Improved Coefficients for the Karagiannidis–Lioumpas Approximations and Bounds to the Gaussian $Q$-Function

Islam M. Tanash and Taneli Riihonen, Member, IEEE

Abstract—We revisit the Karagiannidis–Lioumpas (KL) approximation of the $Q$-function by optimizing its coefficients in terms of absolute error, relative error and total error. For minimizing the maximum absolute-relative error, we describe the targeted uniform error functions by sets of nonlinear equations so that the optimized coefficients are the solutions thereof. The total error is minimized with numerical search. We also introduce an extra coefficient in the KL approximation to achieve significantly tighter and total error at the expense of unbounded relative error. Furthermore, we extend the KL expression to lower and upper bounds with optimized coefficients that minimize the error measures in the same way as for the approximations.

Index Terms—Communication theory, error probability.

I. INTRODUCTION

KARAGIANNIDIS and Lioumpas presented in [1] a relatively tight, yet analytically tractable, approximation for the Gaussian $Q$-function [2] as follows:

$$Q(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} \exp \left( -\frac{t^2}{2} \right) \, dt$$

$$\approx a \exp (-bx^2) \cdot \frac{1 - \exp (-cx)}{x} \triangleq \tilde{Q}(x) \tag{1}$$

for which their original study sets $(a, b, c) = (\frac{1}{\sqrt{2}}, \frac{1}{2}, -\frac{4}{\sqrt{2}})$ and proposes for error minimization example coefficient values $A = 1.98$ and $B = 1.135$ rendering $(a, c) \approx (0.3515, 1.4001)$.

Despite drawing some criticism [3] shortly after publication, the ‘Karagiannidis–Lioumpas (KL) approximation’ has gradually established itself as one of the most usable substitutes for the Gaussian $Q$-function in communication theory problems and the letter [1] has received a large number of citations; it is only fitting to begin calling the expression after its inventors.

A diverse set of applications for the KL approximation can be found in [4]–[8] to name but a few prominent articles. In general, the approximation is often used in the calculation of average bit or symbol error probability as a tractable replacement for the Gaussian $Q$-function such that analysis can be carried out and completed in a closed form at the cost of making results tight approximations instead of exact ones. This usually involves integrating something like, e.g., $f(Q(x(\gamma)))$, where even simple functions $f(q)$ and $x(\gamma)$, which are derived from the communication system under study, may forbid exact analysis using the actual $Q$-function [9]–[11]. One should note especially that the approximation is always used in an intermediate step of analytical derivations and it is not meant for numerical probability computations per se — instead, rational Chebyshev functions [12] are perfect to that end.

This Letter is inspired by the fact that the original study [1] presents explicit values of $a$ and $c$ for only one approximation (which has low integrated total error when $b = \frac{3}{2}$, to be exact). However, the KL approximation family is actually much more versatile, whereby new coefficients can be acquired in terms of other criteria for better accuracy depending on the application. The KL expression can be also repurposed to achieve lower and upper bounds (that are also tight approximations) and, in certain cases, coefficients admit explicit values. Furthermore, by introducing the extra coefficient $b$ in (1) that originally was $b = \frac{3}{2}$ and permitting $b < \frac{3}{2}$, we achieve significantly improved accuracy in terms of absolute and total error.

The objective of this Letter is to apply the KL expression of the $Q$-function to derive improved approximations and bounds which are global and tight over $x \geq 0$ by optimizing the coefficients $(a, b, c)$ in respect to their minimum global absolute or relative error or minimum integrated total error. Like [11] for another popular expression [9], we present new formulation that minimizes the maximum global error of (1) by constructing a set of equations, which describes the corresponding error function, and solve them numerically to find the optimized coefficients. The total error is optimized with exhaustive search for reference. In general, when optimizing one of the three criteria, better performance will be achieved at the expense of decreased accuracy in terms of the others.

