ON THE VON MISES APPROXIMATION FOR THE DISTRIBUTION OF THE PHASE ANGLE BETWEEN TWO INDEPENDENT COMPLEX GAUSSIAN VECTORS

Nick Letzepis

Defence Science and Technology Organisation
West Avenue, Edinburgh SA 5111
Email: `nick.letzepis@ieee.org`

ABSTRACT

This paper analyses the von Mises approximation for the distribution of the phase angle between two independent complex Gaussian vectors. By upper bounding the Kullback-Leibler divergence, it is shown that when their circular means and variances coincide, the distribution converges to a von Mises distribution both in the low and high signal-to-noise ratio regimes.

Index Terms— von Mises, Tikhonov, Kullback-Leibler, differential phase shift keying.

1. INTRODUCTION

The distribution of the phase angle between two complex Gaussian vectors is of particular importance in communication systems that employ differential modulation schemes [1]. The distribution has been derived under various assumptions by a number of researchers. One of the earliest works is that of Tsvetnov [2], who derived the distribution for the case when both vectors have the same signal-to-noise ratio (SNR). Tsvetnov later extended this result to when they have different SNRs [3]. Fleck and Trabka [4] also derive the equal-SNR case distribution for analysis of the uncoded *symbol error probability* (SEP) of *differential phase shift keying* (DPSK). Pawula *et al.* [5] derived the distribution for both equal and non-equal SNR cases as well as when the Gaussian vectors are correlated. Later, Pawula [6] restates these results such that the integrand contains only exponential functions. In all of the aforementioned works, the *probability density function* (PDF) and *cumulative density function* (CDF) are only expressible in an integral form.

The *von Mises* (also known as the *Tikhonov*) distribution is well known for its use in directional statistics [7]. In [8], Shmaliy proposed using the von Mises distribution to approximate the distribution of the phase angle of a single complex Gaussian random variable. To find the best fit in a *least mean squared error* (LMSE) sense, Shmaliy proposed modelling the concentration parameter of a von Mises distribution as a function of the SNR. The distribution of the angle between two Gaussian vectors is then approximated by that of the difference of two independent von Mises random variables with parameters optimised to minimise the LMSE.

In this paper, the validity of the von Mises PDF approximation for the phase of a complex Gaussian vector is analysed using the *Kullback-Leibler* (KL) divergence (or relative entropy) [9]. The KL divergence is a non-symmetric, information theoretic measure of the difference between two PDFs of the same support. An upper bound on the KL divergence between the PDF of the phase of a complex Gaussian vector and a von Mises PDF is derived. Asymptotically in the high SNR regime, when their circular means and variances coincide, it is shown that the upper bound is inversely proportional to the SNR. In the low SNR regime, the upper bound is proportional to the square root of the SNR. These results show that, asymptotically in both the high and low SNR regimes, the actual PDF converges to a von Mises PDF with the same circular mean and variance.

Whilst the PDF of the (wrapped) sum of two independent von Mises distributed random variables is known in closed form, it can also be approximated by a von Mises PDF [7]. To investigate this further, an upper bound on their KL divergence is derived. When their circular mean and variances coincide, and their concentration parameters increase, the upper bound is shown asymptotically to be proportional to the sum of their reciprocal concentration parameters. Similarly, as their concentration parameters decrease the upper bound is shown to be proportional to the square of the product of their concentration parameters. Thus, asymptotically in these two regimes, the PDF of the sum of two independent von Mises random variables converges to a von Mises PDF with the same circular mean and variance. This implies the PDF of the phase difference between two independent Gaussian random vectors also converges to a von Mises PDF.

