Bias reduction in the closed-form maximum likelihood estimator for the Nakagami-m Fading Parameter

Pedro Luiz Ramos, Francisco Louzada and Eduardo Ramos

***Abstract*—In this work, we discuss a corrective bias for the modified maximum likelihood estimator of the Nakagami-m fading parameter. The obtained corrected estimator has a closed-form expression, which plays an essential role in describing fading effects in wireless communication. Numerical results show that the bias-corrected estimator outperforms the existing methods and returns unbiased estimates for the fading parameter even for very small sample sizes.**

***Index Terms*—Nakagami-m fading parameter, maximum likelihood estimator, closed-form estimator, unbiased estimator.**

I. Introduction

The Nakagami-m (NK) distribution [1] has been widely used for modeling radio links. Let X be a random variable with Nakagami-m distribution then its probability density function (PDF) is given by

$$f(t|m,\Omega) = \frac{2}{\Gamma(m)}\left(\frac{m}{\Omega}\right)^m t^{2m-1}\exp\left(\frac{m}{\Omega}t^2\right), \qquad (1)$$

where $\Gamma(m) = \int_0^\infty x^{m-1}e^{-mx}dx$ for all $t>0$, $m \geq 0.5$, and $\Omega > 0$ are the fading and scale parameters, respectively.

Inferential procedures for the parameters of the NK distribution have been discussed earlier. The unbiased estimator for the scale parameter has been presented by Nakagami [1] and is given by $\hat{\Omega} = \frac{1}{n}\sum_{i=1}^n t_i^2$. On the other hand, the moment estimator for the fading parameter has a significant bias. The maximum likelihood estimator (MLE) is also biased for small and moderate samples as well as does not have closed-form expression [2], which is undesirable. Considerable effort has been made to achieved improved estimators with closed-form expression for the fading parameter [3], [4], [5] due to its applicability in embed technology. Ramos et al. [6] proposed a modified maximum likelihood estimator closed-form estimators showing that the proposed approach outperforms the moment estimator and returns similar results when compared with the standard MLE.

In this letter, we derive a closed-form expression for the bias of the modified maximum likelihood estimator using the bias-corrected approach discussed by Cox and Snell [7]. We prove that such a bias-corrected estimator outperforms the existing inferential procedures in terms of minimum Bias and mean square error and returned improved estimates even for very small sample sizes ($n \geq 3$).

P.L. Ramos, E. Ramos and F. Louzada are in the Institute of Mathematical Science and Computing at University of São Paulo, São Carlos, Brazil.

Manuscript received April 11, 2020; revised May 22, 2020.

II. Modified maximum likelihood estimator

Ramos et al. [6] proposed a different approach based on the modified maximum likelihood (MML) to achieve closed-form estimators. The estimators are given by

$$\hat{\Omega}_{\text{MML}} = \frac{1}{n}\sum_{i=1}^n t_i^2, \qquad (2)$$

$$\hat{m}_{\text{MML}} = \frac{\sum_{i=1}^n t_i^2}{\left(\sum_{i=1}^n t_i^2\log\left(t_i^2\right) - \frac{1}{n}\sum_{i=1}^n t_i^2\sum_{i=1}^n \log\left(t_i^2\right)\right)}. \qquad (3)$$

The asymptotic variances Var(m_{MML}) and Var(Ω_{MML}) are given by

$$Var(m_{\text{MML}}) = m^2 + m^3\psi^{(1)}(m+1),$$

$$Var(\Omega_{\text{MML}}) = \frac{\Omega^2}{nm},$$

where the polygamma function is given by

$$\psi^{(m)}(x) = \frac{\partial^{m+1}}{\partial x^{m+1}}\log\Gamma(x) = -\int_0^1 \frac{t^{x-1}}{1-t}\log(t)^m dt.$$

