On Optimality of Gamma Approximation for Lognormal Shadowing Models

Shuping Dang, *Member, IEEE*, Jia Ye, *Member, IEEE*, Mark A. Beach, *Senior Member, IEEE*, and Marwa Chafii, *Member, IEEE*

***Abstract*—In this letter, we study the information-theoretic optimality of the gamma approximation for lognormal shadowing models in order to provide a rigorous mathematical ground for this useful technique. Specifically, we adopt the Kullback-Leibler (KL) divergence as the metric quantifying the distance between the original lognormal and the approximate gamma distributions. The KL divergence resulted from the moment matching criterion, the minimum achievable KL divergence of the gamma approximation, and the statistical parameter mapping relations are all derived in closed form. By these closed-form analytical expressions, we are able to rigorously examine the utility and optimality of the gamma approximation with a benchmark. Comparing the closed-form expressions of the KL divergence by moment matching and the minimum achievable benchmark, we find that the moment matching criterion, as a heuristic method, cannot guarantee the information-theoretic optimality. We also present and discuss the relevant results to substantiate the information-theoretic optimality achieved by our proposed statistical parameter mapping relations and the corresponding analytical insights.**

***Index Terms*—Channel model substitution, lognormal shadowing, gamma approximation, optimality analysis, KL divergence.**

I. Introduction

Shadowing is a common phenomenon attenuating the strength of radio frequency (RF) signals propagating through wireless fading channels, which is caused by obstacles and unfavorable terrains between wireless transmitters and receivers [1]. Empirical studies have found that the severity and impacts of shadowing can be well characterized by the lognormal distribution model [2]. However, the lognormal distribution model of shadowing is less mathematically tractable and could result in processing difficulties when analyzing the performance of wireless systems, albeit with the best fit to the empirical fading data. To attain a good compromise between accuracy and mathematical tractability, Abdi and Kaveh in [3] proposed a gamma approximation technique relying on the moment matching criterion to substitute the original lognormal distribution model. Following this pioneering work, Kostić in [4] carried out an in-depth study to numerically examine the utility of the gamma distribution by moment matching and proposed a composite fading channel model that integrates path loss, shadowing, and multi-path fading.

The gamma approximation technique has then been well recognized and widely applied in many following milestones of performance analysis and optimization for the fourth and fifth generation (4G, 5G) wireless communication systems [5]. In the context of sixth generation (6G) communications [6], the gamma approximation technique still plays a crucial role in facilitating performance analysis and optimization of a series of state-of-the-art communication systems, including terahertz (THz) networks [7], reconfigurable intelligent surface (RIS)-aided wireless networks [8], and three-dimensional (3D) massive multiple-input multiple-output (MIMO) networks [9]. Although the gamma approximation technique provides an effective approach to considerably simplify the performance analysis of wireless communication systems subject to lognormal shadowing, it mainly relies on the moment matching criterion for statistical parameter mapping, which might not necessarily produce the best fit. Specifically, it is shown in [10] that the moment matching based gamma approximation fails to capture the shape of the lognormal distribution when the shadowing severity parameter is large. The information-theoretic optimality of the moment matching criterion, mapping the statistical parameters from the lognormal distribution to the gamma distribution, has not been properly studied and proven yet. The lack of information-theoretic optimality studies leaves the gamma approximation in uncertainty. It is thus critical to uncover the accuracy of the gamma approximation technique to check whether the replacement performs well in scenarios of interest.

To bridge this long-standing knowledge gap and provide a rigorous mathematical ground, we analyze the information-theoretic optimality of the gamma approximation for lognormal shadowing models in this letter. Following [11], we adopt the Kullback-Leibler (KL) divergence as the metric measuring the distance between the lognormal and the gamma distributions. The explicit statistical parameter mapping relations derived in this letter lead to the closed-form expression of the minimum achievable KL divergence of gamma approximation. We further prove its information-theoretic optimality through the second partial derivative test and corroborate it with numerical simulations.