2. PRELIMINARIES

Let $X \sim \mathcal{CN}\left(\sqrt{\gamma}e^{\jmath\theta}, 1\right)$ denote a complex circularly symmetric random variable with mean $\sqrt{\gamma}e^{\jmath\theta}$ and unit variance, where γ denotes the SNR. Furthermore, let $P = |X|$ and $\Phi = \arctan(\Im\{X\}, \Re\{X\})$ denote the magnitude and phase of X, where $\arctan(b, a)$ is the two argument arctangent function. The joint and marginal PDFs were derived by Bennett [10],

$$f_{P,\Phi}(\rho, \phi; \gamma, \theta) = \frac{\rho}{\pi} e^{-\left(\rho^2+\gamma-2\sqrt{\gamma}\rho\cos(\phi-\theta)\right)} \tag{1}$$

$$f_P(\rho; \gamma) = 2\rho e^{-\left(\rho^2+\gamma\right)} I_0\left(2\sqrt{\gamma}\rho\right) \tag{2}$$

$$f_\Phi(\phi; \gamma, \theta) = \frac{e^{-\gamma}}{2\pi} + \frac{1}{2}\sqrt{\frac{\gamma}{\pi}} \cos(\phi-\theta) e^{-\gamma \sin^2(\phi-\theta)} \times \left[1 + \operatorname{erf}\left(\sqrt{\gamma}\cos(\phi-\theta)\right)\right], \tag{3}$$

where $I_\nu(z)$ denotes the modified Bessel function of the first kind [11, p.374], and $\operatorname{erf}(z)$ denotes the error function [11, p.228].

Let $V \sim \mathcal{VM}(\theta, \kappa)$ denote a von Mises distributed random variable with mean θ and concentration κ. Its PDF, circular variance, and entropy are given by [7],

$$f_V(\phi; \theta, \kappa) = \frac{e^{\kappa\cos(\phi-\theta)}}{2\pi I_0(\kappa)} \tag{4}$$

$$\varsigma_V(\kappa) = 1 - \mathbb{E}\left[\cos(V)\right] = 1 - \frac{I_1(\kappa)}{I_0(\kappa)} \tag{5}$$

$$h_V(\kappa) = \log(2\pi I_0(\kappa)) - \kappa\left(1 - \varsigma_V(\kappa)\right), \tag{6}$$

respectively. In [12], Shmaliy showed that, conditioned on P, the distribution of Φ is a von Mises distribution, i.e.

$$f_{\Phi|P}(\phi|\rho; \gamma, \theta) = f_V\left(\phi; \theta, 2\sqrt{\gamma}\rho\right). \tag{7}$$

Thus (7) heuristically implies that as the SNR increases, the density of P becomes more concentrated about its mean value, and therefore f_Φ will approach a von Mises density. On the other hand, as SNR decreases, the density of P becomes dispersed from its mean, and f_Φ will approach a uniform density, also a special case of the von Mises density.

Now let Φ_1 and Φ_2 be two independent random variables with densities $f_\Phi(\phi; \gamma_1, \theta_1)$ and $f_\Phi(\phi; \gamma_2, \theta_2)$ respectively as given in (3). Furthermore, let $\Psi = (\Phi_2 - \Phi_1)_{2\pi}$ denote the wrapped difference of Φ_2 and Φ_1. The PDF of Ψ is given by [2–6]

$$f_\Psi(\psi; \gamma_1, \gamma_2, \Delta\theta) = \frac{1}{4\pi}\int_{-\pi/2}^{\pi/2} e^{-S(t,\psi-\Delta\theta)} \times \left[1 + 2\Upsilon - S(t, \psi - \Delta\theta)\right]\cos t\, dt, \tag{8}$$

where $\Delta\theta = (\theta_2 - \theta_1)_{2\pi}$ is the phase difference, $S(t, \psi) = \Upsilon - \Lambda \sin t - \Omega \cos\psi \cos t$, $\Upsilon = \frac{\gamma_2+\gamma_1}{2}$, $\Lambda = \frac{\gamma_2-\gamma_1}{2}$, and $\Omega = \sqrt{\gamma_1\gamma_2}$. Similar to (7), it is not difficult to show that [13]