The new coefficients solved herein are applicable as one-to-one replacements for the original ones of [1] adopted into the analysis of [4]–[8] and many other studies. Literature is rich in approximations/bounds for the $Q$-function and, typically, the application’s mathematics define, which ones are tractable for it. Whenever (1) is preferred, our coefficients offer variety to tailor accuracy for the application or to use bounds.

The tractable series expansion of the KL expression proposed in [13] can likewise be used with these substitutes (but without guarantee that bounds remain global). Consequently, they are useful also in the contexts of, e.g., [14]–[20] and many others.

II. PRELIMINARIES

The case $x \geq 0$ is presumed throughout this Letter with little loss of generality because the relation $Q(x) = 1 - Q(-x)$ extends all the considered functions to the negative real axis. In fact, this is the main motive for optimizing approximations and bounds also subject to an additional constraint $Q(0) = \frac{1}{2}$ that makes their extensions continuous at the origin like $Q(x)$.

This study solves optimized approximations and bounds for three criteria and for combinations thereof, viz. $\min_{(a,b,c)} \delta_{\text{max}}$.

Manuscript received December 18, 2020; revised January 8, 2021; accepted January 9, 2021. Date of publication January 18, 2021; date of current version May 6, 2021.

The authors are with Tampere University, 33720 Tampere, Finland (e-mail: islam.tanash@tuni.fi; taneli.riihonen@tuni.fi).

Digital Object Identifier 10.1109/LCOMM.2021.3052257
and the error functions are defined as
\[ d(x) \triangleq \tilde{Q}(x) - Q(x), \]
\[ r(x) \triangleq \frac{d(x)}{Q(x)} = \frac{\tilde{Q}(x)}{Q(x)} - 1. \]

For baseline reference, the coefficients originally given in [1] render \( d_{\text{max}} \approx 0.00789 \), \( r_{\text{max}} \approx 0.119 \), and \( d_{\text{tot}} \approx 0.00385 \).

As implied above, the presented approximations and bounds will be global ones, i.e., tight over the whole non-negative real axis (for all \( x \geq 0 \)). The error functions converge to explicit values, which may be local extrema, at both ends of this range:
\[ \lim_{x \to 0} d(x) = ac - \frac{1}{2} \lim_{x \to \infty} r(x) = 2ac - 1, \]
\[ \lim_{x \to \infty} d(x) = 0, \lim_{x \to \infty} r(x) = \begin{cases} \infty, & \text{if } b < \frac{1}{2}, \\ a\sqrt{2\pi} - 1, & \text{if } b = \frac{1}{2}, \\ -1, & \text{if } b > \frac{1}{2}. \end{cases} \]

The last limit shows especially that global approximations and bounds in terms of relative error exist if and only if we set \( b = \frac{1}{2} \). However, as a novel fact, our study demonstrates that absolute error and total error can be instead significantly reduced by permitting \( b < \frac{1}{2} \). Therefore, two scenarios of approximations for the absolute and total error are considered in this Letter, i.e., approximations with \( b = \frac{1}{2} \) or \( b < \frac{1}{2} \).

Local error extrema may occur also at critical points, where the derivatives of the continuous error functions vanish. Denoting differentiation with an apostrophe, they are given by
\[ d'(x) = \tilde{Q}'(x) - Q'(x), \]
\[ r'(x) = \frac{\tilde{Q}'(x)Q(x) - Q'(x)\tilde{Q}(x)}{(Q(x))^2}, \]
where
\[ \tilde{Q}'(x) = -\frac{a}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right), \]
\[ Q'(x) = -\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right). \]

Two variations of approximations are considered herein: \( d(0) = r(0) = 0 \) and \( d(0) = -d_{\text{max}} \) (resp. \( r(0) = -r_{\text{max}} \)). The former case maintains the continuity of the \( Q \)-function when extending to \( x < 0 \) and results in \( c = \frac{1}{\sqrt{2\pi}} \), when substituted in \( \lim_{x \to 0} d(x) \) (resp. \( \lim_{x \to 0} r(x) \)) that is given in (4). The latter case provides slightly better accuracy at the cost of a discontinuity occurring at \( x = 0 \) and results in \( c = \sqrt{\frac{2}{\pi}} \) in the cases of relative error, by solving \( \lim_{x \to 0} r(x) = -r_{\text{max}} \) with \( \lim_{x \to \infty} r(x) = -r_{\text{max}} \) that are defined in (4).

