecs.vuw.ac.nz).

Digital Object Identifier 10.1109/LWC.2018.2826001

central moments of X. When the CDF of X is approximated by $\Phi_X(x)$, the magnitude of the leading error term can be defined based on the terminating expression in [3, eq. 12.61] where only the first three terms of (2) are used as

$$\Xi = \left| \frac{\mu_3}{6}(x^2-1)\phi_X(x) + \frac{(\mu_4-3)}{24}(x^3-3x)\phi_X(x) + \frac{\mu_3^2}{72}(x^5-10x^3+15x)\phi_X(x) \right|. \quad (3)$$

The terminating expression in [3, eq. 12.61] is used in many applications such as [2] and [5], and is motivated by the fact that X is a standard random variable. Specifically, the higher order terms of X, accompanied by large factorials in the denominator, take smaller values when X is small and thus can be ignored. Also, the leading error term in (3) goes to zero when X is exactly Gaussian. However, there is no precise analysis about the truncation error available in general [6, p. 187]. Note that in [2], (3) is used as a numerical measure of the Gaussian approximation accuracy. However, (3) is a complex function of x. Thus, we take a simple bounded approach and derive a more general and easy-to-evaluate approximation.

Note that (3) consists of the constants μ_3 and μ_4, and simple functions of a single argument, x. We find that $(x^2-1)\phi_X(x)$ has two maxima at $x=\pm\sqrt{3}$ and a minimum at $x=0$. Thus, $|(x^2-1)\phi_X(x)| \leq 0.4$. Similarly, we find $|(x^3-3x)\phi_X(x)| \leq 0.5506$ and $|(x^5-10x^3+15x)\phi_X(x)| \leq 2.3071$. Substituting these bounds into (3), Ξ can be upper bounded as

$$\Xi \leq 0.067|\mu_3| + 0.023|\mu_4-3| + 0.032|\mu_3^2|. \quad (4)$$

Based on (4) we can upper bound the correction term Ξ by simply calculating the constants, μ_3 and μ_4.

Discussion: In assessing accuracy, ideally the difference between the true CDF and the approximate CDF can be evaluated or bounded. However, since approximations are often used when the underlying distribution is intractable, this may be impractical. Hence, expansions such as (2) are useful as the leading order correction terms can be identified, recognizing that the full infinite series is usually intractable. The size of leading order corrections such as (3) are a useful general purpose tool in assessing accuracy. In this letter, we take a step further and upper bound such terms over all x-values as in (4) giving a single expression, solely in terms of μ_3 and μ_4. This upper bound can be further refined for a particular range of interest for x. The motivation here is that (4) is genuinely easy to use and links system parameters which control μ_3 and μ_4 with accuracy. Hence, the researcher is able to explore the critical question: *when is the approximation accurate?* without lengthy simulation studies.

Also note that existing approaches, such as the Berry-Esseen theorem [6, Ch. 11] provides a useful alternative to assess the accuracy in Gaussian approximations when structure exists, i.e., sums and in particular sums of independent variables. Our approach can be used even when no such structure exists in the underlying variables.

III. Accuracy of the Gamma Approximation

Let V denote the original random variable for which the distribution is approximated by a gamma distribution. We define a standardized random variable $X=\beta V$ such that $\beta=\nu_1/\nu_2$. If preliminary results suggest a gamma distribution is a good approximation for X, based on the Laguerre series expansion in [7] the PDF of X can be written as

$$f_X(x) = \frac{x^{m-1}e^{-x}}{\Gamma(m)}\left\{1+\xi_3 L_3^{(m)}(x)+\xi_4 L_4^{(m)}(x)+\ldots\right\}, \quad (5)$$

