

Constant Power Signaling in Rayleigh Fading Channels: Joint Output Probability Distribution and Information Rate Bounds

Rauf Iqbal, Parastoo Sadeghi, and Thushara D. Abhayapala
 Research School of Information Sciences and Engineering
 The Australian National University, ACT 200, AUSTRALIA
 Email: parastoo.sadeghi@anu.edu.au

Abstract

In this paper, we consider a wireless communication scenario in which the channel output is marginally Gaussian, but not jointly Gaussian. In particular, we study the joint probability distribution of channel outputs in correlated Rayleigh fading channels in response to constant power signaling, such as M -ary phase shift keying (MPSK). We show that the distribution of the channel output at any given sampling time is marginally Gaussian. However, the joint distribution of a sequence of channel outputs cannot be jointly Gaussian. A consequence of this result is that the information rates stated to be exact in two recent contributions, are strict upper bounds to the achievable data rates. We examine the tightness of these upper bounds by comparing them with the MPSK upper bound under perfect channel state information (CSI) assumption. We find that the CSI upper bound is considerably tighter in slow fading channels, high signal-to-noise ratios, and low-dimension (such as binary) PSK signaling.

1. Introduction

1.1. Motivation and Background

Constant power signaling, such as M -ary phase shift keying (MPSK) [9] is one of the most widely-used signaling schemes among modern digital modulation techniques. This is mainly due to its relative robustness to noise and ease of implementation with inexpensive non-linear power amplifiers [10]. In wireless communication systems, transmitted signals experience random and time-varying multipath fading effects, which can be characterized by frequency-flat and time-correlated Rayleigh fading [9]. It is of information-theoretic interest to determine the maximum rate at which information can be transferred reliably over a Rayleigh fading channel using constant power inputs.

The information capacity of the fading channels strongly depends on the availability of the channel state information (CSI) at the transmitter and the receiver sides. In many practical situations, the exact state of the fading channel is not available at either side, due to the random and time-varying nature of the channel. The capacity problem without CSI has been studied for some simplified channel models e.g., memoryless Rayleigh fading channel [1], block-fading Rayleigh channels [5, 8], and finite-state Markov channels [4, 12].

In another approach, upper and lower bounds to the achievable information rates were derived in [3] under no CSI assumption for independent and identically distributed (i.i.d.) constant power and Gaussian inputs. The upper bound provided in [3] is correctly based on the premise that the entropy of any random process with a fixed covariance matrix is maximized when the process is jointly Gaussian. However, the *marginal* Gaussian distribution property of the channel output in constant power signaling has further led some authors [13, 14] to tacitly conclude that the channel output sequence is also *jointly* Gaussian and hence, to report an *exact* information rate for orthogonal frequency division multiplexing (OFDM) systems.

1.2. Contributions

Motivated to further clarify rather mixed results in the literature, we propose to rigorously study the marginal and joint probability distributions of the channel outputs in response to constant power signaling in correlated Rayleigh fading channels. This is carried out through mathematical analysis and computer simulations. We show that the system model under consideration is a clear example of the case where marginal Gaussianity does not imply joint Gaussianity. As a result, we show that the information rates reported in [13, 14] are, in fact, strict upper bounds. We then examine the tightness of these bounds by comparing them with the perfect CSI upper bound for MPSK inputs. Interestingly, the comparison reveals that the perfect CSI upper bound can provide tighter bounds than those in [3, 13] in

slow fading channels and high signal-to-noise ratio (SNR) conditions, especially when the MPSK dimension is low (such as binary signaling).