II. Fundamentals

By integrating the Friis transmission formula, it was reported in [2] that the long-term variability of the average signal power at the receiver caused by shadowing can be modeled by the lognormal probability density function (PDF):

$$f_\Omega(\Omega) = \frac{1}{\sqrt{2\pi}\sigma\Omega}\exp\left(-\frac{1}{2\sigma^2}\left(\log\left(\frac{L^\alpha\Omega}{P_t}\right)\right)^2\right),\ \forall\ \Omega > 0, \quad (1)$$

where σ is the shadowing severity parameter; L is the distance between the transmitter and the receiver; $\alpha > 0$ is the path loss exponent; P_t is the transmit power. Due to the poor mathematical

This work was supported by the UKRI/EPSRC SWAN Prosperity Partnership (EP/T005572/1).

Shuping Dang and Mark A. Beach are with Department of Electrical and Electronic Engineering, University of Bristol, Bristol BS8 1UB, U.K. (e-mail: shuping.dang@bristol.ac.uk, m.a.beach@bristol.ac.uk).

Jia Ye is with Computer, Electrical and Mathematical Sciences and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia (e-mail: jia.ye@kaust.edu.sa).

Marwa Chafii is with the Engineering Division, New York University (NYU) Abu Dhabi, 129188, UAE, and also with NYU WIRELESS, NYU Tandon School of Engineering, Brooklyn, 11201, NY, USA (e-email: marwa.chafii@nyu.edu).

tractability of (1) in wireless performance analysis, it was proposed in [4] to utilize the gamma distribution with the following PDF:

$$\hat{f}_\Omega(\Omega) = \left(\frac{\theta}{\nu}\right)^\theta \frac{\Omega^{\theta-1}}{\Gamma(\theta)} \exp\left(-\frac{\theta\Omega}{\nu}\right),\ \forall\ \Omega > 0 \tag{2}$$

to approximate the original lognormal distribution, where $\Gamma(z) = \int_0^\infty t^{z-1}\exp(-t)\mathrm{d}t$ denotes the gamma function solving the Euler integral of the second kind; θ and ν are free statistical parameters subject to adaptation to fit (1).

The accuracy of the gamma approximation depends on the values of the two free statistical parameters, i.e., θ and ν, which can be quantified from the information-theoretic perspective by the KL divergence given as

$$\mathsf{KL}(\theta,\nu) = \int_0^\infty f_\Omega(\Omega)\log\left(\frac{f_\Omega(\Omega)}{\hat{f}_\Omega(\Omega)}\right)\mathrm{d}\Omega = I_1 - I_2(\theta,\nu) > 0, \tag{3}$$

where $I_1 = \int_0^\infty f_\Omega(\Omega)\log\left(f_\Omega(\Omega)\right)\mathrm{d}\Omega$ and $I_2(\theta,\nu) = \int_0^\infty f_\Omega(\Omega)\log\left(\hat{f}_\Omega(\Omega)\right)\mathrm{d}\Omega$.

Remark 1: From (3), it becomes obvious that the first term I_1 is fixed and only related to the original distribution, whereas the second term $I_2(\theta,\nu)$ involves the approximate distribution and depends on free statistical parameters θ and ν. A lower KL divergence $\mathsf{KL}(\theta,\nu)$ indicates a more accurate distribution substitute, equivalently resulted from a greater $I_2(\theta,\nu)$.

III. KL Divergence by Moment Matching

Let us first focus on the derivation of I_1, which can be expanded and computed as

$$\begin{aligned} I_1 &= \int_0^\infty \frac{\log\left(\frac{1}{\sqrt{2\pi}\sigma\Omega}\right)}{\sqrt{2\pi}\sigma\Omega}\exp\left(-\frac{1}{2\sigma^2}\left(\log\left(\frac{L^\alpha\Omega}{P_t}\right)\right)^2\right)\mathrm{d}\Omega \\ &\quad - \int_0^\infty \frac{\left(\log\left(\frac{L^\alpha\Omega}{P_t}\right)\right)^2}{2\sqrt{2\pi}\sigma^3\Omega}\exp\left(-\frac{1}{2\sigma^2}\left(\log\left(\frac{L^\alpha\Omega}{P_t}\right)\right)^2\right)\mathrm{d}\Omega \\ &= \log\left(\frac{L^\alpha}{\sqrt{2\pi}\sigma P_t}\right) - \frac{1}{2}. \end{aligned} \tag{4}$$