where $m=\frac{\nu_1^2}{\nu_2}$, $\xi_3=\frac{\Gamma(m)}{3!\Gamma(m+3)}(\mu_3-2m)$, $\xi_4=\frac{\Gamma(m)}{4!\Gamma(m+4)}(\mu_4-12\mu_3-3m^2+18m)$, $L_3^{(m)}(x)=x^3-3(m+2)x^2+3(m+2)(m+1)x-(m+2)(m+1)m$ and $L_4^{(m)}(x)=x^4-4(m+3)x^3+6(m+3)(m+2)x^2-4(m+3)(m+2)(m+1)x+(m+3)(m+2)(m+1)m$. Integrating (5) over x we can derive the CDF of X as in (6) shown at the top of the next page given at the top of the next page, with $F_m(x)=\int_0^x \frac{u^{m-1}e^{-u}}{\Gamma(m)}du$ denoting the CDF of the gamma distribution with shape parameter m and scale parameter one. When the CDF of X is approximated by $F_m(x)$, the magnitude of the leading error term can be defined based on the first six terms of (6) that contains $F_m(x),\ldots,F_{m+4}(x)$, as illustrated in (7) as shown at the top of the next page. Substituting the expressions for ξ_3 and ξ_4, following mathematical manipulation (7) can be expressed as

$$\Xi = |\varphi\Theta_1 + \vartheta\Theta_2|, \quad (8)$$

where $\Theta_1=[-F_m(x)+3F_{m+1}(x)-3F_{m+2}(x)+F_{m+3}(x)]$, $\Theta_2=[F_m(x)-4F_{m+1}(x)+6F_{m+2}(x)-4F_{m+3}(x)+F_{m+4}(x)]$, $\varphi=\frac{\mu_3-2m}{3!}$ and $\vartheta=\frac{\mu_4-12\mu_3-3m^2+18m}{4!}$. Note that unlike the Gaussian case in (3), (8) depends on an additional argument m, which is a function of ν_1 and ν_2. Thus, deriving a general upper bound for the correction term is mathematically challenging. For a given value of m, we can find the maximum values of $|\Theta_1|$ and $|\Theta_2|$ and derive an approximation on the correction term. In Table I we present the maximum values of $|\Theta_1|$ and $|\Theta_2|$ for different values of m and numerically show that for $m\geq 10$, $|\Theta_1|\leq 0.01$ and $|\Theta_2|<0.005$. From (8) and Table I, we can upper bound the correction term using a gamma approximation by simply calculating the third and the fourth central moments of X.

TABLE I
UPPER BOUND OF $|\Theta_1|$ AND $|\Theta_2|$ FOR DIFFERENT VALUES OF m

m	1	2	4	6	8	$\geq$10
$\lvert\Theta_1\rvert$	0.168	0.077	0.030	0.019	0.013	$\leq$ 0.01
$\lvert\Theta_2\rvert$	0.132	0.051	0.021	0.011	0.007	$\leq$ 0.005

IV. Applications

To illustrate the importance of our accuracy measure, we now discuss two topical applications. We assess the accuracy of the Gaussian approximation to a 3rd Generation Partnership Project (3GPP) mmWave channel model. We also consider a distributed massive MIMO system where the received signal-to-noise ratio (SNR) is approximated by the gamma distribution. We provide interesting insights into the accuracy of these approximations, which cannot be gained via simulations.

A. MmWave Communications

MmWave frequencies (30-300 GHz) are severely attenuated by path loss and shadowing. Recently the 3GPP released its study on channel models for frequencies above 6 GHz, in

$$F_X(x) = F_m(x) + m(m+1)(m+2)[(m+3)\xi_4 - \xi_3]F_m(x) + m(m+1)(m+2)[3\xi_3 - 4(m+3)\xi_4]F_{m+1}(x) + m(m+1)(m+2)[-3\xi_3 + 6(m+3)\xi_4]F_{m+2}(x) + m(m+1)(m+2)[\xi_3 - 4(m+3)\xi_4]F_{m+3}(x) + m(m+1)(m+2)[(m+3)\xi_4]F_{m+4}(x) + \dots \quad (6)$$

$$\Xi = m(m+1)(m+2)|((m+3)\xi_4 - \xi_3)F_m(x) + (3\xi_3 - 4(m+3)\xi_4)F_{m+1}(x) + (-3\xi_3 + 6(m+3)\xi_4)F_{m+2}(x) + (\xi_3 - 4(m+3)\xi_4)F_{m+3}(x) + ((m+3)\xi_4)F_{m+4}(x)|. \quad (7)$$

