method to compute LLRs with low complexity for soft-input channel decoding. Analysis shows that the proposed algorithm can reduce the complexity from $\mathcal{O}(K^3)$ to $\mathcal{O}(K^2)$. It is verified that the proposed algorithm outperforms the conventional method and achieves the near-optimal performance of the classical MMSE algorithm with a small number of iterations. Additionally, the idea of using the GS method to efficiently solve the complicated matrix inversion can be applied to other signal processing problems involving matrix inversion of large size in wireless communications, such as downlink precoding in large-scale MIMO systems.

REFERENCES

[1] B. Dongwoon *et al.*, "LTE-advanced modem design: Challenges and perspectives," *IEEE Commun. Mag.*, vol. 50, no. 2, pp. 178–186, Feb. 2012.
[2] D. Skordoulis *et al.*, "IEEE 802.11n MAC frame aggregation mechanisms for next generation high-throughput WLANs," *IEEE Wireless Commun. Mag.*, vol. 15, no. 2, pp. 40–47, Feb. 2008.
[3] R. W. Health, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.
[4] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
[5] H. Ngo, E. Larsson, and T. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2012.
[6] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.
[7] L. G. Barbero and J. S. Thompson, "Fixing the complexity of the sphere decoder for MIMO detection," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2131–2142, Jun. 2008.
[8] T. Datta, N. Srinidhi, A. Chockalingam, and B. S. Rajan, "Random-restart reactive tabu search algorithm for detection in large-MIMO systems," *IEEE Commun. Lett.*, vol. 14, no. 12, pp. 1107–1109, Dec. 2010.
[9] B. Yin, M. Wu, C. Studer, J. R. Cavallaro, and C. Dick, "Implementation trade-offs for linear detection in large-scale MIMO systems," in *Proc. IEEE ICASSP*, May 2013, pp. 2679–2683.
[10] A. Björck, *Numerical Methods for Least Squares Problems*, Philadelphia, PA, USA: SIAM, 1996.
[11] J. Hoydis, S. T. Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.
[12] L. Dai, Z. Wang, and Z. Yang, "Spectrally efficient time–frequency training OFDM for mobile large-scale MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 251–263, Feb. 2013.
[13] Z. Gao, L. Dai, and Z. Wang, "Structured compressive sensing based superimposed pilot design in downlink large-scale MIMO systems," *Electron. Lett.*, vol. 50, no. 12, pp. 896–898, Jun. 2014.
[14] M. Wu *et al.*, "Large-scale MIMO detection for 3GPP LTE: Algorithms and FPGA implementations," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 916–929, Oct. 2014.
[15] B. Yin, M. Wu, J. R. Cavallaro, and C. Studer, "Conjugate gradient-based soft-output detection and precoding in massive MIMO systems," *arXiv preprint*:1404.0424v1, Apr. 2014.
[16] X. Gao, L. Dai, Y. Ma, and Z. Wang, "Low-complexity near-optimal signal detection for uplink large-scale MIMO systems," *Electron. Lett.*, vol. 50, no. 18, pp. 1326–1328, Aug. 2014.
[17] B. E. Godana and T. Ekman, "Parametrization based limited feedback design for correlated MIMO channels using new statistical models," *IEEE Trans. Wireless Commun.*, vol. 12, no. 11, pp. 5172–5184, Oct. 2013.

On the Accuracy of the High-SNR Approximation of the Differential Entropy of Signals in Additive Gaussian Noise: Real and Complex Cases

Ramy H. Gohary and Halim Yanikomeroglu

***Abstract*—One approach to the analysis of the high signal-to-noise ratio (SNR) capacity of noncoherent wireless communication systems is to ignore the noise component of the received signal in the computation of its differential entropy. In this paper, we consider the error incurred by this approximation when the transmitter and the receiver have one antenna each and when the noise has a Gaussian distribution. We consider the complex and real cases, and we show that when the probability density function (pdf) of the signal component of the received signal is piecewise differentiable, the approximation error decays as 1/SNR, which tightens the available result that the error decays as $o(1)$. In addition, we consider the special instance in which the signal component of the received signal corresponds to a signal transmitted over a channel with a Gaussian fading coefficient. For that case, we provide explicit expressions for the first nonconstant term of the Taylor expansion of the differential entropy, and we invoke Schwartz's inequality to obtain an efficiently computable bound on it. Our results are supported by numerical examples.**

***Index Terms*—Differential entropy, Lebesgue dominated convergence, sum and product of random variables.**

I. INTRODUCTION

The capacity of a given channel provides a fundamental limit on the maximum data rate that can be reliably communicated over that channel. In addition, the derivation and structure of capacity expressions usually yield insight into the optimal signaling strategy. Unfortunately, providing exact expressions for the capacity of various channels constitutes a generally difficult task, and only the capacity of a relatively small number of channel instances is available. Among channels of practical interest are those in which the channel state information is not available at the receiver, which are usually referred to as being noncoherent [1] and those in which the transmitted signals are contaminated with phase noise [2]. These channels arise when the channel undergoes fast fading conditions, resulting in the receiver being unable to acquire reliable channel and phase estimates. As such, these channels are likely to arise more frequently in future high-mobility communication scenarios. Unfortunately, capacity expressions for these channels are not known, and only asymptotic results are available. In this paper, we will focus on the accuracy of the differential entropy approximations that underlie the derivation of the asymptotic results.