$$f_{\Psi|P_1,P_2}(\psi|\rho_1, \rho_2; \gamma_1, \gamma_2, \Delta\theta) = f_W(\psi; \Delta\theta, 2\sqrt{\gamma_1}\rho_1, 2\sqrt{\gamma_2}\rho_2) \tag{9}$$

where [7, p.44],

$$f_W(w; \Delta\theta, \kappa_1, \kappa_2) = \frac{1}{2\pi I_0(\kappa_1) I_0(\kappa_2)} \times I_0\left(\sqrt{\kappa_1^2 + \kappa_2^2 + 2\kappa_1\kappa_2\cos(w - \Delta\theta)}\right), \tag{10}$$

is the PDF of the sum of two indepenent von Mises distributed random variables. In [8] Shmaliy proposes approximating (8) by (10) using heuristically chosen functions for κ_1 and κ_2 that are then optimised to minimise the LMSE.

3. THE VON MISES APPROXIMATION OF F_Φ

The conditional PDF (7) highlights an important connection between the von Mises distribution and f_Φ. This result suggests that there is some further underlying analytical reasoning for approximating f_Φ by a von Mises density, which is a much simpler and convenient density to work with. The natural question that arises is: how "good" does this approximation match the true density? To answer this question, the KL divergence is used as a measure of how well the approximation matches the true density. The KL divergence between the PDFs of two random variables X and Y, both defined on the same support $\mathcal{S}$ is [9]

$$D_{\mathrm{KL}}(X \parallel Y) \triangleq H_{\mathrm{cross}}(X, Y) - H(X), \tag{11}$$

where $H_{\mathrm{cross}}(X, Y) \triangleq -\int_{\mathcal{S}} f_X(x)\log f_Y(x)\,dx$ and $H(X) \triangleq -\int_{\mathcal{S}} f_X(x)\log f_X(x)\,dx$ are the cross entropy between X and Y, and the entropy of X respectively. Given the distribution of Φ in (3), a closed form expression for its KL divergence from the von Mises density in (4) is intractable. Instead, one must resort to bounding approaches. Since the KL divergence is an asymmetric measure, one may either consider bounding $D_{\mathrm{KL}}(\Phi \parallel V)$ or $D_{\mathrm{KL}}(V \parallel \Phi)$. It turns out that bounding the former yields a much simpler approach resulting in a closed form expression. To this end, consider the following results.

Lemma 1. *Let Φ and P denote two random variables with joint PDF given by* (1). *Their joint entropy is given by*

$$H(\Phi, P) = 1 + \log(\pi) - \tfrac{1}{2}\log(\gamma) - \tfrac{1}{2}\,\mathrm{E}_1(\gamma) \tag{12}$$

where $\mathrm{E}_1(z)$ denotes the exponential integral [11, p.228].

Proof. The proof follows via [14, Lemma 10.1] and [15, eq. 2.15.5.4]. □

Lemma 2. *Define Φ and P as in Lemma 1. The entropy of P is upper bounded according to*

$$H(P) \leq 1 + 2\gamma - \log(2\sqrt{\gamma}) - \tfrac{1}{2}\,\mathrm{E}_1(\gamma) - \log I_0(2\sqrt{\gamma}\mu_P(\gamma)) \tag{13}$$

where

$$\mu_P(\gamma) = \frac{\sqrt{\pi}}{2} e^{-\frac{\gamma}{2}} I_0\left(\frac{\gamma}{2}\right) + \sqrt{\gamma}\left(1 - \varsigma_\Phi(\gamma)\right), \tag{14}$$

is the mean value of P, and

$$\varsigma_\Phi(\gamma) = 1 - \frac{1}{2}\sqrt{\pi\gamma} e^{-\frac{\gamma}{2}} \left[I_0\left(\frac{\gamma}{2}\right) + I_1\left(\frac{\gamma}{2}\right)\right]. \tag{15}$$

is the circular variance of Φ.