3GPP TR 38.900 V14.2.0 (Rel. 14) [8]. According to [8], for non-line-of-sight (NLOS) links the channel matrix with small-scale fading gains of a system with n_{rx} receive antennas and n_{tx} transmit antennas can be modeled as

$$\mathbf{H} = \sum_{r=1}^{R}\sum_{s=1}^{S}\sqrt{\frac{P_r}{S}}h_{rs}\mathbf{a}_{rx}(\theta_{rs}^{AoA}, \delta_{rs}^{AoA})\mathbf{a}_{tx}^*(\theta_{rs}^{AoD}, \delta_{rs}^{AoD}), \quad (9)$$

where R is the number of clusters, S is the number of subpaths in each cluster, h_{rs} is the complex small-scale fading gain on the s-th subpath of the r-th cluster, P_r is the r-th cluster power normalized so that $\sum_{r=1}^{R} P_r = 1$, $\mathbf{a}_{rx}(\cdot) \in \mathcal{C}^{n_{rx}\times 1}$ and $\mathbf{a}_{tx}(\cdot) \in \mathcal{C}^{n_{tx}\times 1}$ are the vector response functions for the receive and transmit antenna arrays to the angular arrivals and departures. Each subpath contains horizontal and vertical angle of arrivals (AoAs), θ_{rs}^{AoA}, δ_{rs}^{AoA}, respectively and horizontal and vertical angle of departures (AoDs), θ_{rs}^{AoD}, δ_{rs}^{AoD}, respectively. The (i,j)-th element of $\mathbf{H}$ is given by

$$\mathbf{H}(i,j) = \sum_{k=1}^{K} h_k\omega_k, \quad (10)$$

where $K = RS$, h_k and ω_k are the k-th elements in $\{h_{11}, h_{12}, \dots, h_{RS}\}$ and $\{\omega_{11}, \omega_{12}, \dots, \omega_{RS}\}$, respectively. Also, $\omega_{rs} = \sqrt{\frac{P_r}{S}}\mathbf{a}_{rx,i}(\theta_{rs}^{AoA}, \delta_{rs}^{AoA})\mathbf{a}_{tx,j}^*(\theta_{rs}^{AoD}, \delta_{rs}^{AoD})$ with $\mathbf{a}_{rx,i}$ and $\mathbf{a}_{tx,j}$ denoting the i-th and the j-th elements of $\mathbf{a}_{rx}$ and $\mathbf{a}_{tx}$. Consider a simple case of single antenna polarization [8], where $h_k = e^{j\upsilon_k}$ with $\upsilon_k \sim \mathcal{U}[0, 2\pi]$. Even for such a simple model deriving the exact distribution of $\mathbf{H}(i,j)$ is mathematically challenging. Therefore, in the following we assess the accuracy of a Gaussian approximation.

Based on Section II, we compute the first four central moments of the real and imaginary parts of $\mathbf{H}(i,j)$. Focusing only on the real part we define $V = \mathrm{Re}\{\mathbf{H}(i,j)\}$ and write

$$V = \sum_{k=1}^{K}(\mathrm{Re}(h_k)\mathrm{Re}(\omega_k) - \mathrm{Im}(h_k)\mathrm{Im}(\omega_k)), \quad (11)$$