Wang et al. [8] proposed the same closed-form estimator but based on the fractional moment estimator that is given by

$$\hat{\mu}_k = \begin{cases} \dfrac{\hat{m}_{1/p}\hat{m}_2}{2p\left(\hat{m}_{2+1/p} - \hat{m}_{1/p}\hat{m}_2\right)}, & k > 0 \\ \dfrac{\hat{m}_2}{\left(\frac{1}{n}\sum_{i=1}^n t_i^2\log\left(t_i^2\right) - \frac{1}{n}\hat{m}_2\sum_{i=1}^n \log\left(t_i^2\right)\right)}, & k = 0 \end{cases} \qquad (4)$$

where

$$m_k = \frac{\Gamma(\mu + k/2)}{\Gamma(\mu)}\left(\frac{\Omega}{\mu}\right)^{k/2}.$$

The authors argued that $\hat{\mu}_0$ provides better estimates when compared with another fractional moment-based estimator family and should be used. Therefore, the case where $k = 0$ is the best fractal moment estimator and has the same form as the MML. It is important to point out that, we have presented the estimator in terms of the MML approach in order to derive the bias correction which is discussed next.

III. Bias Correction

Here we consider a different approach to correct the Bias of the MML. Let $L(\boldsymbol{\theta}; \boldsymbol{t})$ be the likelihood function with a vector $\boldsymbol{\theta}$ of size p. The joint cumulants of the derivatives of $l(\boldsymbol{\theta}; \boldsymbol{t})$ are given by

$$h_{ij}(\boldsymbol{\theta}) = E\left(\frac{\partial^2 l(\boldsymbol{\theta}; \boldsymbol{t})}{\partial\theta_i\theta_j}\right), \quad h_{ijk}(\boldsymbol{\theta}) = E\left(\frac{\partial^3 l(\boldsymbol{\theta}; \boldsymbol{t})}{\partial\theta_i\partial\theta_j\partial\theta_k}\right) \text{ and}$$
$$h_{ij,k}(\boldsymbol{\theta}) = E\left(\frac{\partial^2 l(\boldsymbol{\theta}; \boldsymbol{t})}{\partial\theta_i\partial\theta_j} . \frac{\partial l(\boldsymbol{\theta}; \boldsymbol{t})}{\partial\theta_k}\right), \quad \text{for} \quad i, j, k = 1, \ldots, p.$$

Consequently, the derivatives of such cumulants are given by

$$h_{ij}^{(k)}(\boldsymbol{\theta}) = \frac{\partial h_{ij}(\boldsymbol{\theta})}{\partial\theta_k}, \quad \text{for} \quad i, j, k = 1, \ldots, p.$$

The Bias of θ_m studied by Cox and Snell [7] for an independent sample without necessarily be identically distributed can be written by

$$\begin{aligned} Bias(\hat{\theta}_m) = &\sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{k=1}^{p} s_{mi}(\boldsymbol{\theta}) s_{jk}(\boldsymbol{\theta}) \left(h_{ij,k}(\boldsymbol{\theta}) + 0.5 h_{ijk}(\boldsymbol{\theta})\right) \\ &+ O(n^{-2}), \end{aligned} \tag{5}$$

where s^{ij} is the (i, j)-th element of the inverse of Fisher's information matrix of $\hat{\boldsymbol{\theta}}$.

The likelihood function for the Nakagami is given by

$$L(\Omega, m; \boldsymbol{t}) = \frac{2^n}{\Gamma(m)^n}\left(\frac{m}{\Omega}\right)^{nm}\left\{\prod_{i=1}^{n} t_i^{2m-1}\right\}\exp\left(-\frac{m}{\Omega}\sum_{i=1}^{n} t_i^2\right). \tag{6}$$

The Bias correction has been presented by Schwartz et al. [9] for the MLE of m, which is given by

$$Bias(\hat{m}) = \frac{\hat{m}\psi^{(1)}(\hat{m}) - \hat{m}^2\psi^{(2)}(\hat{m}) - 2}{2n\left(\hat{m}\psi^{(1)}(\hat{m}) - 1\right)^2} + O(n^{-2}). \tag{7}$$

It follows that

$$\hat{m}_{BC} = \hat{m}_{\text{MML}} - Bias(\hat{m}_{\text{MML}}), \tag{8}$$

hereafter, BC estimator. Unfortunately the solution of this estimator involves the computation of transcendental functions, increasing considerably the computational time as well as the complexity of the estimator. The following theorem provides a useful closed-form approximated Bias corrected (ABC) estimator.

