Simple and Accurate Low SNR Ergodic Capacity Approximations

Bitan Banerjee, *Student Member, IEEE*, Ahmad Abu Al Haija, *Member, IEEE*, Chintha Tellambura, *Fellow, IEEE*, and Himal A. Suraweera, *Senior Member, IEEE*

***Abstract*— Some of the existing ergodic capacity approximations for the low signal-to-noise ratio (SNR) region may lack accuracy. To overcome this, we derive two simple yet accurate Padé approximations for the low-SNR ergodic capacity. These approximations utilize the channel moments, which need not be updated for each distinct SNR value. The moments can be derived from the probability density function or from the moment generating function. For instance, we derive the general expressions for the moments of multiple input single output and multiple input multiple output channels. Numerical results demonstrate the superior accuracy of our approximations over some of the existing approximations. Moreover, we demonstrate how our approximations can be efficiently used to analyze several types of wireless links.**

***Index Terms*— Ergodic capacity, low-SNR, padé approximation, MIMO channels, wireless relaying.**

I. Introduction

The proliferation of low-power wireless devices such as Internet-of-Things (IoT) results in wireless links operating at the low signal-to-noise ratio (SNR) region. Hence, this region is becoming important in terms of characterizing the performance of modern wireless networks. Among all performance measures, perhaps the most important one is the ergodic capacity, which refers to the rate at which information can be reliably transmitted over a wireless channel [1]. The ergodic capacity is defined as

$$\mathbb{E}[C(\rho)] = \mathbb{E}[\log_2(1 + \rho X)], \tag{1}$$

where ρ is the average SNR, $\mathbb{E}[\cdot]$ is the expectation, and X is a non-negative random variable characterizing the randomness of the link [2], [3].

Efficient capacity approximations can be utilized for optimal power allocation, network design and other purposes [4]. However, to the best of our knowledge, low-SNR capacity approximations that are simple, tight and that do not require the exact probability density function (PDF) of the channel are not readily available.

To characterize the low-SNR capacity, Verdú in his seminal letter introduced the concept of wideband slope [1].

Manuscript received September 28, 2017; revised November 12, 2017; accepted November 17, 2017. Date of publication November 21, 2017; date of current version February 9, 2018. The associate editor coordinating the review of this paper and approving it for publication was M. Rodrigues. *(Corresponding author: Bitan Banerjee.)*

B. Banerjee and C. Tellambura are with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 1H9, Canada (e-mail: bitan@ualberta.ca; ct4@ualberta.ca).

A. Abu Al Haija is with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada (e-mail: abualhai@ualberta.ca).

H. A. Suraweera is with the Department of Electrical and Electronic Engineering, University of Peradeniya, Peradeniya 20400, Sri Lanka (e-mail: himal@ee.pdn.ac.lk).

Digital Object Identifier 10.1109/LCOMM.2017.2776225

TABLE I
Capacity (bps/Hz) for 5 Antenna MRC

Exact	[5]	[6], [7]	[8]	[9] (7 terms)
0.76	0.70	0.70	0.93	0.86

Thereafter, further approximations were developed in [5]–[11]. The most common ones are based on the Taylor series expansion (TSE) of $f(x) = \log(1 + x)$ around $x = 0$. The accuracy can be improved when $f(x)$ is expanded around $x = \mathbb{E}[X]$ [6], [7]. Although such approximations [5]–[7] are suitable for single antenna systems, and do not require the exact PDF of the SNR, they are inaccurate for use in multiple input multiple output (MIMO) systems. To overcome this, [8] developed two curve-fitting approaches for MIMO channels by letting $f(x) \approx a\sqrt{x}$ and $f(x) \approx ax^b$, where a and b are determined numerically using ρ and the exact PDF of X. However, these methods are less accurate than the TSE approximations [5]–[7] but may be more suitable for power allocation problems. Reference [9] derived a truncated infinite series for the ergodic capacity. However, truncation analysis is cumbersome and oscillating nature of this series affects the accuracy. To study the MIMO multiple access channel capacity, [10] used the wideband slope method in [1].