### III. ALTERNATIVE IMPROVED COEFFICIENTS FOR (1)

In this section, we describe the methodologies to solve the new coefficients \((a, b, c)\) for the KL expression. They are optimized either in the minimax sense or in terms of the integrated total error to yield an approximation, an upper bound or a lower bound. All the 17 thus-obtained improved/alternative coefficient sets and accuracy thereof are listed in Table I.

### A. Global Uniform Approximations and Bounds

The minimax optimization problems are solved in terms of both absolute and relative errors defined in (2) and (3), respectively, by constructing a set of nonlinear equations. This set describes the resulting error function, which should be uniform with equal values for all the extremal points. Each extremum point yields two equations, where one expresses its value and the other sets the derivative of the error function to zero at that point. In addition, one equation (for \( d(x) \)) or two equations (for \( r(x) \)) is/are obtained from evaluating the limits at the two endpoints of the considered range, \([0, \infty)\), per (4).

The resulting sets of equations, which have equal number of equations and unknowns, can be solved straightforwardly by any numerical tool for the considered variations to find the optimized sets of coefficients that satisfy \( \min_{(a, b, c)} d_{\text{max}} \) for the absolute error and \( \min_{(a, b, c)} r_{\text{max}} \) for the relative error. We used iteratively random initial guesses for the unknowns in this approach, namely \((a, b, c)\), \( d_{\text{max}} \) or \( r_{\text{max}} \), and the location of the extrema \((x_k)\), until \texttt{fsolve} in Matlab converged to the solution, which is confirmed by substitution. The formulations for the minimax approximations/bounds are described below.

1) **Approximations in Terms of Absolute Error:** The coefficients \((a, b, c)\) are optimized for approximations in terms of the absolute error by formulating a set of equations as
\[ d'(x_k) = 0, \quad \text{for } k = 1, 2 \text{ or } 1, 2, 3, \]
\[ d(x_k) = (-1)^{k+1}d_{\text{max}}, \quad \text{for } k = 1, 2 \text{ or } 1, 2, 3, \]
\[ \left\{ \begin{array}{l} a = 1 \quad \text{when } d(0) = 0, \\ c = \frac{1}{2} - d_{\text{max}}, \quad \text{when } d(0) = -d_{\text{max}}, \end{array} \right. \]
where \( x_k \) is an extremum point. The number of the error function’s extrema depends on the value of \( b \); if \( b \) is fixed to \( \frac{1}{2} \), then we have two extrema, whereas if \( b \) is allowed to be any positive value, then we need three separate extremum.

2) **Lower Bounds in Terms of Absolute Error:** For the lower bounds, we need to find the optimized coefficients which minimize the global absolute error for \( d(x) \leq 0 \) when \( x \geq 0 \). The value of \( b \) must always equal to \( \frac{1}{2} \). The tightest resulting uniform error function will start from \( d(0) = -d_{\text{max}} \), with its