Manuscript received June 11, 2014; revised September 18, 2014; accepted October 29, 2014. Date of publication November 3, 2014; date of current version October 13, 2015. An earlier version of this paper was presented at the 2014 IEEE International Conference on Acoustics, Speech, and Signal Processing, Florence, Italy. The review of this paper was coordinated by Dr. I. Krikidis.

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada; by Huawei Canada Company, Ltd.; and by the Ontario Ministry of Economic Development and Innovations through the Ontario Research Fund Research Excellence Program.

The authors are with the Department of Systems and Computer Engineering, Carleton University, Ottawa, ON K1S 5B6, Canada.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2014.2366911

Evaluating the capacity of a given channel involves the search for input distributions that maximize the mutual information between the random variables representing the transmitted signals and the signals observed by the receiver. When these signals assume continuous values, the mutual information is expressed as the difference between the nonconditional differential entropy of the received signals and their differential entropy conditioned on the transmitted signals. In many cases, including noncoherent communication ones [1]–[4], the received signals are contaminated by additive noise, and computing their conditional entropy is straightforward. In contrast, computing the nonconditional counterpart of this entropy is rather difficult.

Computing, let alone maximizing, the nonconditional differential entropy of the received signal is usually difficult because this signal is composed of the sum of random variables of potentially unknown distributions. An exception is the case in [5], wherein a closed-form expression is given for the optimal distribution that maximizes the sum of independent random variables with finite supports and distributions symmetric around zero. In more general situations, bounds are used to gain insight into the distributions that maximize the nonconditional differential entropy. For instance, the entropy power inequality was used in [6] and [7] to obtain a lower bound on the differential entropy of the sum of independent random variables. In a complementary fashion, upper bounds on the entropy of the sum of random variables have been derived under various assumptions. For instance, a tight upper bound on the differential entropy of the sum of two dependent random variables with log-concave distributions was provided in [8]. In [9] and [10], expressions were derived for the difference and ratio of the entropy of the sum and difference of independent identically distributed (i.i.d.) random variables. Upper bounds on the differential entropy of the sum of independent random variables are provided in [11].

The difficulty of computing and maximizing the nonconditional entropy of the received signal arises in the evaluation of the noncoherent capacity of multiple-input–multiple-output (MIMO) communication systems [1], [4] and in systems with phase noise [2]. For noncoherent systems, the high-SNR analysis relies on ignoring the noise contribution to the nonconditional entropy of the received signal and on approximating it with the entropy of the product of the input signal and the channel coefficients. This approach enables an approximate expression for the noncoherent capacity to be obtained and the asymptotically optimal input distribution to be determined.

Given the central role of ignoring the noise contribution in computing the nonconditional differential entropy at high SNRs, in this paper, we investigate the accuracy of this approximation when the received signal is complex, with potentially correlated real and imaginary components. As such, the results reported herein complement those reported in [12] for the real case. The potential correlation between the real and imaginary components of the received signal renders the analysis in [12] inapplicable in the complex case. This is because studying the limiting behavior of the differential entropy in this case requires the joint manipulation of bivariate random variables. To assess how the approximation error decays with SNR in this case, we consider an expression of the differential entropy of the sum of an arbitrary and a zero-mean Gaussian-distributed random variable with variance given by the inverse of the SNR. We derive explicit expressions for the first nonconstant term of the Taylor expansion of the differential entropy of the received signal, and we use these expressions to show that, when the probability density function (pdf) of the signal component of the received signal is continuous, the approximation error decays as 1/SNR. Next, we consider a noncoherent wireless communication system in which the signal component of the received signal is the product of a random variable with an unknown distribution and a Gaussian-distributed random variable representing the channel fading coefficient. For that case, we provide an explicit expression for the first nonconstant term of the Taylor expansion of the differential entropy, and to circumvent the difficulty that arises in computing that term, we invoke Schwartz's inequality to obtain a computable bound on it.

II. Preliminaries

The forthcoming analysis focuses on scenarios with SNRs greater than 0 dB. For those scenarios, the Taylor expansion is used to express the differential entropy of the received signal as a function of the SNR. In particular, SNR is used to define a sequence of pdfs of the received signal and, subsequently, a sequence of differential entropy values. It is generally not known whether the limit of the sequence of differential entropy values corresponding to the sequence of pdfs is equal to the differential entropy of the limiting pdf. However, sufficient conditions under which this holds are provided in the following theorem [13].

Theorem 1: Let $\{Y_i \in \mathbb{C}\}$ be a sequence of continuous random variables with pdfs $\{p_i\}$, and let $Y \in \mathbb{C}$ be a continuous random variable with pdf p such that $p_i \to p$ pointwise. If $\max\{\sup_y p(y), \sup_y(p_i(y)\} < \infty$, and $\int |y|^\kappa p(y)\, dy < \infty$ for some $\kappa > 1$, then the limit of the sequence of differential entropy values is equal to the differential entropy of the limiting pdf.

Proof: See [13, Theorems 2 and 3]. ■

In addition to satisfying the conditions of this theorem, we will assume that, for the considered pdfs, all the derivatives of the differential entropy are finite, and subsequently, their Taylor series expansion in terms of SNR converges. Instances of such pdfs will be considered in Section V.