Table I compares the relative accuracy levels of several existing capacity approximations [5]–[9]. Results are generated for maximal ratio combining (MRC) in Rayleigh fading at $\rho = -5$ dB. Note that the absolute error can be as high as 22% (for [8]), which suggests that [8] might not be effective as a generic approximation. Moreover, Table I shows that the approximation [9] is less accurate than the TSE approximations with a fewer number of terms.

Hence, in this letter we aim to find new approximations for ergodic capacity in the low-SNR region that 1) are tight, 2) require statistical moments, which do not require updating for each SNR value, and 3) facilitate simple analysis of wireless systems.

The main contributions are as follows:

1) We derive two simple yet accurate low-SNR ergodic capacity approximations, utilizing first two or four moments of X. The moments are used in the form of a PA. The PA $[p/q]_h$ of function $h(x)$ is a rational function, with a numerator of degree at most p and a denominator of degree at most q, and whose power series expansion matches the Macalurin series of $h(x)$ up to the degree $p + q$. It turns out that PA of $f(x)$ is accurate even for $|x| > 1$, while the TSE of $f(x)$ is divergent for $|x| > 1$ [12]. Previously, PAs were used to develop the moment generating function (MGF) in error performance analysis [3], [13]. To the best of our

knowledge, this is the first letter to use PA technique for capacity approximation.

2) We derive the general expressions for the moments of x for multiple input single output (MISO), single input multiple output (SIMO) and MIMO output channels. Since the moments of X are invariant of the SNR, they are computed only once. Moreover, they can be derived from the PDF or MGF of X, which is typically available for many diversity receiver structures and propagation channels [2], [3].
3) Finally, we provide numerical results to compare the relative accuracy of several approximations and establish the high accuracy of the two proposed approximations.

II. New Ergodic Capacity Approximations

We now provide Theorem 1 to express the ergodic capacity approximations:

Theorem 1: At low SNR, the ergodic capacity given in (1) can be approximated as:

$$\text{App. 1}: \ \mathbb{E}[C(\rho)] \equiv \frac{\rho\mathbb{E}[X]/\ln 2}{1+\frac{1}{2}\mu_2\rho} + R_1(\rho), \tag{2}$$

$$\text{App. 2}: \ \mathbb{E}[C(\rho)] \equiv \frac{\rho\mathbb{E}[X]}{\ln 2}\left[\frac{1+\alpha_1\rho}{1+\alpha_2\rho+\alpha_3\rho^2}\right] + R_2(\rho), \tag{3}$$

where $R_j(\rho)$, $j \in \{1,2\}$*, are error terms that vanish as* $\rho \to 0$, $\alpha_1 = \frac{(18\,\mu_4 - 24\,\mu_2\mu_3 + 9\,\mu_2{}^3)}{24\,\mu_3 - 18\,\mu_2{}^2}$, $\alpha_2 = \frac{(18\,\mu_4 - 12\,\mu_2\mu_3)}{24\,\mu_3 - 18\,\mu_2{}^2}$, $\alpha_3 = \frac{(9\,\mu_2\mu_4 - 8\,\mu_3^2)}{24\,\mu_3 - 18\,\mu_2{}^2}$ *and* $\mu_k = \mathbb{E}[X^k]/\mathbb{E}[X]$.

Proof: App. 1 (App. 2) is obtained using the PA of [1/1] and [2/2] order respectively (see Appendix A). The error terms, $R_j(\rho)$, $j = 1, 2$, can be bounded. The details of error bound related analysis are given in Appendix B. ☐

Remarks:

1) It is known that the PA of $f(x)$ is accurate far beyond the radius of convergence of the Taylor series [12]. The radius of convergence of a TSE of any function $g(x)$ around point x_0 depends on the distance of closest singularity of $g(x)$ from x_0 [14]. Thus, the TSE of $f(x) = \log(1+x) = \sum_{k=1} \frac{(-1)^{k-1}}{k} x^k$ around $x = 0$ has a maximum radius of convergence of one. Thus, the TSE is accurate only for $|x| < 1$. In contrast, a PA for $f(x)$ can be accurate for $|x| < 10$ [12]. This example shows that the conversion from the TSE to the Padé form usually allows good accuracy even outside the radius of convergence of the TSE. Moreover, the poles of PA in App. 1 and 2 lie on the negative real axis [15, Eq. (8.3)]. These reasons suggest that App. 1 and App. 2 may outperform the TSE approximations for the capacity.
2) The TSE of App. 1 (cf. (2)) around zero gives identical expression till two (four) terms of the TSE of $\mathbb{E}[C(\rho)]$ around zero, and some additional terms. These additional terms help PA to provide a tighter error bound than TSE approximations [5]–[7].
3) Similar approximations can be developed using more than four moments. App. 2 is more accurate than App. 1 since it utilizes more moments. However, the latter is simpler and useful for cases where derivation of higher order channel moments are tedious, such as [16].

III. Applications of Capacity Approximations

A. Applications in MIMO Channels

To apply Theorem 1 to multi-antenna channels, we need $\mathbb{E}[X^k]$. We thus derive these for several important cases next. The list is by no means exhaustive. In all cases, the number of transmit and receive antennas are N_t and N_r respectively.

1) MRC With SIMO and MISO Channels: In a cellular network with a multiple antenna base station and single antenna user terminals, the MISO (SIMO) channel resembles the reception (transmit) diversity in downlink (uplink) transmission. We consider MRC reception with N diversity branches of correlated and non-identical Nakagami-m fading channels. The capacity of MISO (SIMO) channel is given by (1) with $X = \sum_{i=1}^{N} X_i(\alpha, \beta_n)$ where X_i's are Gamma distributed with shape and rate parameter α and β_n. MGF of X is given in [17]. Thereafter, applying Leibniz rule over the MGF, k-th moment of X follows,

$$\mathbb{E}[X^k] = \sum_{i_1+\cdots+i_N=k} \binom{k}{i_1,\ldots,i_N} \prod_{1\le j\le N} (\alpha\lambda_j)^{i_j}, \quad k = 0, 1, \ldots \tag{4}$$

where λ_js are the eigenvalues of the matrix $\mathbf{DC}$, where $\mathbf{D}$ is the $N \times N$ diagonal matrix with entries β_n and $\mathbf{C}$ is the $N \times N$ correlation matrix. This result (4) covers many systems. The details are omitted for brevity.

2) MIMO With $N_t \gg N_r$: We consider a MIMO case where $N_t \gg N_r$ as the use of massive arrays has been recently promoted for example use in 5G. When N_r represents users with single antennas, it can be considered as a massive MIMO system. The ergodic capacity for MIMO channel is given by $C = \mathbb{E}[\log_2(\det(\mathbf{I}+\rho\mathbf{W}))]$, where $\mathbf{W}$ is the Wishart matrix associated with the MIMO channel. Hence, to use Theorem 1 to compute this capacity, we need $\mathbb{E}[\mathrm{Tr}(\mathbf{W}^k)]$, where $\mathrm{Tr}(\cdot)$ is the trace operator. The trace of $\mathbf{W}$ can be expanded as a series of Laguerre polynomials [18]. Using that expansion, we derive a closed-form expression as follows:

$$\mathbb{E}[\mathrm{Tr}(\mathbf{W}^k)] = \sum_{l=0}^{N_r-1}\sum_{i=0}^{l}\sum_{j=0}^{l}(-1)^{i+j} \times \frac{l!(l+N_t-N_r)!\Gamma(i+j+N_t-N_r+k+1)}{(l-i)!(l-j)!(N_t-N_r+i)!(N_t-N_r+j)!i!j!}, \tag{5}$$

where $k = 0, 1, 2, \ldots$ and $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$ is the gamma function. One important aspect of MIMO is precoder design and optimization. Using the moments from (5) and App. 1 or App. 2, the low-SNR ergodic capacity can be expressed as a function of precoder matrix entries, which enables the optimization.