where $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ denote the real and the imaginary parts of a complex number. Note that $\mathrm{Re}(h_k) = \cos\upsilon_k$ and $\mathrm{Im}(h_k) = \sin\upsilon_k$. Thus, $\mathrm{E}[\mathrm{Re}(h_k)] = \mathrm{E}[\mathrm{Im}(h_k)] = 0$ and $\mathrm{E}[\mathrm{Re}(h_k)^2] = \mathrm{E}[\mathrm{Im}(h_k)^2] = 1/2$, based on which we can show that $\mathrm{E}[V] = 0$ and $\mathrm{E}[V^2] = \frac{1}{2}\sum_{k=1}^{K}|\omega_k|^2$. Whilst not shown here due to page limitations, following lengthy mathematical manipulations we can also show that $\mathrm{E}[V^3] = 0$, and $\mathrm{E}[V^4] = \frac{3}{4} - \frac{3\sum_{r=1}^{R}P_r^2}{8S}$. According to (1), we next define X by standardizing V and proceed to derive the third and the fourth central moments of X. Based on the results for $\mathrm{E}[V^3]$ and $\mathrm{E}[V^4]$ we find that $\mu_3 = \mathrm{E}[X^3] = 0$ and $\mu_4 = \mathrm{E}[X^4] = 3 - \frac{3\sum_{r=1}^{R}P_r^2}{2S}$. Finally, by substituting μ_3 and μ_4 into (4) the leading error term of the Gaussian approximation can be upper bounded as

$$\Xi \le \frac{0.0344\sum_{r=1}^{R}P_r^2}{S}. \quad (12)$$

Noting that, for a given R and S, $\frac{0.0344\sum_{r=1}^{R}P_r^2}{S}$ is maximum when $P_1 = 1, P_2 = \dots = P_R = 0$ and minimum when $P_1 = P_2 = \dots = P_R = \frac{1}{R}$, we can further bound (12) as

$$\frac{0.0344}{RS} \le \frac{0.0344\sum_{r=1}^{R}P_r^2}{S} \le \frac{0.0344}{S}. \quad (13)$$

Based on this interesting result, we can draw more specific observations. For the NLOS urban micro environment considered in [8, Table 7.5-6] where $R = 19$ and $S = 20$ the approximate error of the correction term is

$$9 \times 10^{-5} \le \frac{0.0344\sum_{r=1}^{R}P_r^2}{S} \le 1.7 \times 10^{-3}, \quad (14)$$

which shows that the channel is remarkably Gaussian. Interestingly, we also note that the upper bound in (12) is a function of the cluster power spread. This can be clearly seen by rearranging the summation of P_r^2 in (12) as

$$\sum_{r=1}^{R}P_r^2 = (R-1)\Delta_p + 1/R, \quad (15)$$

where $\Delta_p = \frac{\sum_{r=1}^{R}(P_r - \frac{\sum_{q=1}^{R}P_q}{(R-1)})^2}{R}$, is the sample variance of the cluster powers. As such, the smaller the Δ_p the more accurate the Gaussian approximation.

Apart from the channel model in [8], there are other models adopted in the literature for mmWave communications such as the channel model in [9] where $h_k \sim \mathcal{CN}(0,1)$. If we apply our Gaussian accuracy analysis to this model we see that the correction term is zero as the model is exactly Gaussian.

B. Distributed Massive MIMO

As the second application, let us now consider the uplink of a massive MIMO network with n_{tx} transmitters and n_{rx} cooperating receivers where $n_{rx} \gg n_{tx}$. The transmitters and the receivers, all equipped with a single antenna, can be located anywhere in the network, i.e., co-located, fully distributed, in clusters, etc. We assume that the receivers are connected to a central base station, at which all transmitted bits are jointly detected based on a zero forcing (ZF) linear receiver. The $\mathcal{C}^{n_{rx}\times 1}$ received vector at the central base station is given by

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{w}, \quad (16)$$

where $\mathbf{s} = \{s_1, s_2, \ldots, s_{n_{tx}}\}^T$ is the $\mathcal{C}^{n_{tx}\times 1}$ data vector which contains transmitted symbols from the n_{tx} transmitters normalized such that $E\{|s_k|^2\} = 1, \forall k \in \{1, 2, \ldots n_{tx}\}$. The variable $\mathbf{w}$ represents the $\mathcal{C}^{n_{rx}\times 1}$ additive white Gaussian noise vector at the n_{rx} receive antennas which has independent entries with $E\{|w_n|^2\} = \sigma^2, \forall n \in \{1, 2, \ldots n_{rx}\}$. The channel matrix $\mathbf{H}$ is assumed to have full column rank with independent elements $h_{nk} \sim \mathcal{CN}(0, P_{nk})$, where $E\{|h_{nk}|^2\} = P_{nk}$ is the received power at antenna n from transmitter k. The geographical spread of transmitters and receivers creates a channel matrix $\mathbf{H}$, which has independent entries with different mean power values, P_{nk}.