Proof. The proof follows via the convexity of $\log I_0(x)$ and Jensen's inequality [9] and application of [15, eq. 2.15.5.4]. □

Lemma 3. *Define Φ and P as in Lemma 1. The entropy of Φ is lower bounded according to*

$$H(\Phi) \geq \log(2\pi I_0(2\sqrt{\gamma}\mu_P(\gamma))) - 2\gamma. \tag{16}$$

Proof. The proof follows using the property that conditioning reduces entropy [9], i.e. $H(\Phi) \geq H(\Phi|P) = H(\Phi, P) - H(P)$, and Lemmas 1 and 2. □

Lemma 4. *Define Φ and P as in Lemma 1, and $V \sim \mathcal{VM}(\theta, \kappa)$. Then,*

$$H_{\text{cross}}(\Phi, P) = \log(2\pi I_0(\kappa)) - \kappa\left(1 - \varsigma_\Phi(\gamma)\right). \tag{17}$$

Proof. The proof follows via (6) and (7). □

Theorem 1. *Define P, Φ and V as given in Lemma 4. The Kullback-Leibler divergence is upper bounded according to*

$$\begin{aligned} D_{\text{KL}}(\Phi \parallel V) \leq D_{\text{ub}}(\gamma; \kappa) &\triangleq \log I_0(\kappa) \\ &- \log I_0\left[2\sqrt{\gamma}\mu_P(\gamma)\right] + 2\gamma - \kappa\left[1 - \varsigma_\Phi(\gamma)\right]. \end{aligned} \tag{18}$$

Proof. The proof follows via (11), and Lemmas 3 and 4. □

Since the von Mises density is the maximum entropy distribution under circular mean and variance constraints [7], the KL divergence between any wrapped distribution and a von Mises density is minimised when their circular means and variances coincide, i.e. the solution to $\varsigma_V(\kappa) = \varsigma_\Phi(\gamma)$. This is also evident upon closer inspection of (17). To examine this further, consider the following corollary.

Corollary 1. *For large γ,*

$$\varsigma_\Phi(\gamma) - \varsigma_V(2\gamma) = \frac{1}{16\gamma^2} + \mathcal{O}(\gamma^{-3}), \tag{19}$$

and for small γ,

$$\varsigma_\Phi(\gamma) - \varsigma_V(\sqrt{\pi\gamma}) = \left[\frac{1}{8\pi} - \frac{1}{16}\right](\pi\gamma)^{\frac{3}{2}} + \mathcal{O}((\pi\gamma)^{\frac{5}{2}}). \tag{20}$$

Proof. These results follow via application of [11, eq. 9.6.10 and eq. 9.7.1], and [7, eq. 3.5.32 and eq. 3.5.34]. □

Thus, Corollary 1 implies that at low SNR, the optimal κ value that minimises the KL divergence converges to $\sqrt{\pi\gamma}$, whereas for high SNR it converges to 2γ. Moveover, asymptotically, at these optimal κ values, the upper bound behaves as follows.

Corollary 2. *For large γ,*

$$D_{\text{ub}}(\gamma; 2\gamma) = \frac{1}{4\gamma} + \mathcal{O}(\gamma^{-2}). \tag{21}$$

For small γ,

$$D_{\text{ub}}(\gamma; \sqrt{\pi\gamma}) = (2 - \tfrac{\pi}{2})\gamma + \mathcal{O}(\gamma^2). \tag{22}$$

Hence, the asymptotic results of Corollary 2 imply that: as $\gamma \to \infty$, the density of Φ approaches a von Mises density with mean θ and concentration 2γ; and on the other hand, as $\gamma \to 0$, the density of Φ also approaches a von Mises density with mean θ and concentration $\sqrt{\pi\gamma}$.