Likewise, we can also expand and compute $I_2(\theta,\nu)$ as

$$\begin{aligned} I_2(\theta,\nu) &= \int_0^\infty \frac{\theta\log\left(\frac{\theta}{\nu}\right)}{\sqrt{2\pi}\sigma\Omega}\exp\left(-\frac{1}{2\sigma^2}\left(\log\left(\frac{L^\alpha\Omega}{P_t}\right)\right)^2\right)\mathrm{d}\Omega \\ &\quad + \int_0^\infty \frac{\log\left(\frac{\Omega^{\theta-1}}{\Gamma(\theta)}\right)}{\sqrt{2\pi}\sigma\Omega}\exp\left(-\frac{1}{2\sigma^2}\left(\log\left(\frac{L^\alpha\Omega}{P_t}\right)\right)^2\right)\mathrm{d}\Omega \\ &\quad - \int_0^\infty \frac{\theta}{\sqrt{2\pi}\sigma\nu}\exp\left(-\frac{1}{2\sigma^2}\left(\log\left(\frac{L^\alpha\Omega}{P_t}\right)\right)^2\right)\mathrm{d}\Omega \\ &= \theta\log\left(\frac{\theta P_t}{\nu L^\alpha}\right) - \log\left(\frac{P_t\Gamma(\theta)}{L^\alpha}\right) - \frac{\theta P_t}{\nu L^\alpha}\exp\left(\frac{\sigma^2}{2}\right). \end{aligned} \tag{5}$$

Substituting (4) and (5) back into (3) yields the general form of the KL divergence of the gamma approximation infra:

$$\mathsf{KL}(\theta,\nu) = \log\left(\frac{\Gamma(\theta)}{\sqrt{2\pi}\sigma}\right) - \frac{1}{2} + \theta\log\left(\frac{\nu L^\alpha}{\theta P_t}\right) + \frac{\theta P_t}{\nu L^\alpha}\exp\left(\frac{\sigma^2}{2}\right). \tag{6}$$

Heuristically, by matching the first and second moments of the lognormal and gamma distributions, i.e., $\underset{\Omega\sim f_\Omega}{\mathbb{E}}\{\Omega\} = (P_t/L^\alpha)\exp(\sigma^2/2) = \underset{\Omega\sim\hat{f}_\Omega}{\mathbb{E}}\{\Omega\} = \nu$ and $\underset{\Omega\sim f_\Omega}{\mathbb{E}}\{\Omega^2\} = (P_t/L^\alpha)^2\exp(2\sigma^2) = \underset{\Omega\sim\hat{f}_\Omega}{\mathbb{E}}\{\Omega^2\} = (1+\theta)\nu^2/\theta$, free statistical parameters θ and ν can be determined by the moment matching criterion as [4]

$$\theta_{\mathrm{MM}} = (\exp(\sigma^2)-1)^{-1} \ \text{ and }\ \nu_{\mathrm{MM}} = \frac{P_t}{L^\alpha}\exp\left(\frac{\sigma^2}{2}\right). \tag{7}$$

By substituting (7) into (6), we can obtain the closed-form expression of the KL divergence below by applying the moment matching criterion:

$$\begin{aligned} \mathsf{KL}(\theta_{\mathrm{MM}},\nu_{\mathrm{MM}}) &= \log\left(\frac{1}{\sqrt{2\pi}\sigma}\Gamma\left(\frac{1}{\exp(\sigma^2)-1}\right)\right) - \frac{1}{2} \\ &\quad + \frac{1+\sigma^2/2+\log\left(\exp(\sigma^2)-1\right)}{\exp(\sigma^2)-1}. \end{aligned} \tag{8}$$

IV. Minimum Achievable KL Divergence

To obtain the optimal set of free parameters leading to the minimum KL divergence, we first formulate the KL divergence minimization problem as follows:

$$\begin{aligned} \{\theta_{\mathrm{KLD}},\nu_{\mathrm{KLD}}\} &= \underset{\{\theta,\nu\}}{\arg\min}\{\mathsf{KL}(\theta,\nu)\} \overset{(a)}{=} \underset{\{\theta,\nu\}}{\arg\min}\{-I_2(\theta,\nu)\} \\ &= \underset{\{\theta,\nu\}}{\arg\max}\{I_2(\theta,\nu)\} \end{aligned} \tag{9}$$

where transfer (a) is equivalent because the first term I_1 is irrespective of θ and ν (c.f. (4)).