### TABLE I

| \( n \times \) type | \( a \) | \( b \) | \( c \) | \( d_{\text{max}} \) | \( r_{\text{max}} \) |
|------------------|-------|-------|-------|----------------|----------------|
| 1 | \( Q(0) = 0 \) | 0.3572481 | 1/2 | 1.406071 | 0.007487 || 0.189 | 0.008462 |
| 2 | \( Q(0) = 0 \), \( Q(1) = Q(0) \) | 0.3227222 | 0.474742 | 1.554464 | 0.009356 | 0.001159 |
| 3 | \( Q(0) = 0 \), \( Q(1) = Q(0) \) | 0.319655 | 0.499381 | 1/2(34) | 0.005375 | 0.001350 |
| 4 | \( Q(0) = 0 \), \( Q(1) = Q(0) \) | 0.335549 | 1/2 | 1.444446 | 0.002552 | 0.002552 |
| 5 | \( Q(0) = 0 \), \( Q(1) = Q(0) \) | 0.332160 | 1/2 | 1/2(3) | 0.002552 | 0.001647 |
| 6 | \( \tilde{Q}(0) = 0 \) | 0.380797 | 1/2 | 1/2(3) | 0.007242 | 0.001043 |
| 7 | \( Q(0) = 0 \), \( Q(1) = Q(0) \) | 0.376565 | 1/2 | 1/2(34) | 0.013797 | 0.003475 |
| 8 | \( Q(0) = 0 \), \( Q(1) = Q(0) \) | 0.329781 | 1/2 | 1.506309 | 0.002227 | 0.001974 |
| 9 | \( \tilde{Q}(0) = 0 \) | 0.366201 | 1/2 | 1/2(3) | 0.007485 | 0.001974 |

* some terms \( 0.3572481 \) to \( 0.474742 \) [1], \( 1.406071 \) to \( 1.554464 \), and \( 0.007487 \) to \( 0.008462 \) had to be taken as such values to obtain the minimal error.
maximum equal to zero and its minimum equal to $-d_{\text{max}}$ so that we can formulate a set of equations as
\[
\begin{align*}
d'(x_1) &= d'(x_2) = 0, \\
d(x_1) &= 0, \\ ac &= \frac{1}{2} - d_{\text{max}}.
\end{align*}
\] (8)

When $d(0) = 0$, we get $a = \sqrt{\frac{2}{\pi}}$ and $c = \sqrt{\frac{2}{\pi}}$ by imposing $d'(0) = 0$ (only in this case), which produces $ac^2 = \sqrt{2/\pi}$, and solving with $c = \frac{1}{2}$ that results from setting $d(0) = 0$.

3) Upper Bounds in Terms of Absolute Error: The set of equations becomes
\[
\begin{align*}
d'(x_1) &= d'(x_2) = d'(x_3) = 0, \\
d(x_1) &= d(x_2) = d_{\text{max}}, \\ ac &= \frac{1}{2}.
\end{align*}
\] (9)

In particular, we shape the uniform error function to have three extrema with $d(x) \geq 0$ when $x \geq 0$ in which its maxima are equal to $d_{\text{max}}$ and its minimum is equal to zero. The corresponding error function must always start from $d(0) = 0$.

4) Approximations in Terms of Relative Error: The targeted uniform error function in terms of the relative error consists of only one maximum point and converges to $-r_{\text{max}}$ as $x$ tends to infinity, which results in $-r_{\text{max}} = a\sqrt{\frac{2}{\pi}} - 1$ according to (4). Therefore, we can formulate the set of equations as
\[
\begin{align*}
r'(x_1) &= 0, \\ r(x_1) &= r_{\text{max}}, \\ a c &= \frac{1}{2}, \\ a &= 1 - \frac{r_{\text{max}}}{\sqrt{2 \pi}}.
\end{align*}
\] (10)

5) Lower Bounds in Terms of Relative Error: We need to find the optimized coefficients, $a$ and $c$, in the minimax sense for $r(x) \leq 0$ when $x \geq 0$ which converges to $-r_{\text{max}}$ as $x$ tends to infinity. The resulting error function can either start from $r(0) = -r_{\text{max}}$ to formulate a set of equations as
\[
\begin{align*}
r'(x_1) &= r(x_1) = 0, \\ a c &= \frac{1 - r_{\text{max}}}{2}, \\ a &= \frac{1 - r_{\text{max}}}{\sqrt{2 \pi}}.
\end{align*}
\] (11)

or from $r(0) = 0$ yielding $a = \sqrt{\frac{2}{\pi}}$ and $c = \sqrt{\frac{8}{\pi}}$ like with the corresponding lower bound in terms of absolute error.