4. THE VON MISES APPROXIMATION OF F_Ψ

Whilst the sum of two independent von Mises distributed random variables is known in closed form (see (10)) it has also been shown empirically that it can be approximated by a von Mises PDF [7]. Since the PDF of Φ can be approximated by a von Mises PDF then it seems reasonable that f_Ψ can also be approximated by a von Mises PDF. To investigate this further, consider the following results.

Theorem 2. *Let random variables V and W be distributed according to* (4) *and* (10) *respectively. Their KL divergence is upper bounded by*

$$\begin{aligned} D_{\text{KL}}(W \parallel V) \leq D_{\text{ub},2}(\kappa_1, \kappa_2; \kappa) &\triangleq \log(I_0(\kappa)) \\ &- \kappa\left[1 - \varsigma_W(\kappa_1, \kappa_2)\right] - \log(I_0(\kappa_1) I_0(\kappa_2)) \\ &+ \log I_0\left[\sqrt{\kappa_1^2 + \kappa_2^2 + 2\kappa_1\kappa_2(1 - \varsigma_W(\kappa_1, \kappa_2))}\right], \end{aligned} \tag{23}$$

where

$$\varsigma_W(\kappa_1, \kappa_2) = \varsigma_V(\kappa_1) + \varsigma_V(\kappa_2) - \varsigma_V(\kappa_1)\varsigma_V(\kappa_2) \tag{24}$$

is the circular variance of W.

Proof. Following the same steps as the proof of Lemma 4,

$$H_{\text{cross}}(W, V) = \log(2\pi I_0(\kappa)) - \kappa\left[1 - \varsigma_W(\kappa_1, \kappa_2)\right]. \tag{25}$$

Since $\log(I_0(\sqrt{x}))$ is a concave function, then using Jensen's inequality,

$$\begin{aligned} H(W) &\geq \log\left(2\pi I_0(\kappa_1) I_0(\kappa_2)\right) \\ &- \log I_0\left(\sqrt{\kappa_1^2 + \kappa_2^2 + 2\kappa_1\kappa_2(1 - \varsigma_W(\kappa_1, \kappa_2))}\right). \end{aligned} \tag{26}$$

Subtracting (26) from (25) results in (23) as stated in the theorem. □

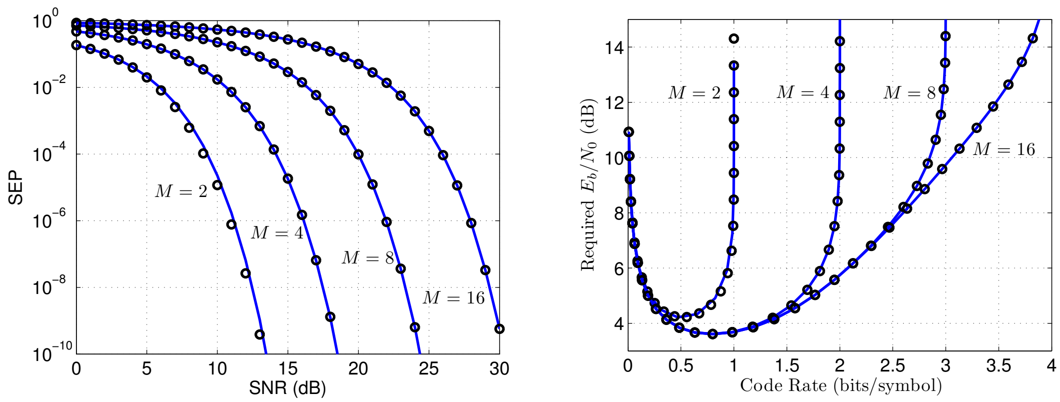

Fig. 1. Uncoded SEP (left) and the achievable rate for M-ary DPSK (right). Solid lines show true values, circles show computation using approximation (31).

As in the case for the PDF of Φ, since the von Mises distribution is the maximum entropy distribution, the KL divergence is minimised when $\varsigma_V(\kappa) = \varsigma_W(\kappa_1, \kappa_2)$, which is also clear by inspection of (25).