Remark 2: Although θ and ν are generally within certain ranges in practice because of the physical ranges of σ, α, L, and P_t, we intend not to put constraints on θ and ν, but, instead, maintain both to be positive free parameters in order to gain the processing flexibility[1].

Observing the explicit form of $I_2(\theta,\nu)$ given in (4), it is straightforward that $I_2(\theta,\nu)$ is defined and continuous when $\theta > 0$ and $\nu > 0$. For the unconstrained two-parameter maximization problem formulated in (9), within the definition domain of objective function $I_2(\theta,\nu)$, we can resort to partial derivatives to discuss its stationary points and boundary conditions to find out the global maximum. We hereby derive its first partial derivatives with respect to θ and ν, respectively:

$$\frac{\partial I_2}{\partial\theta} = 1 - \frac{P_t}{\nu L^\alpha}\exp\left(\frac{\sigma^2}{2}\right) + \log\left(\frac{\theta P_t}{\nu L^\alpha}\right) - \psi_0(\theta), \tag{10}$$

$$\frac{\partial I_2}{\partial\nu} = \frac{\theta P_t}{\nu^2 L^\alpha}\exp\left(\frac{\sigma^2}{2}\right) - \frac{\theta}{\nu}, \tag{11}$$

where $\psi_n(z) = \frac{\mathrm{d}^{n+1}}{\mathrm{d}z^{n+1}}\log(\Gamma(z))$ is the nth-order polygamma function [12]. Based on (10) and (11), we can further obtain four second partial derivatives below:

$$\frac{\partial^2 I_2}{\partial\theta^2} = \frac{1}{\theta} - \psi_1(\theta), \tag{12}$$

$$\frac{\partial^2 I_2}{\partial\theta\nu} = \frac{\partial^2 I_2}{\partial\nu\theta} = \frac{P_t}{\nu^2 L^\alpha}\exp\left(\frac{\sigma^2}{2}\right) - \frac{1}{\nu}, \tag{13}$$

$$\frac{\partial^2 I_2}{\partial\nu^2} = -\frac{2\theta P_t}{\nu^3 L^\alpha}\exp\left(\frac{\sigma^2}{2}\right) + \frac{\theta}{\nu^2}, \tag{14}$$

[1]It is worth noting that the positive condition associated with θ and ν does not need to be specified in the formulated KL divergence minimization problem because as a real-valued optimization problem this condition has already been implicitly ensured by the objective function comprised by logarithmic functions.

which are defined over entire positive open intervals $\theta \in (0, \infty)$ and $\nu \in (0, \infty)$. Solving the equation set consisting of $\frac{\partial I_2}{\partial \theta} = 0$ and $\frac{\partial I_2}{\partial \nu} = 0$ gives the single stationary point:

$$(\theta_{\rm KLD}, \nu_{\rm KLD}) = \left(\dot{\theta}(\sigma), \frac{P_t}{L^\alpha} \exp\left(\frac{\sigma^2}{2}\right)\right), \tag{15}$$

where $\dot{\theta}(\sigma)$ is a solution to transcendental equation $\theta = \exp\left(\frac{\sigma^2}{2} + \psi_0(\theta)\right)$, which is only related to σ. To guarantee the existence and uniqueness of the stationary point, we propose and prove the following lemma apropos of $\dot{\theta}(\sigma)$:

Lemma 1: $\dot{\theta}(\sigma)$ must exist when $\sigma > 0$ and is the unique solution to transcendental equation $\theta = \exp\left(\frac{\sigma^2}{2} + \psi_0(\theta)\right)$.

Proof: Let $\Phi(\theta) = \theta - \exp\left(\frac{\sigma^2}{2} + \psi_0(\theta)\right)$. Because of the basic relation $\psi_n'(\theta) = \psi_{n+1}(\theta)$ [12], the first and second derivatives of $\Phi(\theta)$ with respect to θ can be determined as

$$\Phi'(\theta) = 1 - \exp\left(\frac{\sigma^2}{2} + \psi_0(\theta)\right)\psi_1(\theta), \tag{16}$$

$$\Phi''(\theta) = -\exp\left(\frac{\sigma^2}{2} + \psi_0(\theta)\right)\left[(\psi_1(\theta))^2 + \psi_2(\theta)\right] < 0. \tag{17}$$