The fact that, for the considered pdfs, the derivatives of the differential entropy are finite ensures that the conditions of Lebesgue's dominated convergence theorem are satisfied [14], which ensures the exchangeability of the operations of integration and taking limits. In the forthcoming analysis, we will rely on this theorem to move the limit operator inside the integration symbol.

III. Taylor Expansion of the Entropy of the Sum of Arbitrary and Gaussian Distributed Random Variables—The General Complex Case

Let $X = X_r + \jmath X_i$ be a complex random variable with a twice differentiable pdf $p_X(\cdot,\cdot)$ with support $(-\infty,\infty)\times(-\infty,\infty)$, i.e., $p_X(x_r, x_i) > 0$ for all finite x_r and x_i. Let $Z = Z_r + \jmath Z_i$, where Z_r and Z_i are i.i.d. zero-mean unit variance Gaussian random variables, i.e., $p_Z(z_r, z_i) = \frac{1}{2\pi} e^{-\frac{z_r^2+z_i^2}{2}}$.

In a point-to-point communication channel with additive Gaussian noise, the received signal $Y = X + tZ$, where $t = \frac{1}{\sqrt{\text{SNR}}}$. Our goal is to examine the Taylor expansion of the (nonconditional) differential entropy of Y at SNRs greater than 0 dB, i.e., for values of $t \in (0, 1)$. For notational convenience, we will denote this entropy by $h_t(Y)$. For any $t > 0$, the pdf of tZ is given by $\frac{1}{2\pi t^2} e^{-\frac{z_r^2+z_i^2}{2t^2}}$. To obtain the pdf of Y, we consider the cumulative distribution function (cdf), i.e.,

$$F_Y(y_r, y_i) = \Pr\{-\infty \le tZ_r \le \infty, -\infty \le X_r \le y_r - tZ_r \quad -\infty \le tZ_i \le \infty, -\infty \le X_i \le y_i - tZ_i\} \tag{1}$$

$$= \frac{1}{2\pi t^2} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{u^2+v^2}{2t^2}} \times \int_{-\infty}^{y_i - v}\int_{-\infty}^{y_r - u} p_{X_r X_i}(x_r, x_i)\, dx_r\, dx_i\, du\, dv. \tag{2}$$

Differentiating with respect to y_r and y_i yields

$$p_Y(y_r, y_i; t) = \frac{1}{2\pi t^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_X(y_r - u, y_i - v) e^{-\frac{u^2+v^2}{2t^2}} \, du \, dv. \quad (3)$$

Notice that, in this notation, t parametrizes a family of pdfs. Using this notation, we can write

$$h_t(Y) = -\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_Y(y_r, y_i; t) \log p_Y(y_r, y_i; t) \, dy_r \, dy_i. \quad (4)$$

Evaluating $h_t(Y)$ directly seems intractable. However, an accurate high-SNR approximation of this quantity can be obtained if $h_t(Y)$ were analytic in t, i.e., if all its derivatives with respect to t were finite. A class of practical instances in which the distributions $p_X(\cdot,\cdot)$ yield differential entropy values $h_t(Y)$ that satisfy this condition is identified in Section IV. When this condition is satisfied, the Taylor expansion of $h_t(Y)$ around $t = 0$ converges for all $t \in (0, 1)$. To expand $h_t(Y)$, we note that $p_Y(y_r, y_i; 0) = p_X(y_r, y_i)$, and therefore $h_0(Y) = h(X)$. Next, we note that $h_t(Y)$ in the complex case is a function of t^2, not t as in the real case [12]. This observation will yield a result similar to the one obtained in [12] for the real case, although through a different mechanism. For $t < 1$, the Taylor expansion for the complex case can be expressed as

$$h_t(Y) = h(X) + \left.\frac{dh_t(Y)}{dt^2}\right|_{t^2=0} t^2 + \mathcal{O}(t^4). \quad (5)$$

Using (5), we are now ready to provide our main result for the complex case.

Theorem 2: Let X and Y be complex random variables related by $Y = X + tZ$, where $t \in (0, 1)$ and Z is a zero-mean circularly symmetric unit-variance Gaussian random variable. If the pdf of X, $p_X(\cdot,\cdot)$, is twice piecewise differentiable in each argument and the derivatives of $h_t(Y)$ with respect to t^2 are finite, the Taylor expansion converges, and

$$h_t(Y) = h(X) - \frac{t^2}{2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (1 + \log p_X(y_r, y_i)) \times \left(\frac{\partial^2 p_X(y_r, y_i)}{\partial y_r^2} + \frac{\partial^2 p_X(y_r, y_i)}{\partial y_i^2} \right) dy_r \, dy_i + \mathcal{O}(t^4). \quad (6)$$

Proof: See Appendix A. ■

In Section V, we will provide an example that illustrates the utility of this theorem. However, computing the coefficient of t^2 in (6) is generally difficult. A special case that arises in many practical scenarios is when X is circularly symmetric with independent real and imaginary components. In that case, $h_t(Y) = 2h_t(Y_r) = 2h_t(Y_i)$ and, when $p_X(\cdot)$ is piecewise differentiable, the computation of $h_t(Y)$ reduces to the one corresponding to the real case [12], i.e.,

$$h_t(Y_r) = h(X_r) + \frac{t^2}{2} \left(\int_{-\infty}^{\infty} \frac{1}{p_X(y_r)} \left(\frac{dp_X(y_r)}{dy_r} \right)^2 dy_r - (1 + \log p_X(y_r)) \left.\frac{dp_X(y_r)}{dy}\right|_{-\infty}^{\infty} \right) + \mathcal{O}(t^4). \quad (7)$$

For simplicity, we will henceforth focus on the case in which X is circularly symmetric with independent real and imaginary components, and we will use (7) to derive bounds on the error arising from approximating $h_t(Y)$ with $h(X)$. The following remarks are in order.