Throughout the paper, the following notations will be used: Bold lower (upper) letters denote vectors (matrices). Superscripts $*$ and T denote the conjugate transpose and transpose, respectively. The subscript of the form $(\cdot)_x$ represents the modulo- x operation. In all mathematical equations, $\log(\cdot)$ function is the natural logarithm. However, in figures the information rates are plotted in bits per channel use for easier reference. The notation $E\{\cdot\}$ denotes the mathematical expectation and the matrix \mathbf{I} is the $N \times N$ identity matrix. The operators $\Re(z)$ and $\Im(z)$ represent respectively the real and imaginary parts of a complex number z . The symbol $\mathcal{N}(\mathbf{m}, \Phi)$ is used to represent an N -dimensional complex Gaussian distribution with mean \mathbf{m} and covariance matrix Φ .

2. System Model

The channel between the transmitter and the mobile receiver is modelled as a time-varying and strictly bandlimited flat-fading process. Here, we consider only a single carrier system. However, the results extend naturally to the case of OFDM [13, 14], where the received signal can be made free of intersymbol and intercarrier interference through the use of cyclic prefix. The received $N \times 1$ vector \mathbf{y} at the output of the matched filter¹ is given in complex low-pass form as

$$\mathbf{y} = \mathbf{X}\mathbf{h} + \mathbf{w} \quad (1)$$

$$= \mathbf{s} + \mathbf{w}, \quad (2)$$

where \mathbf{X} is a diagonal matrix with the transmitted symbol vector $\mathbf{x} = [x_1, x_2, \dots, x_N]^T$ along its main diagonal. Vector $\mathbf{h} = [h_1, h_2, \dots, h_N]^T$ represents the samples of a zero-mean complex Gaussian channel fading process. We assume that each element of the vector \mathbf{h} is normalized to have unit variance with independent real and imaginary parts. We also assume that the actual realization of the random channel \mathbf{h} is not known *a priori* at the transmitter or the receiver. However, the statistics of the Gaussian channel are known at both sides. This is a reasonable assumption even in relatively fast fading conditions, whenever the mobile unit is moving at constant or (close to constant) velocity. For Gaussian fading process, channel statistics are fully determined by the Toeplitz positive semidefinite covariance matrix \mathbf{C}_h . Under constant mobile velocity assumption, the covariance matrix \mathbf{C}_h does not change rapidly and hence, can be measured accurately. In the bandlimited Clarke's

¹We focus on the matched filter, because of its widespread use in many practical systems, even though it is not the optimal detection scheme for the system considered here.

fading model [9], the element of matrix \mathbf{C}_h at row k and column j is given by

$$[\mathbf{C}_h]_{k,j} = E\{h_k h_j^*\} = J_0(2\pi f_D T |k - j|), \quad (3)$$

where J_0 is the zeroth-order Bessel function of the first kind, f_D is the maximum Doppler frequency shift, and T is the transmitted symbol period.

In (2), the $N \times 1$ vector $\mathbf{w} = [w_1, w_2, \dots, w_N]^T$ consists of samples from a zero-mean complex additive white Gaussian noise (AWGN) process. The noise samples w_ℓ are independent of each other and have unit variance (the covariance matrix is \mathbf{I}). Equation (2) is normalized in such a way that the average transmitted signal power is $E\{|x_j|^2\} = \mathcal{E}$. The average received SNR is then given by $SNR = \mathcal{E}$.

3. Marginal Distribution of Output

For deriving the marginal distribution of the channel output, we need to consider an arbitrary sample of the channel output at discrete time index ℓ , which is given as

$$y_\ell = \underbrace{x_\ell h_\ell}_{s_\ell} + w_\ell = \sqrt{\mathcal{E}} e^{i\theta_\ell} |h_\ell| e^{i\phi_\ell} + w_\ell, \quad (4)$$