Because of the negativity of the second derivative, $\Phi(\theta)$ is a concave function of θ, reaching its maximum when $\Phi'(\theta) = 0$, i.e., $\exp\left(\frac{\sigma^2}{2} + \psi_0(\theta)\right)\psi_1(\theta) = 1$. We denote $\ddot{\theta}(\sigma) > 0$ as the solution satisfying equality $\exp\left(\frac{\sigma^2}{2} + \psi_0(\theta)\right)\psi_1(\theta) = 1$, which is thus the global maximum of $\Phi(\theta)$. Through the analytic continuation technique, we extend the domain of definition of $\Phi(\theta)$ to $\theta = 0$ and let $\Phi(0) = \lim_{\theta \to 0^+}\{\Phi(\theta)\} = 0$. Due to the concavity of $\Phi(\theta)$, $\Phi(\theta)$ is a monotone increasing function over $(0, \ddot{\theta}(\sigma)]$ and a monotone decreasing function over $[\ddot{\theta}(\sigma), \infty)$. As a result, we have $\Phi(\ddot{\theta}(\sigma)) > 0$. Meanwhile, we can compute $\lim_{\theta \to \infty}\{\Phi(\theta)\} = -\infty < 0$. By the general existence theorem of zero points [13], there must exist and only exist one unique solution $\dot{\theta}(\sigma) \in [\ddot{\theta}(\sigma), \infty)$ to $\Phi(\theta) = 0$ over the entire domain of definition. ■

Based on the proven existence and uniqueness of $\dot{\theta}(\sigma)$, we can apply numerical methods, e.g., the Newton-Raphson method, with an arbitrary positive initial searching point to approach solution $\dot{\theta}(\sigma)$ to transcendental equation $\Phi(\theta) = \theta - \exp\left(\frac{\sigma^2}{2} + \psi_0(\theta)\right) = 0$. Admitting a bit loss of accuracy, we can also resort to the asymptotic expression of digamma function $\psi_0(\theta) \sim \log(\theta) - 1/(2\theta)$ for large θ to obtain an approximate solution $\dot{\theta}(\sigma) \approx 1/\sigma^2$ [12]. We utilize `vpasolve`, an embedded numerical solver on MATLAB, in conjunction with a randomized positive initialization strategy to plot Fig. 1 for different σ. This plot can be referred to identify the values of $\dot{\theta}(\sigma)$ and $\ddot{\theta}(\sigma)$ once σ is given. The data presented in this plot is also in line with the above analysis pertaining to the range, existence, and uniqueness of $\dot{\theta}(\sigma)$.

With (12), (13), and (14) in conjunction with *Lemma 1*, we are now capable of further examining the nature of stationary point $(\theta_{\rm KLD}, \nu_{\rm KLD}) = \left(\dot{\theta}(\sigma), \frac{P_t}{L^\alpha}\exp\left(\frac{\sigma^2}{2}\right)\right)$ by conducting the second partial derivative test:

$$\left.\frac{\partial^2 I_2}{\partial \theta^2}\right|_{\theta = \dot{\theta}(\sigma)} = \frac{1}{\dot{\theta}(\sigma)} - \psi_1(\dot{\theta}(\sigma)) < 0, \ \forall\ \sigma > 0, \tag{18}$$

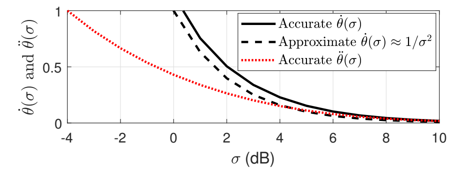

Fig. 1: Plot of $\dot{\theta}(\sigma)$ and $\ddot{\theta}(\sigma)$ versus σ in dB.