Theorem III.1. *The estimator (8) can be approximated in a closed form expression given by*

$$\hat{m}_{\text{ABC}} = \hat{m}_{\text{MML}} - \frac{\hat{m}_{\text{MML}}}{n}\left(\frac{3}{2} + \frac{3}{2}\left(\frac{\hat{m}_{\text{MML}}}{1 + \hat{m}_{\text{MML}}}\right) + \frac{5}{6}\frac{\hat{m}_{\text{MML}}}{(1 + \hat{m}_{\text{MML}})^2}\right), \tag{9}$$

Proof. See Appendix A. □

The asymptotic variance of $\hat{m}_{\text{BC}}$ is also given by

$$Var(m_{\text{BC}}) = m^2 + m^3\psi^{(1)}(m + 1)\,.$$

An important issue that may arise when dealing with an approximation is that the obtained estimator could return a negative value. The following theorem guarantees that $\hat{m}_{\text{ABC}} > 0$ for all $n \geq 3$.

Theorem III.2. *Let $\boldsymbol{t} = (t_1, \ldots, t_n)$ be random sample with size $n \geq 3$ such that $t_1, t_2, \cdots, t_n$ are positive and not all equal, then $\hat{m}_{\text{ABC}} > 0$.*

Proof. See Appendix B. □

IV. Simulation Study

This section is devoted to comparing the quality of our proposed approach in terms of minimum Bias and the root-mean-square error (RMSE). Here we also consider the moment estimator (ME), where $\hat{\Omega_{ME}}$ is given as the same as (2) and $\hat{\mu}_{ME} = \left(\sum_{i=1}^{n} t^2\right)^2\left[\left(n\sum_{i=1}^{n} t^4\right) - \left(\sum_{i=1}^{n} t^2\right)^2\right]^{-1}$ and the standard maximum likelihood estimator (MLE) that is given in [2]. The metrics used in the comparison are computed by

$$\text{Bias}_m = \frac{1}{N}\sum_{i=1}^{N}(\hat{m}_i - m) \quad \text{and} \quad \text{RMSE}_m = \sqrt{\sum_{i=1}^{N}\frac{(\hat{m}_i - m)^2}{N}},$$

where $N = 100{,}000$ is the number of estimates obtained through the ME, MLE, MML, BC, and ABC estimators. Since Ω is unbiased and has the same form the different estimation procedures, we have omitted the results related to this parameter. The moment estimator is not considered here since Ramos et al. [6] have already shown that this inferential method returned inferior estimates when compared with MML estimator (see also, [8]). The most efficient estimation method will return the Bias closer to zero with smaller RMSEs. The software R (R Core Development Team) was used to conduct this simulation study. The codes are available in the supplemental material.

Figures 1-4 displays the Bias and the RMSEs from the estimates of m obtained using the N samples. The horizontal lines in the figures correspond to Bias and RMSEs being one and zero, respectively.

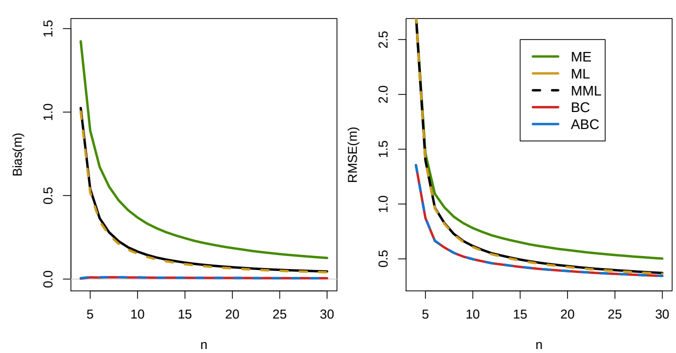

Fig. 1. Bias and RMSEs for m considering $m = 0.5, \Omega = 2$ for $N = 100{,}000$ random samples and $n = (4, 5, \ldots, 30)$.