B. Applications in Multi-User Channels

We now consider decode-and-forward (DF) relaying [19, p. 386] where optimal power allocation is performed

TABLE II
EXISTING CAPACITY APPROXIMATIONS

Legend	Capacity Approximation
Mac [5]	$\frac{1}{\ln 2}\left(\sum_{k=1}^{2}\rho^k\mathbb{E}\left[X^k\right]\right)$
Tay2 (4) [6], [7]	$\frac{1}{\ln 2}\left(\sum_{k=1}^{2(4)}\rho^k\mathbb{E}\left[X^k\right]-\mathbb{E}[X]^k\right)$
Curve [8]	$\sqrt{\rho}\mathbb{E}\left[\sqrt{X}\right]$
Slope [1], [10], [11] (for MIMO)	$\frac{2N_tN_r}{N_t+N_r}\log_2\left(\frac{\rho}{\ln 2/N_r}\right)$

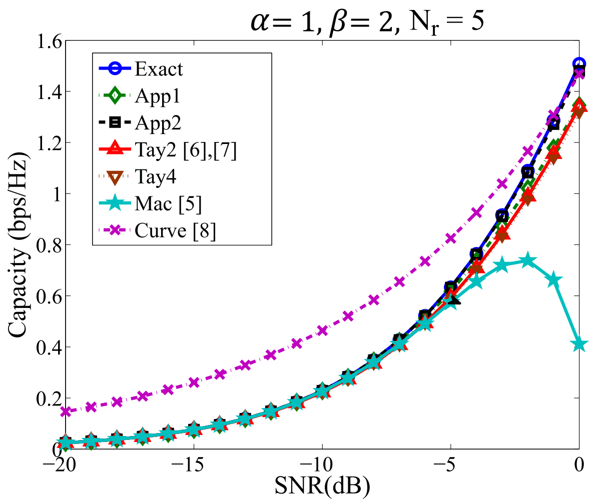

Fig. 1. Comparison of capacity approximations for MRC.

using the proposed capacity approximations. In this scheme, the source ($\mathcal{S}$) deploys superposition block Markov encoding and transmits both its new information and the old information sent in the previous block. The relay ($\mathcal{R}$) decodes the new information in one block and forwards it to the destination ($\mathcal{D}$) in next block. $\mathcal{D}$ performs backward decoding for $\mathcal{S}$ information. Let $\mathcal{S}$ and $\mathcal{R}$ transmit with the same power (ρ). Then, the achievable rate can be expressed as [19]:

$$R = \max_{0<\delta<1} \min\{I_1, I_2\}, \quad I_1 = \mathbb{E}\left[\log(1+g_{rs}^2\delta\rho)\right],$$
$$I_2 = \mathbb{E}\left[\log\left(1+\rho\left(g_{ds}^2+g_{dr}^2+2g_{ds}g_{dr}\sqrt{(1-\delta)}\right)\right)\right], \tag{6}$$

where g_{rs} (g_{ds}) is the fading link from $\mathcal{S}$ to $\mathcal{R}$ ($\mathcal{D}$) while g_{dr} is the link from $\mathcal{R}$ to $\mathcal{D}$. $\mathcal{S}$ allocates part of its power ($\delta \in (0,1)$) to the new information and $1-\delta$ to the old information. The constraint I_1 (I_2) ensures reliable decoding at $\mathcal{R}$ ($\mathcal{D}$). Since I_1 increases with δ while I_2 decreases, the optimal power allocation (δ^*) that maximizes R is obtained by setting $I_1 = I_2$. Using App. 1 or App. 2 to approximate I_1 and I_2, we derive optimal δ^* by solving $I_1 = I_2$.

IV. PERFORMANCE EVALUATION

We compute low-SNR ergodic capacities using App. 1, App. 2, exact numerical value and existing approximations. These and the legends are shown in Table II. We use TSE with central moments as the comparison benchmark.