For such a network, the instantaneous received SNR after linear processing of user k can be approximated by [10]

$$\tilde{\gamma}_k = \varrho \sum_{n=1}^{n_{rx}} |h_{nk}|^2, \tag{17}$$

where $\varrho = \left(\frac{n_{rx}-n_{tx}+1}{n_{rx}\sigma^2}\right)$. While $\tilde{\gamma}_k$ in (17) has a known PDF and CDF, the stable numerical computation of these distributions is quite challenging. Therefore, in the following we analyse the accuracy of a gamma approximation to the CDF of $\tilde{\gamma}_k$. We compare different power profiles and draw insights on certain power profiles that can be accurately approximated by a gamma distribution.

Using lengthy but straightforward calculations we can derive the first four central moments of $\tilde{\gamma}_k$ as $\mathrm{E}[\tilde{\gamma}_k] = \varrho \sum_{n=1}^{n_{rx}} P_{nk}$, $\mathrm{E}[(\tilde{\gamma}_k - \mathrm{E}[\tilde{\gamma}_k])^2] = \varrho^2 \sum_{n=1}^{n_{rx}} P_{nk}^2$, $\mathrm{E}[(\tilde{\gamma}_k - \mathrm{E}[\tilde{\gamma}_k])^3] = 2\varrho^3 \sum_{n=1}^{n_{rx}} P_{nk}^3$, and

$$\mathrm{E}[(\tilde{\gamma}_k - \mathrm{E}[\tilde{\gamma}_k])^4] = \varrho^4 \left[6 \sum_{n=1}^{n_{rx}} P_{nk}^4 + 3 \left(\sum_{n=1}^{n_{rx}} P_{nk}^2 \right)^2 \right]. \tag{18}$$

To use the gamma approximation we define a standardized random variable $X = \beta \tilde{\gamma}_k$ where $\beta = \frac{\sum_{n=1}^{n_{rx}} P_{nk}}{\varrho \sum_{n=1}^{n_{rx}} P_{nk}^2}$, as defined in Section III, and proceed to derive the third and the fourth central moments of X as

$$\mu_3 = 2 \left[\frac{\sum_{n=1}^{n_{rx}} P_{nk}}{\sum_{n=1}^{n_{rx}} P_{nk}^2} \right]^3 \sum_{n=1}^{n_{rx}} P_{nk}^3, \tag{19}$$

$$\mu_4 = \left[\frac{\sum_{n=1}^{n_{rx}} P_{nk}}{\sum_{n=1}^{n_{rx}} P_{nk}^2} \right]^4 \left[6 \sum_{n=1}^{n_{rx}} P_{nk}^4 + 3 \left(\sum_{n=1}^{n_{rx}} P_{nk}^2 \right)^2 \right]. \tag{20}$$

From our analysis we can now show that if the CDF of X is approximated by a gamma distribution with the shape parameter $m = \frac{(\sum_{n=1}^{n_{rx}} P_{nk})^2}{\sum_{n=1}^{n_{rx}} P_{nk}^2}$, then the correction term is approximated by (8) which can be solved exactly for a given set of P_{nk} values. Furthermore, by substituting μ_3 and μ_4 into φ and ϑ and reexpressing them in-terms of the second, third and fourth central moments of $\{P_{1k}, P_{2k}, \ldots P_{n_{rx}k}\}$, we can show that the gamma approximation is very good when these sample moments are small compared to the sample mean.