As shown in Figures 1-2, both Bias and RMSEs related to the estimates of the fading parameter tend to zero when n increases. The MML estimator return estimates with a significant Bias. On the other hand, the corrective approach applied in the MML improves the obtained estimates, the bias correction approach returns precise estimates even for samples

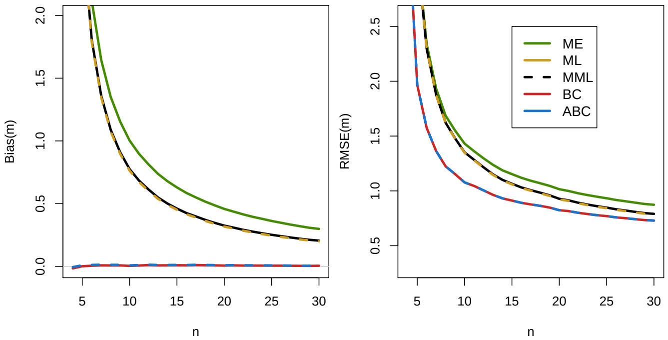

Fig. 2. Bias and RMSEs for m considering $m = 2, \Omega = 4$ for $N = 100,000$ random samples and $n = (4, 5, \ldots, 30)$.

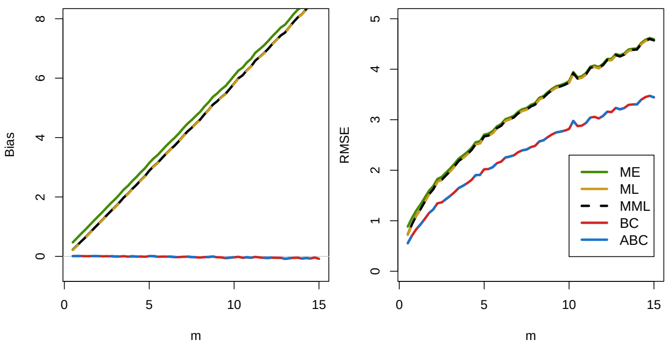

Fig. 3. Bias, RMSEs for different values of $m = (0.5, 0.75, \ldots, 15)$ and $\Omega = 4$ for $N = 100,000$ simulated samples with size $n = 8$.

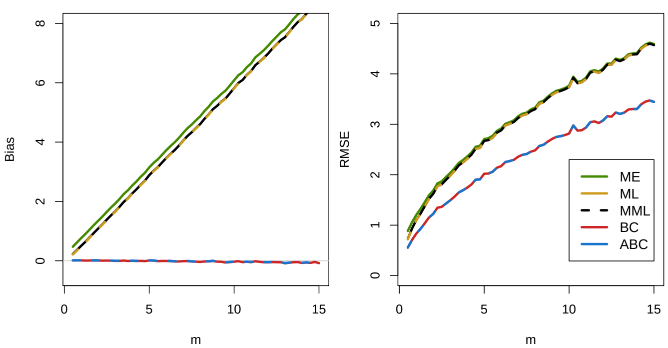

Fig. 4. Bias, RMSEs for $m = (0.5, 0.75, 1.0, 1.5, \ldots, 15)$ considering $\Omega = 80$ for $N = 100,000$ simulated samples and $n = 8$.

with size $n = 4$. Moreover, both BC and ABC returned similar results. However, ABC has a simple structure. From Figures 3 and 4, we observed that the ABC estimator also returned precise estimates for different values of m. More importantly, the RMSE is the same as the BC estimator and smaller than the other currently inferential methods. It is worth mentioning that we considered different values Ω, which does not influence the Bias of m. This is expected since they are independent and orthogonal, i.e., $h_{12}(m, \omega) = h_{21}(m, \omega) = 0$. Overall, these results show that ABC estimators should be considered for estimating the fading parameter of the NK distribution.