Remark 1: Notice that the analysis in [12] implies that, for the coefficient of t to be zero and (7) to be a valid Taylor expansion of $h_t(Y)$, $p_X(\cdot)$ must be piecewise differentiable. □

Remark 2: Although the results presented herein pertain to single-input–single-output (SISO) systems, for MIMO systems with full-rank channel matrices and at least as many transmit antennas as receive ones, the approximation error can be easily seen to decay as 1/SNR. However, computing the analogous of (6) and (7) for the MIMO case appears to be intractable. □

IV. Real Single-Input–Single-Output and Multiple-Input–Single-Output Cases With $X = RS$

A. SISO Case

A special case of interest arises in noncoherent communication scenarios when $X = RS$, where R and S are two independent random variables, and R represents the input signal and S represents the channel. To study this case, we use the following result.

Lemma 1: Let $X = RS$, where R and S are random variables with a joint pdf $p_{R,S}(\cdot,\cdot)$. Then

$$p_X(x) = \int_0^{\infty} \frac{1}{\rho} \left(p_{R,S}\left(\rho, \frac{x}{\rho}\right) + p_{R,S}\left(-\rho, -\frac{x}{\rho}\right) \right) d\rho. \quad (8)$$

In particular, if S is statistically independent of R and is Gaussian distributed with zero mean and unit variance, letting $\bar{p}_R(\rho) = p_R(\rho) + p_R(-\rho)$ yields

$$p_X(x) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \frac{1}{\rho} \bar{p}_R(\rho) e^{-\frac{x^2}{2\rho^2}} \, d\rho. \quad (9)$$

Proof: This lemma can be proved by considering the cdf of X, $F_X(x) = \Pr\{RS \leq x\}$. Expressing this probability in terms of $p_{R,S}(\cdot,\cdot)$ and taking the derivative with respect to ρ yields (8) (see, e.g. [15]). ■

Immediate from (9) is that, when X is the product of two independent random variables, one of which is Gaussian, the support of $p_X(\cdot)$ is the entire real line $(-\infty, \infty)$, and $p_X(\pm\infty) = 0$. Differentiating $p_X(x)$ in (9) with respect to x and using Lebesgue's dominated convergence theorem, it can be verified that $\lim_{x\to\pm\infty} \frac{dp_X(x)}{dx} = 0$, implying that (7) can be expressed as

$$h_t(Y) = h(X) + \frac{t^2}{2} \int_{-\infty}^{\infty} \frac{1}{p_X(y)} \left(\frac{dp_X(y)}{dy} \right)^2 dy + \mathcal{O}(t^4). \quad (10)$$

Remark 3: For piecewise differentiable distributions with $(1 + \log p_X(y_r)) \left.\frac{dp_X(y_r)}{dy}\right|_{-\infty}^{\infty} = 0$, (7) asserts that, at sufficiently high SNRs, additive Gaussian noise strictly increases entropy. □

To bound the error in approximating $h_t(Y)$ with $h(X)$ at high SNRs, we need to bound $\int_{-\infty}^{\infty} \frac{1}{p_X(y)} \frac{dp_X(y)}{dy}^2 dy$. When $p_X(x)$ is given by the expression in (9), the integrand

$$\frac{1}{p_X(y)} \left(\frac{dp_X(y)}{dy} \right)^2 = \frac{y^2}{\sqrt{2\pi}} \frac{\left(\int_0^{\infty} \frac{1}{\rho^3} \bar{p}_R(\rho) e^{-\frac{y^2}{2\rho^2}} \, d\rho \right)^2}{\int_0^{\infty} \frac{1}{\rho} \bar{p}_R(\rho) e^{-\frac{y^2}{2\rho^2}} \, d\rho}. \quad (11)$$

Using Schwartz's inequality, the right-hand side of (11) can be bounded, thereby yielding

$$\frac{1}{p_X(y)}\left(\frac{dp_X(y)}{dy}\right)^2 \leq \frac{y^2}{\sqrt{2\pi}}\int_0^\infty \frac{1}{\rho^5}\bar{p}_R(\rho)e^{-\frac{y^2}{2\rho^2}}\,d\rho \tag{12}$$

and equality holds if ρ assumes a deterministic value in $[0,\infty)$. Substituting this result, the coefficient of $t^2/2$ in (10) can be readily seen to be bounded by

$$\int_{-\infty}^{\infty}\frac{y^2}{\sqrt{2\pi}}\int_0^\infty \frac{1}{\rho^5}\bar{p}_R(\rho)e^{-\frac{y^2}{2\rho^2}}\,d\rho\,dy = \int_0^\infty \frac{1}{\rho^2}\bar{p}_R(\rho)\,d\rho. \tag{13}$$

In the following, we will relate the results obtained here to the multiple-input–single-output (MISO) case. Unfortunately, an analogous study for the MIMO case seems intractable.