For example, consider a massive MIMO network with $n_{rx} = 100$ receivers, distributed according to three scenarios. 1) all the antennas are co-located such that $P_{nk} = 10$ dB, $\forall n \in \{1, 2, \ldots, 100\}$; 2) antennas are located in two cooperating clusters such that $P_{nk} = 10$ dB, $\forall n \in \{1, 2, \ldots, 50\}$ and $P_{nk} = 12$ dB, $\forall n \in \{51, 52, \ldots, 100\}$; 3) user k is located closer to antenna one such that $P_{1k} = 10$ dB and, $P_{nk} = 0$ dB, $\forall n \in \{2, 3 \ldots, 100\}$. As such, in scenarios 1 and 2 the sample variance is small compared to the sample mean while in scenario 3 it is large.

For scenario 1, $\varphi = 0$ and $\vartheta = 0$, giving $\Xi = 0$, which is consistent with the fact that with co-located antennas, $\tilde{\gamma}_k$ has a gamma distribution. For scenario 2, $m = 96$, $\varphi = 1.103$ and $\vartheta = -21$. From (8) we upper bound the correction term as

$$\Xi \leq 8.84 \times 10^{-4}. \tag{21}$$

Similarly, for scenario 3 we get $m = 60$, $\varphi = 40$ and $\vartheta = 87$. Thus, from (8) we upper bound the correction term as

$$\Xi \leq 4.7 \times 10^{-2}. \tag{22}$$

Comparing the three, we note that the gamma approximation is more accurate for scenarios 1 and 2, when the power profile has a small variation when compared to the sample mean.

Whilst not shown here due to page limitations, for a given power profile we could also observe that the gamma approximation becomes more accurate as the number of antennas increases. In fact, for scenario 2, we observe that as we increase the total number of antennas from 60 to 140, the upper bound for Ξ is reduced from 1.8×10^{-3} to 4.91×10^{-4}.

V. Conclusion

A new theoretical method to analyse the accuracy of a Gaussian or a gamma approximation of a random variable is presented, based on an upper bound derived for the leading order correction term. Specifically, we use the Edgeworth series expansion and the Laguerre series expansion to analyse the Gaussian and the gamma approximations, respectively. Our method is simple to analyse and only requires the third and fourth central moments of the random variable. A desirable extension to this letter is to conduct a similar accuracy analysis for other common distributions such as the beta distribution.

References

[1] H. Inaltekin, "Gaussian approximation for the wireless multi-access interference distribution," *IEEE Trans. Signal Process.*, vol. 60, no. 11, pp. 6114–6120, Nov. 2012.

[2] D. Morales-Jimenez, R. H. Y. Louie, M. R. McKay, and Y. Chen, "Analysis and design of multiple-antenna cognitive radios with multiple primary user signals," *IEEE Trans. Signal Process.*, vol. 63, no. 18, pp. 4925–4939, Sep. 2015.

[3] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, 2nd ed. New York, NY, USA: Wiley, 1995.

[4] S. Blinnikov and R. Moessner, "Expansions for nearly Gaussian distributions," *Astron. Astrophys. Suppl.*, vol. 130, pp. 193–205, May 1998.

[5] S. Li, M. R. McKay, and Y. Chen, "Characterizing the mutual information distribution of MIMO systems: Beyond the Gaussian approximation," in *Proc. Asilomar Conf. Signals Syst. Comput.*, Pacific Grove, CA, USA, Nov. 2012, pp. 1758–1762.

[6] A. DasGupta, *Asymptotic Theory of Statistics and Probability*. New York, NY, USA: Springer, 2008.

[7] C. E. Smith, "A Laguerre series approximation for the probability density of the first passage time of the Ornstein–Uhlenbeck process," in *Proc. Int. Conf. Noise Phys. Syst. 1/f Fluctuations*, Tokyo, Japan, 1991.

[8] "Technical specification group radio access network; channel model for frequency spectrum above 6 GHz (rel. 14), 14.2.0," 3rd Gener. Partnership Project, Sophia Antipolis, France, Rep. TR 38.900, pp. 1164–1179, Dec. 2016.

[9] M. R. Akdeniz *et al.*, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, Jun. 2014.

[10] R. Senanayake, P. Smith, P. L. Yeoh, and J. Evans, "An SNR approximation for distributed massive MIMO with zero forcing," *IEEE Commun. Lett.*, vol. 19, no. 11, pp. 1885–1888, Nov. 2015.