V. Discussion

In this paper, we have proposed a bias-corrective approach for the modified maximum likelihood estimator of the Nakagami fading parameter. The bias is obtained under formal rules; however, it does not have closed-form expression. To overcome this problem, we proposed a useful approximation that has closed-form and returns precise estimates. A simulation study is presented to demonstrates that our new method outperforms the existing estimators in terms of Bias and RMSE. Overall, the proposed bias-corrected estimators returned precise estimates even for very small sample sizes and should be used to estimates the fading parameter. As a possible extension of this current work we will consider the approximate form of the distribution obtained by Cheng et al. [10] with our estimator and compare the obtained confidence intervals with other common intervals achieved from asymptotic theory and Bayesian methods.

Appendix A
Proof of Theorem III.1

Let $r(x) = \frac{x\psi^{(1)}(x) - x^2\psi^{(2)}(x) - 2}{2\left(x\psi^{(1)}(x)-1\right)^2}$ be a function that will be approximated via a rational function. Using the asymptotic relations

$$\lim_{x\to 0^+} \frac{\psi^{(1)}(x)}{\frac{1}{x^2}} = 1, \; \psi^{(1)}(x) = \frac{1}{x} + \frac{1}{2x^2} + o\left(\frac{1}{x^2}\right),$$

$$\lim_{x\to 0^+} \frac{\psi^{(2)}(x)}{\frac{1}{x^3}} = -2, \text{ and } \psi^{(2)}(x) = -\frac{1}{x^2} - \frac{1}{x^3} + o\left(\frac{1}{x^3}\right),$$

we obtained

$$\lim_{x\to 0^+} r(x) = 0 \text{ and } \lim_{x\to\infty^+} r(x) - 3x = -\frac{2}{3}.$$

and since $\lim_{x\to\infty^+} 3x - \left(\frac{3x}{2} + \frac{3x^2}{2(1+x)}\right) = \lim_{x\to\infty^+} \frac{3x}{2(x+1)} = \frac{3}{2}$ we obtain

$$\lim_{x\to 0^+} r(x) - \frac{3x}{2} - \frac{3x^2}{2(1+x)} = 0 \text{ and } \lim_{x\to\infty^+} r(x) - \frac{3x}{2} - \frac{3x^2}{2(1+x)} = \frac{5}{6}.$$

Therefore, defining $f(x) = r(x) - \frac{3x}{2} - \frac{3x^2}{2(1+x)}$ and using the reparametrization $x(y) = \frac{y}{1-y}$ we noticed graphically (see Figure 5) that $f(x(y))$ closely resembles a quadratic function ky^2 and since $\lim_{y\to 1} f(x(y)) = \frac{5}{6}$ we choose $k = \frac{5}{6}$. Thus, defining $g(x) = r(x) - \frac{3x}{2} - \frac{3x^2}{2(1+x)} - \frac{5x^2}{6(1+x)^2}$, the global error $\max_{y\in(0,1)} |g(x(y))|$ is expected to be very small, as we can see in Figure 5.

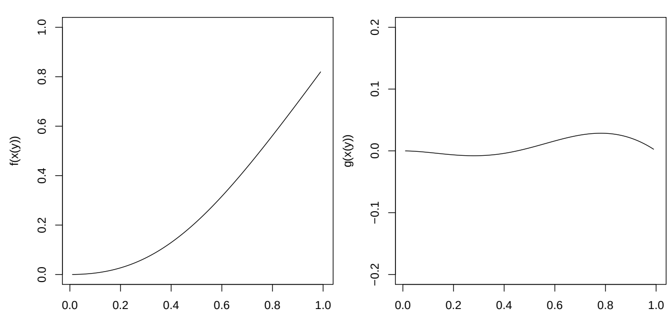

Fig. 5. Left Panel: Graph of $f(x(y))$. Right Panel: Graph of $g(x(y))$.

The following proposition formalizes this fact.

Proposition A.1. *For* $g(x) = r(x) - \left(\frac{3}{2}x + \frac{3x^2}{2(1+x)} + \frac{5x^2}{6(1+x)^2}\right)$, *it follows that* $\max_{x\in(0,\infty)} |g(x)| < 5 \cdot 10^{-2}$.