where fading amplitude $|h_\ell|$ is Rayleigh-distributed and fading phase ϕ_ℓ is uniformly-distributed over $[-\pi, \pi)$. $\sqrt{\mathcal{E}} e^{i\theta_\ell}$ is the transmitted constant power signal. For M -ary phase shift keying, the signal phase can be one of the M possible phases as $\theta_\ell = 2\pi m/M$ for $m \in \{0, \dots, M-1\}$. Since the transmitted signal has constant modulus, the distribution of $|s_\ell| = \sqrt{\mathcal{E}} |h_\ell|$ remains Rayleigh. To prove the Gaussianity of s_ℓ (and hence marginal Gaussianity of y_ℓ), we only need to prove that the phase of s_ℓ (which is $(\theta_\ell + \phi_\ell)_{2\pi}$) is uniformly distributed over $[-\pi, \pi)$. If the input phase was *continuously* uniform between $-\pi$ and π , one could show that $(\theta_\ell + \phi_\ell)_{2\pi}$ remains uniform over $[-\pi, \pi)$ by invoking standard arguments in probability theory [7]. Intuitively, one expects that the distribution of $(\theta_\ell + \phi_\ell)_{2\pi}$ remains uniform for MPSK signals with discrete distributions at $\theta_\ell = 2\pi m/M$. This intuitive reasoning has been used in [13] to show marginal Gaussianity of the output.

In fact, the following theorem from [15] provides even a stronger result.

Theorem 3.1 *If two processes $\alpha(t)$ and $\beta(t)$ have the probability density functions $p_1(\alpha, t)$ and $p_2(\beta, t)$, respectively, such that $\alpha(t)$, modulo q , is statistically independent of $\beta(t)$ and uniformly distributed, i.e.,*

$$p_1(\alpha_q, t) = \frac{1}{q} \{u(\alpha) - u(\alpha - q)\},$$

where $u(x)$ is the unit step function, then the probability density function of $(\alpha_q(t) \pm \beta(t))$, modulo q , is also

uniformly distributed, regardless of the probability density function of $\beta(t)$. That is, if $\gamma(t) = \alpha_q(t) \pm \beta(t)$, then

$$p(\gamma_q, t) = \frac{1}{y} \{u(\gamma) - u(\gamma - q)\}.$$

By directly applying the above theorem to the problem at hand, we formulate the following result:

Theorem 3.2 *The distribution of $\{\psi = (\phi + \theta)_{2\pi}\}$, where ϕ and θ are statistically independent channel and input phase processes, is uniform over $[0, 2\pi)$ (or equivalently over $[-\pi, \pi)$) and is independent of the distribution of the input phase, θ .*

An immediate consequence of the second part of the theorem 3.2 is the following corollary.

Corollary 3.3 *The channel output y_ℓ in a time-selective Rayleigh fading channel is marginally Gaussian with zero-mean for any MPSK input distribution.*

4. Joint Distribution of the Output Sequence

4.1 Mathematical Analysis

In Section 3, we proved that each element s_ℓ of the vector \mathbf{s} in (2) is zero-mean Gaussian. While joint Gaussianity implies marginal Gaussianity, the reverse is not necessarily true [11]. Therefore, \mathbf{s} is not guaranteed to be multivariate Gaussian, although it is marginally Gaussian.

In the following, with the help of analytical arguments and numerical simulations we show that \mathbf{s} is indeed jointly non-Gaussian and, hence, the output vector \mathbf{y} , which is the sum of independent non-Gaussian vector \mathbf{s} and Gaussian vector \mathbf{w} , is non-Gaussian.

Firstly, we know that Gaussianity is preserved under invertible linear transformation [9]. Notice that if \mathbf{X} were known and given, then $\mathbf{s} = \mathbf{X}\mathbf{h}$ would represent an invertible linear transformation of the Gaussian random vector \mathbf{h} and hence, would be jointly Gaussian. For example, in a pilot-aided data transmission scheme [16], \mathbf{X} is known during the pilot transmission phase and the distribution of \mathbf{s} and, hence, the output vector \mathbf{y} in (2) is N -dimensional complex Gaussian. However, in the current scenario, \mathbf{X} is the transmitted data, which is not known. The product $\mathbf{X}\mathbf{h}$ therefore represents a linear, but *random* and hence, non-invertible transformation. Thus, the output \mathbf{y} can not be guaranteed to be Gaussian.