Corollary 3. *Let W be distributed according to* (10) *with $\kappa_1 = \alpha\kappa_2$, $0 < \alpha \le 1$, i.e. $\kappa_1 \le \kappa_2$. For large κ_1,*

$$\varsigma_V\left(\tfrac{\kappa_1}{1+\alpha}\right) - \varsigma_W\left(\kappa_1, \tfrac{\kappa_1}{\alpha}\right) = \frac{\alpha}{2\kappa_1^2} + \mathcal{O}(\kappa_1^{-3}). \tag{27}$$

For small κ_1,

$$\varsigma_V\left(\tfrac{\kappa_1^2}{2\alpha}\right) - \varsigma_W\left(\kappa_1, \tfrac{\kappa_1}{\alpha}\right) = -\left(\frac{1-\alpha^2}{48\alpha^3}\right)\kappa_1^4 + \mathcal{O}(\kappa_1^6). \tag{28}$$

In other words, Corollary 3 implies that as κ_1 and κ_2 grow large (with a fixed ratio α), the circular variance of W approaches that of a von Mises density with concentration $\kappa_1\kappa_2/(\kappa_1 + \kappa_2)$. On the other hand as κ_1 and κ_2 become small, the circular variance of W approaches that of a von Mises density with concentration $\kappa_1\kappa_2/2$.

Corollary 4. *Define W as in Corollary 3. Then for large κ_1*

$$D_{\text{ub},2}\left(\kappa_1, \tfrac{\kappa_1}{\alpha}; \tfrac{\kappa_1}{\alpha+1}\right) = \frac{\alpha}{4(\alpha+1)\kappa_1} + \mathcal{O}(\kappa_1^{-2}). \tag{29}$$

For small κ_1

$$D_{\text{ub},2}\left(\kappa_1, \tfrac{\kappa_1}{\alpha}; \tfrac{\kappa_1^2}{2\alpha}\right) = \frac{\kappa_1^4}{32\alpha^2} + \mathcal{O}(\kappa_1^6) \tag{30}$$

Corollary 4 implies a certain asymmetry between the asymptotic large and small κ_1 regimes. In particular, as $\kappa_1 \to 0$ the bound is proportional to κ_1^4, whereas as $\kappa_1 \to \infty$ the bound is proportional to κ_1^{-1}. This asymmetry was also observed in the actual KL divergence, albeit with even larger exponents than what the upper bound suggests.

Corollary 4 proves that in the asymptotic large and small limit of κ_1 and κ_2, the PDF of W converges to a von Mises density. Moreover, since in the previous section it was established that the PDF of Φ converges to a von Mises density, then one may also conclude that the PDF of Ψ also converges to a von Mises density. More precisely,

$$f_\Psi(\psi; \Delta\theta, \gamma_1, \gamma_2) \approx f_V(\psi; \Delta\theta, \kappa^*(\gamma_1, \gamma_2)), \tag{31}$$

where $\kappa^*(\gamma_1, \gamma_2)$ is the solution to

$$\varsigma_V(\kappa) = \varsigma_\Psi(\gamma_1, \gamma_2) = \varsigma_\Phi(\gamma_1) + \varsigma_\Phi(\gamma_2) - \varsigma_\Phi(\gamma_1)\varsigma_\Phi(\gamma_2). \tag{32}$$

From Corollaries 1 and 3, as $\gamma_1, \gamma_2 \to \infty$, then $\kappa^*(\gamma_1, \gamma_2) \to \frac{2\gamma_1\gamma_2}{\gamma_1+\gamma_2}$, and as $\gamma_1, \gamma_2 \to 0$, then $\kappa^*(\gamma_1, \gamma_2) \to \frac{\pi}{2}\sqrt{\gamma_1\gamma_2}$.