Proof. Let

$$p_1(n,w) = \frac{(n-1)!}{w^n} + \frac{n!}{2w^{n+1}} + \frac{(n+1)!}{12w^{n+2}} - \frac{(n+3)!}{720w^{n+4}},$$
$$p_2(n,w) = \frac{(n-1)!}{w^n} + \frac{n!}{2w^{n+1}} + \frac{(n+1)!}{12w^{n+2}},$$
$$q_1(n,x) = p_1(n,x+1) + \frac{n!}{x^{n+1}},$$
$$q_2(n,x) = p_2(n,x+1) + \frac{n!}{x^{n+1}},$$
$$g_1(x) = \frac{xq_1(1,x) + x^2q_2(2,x) - 2}{2(xq_2(1,x)-1)^2} - \beta(x) \text{ and}$$
$$g_2(x) = \frac{xq_2(1,x) + x^2q_1(2,x) - 2}{2(xq_1(1,x)-1)^2} - \beta(x).$$

In the following, in item *i)* we shall prove that $g_1(x) < g(x) < g_2(x)$ for all $x > 0$ and in *ii)* we shall prove that $-5 \cdot 10^{-3} < g_1(x)$ and $g_2(x) < 5 \cdot 10^{-3}$ for all $x > 0$ which will conclude the proof.

i) Theorem 9 of Alzer [11] states precisely that the inequality
$$p_1(n,w) < (-1)^{n+1}\psi^{(n)}(w) < p_2(n,w) \tag{10}$$
holds for all $n > 0$ and $w > 0$. This inequality gives us an increasingly worse approximation for $(-1)^{n+1}\psi^{(n)}(w)$ the closer w gets to 0. To fix this, using the classical relation $\psi^{(n)}(x+1) = \psi^{(n)}(x) - (-1)^{n+1}\frac{n!}{x^{n+1}}$ combined with the inequality in (10) applied for $w = x+1$ it follows that
$$q_1(n,x) < (-1)^{n+1}\psi^{(n)}(x) < q_2(n,x) \tag{11}$$
for all $n > 0$ and $x > 0$. In special we have
$$\begin{gathered} xq_2(1,x) - 1 > x\psi^{(1)}(x) - 1 > xq_1(1,x) - 1 = \\ \frac{15x^5 + 80x^4 + 175x^3 + 199x^2 + 120x + 30}{30x(x+1)^5} > 0 \end{gathered} \tag{12}$$
for all $x > 0$. Thus, from (11) and (12) we conclude that
$$g_1(x) < g(x) < g_2(x)$$
for all $x > 0$, which proves item *i)*.

ii) Through careful manual computation or any symbolic arithmetic software available one can compute explicitly $2(xq_1(n,x) - 1)^2(5 \cdot 10^{-2} + g_1(x)) = \frac{N_1(x)}{540x^2(x+1)^8}$ and $2(xq_2(1,x) - 1)^2(5 \cdot 10^{-2} - g_2(x)) = \frac{N_2(x)}{13500x^2(x+1)^{12}}$ where
$$\begin{gathered} N_1(x) = 27x^8 + 384x^7 + 1627x^6 + 3072x^5 + \\ 2976x^4 + 1728x^3 + 1044x^2 + 648x + 108, \text{ and} \\ N_2(x) = 675x^{12} + 4800x^{11} + 15125x^{10} + 40610x^9 + \\ 141565x^8 + 423580x^7 + 836018x^6 + 1059426x^5 + \\ 864483x^4 + 448020x^3 + 142920x^2 + 27000x + 2700. \end{gathered}$$
Thus $N_1(x) > 0$ and $N_2(x) > 0$ for $x > 0$ which combined with (12) implies $5 \cdot 10^{-2} + g_1(x) > 0$ and $5 \cdot 10^{-2} - g_2(x) > 0$ for all $x > 0$, which proves item *ii)* and thus the Proposition.