B. MISO Case

For ease of exposition, here, we will derive a bound analogous to the one in (13) for the case of two transmit and one receive antenna. However, the method outlined below is readily extensible to systems with more than two transmit antennas. In the current case, the random variable $X = R_1S_1 + R_2S_2$, where R_i and S_i, $i, 1, 2$, represent the input signal and the Gaussian-distributed channel coefficient of the ith antenna, $i, 1, 2$, respectively. The pdf of X in that case is given by the convolution integral of two pdfs of the form in (9). In particular

$$p_X(x) = \frac{1}{2\pi}\int_{u=-\infty}^{\infty}\int_{\rho_1=0}^{\infty}\int_{\rho_2=0}^{\infty}\frac{1}{\rho_1\rho_2}\bar{p}_{R_1}(\rho_1)\bar{p}_{R_2}(\rho_2) \times e^{-\frac{u^2}{2\rho_1^2}}e^{-\frac{(x-u)^2}{2\rho_2^2}}\,d\rho_2\,d\rho_1\,du. \tag{14}$$

Applying Schwartz's inequality, the analogue of (12) for this case is given by

$$\frac{1}{p_X(y)}\left(\frac{dp_X(y)}{dy}\right)^2 \leq \int_{u=-\infty}^{\infty}\int_{\rho_1=0}^{\infty}\int_{\rho_2=0}^{\infty}\frac{(y-u)^2}{2\pi\rho_1\rho_2^5} \times \bar{p}_{R_1}(\rho_1)\bar{p}_{R_2}(\rho_2)e^{-\frac{u^2}{2\rho_1^2}-\frac{(y-u)^2}{2\rho_2^2}}\,d\rho_2\,d\rho_1\,du. \tag{15}$$

Integrating over y, u, and ρ_1 yields that the coefficient of $t^2/2$ in (10) is bounded by $\int_0^\infty \frac{1}{\rho_2^2}\bar{p}_{R_2}(\rho_2)d\rho_2$. Since convolution is commutative, the role of ρ_1 and ρ_2 can be exchanged, thereby yielding

$$\int_{y=-\infty}^{\infty}\frac{1}{p_X(y)}\left(\frac{dp_X(y)}{dy}\right)^2 dy \leq \min_{i=1,2}\left\{\int_0^\infty \frac{1}{\rho_i^2}\bar{p}_{R_i}(\rho_i)\,d\rho_i\right\}. \tag{16}$$

V. Examples

We provide examples to illustrate the error arising from the high-SNR approximation of $h_t(Y)$.

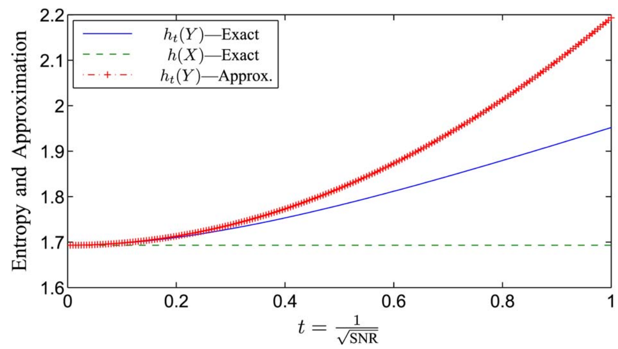

Fig. 1. Quadratic convergence of $h_t(Y)$.

Example 1 (Real SISO Case): Let the input signal R be antipodal, i.e., $R = \pm 1$ with equal probability. In this case, $p_R(\rho) = \frac{1}{2}(\delta(\rho+1)+\delta(\rho-1))$, which yields

$$|h_t(Y) - h(X)| \leq \frac{t^2}{2} + \mathcal{O}(t^4). \tag{17}$$

In this case, the Schwartz's inequality bound is tight. □

Example 2 (Complex SISO Case): In this example, we consider the case in which X is a complex zero-mean Gaussian-distributed random variable, with correlated real and imaginary components with correlation coefficient $\rho \in [0,1]$ and variances σ_r^2 and σ_i^2, respectively, i.e., $p_X(x_r, x_i) = \frac{1}{2\pi\sigma_r\sigma_i\sqrt{1-\rho^2}}e^{-\frac{1}{2(1-\rho^2)}\left(\frac{x_r^2}{2\sigma_r^2}-\frac{2\rho x_r x_i}{\sigma_r\sigma_i}+\frac{x_i^2}{2\sigma_i^2}\right)}$. In this case, after a tedious, but straightforward, computation, the coefficient of $t^2/2$ in (6) can be verified to be

$$\frac{1-\rho^2\log\left(2\pi\sigma_r\sigma_i\sqrt{1-\rho^2}\right)}{(1-\rho^2)^3}\left(\frac{1}{\sigma_r^2}+\frac{1}{\sigma_i^2}\right). \tag{18}$$

In the special case in which the real and imaginary components of X are independent, $\rho = 0$ and the coefficient of $t^2/2$ in (18) simplifies to $\frac{1}{\sigma_r^2}+\frac{1}{\sigma_i^2}$. The validity of this result can be verified by writing the Taylor expansion of $h_t(Y) = \frac{1}{2}\log 2\pi e(\sigma_r^2+t^2) + \frac{1}{2}\log 2\pi e(\sigma_i^2+t^2)$. In addition, note that, for $\sigma_1^2+\sigma_2^2 = 1$, the minimum of $\frac{1}{\sigma_r^2}+\frac{1}{\sigma_i^2}$ is 4 and is achieved when $\sigma_1 = \sigma_2 = \frac{1}{\sqrt{2}}$. □