We construct the distribution of \mathbf{y} using conditional distribution of \mathbf{y} given \mathbf{x} . It can be noticed from (2) that the conditional distribution of \mathbf{y} given \mathbf{x} , $f(\mathbf{y}|\mathbf{x})$, is $\mathcal{N}(\mathbf{0}, \mathbf{C}_{\mathbf{y}|\mathbf{x}})$, where $\mathbf{C}_{\mathbf{y}|\mathbf{x}} = \mathbf{X}\mathbf{C}_h\mathbf{X}^H + \mathbf{I}$ is the conditional covariance of \mathbf{y} given \mathbf{x} . The distribution of \mathbf{y} , $f(\mathbf{y})$, is then given by

$$f(\mathbf{y}) = \sum_{\mathbf{x} \in \mathcal{X}^N} f(\mathbf{y}|\mathbf{x}) p(\mathbf{x}),$$

$$= \frac{1}{(M\pi)^N} \sum_{\mathbf{x} \in \mathcal{X}^N} \frac{1}{|\mathbf{C}_{\mathbf{y}|\mathbf{x}}|} \exp\left[-\mathbf{y}^* \mathbf{C}_{\mathbf{y}|\mathbf{x}}^{-1} \mathbf{y}\right], \quad (5)$$

where \mathcal{X}^N represents the N -dimensional probability space for \mathbf{x} , $p(\mathbf{x})$ is the probability mass function for discrete constant power input and M is the number of points in the MPSK constellation. Equiprobable i.i.d. MPSK inputs have been assumed to arrive at the second equality. Equation (5) shows that $f(\mathbf{y})$ is a summation of Gaussians and is, therefore, non-Gaussian. In the next subsection, we shall verify the non-Gaussianity of the output sequence with the help of numerical simulations.

4.2 Numerical Analysis

In order to prove that the output \mathbf{y} is non-Gaussian, it is sufficient to show that \mathbf{s} in (2) is non-Gaussian for $N = 2$ consecutive times. We proceed step-by-step as follows:

First, we generate a long sequence of channel fading process \mathbf{h} according to the Clarke's model. This can be easily implemented using the L -th order autoregressive (AR) approximation of the process [2] with some high AR order such as $L = 500$. For notational simplicity, let us denote two consecutive channel realizations as h_1 and h_2 . Since the fading process is marginally circularly symmetric complex Gaussian, the contour plot (2D-histogram) of $\Re(h_1)$ and $\Im(h_1)$ consists of concentric circles. This is not shown here for the sake of brevity. However, h_1 is correlated with h_2 . Therefore, the contour plot of the joint distribution of $\Re(h_1)$ and $\Re(h_2)$ consists of concentric ellipses (the same statement applies to imaginary parts). Figure 1 shows one such contour plot for a fading channel with the normalized Doppler frequency shift of $f_D T = 0.2$ and verifies the joint Gaussian distribution of $\Re(h_1)$ and $\Re(h_2)$. The number of samples generated is $P = 10^6$ and the number of bins for each element is 62.

Second, we generate a long sequence of $\mathbf{s} = \mathbf{X}\mathbf{h}$ by multiplying \mathbf{h} with i.i.d. realizations of input \mathbf{X} from the MPSK constellation. For notational simplicity, let us denote two consecutive realizations of \mathbf{s} as s_1 and s_2 , given as

$$\mathbf{s}' = [s_1 s_2] = [e^{i\theta_1} h_1 \ e^{i\theta_2} h_2]^T. \quad (6)$$

Note that s_1 and s_2 are jointly Gaussian if and only if

$$\mathbf{s}'' = \left[\Re(e^{i\theta_1} h_1) \ \Re(e^{i\theta_2} h_2) \ \Im(e^{i\theta_1} h_1) \ \Im(e^{i\theta_2} h_2) \right]^T, \quad (7)$$

is jointly Gaussian [6]. Since any subset of a Gaussian vector is also jointly Gaussian, to prove that \mathbf{s}'' is not Gaussian, we only need to show that at least two elements of \mathbf{s}'' are jointly non-Gaussian. For example, it is sufficient to show that $[\Re(e^{i\theta_1} h_1) \ \Re(e^{i\theta_2} h_2)]^T$ are jointly non-Gaussian.