Exact results in Fig. 1 and Fig. 3 are computed using numerical integration of (1). Whereas exact result for Fig. 2 is not available, thus we randomly generate 10^4 Wishart matrices and compute the low-SNR ergodic capacity.

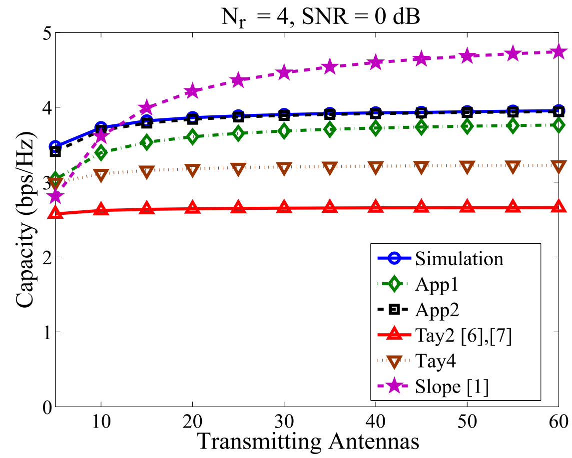

Fig. 2. Comparison of capacity approximations for MIMO.

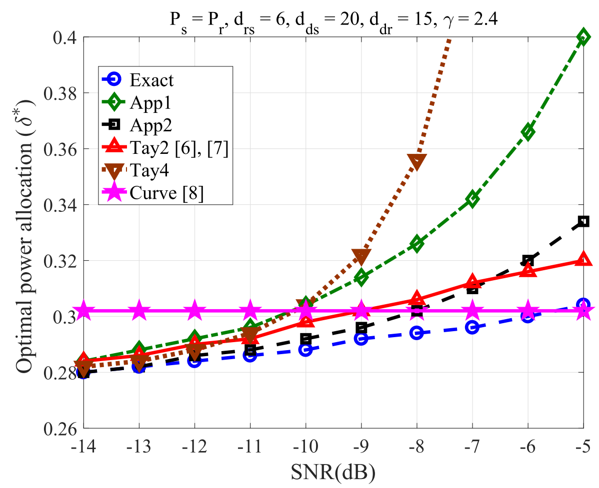

Fig. 3. Optimal power allocation (δ^*) for DF relaying.

Fig. 1 shows the MISO ergodic capacity with MRC reception as given in Section III-A.1, with X_is are i.i.d Gamma distributed. As mentioned earlier, TSE around zero has convergence radius $1/\rho$, thus it diverges when ρ is near 0 dB. In contrast, TSE around $\mathbb{E}[X]$ with two (four) terms does not diverge and perform similar to App. 1 (cf. (2)).

Fig. 2 illustrates the low-SNR ergodic capacity versus N_t for MIMO with fixed SNR $\rho = 0$ dB. This result depicts the divergence of the TSE using the central moments. The MIMO capacity is nearly flat when N_t is much greater than N_r, and we observe that App. 1 and 2 outperform existing approximations.

Fig. 3 provides the optimal resource allocation (δ^*) for DF relaying with inter-node distances (d_{rs}, d_{ds}, d_{dr}). Each link is affected by Rayleigh fading and path loss exponent of $\gamma = 2.4$. We also observe method [8] is better than App.2 for $\rho \geq -8$ dB; however, this crossover depends on several parameters and system details and thus it is difficult to make general conclusions.

V. CONCLUSION

We have proposed two novel, low-SNR approximations for the ergodic capacity. They are generic and require few moments (invariant of the SNR value) only. The moments can be readily derived from the PDF or MGF of X, or even from simulations. Their derivation was illustrated for MISO and MIMO channels. Moreover, we verified the high accuracy of two approximations for MRC channels and MIMO channels.