Example 3 (Convergence With SNR): Our convergence analysis of $h_t(Y)$ leading to (7) implies that, when $p_X(\cdot)$ is piecewise differentiable, $h_t(Y) - h(X)$ tends to zero as t^2, i.e., as 1/SNR. To illustrate the convergence behavior, we consider the case in which X is a real random variable with standard Laplace distribution, i.e., $p_X(x) = \frac{1}{2}e^{-|x|}$. In this case, the second term in the coefficient of $t^2/2$ in (7) is zero, and the first term can be readily evaluated yielding $h_t(Y) = h(X) + t^2/2 + \mathcal{O}(t^4)$. The differential entropy of X is given by $h(X) = 1 + \log(2)$, and the pdf of Y is given by $p_Y(y;t) = \frac{1}{2}e^{-t^2/2}(e^yQ(\frac{y}{t}+t) + e^{-y}Q(-\frac{y}{t}+t))$, where $Q(\cdot)$ is the standard Q-function. In Fig. 1, the parabola $h(X) + t^2/2$ is plotted, and numerical integration is used to evaluate and plot $h_t(Y)$ versus t. For comparison $h(X)$ is also plotted. The quadratic convergence of $h_t(Y)$ to $h(x)$ as t goes to zero can be readily observed from this figure. □

Example 4 (Real MISO Case): If R_1 and R_2 are χ-distributed with m_1 and m_2 degrees of freedom, respectively, their pdfs are given by

$$p_{R_i}(\rho_i) = \frac{2}{2^{m_i/2}\sigma_i^m\Gamma(m_i/2)}\rho^{m_i-1}e^{-\rho_i^2/2\sigma^2}U(\rho_i), \quad i = 1, 2, \tag{19}$$

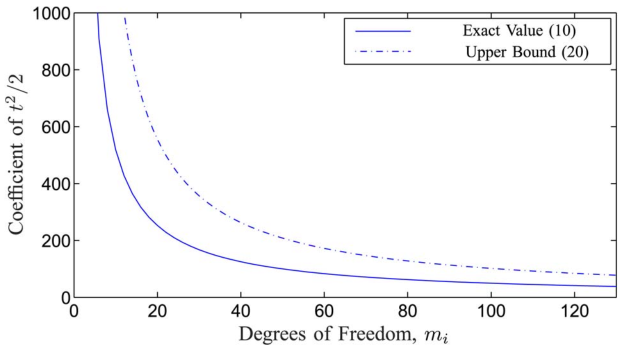

Fig. 2. Convergence of the coefficient of $t^2/2$.

where $\Gamma(\cdot)$ is the standard Gamma function, and both pdfs are assumed to have the same $\sigma > 0$. These pdfs correspond to the distribution of the square root of the sum of the squares of m_1 and m_2 independent zero-mean Gaussian random variables with variance σ^2 each. For $m_i = 2$, (19) reduces to the Rayleigh distribution, and for $m_i = 3$, it reduces to the Maxwell distribution [15]. For $m_i \geq 3$, the right-hand side of (16) can be easily calculated and yields

$$|h_t(Y) - h(X)| \leq \min_{1,2} \left\{ \frac{2\rho_0 t^2}{2^{m_i/2}\sigma^2\Gamma(m_i/2)} \times (m_i - 4)(m_i - 6)\cdots \right\} + \mathcal{O}(t^4), \quad (20)$$

where $\rho_0 = 1$ if m_i is even, and $\rho_0 = \sqrt{\frac{\pi}{2}}$ if m_i is odd. It can be readily verified that the argument of the minimization in (20) decreases monotonically with m_i, which implies that the bound in (20) depends on the component of R with $m^* = \max\{m_1, m_2\}$.

Note that the coefficient of t^2 goes to zero as m_i goes to infinity. A plot illustrating the convergence behavior of the exact value obtained by using numerical integration in (10) and the upper bound in (20) on this coefficient is provided in Fig. 2. □

VI. Conclusion

We have analyzed the dominant error term resulting from ignoring the noise component in computing the differential entropy of the received signal of high-SNR noncoherent SISO communication systems. We have shown that this error decays with 1/SNR, and we have provided an explicit expression and an upper bound for the dominant error term for the case in which the input signal is transmitted over a channel with Gaussian fading coefficients. Our analysis is supported by examples of input signals with various input distributions. The results obtained herein can be readily used to show that, for MIMO systems with full-rank channel matrices and at least as many transmit antennas as receive ones, the error in entropy approximation decays with 1/SNR, which tightens the approximation results in [1] and [4].