Figure 2 shows a contour plot of the joint distribution of $[\Re(e^{i\theta_1} h_1) \ \Re(e^{i\theta_2} h_2)]^T$ for $M = 2$ binary MPSK signaling and confirms its non-Gaussian joint distribution.

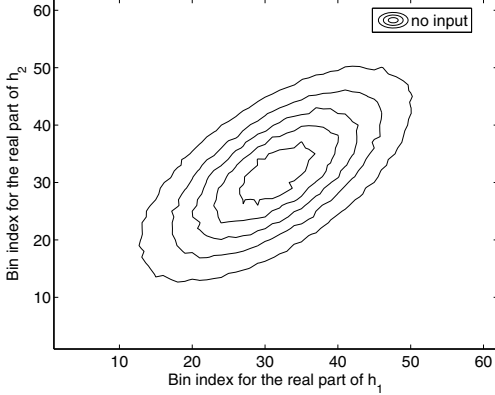


Figure 1. A contour plot of joint distribution of $\Re(h_1)$ and $\Re(h_2)$: $f_D T = 0.2$.

Figure 3 shows a contour plot of the joint distribution of $[\Re(e^{i\theta_1} h_1) \ \Re(e^{i\theta_2} h_2)]^T$ for 16-PSK signaling. Compared with Figure 3, the contour plot is somewhat more ‘rounded’. Nevertheless, it is still not quite circular.

5. Discussion on Information Rates

In Section 4, we showed that a sequence of channel outputs of length N in response to MPSK signaling is jointly non-Gaussian. To see the consequence of this result on information rates, we first write an element of channel output covariance matrix at row k and column j as

$$[\mathbf{C}_y]_{k,j} = E \{ y_k y_j^* \} = E \{ h_k h_j^* \} E \{ x_k x_j^* \} + \delta_{k,j}, \quad (8)$$

which is a consequence of the independence between the input and channel processes. $E \{ h_k h_j^* \}$ was defined in (3). In (8), $\delta_{jk} = 0$ for $j \neq k$. Therefore, we can upper bound the entropy rate of channel output sequence as follows

$$\begin{aligned} \frac{1}{N} h(\mathbf{y}) &< \frac{1}{N} \log \det(\pi e \mathbf{C}_y) \\ &\leq \frac{1}{N} \log \prod_{n=1}^N \pi e [C_y]_{n,n} \\ &= \log \pi e (\mathcal{E} + 1), \end{aligned} \quad (9)$$

where the first inequality holds because the entropy of any random vector is upper bounded by the entropy of a Gaussian random vector with the same covariance matrix. The second inequality follows from Hadamard’s inequality, with equality if and only if $[C_y]_{k,j} = 0$, $k \neq j$, i.e., if the input symbols are uncorrelated. The above procedure is correctly followed in [3]. We note from (8) that the output process becomes uncorrelated for i.i.d MPSK inputs, but since the output is not jointly Gaussian, one cannot

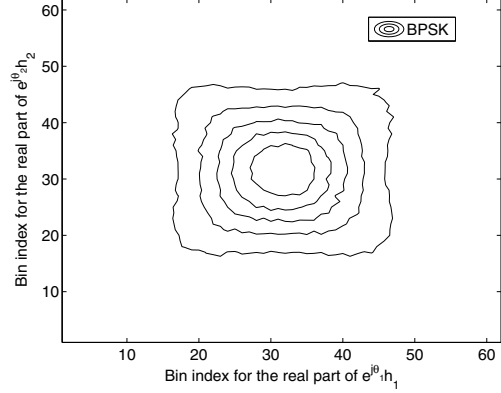


Figure 2. A contour plot of joint distribution of $[\Re(e^{i\theta_1} h_1) \ \Re(e^{i\theta_2} h_2)]^T$: BPSK signaling.

conclude independence. The point of view of [13, 14] that i.i.d. constant power signaling *whitens* the channel process (makes it independent) is not justified.