Appendix A
Proof of Theorem 1

Using the TSE of $\log_2(1 + \rho X)$ around zero, we can rewrite (1) as,

$$\mathbb{E}[C(\rho)] = \tfrac{1}{\ln 2}\mathbb{E}\left[\sum_{k=1}^{\infty} \tfrac{(-1)^{k+1}}{k}\rho^k X^k\right]. \tag{7}$$

To develop the PA, we truncate the series (7) to a finite number of terms. We then apply the Fubini-Tonelli theorem to interchange the summation and expectation [20],

$$\mathbb{E}[C(\rho)] = \tfrac{\mathbb{E}[X]}{\ln 2}\sum_{k=1}^{\infty} \tfrac{(-1)^{k+1}}{k}\mu_k \rho^k. \tag{8}$$

For App. 1 (cf. (2)) we consider first two (four) terms of (8) to determine $C_{[1/1]}(\rho)$ and $C_{[2/2]}(\rho)$. Generally, PA of order $[m/n]$ can be expressed as,

$$C_{[m/n]}(\rho) = \frac{\mathbb{E}[X]}{\ln 2}\frac{\sum_{j=0}^{m} p_j \rho^j}{1 + \sum_{j=1}^{n} q_j \rho^j} + R_{mn}(\rho). \tag{9}$$

Coefficients p_j and q_j are derived by solving following $n + m + 1$ linear equations,

$$\begin{aligned} p_0 &= a_0 \\ p_1 &= a_1 + a_0 q_1 \\ &\vdots \\ p_m &= a_m + a_{m-1}q_1 + \cdots + a_{m-n}q_n \\ 0 &= a_{m+1} + a_m q_1 + \cdots + a_{m-n+1}q_n \\ &\vdots \\ 0 &= a_{m+n} + a_{m+n-1}q_1 + \cdots + a_m q_n, \end{aligned} \tag{10}$$

where a_j is the coefficient of ρ^j in the summation of (8). For App I, $m = 1$ and $n = 1$, and corresponding coefficients are, $a_0 = 0$, $a_1 = 1$, and $a_2 = -\frac{1}{2}\mu_2$, respectively. Solving the equations we obtain $q_1 = \frac{1}{2}\mu_2$, $p_0 = 0$, and $p_1 = 1$. Thereafter, replacing them in (9) we obtain App. 1. Similarly, using $a_0 = 0$, $a_1 = 1$, $a_2 = -\frac{1}{2}\mu_2$, $a_3 = \frac{1}{3}\mu_3$, and $a_4 = -\frac{1}{4}\mu_4$ we obtain the coefficients for App. 2.

Appendix B
Error Bound Analysis

A general error bound for PA is given in [21], which, however, is cumbersome to handle. Instead, we compare the truncation errors of App. 1 and TSE approximations. Thus, only considering the first two terms of (9), the truncation error can be expressed as

$$R_{TSE}(\rho) = \frac{\mathbb{E}[X]}{\ln 2}\sum_{k=3}^{\infty}\frac{(-1)^{k+1}}{k}\mu_k\rho^k. \tag{11}$$

Similarly, the truncation error for App. 1 is given by

$$R_1(\rho) = \frac{\mathbb{E}[X]}{\ln 2}\left(\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}\mu_k\rho^k - \frac{\rho}{1 + \frac{1}{2}\mu_2\rho}\right). \tag{12}$$

Therefore, the relative error can be written as

$$\begin{aligned} R_{TSE}(\rho) &- R_1(\rho) \\ &= \frac{\mathbb{E}[X]}{\ln 2}\left(\frac{\rho}{1 + \frac{1}{2}\mu_2\rho} - \rho + \frac{1}{2}\mu_2\rho^2\right) \\ &= \frac{\mathbb{E}[X]}{\ln 2}\left(\frac{1}{2}\mu_2\rho^2 - \frac{\frac{1}{2}\mu_2\rho^2}{1 + \frac{1}{2}\mu_2\rho}\right) > 0. \end{aligned} \tag{13}$$

Thus, the TSE approximation has a larger truncation error than that of App. 1. Similarly, we can show that App. 2. has smaller truncation error than TSE approximation with four terms.

References

[1] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1319–1343, Jun. 2002.

[2] Y. Dhungana and C. Tellambura, "New simple approximations for error probability and outage in fading," *IEEE Commun. Lett.*, vol. 16, no. 11, pp. 1760–1763, Nov. 2012.