Appendix A
Proof of Theorem 2

We use (3) and (4) to evaluate $\frac{dh_t(Y)}{dt^2}\big|_{t^2=0}$ in (5). Since we consider cases in which $h_t(Y)$ is analytic, evaluating $\frac{dh_t(Y)}{dt^2}\big|_{t^2=0}$ is equivalent to evaluating $\lim_{t^2\searrow 0} \frac{dh_t(Y)}{dt^2}$. Using (3) and (4), and invoking Lebesgue's dominated convergence theorem, we can write

$$\frac{dh_t(Y)}{dt^2}\bigg|_{t^2=0} = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (1 + \log p_X(y_r, y_i)) \times \lim_{t^2\searrow 0} \frac{\partial p_Y(y_r, y_i; t)}{\partial t^2} \, dy_r \, dy_i. \quad (21)$$

From (3), $\lim_{t^2\searrow 0} \log p_Y(y_r, y_i; t) = \log p_X(y_r, y_i)$, and

$$\frac{\partial p_Y(y_r, y_i; t)}{\partial t^2} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_X(y_r - u, y_i - v) \times \left(\frac{u^2 + v^2}{2t^6} - \frac{1}{t^4}\right) e^{-\frac{u^2+v^2}{2t^2}} \, du \, dv. \quad (22)$$

To evaluate $\lim_{t\searrow 0} \frac{\partial p_Y(y_r, y_i; t)}{\partial t}$, we use the following result.

Lemma 2:

$$\lim_{\epsilon\searrow 0} \frac{u^2 + v^2}{4\epsilon^3} e^{-\frac{u^2+v^2}{2\epsilon}} = \lim_{\epsilon\searrow 0} \frac{1}{\epsilon^2} e^{-\frac{u^2+v^2}{2\epsilon}}. \quad (23)$$

Proof: The proof follows from applying l'Hôpital's rule to the right-hand side of (23). ■

Using Lemma 2 in (22) with $\epsilon = t^2$ and taking the limit as $t^2 \searrow 0$ yields

$$\lim_{t^2\searrow 0} \frac{\partial p_Y(y_r, y_i; t)}{\partial t^2} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_X(y_r - u, y_i - v) \times \lim_{t^2\searrow 0} \frac{1}{t^4} e^{-\frac{u^2+v^2}{2t^2}} \, du \, dv. \quad (24)$$

To evaluate this integral, we will change the coordinates from Cartesian to polar. With this change of coordinates, we have $(u, v) \mapsto (r, \theta)$, where $u = r\cos\theta$ and $v = r\sin\theta$, and the area element $du\, dv \mapsto r\, dr\, d\theta$. Using this change of coordinates, we can write

$$\lim_{t^2\searrow 0} \frac{\partial p_Y(y_r, y_i; t)}{\partial t^2} = \frac{1}{2\pi}\int_{\theta=-\pi}^{\pi}\int_{r=0}^{\infty} \lim_{t^2\searrow 0} \frac{r}{t^4} e^{-\frac{r^2}{2t^2}} \times p_X(y_r - r\cos\theta, y_i - r\sin\theta) \, dr \, d\theta. \quad (25)$$

To evaluate this integral, we will use the following lemma.

Lemma 3: Let $\delta''(r) = \frac{d^2}{dr^2}\delta(r)$ be the unit triplet function [16]. Then

$$\lim_{t^2\searrow 0} \frac{r}{t^4} e^{-\frac{r^2}{2t^2}} = \delta''(r). \quad (26)$$

Proof: See Appendix B. ■

Substituting from (26) in (25) yields

$$\lim_{t^2\searrow 0} \frac{\partial p_Y(y_r, y_i; t)}{\partial t^2} = \frac{1}{2\pi} \times \int_{\theta=-\pi}^{\pi}\int_{r=0}^{\infty} \delta''(r) p_X(y_r - r\cos\theta, y_i - r\sin\theta) \, dr \, d\theta. \quad (27)$$

Noting that $\cos(\theta+\pi) = -\cos(\theta)$, $\sin(\theta+\pi) = -\sin(\theta)$, and $\delta''(r)$ is an even function, i.e., $\delta''(r) = \delta''(-r)$, (27) can be expressed as

$$\lim_{t^2 \searrow 0} \frac{\partial p_Y(y_r, y_i; t)}{\partial t^2} = \frac{1}{2\pi} \int_{\theta=0}^{\pi} \int_{r=-\infty}^{\infty} \delta''(r) \times p_X(y_r - r\cos\theta, y_i - r\sin\theta)\, dr\, d\theta. \quad (28)$$

To evaluate (28), we use integration by parts and the fact that $\delta'(\pm\infty) = 0$ to write the inner integral as $\cos\theta \int_{r=-\infty}^{\infty} \delta'(r) \frac{\partial p_X(y_r - r\cos\theta, y_i - r\sin\theta)}{\partial(y_r - r\cos\theta)} dr + \sin\theta \int_{r=-\infty}^{\infty} \delta'(r) \frac{\partial p_X(y_r - r\cos\theta, y_i - r\sin\theta)}{\partial(y_i - r\sin\theta)} dr$. Using a similar procedure, this expression can be expressed as $\cos^2\theta \frac{\partial^2 p_X(y_r, y_i)}{\partial y_r^2} + \sin 2\theta \frac{\partial^2 p_X(y_r, y_i)}{\partial y_r \partial y_i} + \sin^2\theta \frac{\partial^2 p_X(y_r, y_i)}{\partial y_i^2}$. Substituting in the integration in (27) yields

$$\lim_{t^2 \searrow 0} \frac{\partial p_Y(y_r, y_i; t)}{\partial t^2} = \frac{1}{2} \frac{\partial^2 p_X(y_r, y_i)}{\partial y_r^2} + \frac{1}{2} \frac{\partial^2 p_X(y_r, y_i)}{\partial y_i^2}. \quad (29)$$

Hence, substituting in (21), we have

$$\left.\frac{dh_t(Y)}{dt^2}\right|_{t^2=0} = -\frac{1}{2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (1 + \log p_X(y_r, y_i)) \times \left(\frac{\partial^2 p_X(y_r, y_i)}{\partial y_r^2} + \frac{\partial^2 p_X(y_r, y_i)}{\partial y_i^2} \right) dy_r\, dy_i. \quad (30)$$

Substituting (30) in (5) yields the statement of the theorem.