6) Upper Bound in Terms of Relative Error: We must ensure that $r(x) \geq 0$ when $x \geq 0$ for the uniform error function. The resulting error function has only one maximum
point and converges to zero as $x$ tends to infinity. Therefore, $a = \frac{1}{\sqrt{2\pi}}$ and $c = \frac{\sqrt{\pi}}{2}$ as proposed earlier in [21] and $b$ is known to be equal to $\frac{1}{2}$. The optimized upper bound in terms of relative error is also optimal in terms of absolute error and integrated total error for the case where $b = \frac{1}{2}$.

B. Numerical Optimization in Terms of Total Error

Instead of defining $d(x) = \int_0^R |d(x)| \, dx$ like in [1] and so making optimized coefficients specific to the value chosen for $R$ and limited to the range $[0, R]$, we measure total error with $R \to \infty$ and obtain globally optimized approximations and bounds. In particular, we optimized the coefficients for the two variations of the approximations with or without setting $b = \frac{1}{2}$ by performing an extensive search, where we evaluated the target metric ($d(x)$) over wide one/two/three-dimensional grids for the unknowns $a, (a, b), (a, c)$, or $(a, b, c)$ with granularity of $0.000001$ and selected the grid point with the minimum total error for each variation. This renders four sets of optimized coefficients. Extra constraint checks guarantee $d(x) < 0$ for the lower bound and $d(x) > 0$ for the upper bound.

IV. NUMERICAL RESULTS AND CONCLUSION

We summarize the improved coefficients for the minimax approximations and bounds and for the total absolute error in Table I and illustrate their error functions in Fig. 1, together with the original KL approximation from [1] and reference approximations and bounds from [9] and [10].

The numerical results show that the improved coefficients of the proposed KL approximations and bounds are optimal subject to their optimization targets, yet expressed precisely in implicit form as solutions to systems of nonlinear equations as opposed to relying on numerical search to minimize error measures. In some specific cases, a part or even all of the three coefficients can be expressed as explicit constants. The best approximation/bound from Table I for a specific application is chosen by contrasting requirements against Fig. 1, provided that the KL expression (1) is suitable for it to begin with.

As an ultimate conclusion, the presented data suggests good alternatives to the original coefficients given in [1] for the case of $b = \frac{1}{2}$. In some applications, the accuracy of the KL approximation might be improved by choosing instead $A = 1.95$, $B = 1.113$ (a compromise between all Ayh-1) for decreasing both absolute and relative error by round 15% at the cost of increasing total error by round 65%; or $A = 2.03$, $B = 1.162$ (At-4) for decreasing absolute error and total error by round 10% and 25%, respectively, at the cost of increasing relative error by round 15%. Sometimes it may also be useful to choose $A = B/\sqrt{\pi} \approx 1.88$, $B = 1.061$ (Ar-5) for minimizing relative error (with round 50% reduction) subject to zero error at the origin. In contrast, when primarily minimizing absolute error, accuracy can be improved significantly by generalizing the KL approximation to allow any positive $b$; namely, the choice $a = 0.32$, $b = 0.4703$, $c = 1.5625$ (Ad-2) guarantees zero error at the origin while decreasing absolute error and total error as much as round 90% and 65%, respectively, at the cost of making relative error unbounded for large arguments.