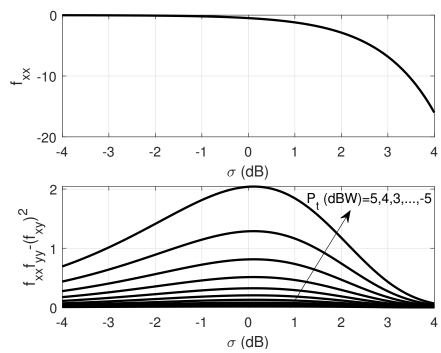

Fig. 2: Plot of (18) and (19) versus σ in dB with normalized transmitter-receiver distance.

$$\begin{aligned}&\left[\frac{\partial^2 I_2}{\partial \theta^2}\frac{\partial^2 I_2}{\partial \nu^2} - \left(\frac{\partial^2 I_2}{\partial \theta \partial \nu}\right)^2\right]_{\theta = \dot{\theta}(\sigma),\ \nu = \frac{P_t}{L^\alpha}\exp\left(\frac{\sigma^2}{2}\right)} \\ &= \frac{L^{2\alpha}}{P_t^2}\exp(-\sigma^2)\left(\dot{\theta}(\sigma)\psi_1(\dot{\theta}(\sigma)) - 1\right) > 0,\ \forall\ \sigma > 0.\end{aligned} \tag{19}$$

Proof: According to the key results proven in [14], the following inequalities hold, $\forall\ \theta > 0$, $n \in \mathbb{N}$:

$$\frac{(n-1)!}{\theta^n} + \frac{n!}{2\theta^{n+1}} < (-1)^{n+1}\psi_n(\theta) < \frac{(n-1)!}{\theta^n} + \frac{n!}{\theta^{n+1}}, \tag{20}$$

Letting $n = 1$ and manipulating (20) result in

$$\frac{1}{\theta} - \psi_1(\theta) < \frac{1}{\theta} - \left(\frac{1}{\theta} + \frac{1}{2\theta^2}\right) = -\frac{1}{2\theta^2} < 0,\ \forall\ \theta > 0, \tag{21}$$

which completes the proof of (18). Likewise, with the help of (20), we can determine the following relation:

$$\theta\psi_1(\theta) - 1 > \theta\left(\frac{1}{\theta} + \frac{1}{2\theta}\right) - 1 = \frac{1}{2} > 0,\ \forall\ \theta > 0. \tag{22}$$

In addition, since $\frac{L^{2\alpha}}{P_t^2}\exp(-\sigma^2) > 0$ always holds, (19) is thereby proven by multiplying both always-positive terms. ■

Both (18) and (19) are plotted in Fig. 2 with normalized transmitter-receiver distance for verification and illustration purposes. Through the second partial derivative test presented above, the stationary point given in (15) is a local maximum of $I_2(\theta, \nu)$. Moreover, because this is the only local maximum over open intervals $\theta \in (0, \infty)$ and $\nu \in (0, \infty)$, it must also be the global maximum of $I_2(\theta, \nu)$ since there do not exist boundary points and points where partial derivatives do not exist. Finally, substituting (15) into (6) gives the minimum achievable KL divergence by applying the gamma approximation technique:

$$\begin{aligned}\mathsf{KL}(\theta_{\rm KLD}, \nu_{\rm KLD}) = &\log\left(\frac{\Gamma(\dot{\theta}(\sigma))}{\sqrt{2\pi}\sigma}\right) - \frac{1}{2} \\ &+ \dot{\theta}(\sigma)\left(1 + \log\left(\frac{1}{\dot{\theta}(\sigma)}\exp\left(\frac{\sigma^2}{2}\right)\right)\right).\end{aligned} \tag{23}$$

Remark 3: Jointly observing (8) and (23), we can easily find that both the KL divergence yielded by the moment matching criterion and the minimum achievable KL divergence are setup-independent and will not be affected by P_t, α, and L. This slightly counter-intuitive outcome signifies the outstanding generality and scalability of the gamma approximation technique for lognormal shadowing models.

V. Results and Discussion

To examine the previous analyses, PDFs and cumulative distribution functions (CDFs) vis-à-vis the original lognormal distribution and the approximate gamma distributions with different statistical mapping criteria are presented in Fig. 3 with normalized shadowing parameter σ. In addition, as maximum likelihood estimation (MLE) is a widely used numerical method for solving parameter estimation problems [15], it is also adopted as a benchmark. Given S independent random realization samples (a.k.a. the empirical data) of the specific lognormal shadowing model with pre-defined P_t/L^α, and σ, denoted as $\{\Omega_s\}_{s=1}^S$, free statistical parameters θ_{MLE} and ν_{MLE} can be obtained according to the following MLE criterion: $\{\theta_{\text{MLE}}, \nu_{\text{MLE}}\} = \underset{\theta_{\text{MLE}}, \nu_{\text{MLE}}}{\arg\max} \left\{\sum_{s=1}^S \log \hat{f}_\Omega(\Omega_s)\right\}$, s.t., $\theta_{\text{MLE}} \geq 0$, $\nu_{\text{MLE}} \geq 0$. This optimization problem can be easily solved by some built-in functions in software, e.g., `fmincon` in MATLAB.