APPENDIX B
PROOF OF LEMMA 3

Consider the limiting case of the zero-mean circularly symmetric 2-D Gaussian distribution as the variance, $\epsilon \to 0$. In this case, we have $\lim_{\epsilon \searrow 0} \frac{1}{2\pi\epsilon} e^{-\frac{u^2+v^2}{2\epsilon}} = \delta(u)\delta(v)$. Using l'Hôpital's rule, we obtain an alternate form of the limit on the left-hand side. In particular

$$\lim_{\epsilon \searrow 0} \frac{1}{2\pi\epsilon} e^{-\frac{u^2+v^2}{2\epsilon}} = \lim_{\epsilon \searrow 0} \frac{1}{2\pi} \frac{u^2+v^2}{2\epsilon^2} e^{-\frac{u^2+v^2}{2\epsilon}}. \quad (31)$$

Expressing $\delta(u)\delta(v)$ in the polar coordinates, we have [17]

$$\delta(u)\delta(v) = \frac{1}{\pi|r|} \delta(r). \quad (32)$$

Since $u^2 + v^2 = r^2$, we can use that $\lim_{\epsilon \searrow 0} \frac{1}{2\pi\epsilon} e^{-\frac{u^2+v^2}{2\epsilon}} = \delta(u)\delta(v)$ with (31) and (32) to write

$$\lim_{\epsilon \searrow 0} \frac{1}{2\pi} \frac{r^2}{2\epsilon^2} e^{-\frac{r^2}{2\epsilon}} = \frac{1}{\pi|r|} \delta(r).$$

Invoking the identity [17] $r^2\delta''(r) = 2\delta(r)$ yields $\lim_{\epsilon \searrow 0} \frac{1}{4\pi\epsilon^2} e^{-\frac{r^2}{2\epsilon}} = \frac{1}{2\pi|r|} \delta''(r)$, which, by substituting $\epsilon = t^2$, yields the statement of the lemma.

REFERENCES

[1] L. Zheng and D. N. C. Tse, "Communication on the Grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel," *IEEE Trans. Inf. Theory*, vol. 48, no. 2, pp. 359–383, Feb. 2002.

[2] G. Durisi, "On the capacity of the block-memoryless phase-noise channel," *IEEE Commun. Lett.*, vol. 16, no. 8, pp. 1157–1160, Aug. 2012.

[3] R. H. Gohary and T. N. Davidson, "Non-coherent MIMO communication: Grassmannian constellations and efficient detection," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, pp. 1176–1205, Mar. 2009.

[4] W. Yang, G. Durisi, and E. Riegler, "On the capacity of large-MIMO block-fading channels," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 117–132, Feb. 2013.

[5] E. Ordentlich, "Maximizing the entropy of a sum of independent bounded random variables," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2176–2181, May 2006.

[6] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY, USA: Wiley, 1991.

[7] M. Madiman and F. Ghassemi, "The entropy power of a sum is fractionally superadditive," in *Proc. IEEE Int. Symp. Inf. Theory*, Seoul, Korea, Jun. 2009, pp. 295–298.

[8] T. M. Cover and Z. Zhang, "On the maximum entropy of the sum of two dependent random variables," *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1244–1246, Apr. 1994.

[9] A. Lapidoth and G. Pete, "On the entropy of the sum and of the difference of independent random variables," in *Proc. IEEE Conv. IEEEI*, Eilat, Israel, Nov. 2008, pp. 623–625.

[10] M. Madiman and I. Kontoyiannis, "The entropies of the sum and the difference of two IID random variables are not too different," in *Proc. IEEE Int. Symp. Inf. Theory*, Austin, TX, USA, Jun. 2010, pp. 1369–1372.

[11] M. Madiman, "On the entropy of sums," in *Proc. IEEE ITW*, Porto, Portugal, May 2008, pp. 303–307.

[12] R. H. Gohary and H. Yanikomeroglu, "On the accuracy of the high-SNR approximation of the differential entropy of signals in additive Gaussian noise," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Florence, Italy, May 2014, pp. 5735–5738.

[13] M. Godavarti and A. Hero, "Convergence of differential entropies," *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 171–176, Jan. 2004.

[14] W. Rudin, *Real and Complex Analysis*, 3rd ed. New York, NY, USA: McGraw-Hill, 1987.

[15] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, 2nd ed. New York, NY, USA: McGraw-Hill, 1984.

[16] C. R. Wylie and L. C. Barrett, *Adacned Engineering Mathematics*. McGraw-Hill, 5th ed., 1982.

[17] R. N. Bracewell, *The Fourier Transform and Its Applications*, 3rd ed. New York, NY, USA: McGraw-Hill, 2000.