REFERENCES

[1] G. K. Karagiannidis and A. S. Lioumpas, “An improved approximation for the Gaussian Q-function,” IEEE Commun. Lett., vol. 11, no. 8, pp. 644–646, Aug. 2007.
[2] S. Aggarwal, “A survey-cum-tutorial on approximations to Gaussian Q-function for symbol error probability analysis over Nakagami-m fading channels,” IEEE Commun. Surveys Tuts., vol. 21, no. 3, pp. 2195–2223, 3rd Quart., 2019.
[3] S. Dyer, S. A. Dyer, “Corrections to, and comments on, ‘an improved approximation for the Gaussian Q-function’,” IEEE Commun. Lett., vol. 12, no. 4, p. 231, Apr. 2008.
[4] C. Potter, G. K. Venayagamoorthy, and K. Kosbar, “RNN based MIMO channel prediction,” Signal Process., vol. 90, no. 2, pp. 440–450, Feb. 2010.
[5] T. Tan and L. B. Le, “Distributed MAC protocol for cognitive radio networks: Design, analysis, and optimization,” IEEE Trans. Veh. Technol., vol. 60, no. 8, pp. 3990–4003, Oct. 2011.
[6] J. Wu, N. B. Mehta, A. F. Molisch, and J. Zhang, “Unified spectral efficiency analysis of cellular systems with channel-aware schedulers,” IEEE Trans. Commun., vol. 59, no. 12, pp. 3463–3474, Dec. 2011.
[7] D. Malak, M. Al-Shalash, and J. G. Andrews, “Optimizing content caching to maximize the density of successful receptions in device-to-device networking,” IEEE Commun. Lett., vol. 64, no. 10, pp. 4365–4380, Oct. 2016.
[8] S. Lin et al., “Rayleigh fading suppression in one-dimensional optical scatterers,” IEEE Access, vol. 7, pp. 17125–17132, Jan. 2019.
[9] M. Chiani, D. Dardari, and M. K. Simon, “New exponential bounds and approximations for the computation of error probability in fading channels,” IEEE Trans. Wireless Commun., vol. 24, no. 5, pp. 840–845, May 2003.
[10] M. Lopez-Benitez and F. Casadevall, “Versatile, accurate, and analytically tractable approximation for the Gaussian Q-function,” IEEE Trans. Commun., vol. 59, no. 4, pp. 917–922, Apr. 2011.
[11] I. M. Tanash and T. Riikonen, “Global minimax approximations and bounds for the Gaussian Q-function by sums of exponentials,” IEEE Trans. Commun., vol. 68, no. 10, pp. 6514–6524, Oct. 2020.
[12] W. J. Cody, “Rational Chebyshev approximations for the error function,” Math. Comput., vol. 23, no. 107, pp. 631–637, Jul. 1969.
[13] Y. Isikapalli and B. D. Rao, “An analytically tractable approximation for the Gaussian Q-function,” IEEE Commun. Lett., vol. 12, no. 9, pp. 669–671, Sep. 2008.
[14] Y. Isikapalli and B. D. Rao, “Packet error probability of a transmit beamforming system with imperfect feedback,” IEEE Trans. Signal Process., vol. 58, no. 4, pp. 2298–2314, Apr. 2010.
[15] Q. F. Zhou, Y. Li, F. C. M. Lau, and B. Vucetic, “Decode-and-forward two-way relaying with network coding and opportunistic relay selection,” IEEE Trans. Commun., vol. 58, no. 11, pp. 3070–3076, Nov. 2010.
[16] M. Seyf, S. Muhaidat, and J. Liang, “Performance analysis of relay selection with feedback delay and channel estimation errors,” IEEE Signal Process. Lett., vol. 18, no. 1, pp. 67–70, Jan. 2011.
[17] M. Seyf, S. Muhaidat, J. Liang, and M. Dianati, “Effect of feedback delay on the performance of cooperative networks with relay selection,” IEEE Trans. Wireless Commun., vol. 10, no. 12, pp. 4161–4171, Dec. 2011.
[18] S. A. Banani and R. G. Vaughan, “Capacity maximisation in eigen-multiple-input multiple-output using adaptive modulation and Reed–Solomon coding,” IET Commun., vol. 6, no. 15, pp. 2413–2424, Oct. 2012.
[19] B. Choi and L. Hanzo, “Adaptive WHIT aigned QAM for fading channels subjected to impulsive noise,” IEEE Commun. Lett., vol. 17, no. 7, pp. 1317–1320, Jul. 2013.
[20] D. A. Basnayaka and H. Haas, “A new degree of freedom for energy efficiency of digital communication systems,” IEEE Trans. Commun., vol. 65, no. 7, pp. 3023–3036, Jul. 2017.
[21] W. M. Jang, “A simple upper bound of the Gaussian Q-function with closed-form error bound,” IEEE Commun. Lett., vol. 15, no. 2, pp. 157–159, Feb. 2011.