[3] Y. Dhungana and C. Tellambura, "Uniform approximations for wireless performance in fading channels," *IEEE Trans. Commun.*, vol. 61, no. 11, pp. 4768–4779, Nov. 2013.

[4] N. Jindal and A. Goldsmith, "Capacity and optimal power allocation for fading broadcast channels with minimum rates," *IEEE Trans. Inf. Theory*, vol. 49, no. 11, pp. 2895–2909, Nov. 2003.

[5] L. Hanlen and A. Grant, "On capacity of ergodic multiple-input multiple-output channels," in *Proc. 6th Austral. Commun. Theory Workshop*, Brisbane, QLD, Australia, Feb. 2005, pp. 130–134.

[6] J. Pérez, J. Ibanez, L. Vielva, D. J. Perez-Blanco, and I. Santamaria, "Tight closed-form approximation for the ergodic capacity of orthogonal STBC," *IEEE Trans. Wireless Commun.*, vol. 6, no. 2, pp. 452–457, Feb. 2007.

[7] D. B. da Costa and S. Aïssa, "Capacity analysis of cooperative systems with relay selection in Nakagami-m fading," *IEEE Commun. Lett.*, vol. 13, no. 9, pp. 637–639, Sep. 2009.

[8] M. Dohler and H. Aghvami, "On the approximation of MIMO capacity," *IEEE Trans. Wireless Commun.*, vol. 4, no. 1, pp. 30–34, Jan. 2005.

[9] A. Laourine, A. Stéphenne, and S. Affes, "On the capacity of log-normal fading channels," *IEEE Trans. Commun.*, vol. 57, no. 6, pp. 1603–1607, Jun. 2009.

[10] X. Li, S. Jin, M. R. McKay, X. Gao, and K.-K. Wong, "Capacity of MIMO-MAC with transmit channel knowledge in the low SNR regime," *IEEE Trans. Wireless Commun.*, vol. 9, no. 3, pp. 926–931, Mar. 2010.

[11] M. Matthaiou, N. D. Chatzidiamantis, and G. K. Karagiannidis, "A new lower bound on the ergodic capacity of distributed MIMO systems," *IEEE Signal Process. Lett.*, vol. 18, no. 4, pp. 227–230, Apr. 2011.

[12] H. S. Yamada and K. S. Ikeda, "A numerical test of Padé approximation for some functions with singularity," *Int. J. Comput. Math.*, vol. 2014, Nov. 2014, Art. no. 587430.

[13] M. H. Ismail and M. M. Matalgah, "Performance of dual maximal ratio combining diversity in nonidentical correlated Weibull fading channels using Padé approximation," *IEEE Trans. Commun.*, vol. 54, no. 3, pp. 397–402, Mar. 2006.

[14] P. Dienes, *The Taylor Series: An Introduction to the Theory of Functions of a Complex Variable*. Oxford, U.K.: Clarendon, 1931.

[15] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers I*. Dordrecht, The Netherlands: Springer, 1994.

[16] J. Zhang, L. Dai, Y. Han, Y. Zhang, and Z. Wang, "On the ergodic capacity of MIMO free-space optical systems over turbulence channels," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 9, pp. 1925–1934, Sep. 2015.

[17] M. S. Alouini, A. Abdi, and M. Kaveh, "Sum of gamma variates and performance of wireless communication systems over Nakagami-fading channels," *IEEE Trans. Veh. Technol.*, vol. 50, no. 6, pp. 1471–1480, Nov. 2001.

[18] İ. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov. 1999.

[19] A. El Gamal and Y.-H. Kim, *Network Information Theory*. New York, NY, USA: Cambridge Univ. Press, 2012.

[20] S. Saks and L. Young, *Theory of the Integral* (Dover Books on Mathematics). New York, NY, USA: Dover, 2005.

[21] D. Elliott, "Truncation errors in Padé approximations to certain functions: An alternative approach," *Math. Comput.*, vol. 21, no. 99, pp. 398–406, 1967.