The mutual information rate $I(\mathbf{y}; \mathbf{x})$ between the output \mathbf{y} and the input \mathbf{x} with no CSI is given by

$$I(\mathbf{y}; \mathbf{x}) = \frac{1}{N} (h(\mathbf{y}) - h(\mathbf{y}|\mathbf{x})). \quad (10)$$

Using (9), one can upper bound the information rate as

$$I(\mathbf{y}; \mathbf{x}) < \log \pi e (\mathcal{E} + 1) - \frac{1}{N} h(\mathbf{y}|\mathbf{x}). \quad (11)$$

Therefore, the use of ‘‘capacity’’ in [13, 14] for the RHS of (11) is misleading and should be considered as a *strict upper bound* to the information rates achievable with M -PSK signaling.

It is reasonable to ask how tight the information rate upper bounds on the RHS of (11) are for practical MPSK schemes. We know trivially that $\log_2(M)$ bits per channel use is an upper bound for MPSK information rates. With this in mind, we compare (11) with the perfect CSI upper bound $I(\mathbf{y}; \mathbf{x}|\mathbf{h}) = 1/N [h(\mathbf{y}|\mathbf{h}) - h(\mathbf{y}|\mathbf{x}, \mathbf{h})]$. To compute the RHS of (11), we use a closed-form expression for $1/N h(\mathbf{y}|\mathbf{x})$ provided in [3]. Figure 4 shows two information upper bounds using (11) for normalized Doppler frequency shifts of $f_D T = 0.05$ and $f_D T = 0.2$. Relatively, we refer to the former as slow fading and to the latter as fast fading. It is clear that for low MPSK dimensions (such as $M = 2$ or BPSK), (11) goes rapidly above the trivial upper bound of $\log_2(M) = 1$ bits per channel use. For slow fading channels, BPSK and 16-PSK CSI upper bounds provide tighter results for almost all SNRs. For the faster fading conditions, (11) provides relatively tighter bounds than the CSI upper bound for 16-PSK signaling for a wide range of SNR, but is still loose for BPSK signaling unless for very low SNR.

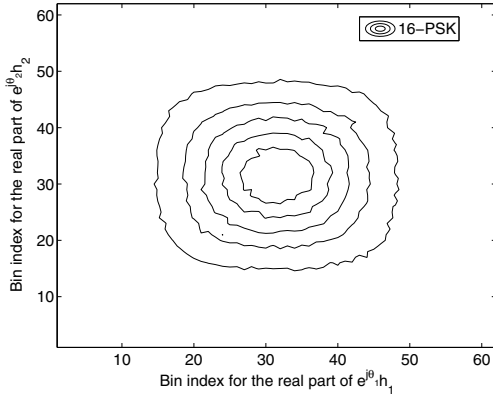


Figure 3. A contour plot of joint distribution of $[\Re(e^{i\theta_1} h_1) \Re(e^{i\theta_2} h_2)]^T$: 16-PSK signaling.

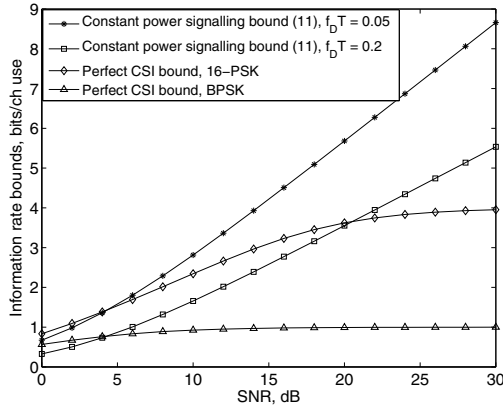


Figure 4. Comparison of information rate upper bounds.