Under the parameter settings in Fig. 3, the values of parameter θ, solely depending on the shadowing severity parameter, are 1.138 and 0.582 respectively under the KL divergence minimization and moment matching criterion, but fluctuate around 1.1 when applying the MLE method. Parameter ν is 0.149 with $P_t/L^\alpha = -10$ dB for both KL divergence minimization and moment matching criterion as they share the same analytical expression. Moreover, a 10 dB gain on the P_t/L^α results in a 10-fold increment on ν, which is around 1.49 and 14.9 when P_t/L^α is set to 0 and 10 dB. Similarly, the value of ν obtained through the MLE method with different P_t/L^α fluctuates as well. It can be observed from Fig. 3 that the gamma distributions yielded by minimizing KL divergence and MLE have a better fit to the original lognormal distribution than the heuristic counterpart returned by the moment matching criterion. Moreover, the MLE mapping gives almost the same distributions as the ones adopting the KL divergence minimization mapping. However, our measured results indicate that the KL divergence minimization mapping criteria only requires 0.0473 seconds to produce the values of two free statistical parameters, while the MLE mapping method requires 0.2622 seconds using $S = 10,000$ empirical samples.

To be rigorous, we also conduct a quantitative comparison between these two statistical parameter mapping criteria and present their KL divergences versus shadowing severity parameter σ in Fig. 4. As the data shown in this figure, both criteria converge when $\sigma < -6$ dB (mildly shadowed scenarios), whereas there exists a considerable discrepancy between both criteria when σ goes large. It should be noted that the plotted results in Figs. 3-4 about the moment matching method are generally in agreement with the statement in the published works [3], [4], [10], but give us more details. In particular, the KL divergence yielded by the moment matching criterion can be about one hundred times higher than the minimum achievable KL divergence when $\sigma = 10$ dB (severely shadowed scenarios). The much lower minimum achievable KL divergence in the high σ region verifies the accuracy advantage of the KL divergence minimization criterion over the moment matching criterion, resulting in a more accurate gamma distribution substitute of the original lognormal distribution. For practical applications knowing the shadowing severity parameter, it is always better to evaluate the approximation accuracy in advance using the obtained closed form.

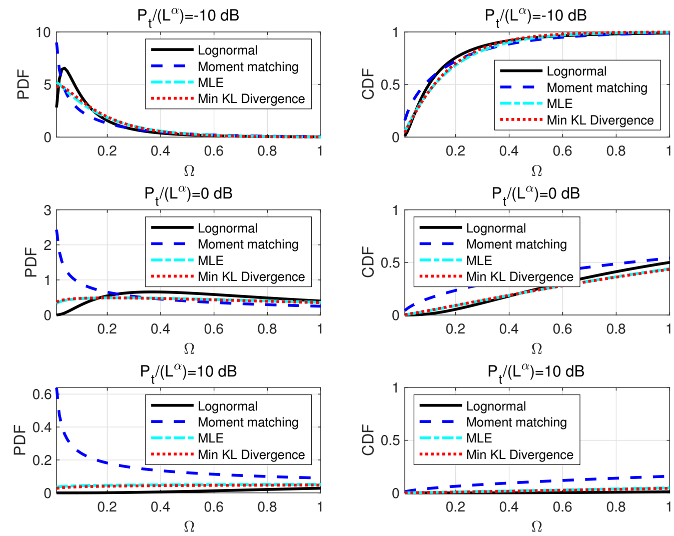

Fig. 3: PDF and CDF yielded by different methods corresponding to different system setups, given $\sigma = 1$ (normalized).