Fig. 1 illustrates a simple application of (31), which shows the uncoded SEP (left) and the capacity of M-ary DPSK (right) using the method of Ungerboeck [16] assuming a phase-difference receiver [17]. The solid lines show these quantities using the actual PDF (8), and the circle markers show their computation using approximation (31). As expected, it can be seen that the approximation is virtually indistingishable from the true value.

5. CONCLUSION

In this paper, using the KL divergence, it was shown that the PDF of the phase angle of a Gaussian vector, and the PDF of the phase angle between two independent Gaussian vectors converges to a von Mises PDF in both the asymptotic high and low SNR regimes. The latter case is of particular importance as it allows its PDF, not expressible in closed form, to be approximated by the more convenient von Mises PDF.

6. REFERENCES

[1] J. G. Proakis and M. Salehi, *Digital Communications*, McGraw Hill, 5th edition, 2008.

[2] V. V. Tsvetnov, "Statistical properties of signals and noises in two-channel phase systems," *Radiotekhnika*, vol. 12, no. 5, pp. 12–29, May 1957.

[3] V. V. Tsvetnov, "Unconditional statistical characteristics of signals and uncorrelated Gaussian noises in two-channel phase systems," *Radiotekh. Elektron.*, vol. 14, no. 12, pp. 2147–2159, 1969.

[4] J. T. Fleck and E. A. Trabka, "Error probabilities of multiple-state differentially coherent phase-shift keyed systems in the presence of white, Gaussian noise," in *Investigation of Digital Data Communication Systems, Rep. UA-1420-S-1*, Buffalo, NY, Jan. 1961, Cornell Aeronaut. Lab., Inc., Detect Memo 2A, available as NTIS Doc. AD256584.

[5] R. F. Pawula, S. O. Rice, and J. H. Roberts, "Distribution of the phase angle between two vectors perturbed by Gaussian noise," *IEEE Trans. on Commun.*, vol. COM-30, no. 8, pp. 1828–1841, Aug. 1982.

[6] R. F. Pawula, "Distribution of the phase angle between two vectors perturbed by Gaussian noise II," *IEEE Trans. Vehicular Technol.*, vol. 50, no. 2, pp. 576–583, Mar 2001.

[7] K. V. Mardia and P. E. Jupp, *Directional Statistics*, Wiley, 1999.

[8] Y. S. Shmaliy, "von Mises/Tikhonov-base distributions for systems with differential phase measurement," *Signal Processing*, vol. 85, pp. 693–703, 2005.

[9] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley, 2nd edition, 2006.

[10] W. R. Bennet, "Methods of solving noise problems," *Proc. of the IRE*, vol. 44, no. 5, pp. 609–638, May 1956.

[11] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 55. National Bureau of Standats, Applied Mathematics Series, 1972.

[12] Y. S. Shmaliy, "On the multivariate conditional probability density of a vector perturbed by Gaussian noise," *IEEE Trans. on Info. Theory*, vol. 53, no. 12, pp. 4792–4797, Dec. 2007.

[13] N. O'Donoughue and J. M. Moura, "On the product of independent complex Gaussians," *IEEE Trans. on Sig. Proc.*, vol. 60, no. 3, pp. 1050–1063, Mar. 2012.

[14] A. Lapidoth and S. M. Moser, "Capacity bounds via duality with applications to multiple-antenna systems on flat-fading channels," *IEEE Trans. on Info. Theory*, vol. 49, no. 10, pp. 2426–2467, Oct. 2003.

[15] A. P. Prudnikov, Yu. A. Brychkov, and O. I. Marichev, *Integrals and Series, Volume 2: Special Functions*, Gordon and Breach Science Publishers, 1986.

[16] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inf. Theory*, vol. 28, no. 1, Jan. 1982.

[17] G. Kaplan and S. Shamai, "On the achievable information rates of DPSK," *IEE Proceedings-I*, vol. 139, no. 3, Jun. 1991.