□

Appendix B
Proof of Theorem III.2

Ramos et al. [6] has proved that $\hat{m}_{MML} > 0$ for all $n \geq 1$, and since $\hat{m}_{ABC} = \beta(\hat{m}_{MML})$ where $\beta(x) = x - \frac{1}{n}\left(\frac{3}{2}x + \frac{3}{2}\left(\frac{x^2}{1+x}\right) + \frac{5}{6}\frac{x}{(1+x)^2}\right)$, it follows that in order to prove that $\hat{m}_{ABC} > 0$ it is enough to prove that $\beta(x) > 0$ for all $x > 0$ and $n \geq 3$.

Under the reparametrization $x(y) = \frac{y}{(1-y)}$, to prove that $\beta(x) > 0$, is equivalent to proving that $\delta(y) = \beta(x(y)) = \frac{y}{(1-y)} - \frac{1}{n}\left(\frac{3y}{2(1-y)} + \frac{3y^2}{2(1-y)} + \frac{5}{6}y^2\right) > 0$ for $y \in (0,1)$. But since
$$\delta(y) = \frac{6n + 5y^2 - 14y - 9}{6n(1-y)},$$
and $\frac{d}{dy}(-5y^2 + 14y + 9) = -10y + 14y > 0$ for $0 \leq y \leq 1$ it follows that $-5y^2 + 14y + 9 \leq -5(1^2) - 14(1) + 9 = 18 \leq 6n$ for $n \geq 3$ and $y \in (0,1)$ an thus $\delta(y) > 0$ for $y \in (0,1)$. Hence, it follows that $\beta(x) > 0$ for all $x > 0$ and $n \geq 3$ and the proof is completed.

Acknowledgment

Pedro L. Ramos is grateful to the São Paulo State Research Foundation (FAPESP Proc. 2017/25971-0), F. Louzada is grateful to CNPq and FAPESP. We also thank Francisco Rodrigues for useful comments.

References

[1] N. Nakagami, “The m-distribution a general formulation of intensity distribution of rapid fading,” *Statistical Methods in Radio Wave Propagation: Proceedings of a Symposium*, pp. 3–36, 1960.
[2] J. Cheng and N. C. Beaulieu, “Maximum-likelihood based estimation of the nakagami m parameter,” *IEEE Communications letters*, vol. 5, no. 3, pp. 101–103, 2001.
[3] Q. Zhang, “A note on the estimation of nakagami-m fading parameter,” *IEEE communications letters*, vol. 6, no. 6, pp. 237–238, 2002.
[4] J. Cheng and N. C. Beaulieu, “Generalized moment estimators for the nakagami fading parameter,” *IEEE communications letters*, vol. 6, no. 4, pp. 144–146, 2002.
[5] J. Gaeddert and A. Annamalai, “Further results on nakagami-m parameter estimation,” in *Vehicular Technology Conference, 2004. VTC2004-Fall. 2004 IEEE 60th*, vol. 6. IEEE, 2004, pp. 4255–4259.
[6] P. L. Ramos, F. Louzada, and E. Ramos, “An efficient, closed-form map estimator for nakagami-m fading parameter,” *IEEE Communications Letters*, vol. 20, no. 11, pp. 2328–2331, 2016.
[7] D. R. Cox and E. J. Snell, “A general definition of residuals,” *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 248–275, 1968.
[8] N. Wang, X. Song, and J. Cheng, “Generalized method of moments estimation of the nakagami-m fading parameter,” *IEEE Transactions on Wireless Communications*, vol. 11, no. 9, pp. 3316–3325, 2012.
[9] J. Schwartz, R. T. Godwin, and D. E. Giles, “Improved maximum-likelihood estimation of the shape parameter in the nakagami distribution,” *Journal of Statistical Computation and Simulation*, vol. 83, no. 3, pp. 434–445, 2013.
[10] J. Cheng, N. Wang, and C. Tellambura, “Probability density function of logarithmic ratio of arithmetic mean to geometric mean for nakagami-m fading power,” in *2010 25th Biennial Symposium on Communications*. IEEE, 2010, pp. 348–351.
[11] H. Alzer, “On some inequalities for the gamma and psi functions,” *Mathematics of computation*, vol. 66, no. 217, pp. 373–389, 1997.