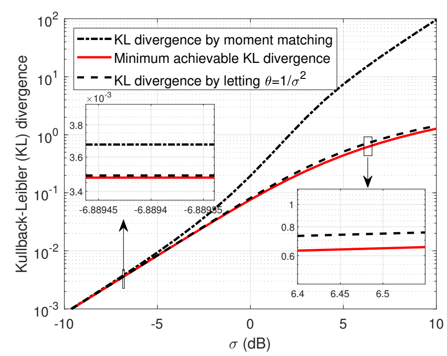

Fig. 4: KL divergence yielded by both statistical parameter mapping criteria versus σ in dB.

VI. Conclusion

In this letter, we carried out the information-theoretic analysis of the gamma approximation technique for substituting lognormal shadowing models. We derived the closed-form expressions of KL divergence and the statistical parameter mapping relations by both the moment matching criterion and the proposed KL divergence minimization criterion, which help researchers to examine the gamma approximation performance with a given shadowing severity parameter before implementation. Model substitution results illustrated that the gamma approximation technique applying the KL divergence minimization criterion can better capture the statistical features of the original lognormal distribution, whilst harvesting an even higher accuracy gain when shadowing is severe.

References

[1] T. Rappaport, *Wireless Communications: Principles and Practice*, ser. Prentice Hall communications engineering and emerging technologies series. Prentice Hall PTR, 2002.

[2] A. Coulson, A. Williamson, and R. Vaughan, "A statistical basis for lognormal shadowing effects in multipath fading channels," *IEEE Transactions on Communications*, vol. 46, no. 4, pp. 494–502, 1998.

[3] A. Abdi and M. Kaveh, "On the utility of gamma PDF in modeling shadow fading (slow fading)," in *Proc. IEEE Vehicular Technology Conference*, vol. 3, Houston, TX, USA, 1999, pp. 2308–2312.

[4] I. Kostić, "Analytical approach to performance analysis for channel subject to shadowing and fading," *IEE Proceedings-Communications*, vol. 152, no. 6, pp. 821–827, 2005.

[5] S. Al-Ahmadi, "The gamma-gamma signal fading model: A survey," *IEEE Antennas and Propagation Magazine*, vol. 56, no. 5, pp. 245–260, 2014.

[6] S. Dang, O. Amin, B. Shihada, and M.-S. Alouini, "What should 6G be?" *Nature Electronics*, vol. 3, no. 1, pp. 20–29, 2020.

[7] J. Ye, S. Dang, G. Ma, O. Amin, B. Shihada, and M.-S. Alouini, "On outage performance of terahertz wireless communication systems," *IEEE Transactions on Communications*, vol. 70, no. 1, pp. 649–663, 2022.

[8] T. N. Do, G. Kaddoum, T. L. Nguyen, D. B. da Costa, and Z. J. Haas, "Multi-RIS-aided wireless systems: Statistical characterization and performance analysis," *IEEE Transactions on Communications*, vol. 69, no. 12, pp. 8641–8658, 2021.

[9] C. F. López, C.-X. Wang, and Y. Zheng, "A 3D non-stationary wideband massive MIMO channel model based on ray-level evolution," *IEEE Transactions on Communications*, vol. 70, no. 1, pp. 621–634, 2022.

[10] R. Agrawal, "On efficacy of rayleigh-inverse gaussian distribution over K-distribution for wireless fading channels," *Wireless Communications and Mobile Computing*, vol. 7, no. 1, pp. 1–7, 2007.

[11] S. Atapattu, C. Tellambura, and H. Jiang, "A mixture gamma distribution to model the SNR of wireless channels," *IEEE Transactions on Wireless Communications*, vol. 10, no. 12, pp. 4193–4203, 2011.

[12] O. Espinosa and V. H. Moll, "A generalized polygamma function," *Integral Transforms and Special Functions*, vol. 15, no. 2, pp. 101–115, 2004.

[13] P. J.-J. Herings, G. A. Koshevoy, A. Talman, and Z. Yang, "General existence theorem of zero points," *Journal of optimization theory and applications*, vol. 120, no. 2, pp. 375–394, 2004.

[14] B.-N. Guo and F. Qi, "Refinements of lower bounds for polygamma functions," *Proceedings of the American Mathematical Society*, pp. 1007–1015, 2013.

[15] R. Millar, *Maximum Likelihood Estimation and Inference: With Examples in R, SAS and ADMB*, ser. Statistics in Practice. Wiley, 2011.