6. Conclusions

We investigated the statistical properties of the channel output in response to constant power signaling in Rayleigh fading channels. We showed that although the channel output at any given time is marginally Gaussian, a sequence of channel outputs are not jointly Gaussian. Therefore, one can only obtain upper bounds on the achievable information rates using standard entropy inequalities. We then compared these information rate upper bounds with the perfect CSI upper bound. The comparison showed that the obtained upper bounds from (11) become loose for high SNR and slow fading channels, especially for low-dimension MPSK signalling, such as BPSK.

7. Acknowledgement

The work of P. Sadeghi was supported under Australian Research Councils Discovery Projects funding scheme (project no. DP0773898).

References

- [1] I. C. Abou-Faycal, M. D. Trot, and S. Shamai. The capacity of discrete-time memoryless rayleigh fading channels. In *IEEE Trans. Inform. Theory*, volume 47, pages 1290–1301, May 2001.
- [2] K. E. Baddour and N. C. Beaulieu. Autoregressive models for fading channel simulation. In *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, pages 1187–1192, San Antonio, TX, Nov. 2001.
- [3] X. Deng and A. M. Haimovich. Information rates of time varying Rayleigh fading channels. In *Proc. IEEE Int. Conf. Commun. (ICC)*, pages 573–577, Paris, France, June 2004.
- [4] A. J. Goldsmith and P. Varaiya. Capacity, mutual information, and coding for finite-state Markov channels. *IEEE Trans. Inform. Theory*, 42(3):868–886, May 1996.
- [5] T. L. Marzetta and B. M. Hochwald. Capacity of a mobile multiple-antenna communication link in rayleigh flat fading. In *IEEE Trans. Inform. Theory*, volume 45, pages 139–157, Jan. 1999.
- [6] K. S. Miller. Complex Gaussian processes. *SIAM Review*, 11(4):544–547, Oct 1969.
- [7] A. Papoulis and S. U. Pillai. *Probability, Random Variables and Stochastic Processes*. Mc Graw Hill, New York, fourth edition, 2002.
- [8] M. Peleg and S. Shamai. On the capacity of blockwise incoherent MPSK channel. In *IEEE Trans. Commun.*, volume 46, pages 603–609, May 1998.
- [9] J. G. Proakis. *Digital Communications*. Mc Graw Hill, New York, fourth edition, 2000.
- [10] T. S. Rappaport. *Wireless Communications, Principles and Practice*. Prentice Hall, Upper Saddle River, second edition, 2002.
- [11] J. P. Romano and A. F. Siegel. *Counterexamples in Probability and Statistics*. CRC Press, 1986.
- [12] P. Sadeghi and P. Rapajic. On information rates of time-varying fading channels modeled as finite-state Markov channels. *IEEE Trans. Commun.*, 56(8):1268–1278, Aug. 2008.
- [13] D. Schafhuber. *Wireless OFDM systems: Channel prediction and system capacity*. PhD thesis, Vienna Univ. of Technology, Austria, March 2004.
- [14] D. Schafhuber, H. Bolsckei, and G. Matz. System capacity of wideband OFDM communications over fading channels without channel knowledge. In *Proc. Int. Symp. on Inform. Theory, ISIT*, page 389, 27 June– 2 July 2004.
- [15] F. J. Scire. A probability density function theorem for the modulo y values of the sum of two statistically independent processes. *Proc. IEEE*, 56(2):204–205, Feb. 1968.
- [16] L. Tong, B. M. Sadler, and M. Dong. Pilot assisted wireless transmission: General model, design criteria, and signal processing. *IEEE Signal Processing Mag.*, 21(6):12–25, Nov. 2004.