The Noncoherent Rician Fading Channel—Part I: Structure of the Capacity-Achieving Input

Mustafa Cenk Gursoy, *Member, IEEE*, H. Vincent Poor, *Fellow, IEEE*, and Sergio Verdú, *Fellow, IEEE*

***Abstract*—Transmission of information over a discrete-time memoryless Rician fading channel is considered, where neither the receiver nor the transmitter knows the fading coefficients. First, the structure of the capacity-achieving input signals is investigated when the input is constrained to have limited peakedness by imposing either a fourth moment or a peak constraint. When the input is subject to second and fourth moment limitations, it is shown that the capacity-achieving input amplitude distribution is discrete with a finite number of mass points in the low-power regime. A similar discrete structure for the optimal amplitude is proven over the entire signal-to-noise ratio (SNR) range when there is only a peak-power constraint. The Rician fading with the phase-noise channel model, where there is phase uncertainty in the specular component, is analyzed. For this model, it is shown that, with only an average power constraint, the capacity-achieving input amplitude is discrete with a finite number of levels. For the classical average-power-limited Rician fading channel, it is proven that the optimal input amplitude distribution has bounded support.**

***Index Terms*—Capacity-achieving input, channel capacity, fading channels, memoryless fading, peak constraints, phase noise, Rician fading.**

I. Introduction

RECENTLY, the information-theoretic analysis of fading channels is receiving much attention. This interest is motivated by the rapid advances in wireless technology and the need to use scarce resources such as bandwidth and power as efficiently as possible under severe fading conditions. Providing the ultimate performance, information-theoretic measures such as capacity, spectral efficiency, and error exponents can be used as benchmarks to which the performance of practical communication systems can be compared. Furthermore, with the recent discovery of codes that operate very close to the Shannon capacity, information-theoretic limits have gained practical relevance. Although the capacity and other information-theoretic measures of fading channels were investigated in as early as the 1960s [4], [5], it is only recently that many interesting fading channel models have been considered under various practically related input and channel constraints.

A significant amount of effort has been expended to study fading channel models where side information about the fading is available at either the receiver or the transmitter or both (see [24]–[27]). However, under fast fading conditions, noncoherent communications, where neither party knows the fading, often become the only available alternative. Richters [5] considered the problem of communicating over an average-power-limited discrete-time memoryless Rayleigh fading channel without any channel side information. He conjectured that the capacity-achieving amplitude distribution is discrete with a finite number of mass points. Recently, Abou-Faycal *et al.* [8] gave a rigorous proof of Richters' conjecture. This result shows that when the fading is known by neither the transmitter nor the receiver, the optimal amplitude distribution has a notably different character than that of unfaded Gaussian channels. A similar discrete structure for the optimal input was also shown in [11] for the pulse amplitude modulated direct detection photon channel when average- and peak-power limitations are imposed on the intensity of a photon emitting source. Katz and Shamai [9] considered the noncoherent additive Gaussian white noise (AWGN) channel and proved that the optimal input amplitude is discrete with an infinite number of mass points (see also [10]). Lapidoth [16] recently analyzed the effects of phase noise over the AWGN channel, characterizing the high-signal-to-noise-ratio (high-SNR) asymptotics of the channel capacity for a general class of phase noise distributions with memory (see also [17]). An extensive study of the capacity of multiantenna fading channels at high SNR was conducted in [18].

Kennedy [4] showed that the infinite bandwidth capacity of fading multipath channels is the same as that of the unfaded Gaussian channel. Although any set of orthogonal signals achieves this capacity for the unfaded Gaussian channel, orthogonal signals that are peaky both in time and frequency are needed in the presence of fading [29, Sec. 8.6]. Indeed, for a general class of fading channels, Verdú [3] has recently shown that if there are no constraints other than average power, flash signaling, a class of unbounded peak-to-average ratio inputs defined in [3], is necessary to achieve the capacity as $\mathrm{SNR} \to 0$ when the channel realization is unknown at the receiver. Flash signaling can be practically employed in systems where sudden discharge of energy (e.g., using capacitors) is allowed, thus sidestepping the use of RF amplifiers. However, these peaky signals are not feasible in communication systems subject to strict peak-to-average ratio requirements. Furthermore, in some

Manuscript received January 31, 2003; revised January 4, 2004 and June 17, 2004; accepted June 23, 2004. The editor coordinating the review of this paper and approving it for publication is G. Leus. This work was supported by the U.S. Army Research Laboratory under Contract DAAD 19-01-2-0011. The material in this paper was presented in part at the Fortieth Annual Allerton Conference on Communication, Control, and Computing, Monticello, IL, October 2002, and the Canadian Workshop on Information Theory, Waterloo, ON, May 18–21, 2003.

M. C. Gursoy was with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA. He is now with the Department of Electrical Engineering, University of Nebraska-Lincoln, Lincoln, NE 68588 USA.

H. V. Poor and S. Verdú are with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: verdu@princeton.edu).

Digital Object Identifier 10.1109/TWC.2005.853970

systems, code division multiple access (CDMA)-type white signals, which spread their energy over the available bandwidth, are used because of their antijamming and low probability of intercept capabilities. Hence, it is of interest to investigate the effect upon the capacity of imposing peakedness constraints, especially in the low-power regime. Médard and Gallager [19] considered noncoherent broadband fading channels with no specular component and limited the peakedness of the input signals by imposing a fourth moment constraint. Then, they showed that such a constraint forces the mutual information to zero inversely with increasing bandwidth. Since CDMA-type signals spread their energy over the available bandwidth, they satisfy the above fourth moment constraint. Therefore, Médard and Gallager conclude that CDMA-type white signals cannot efficiently utilize fading multipath channels at extremely large bandwidth. Other results on this theme are obtained by Telatar and Tse [21] and Subramanian and Hajek [20].

In this paper, the noncoherent discrete-time memoryless Rician fading channel is considered and the capacity and the optimal input structure are studied. In a companion paper [Part II], the spectral-efficiency/bit-energy tradeoff in the low-power regime is further investigated. The organization of the paper is as follows. In Section II, the Rician fading channel model is introduced. In Section III, the structure of the capacity-achieving input distribution in the low-power regime when the channel input is constrained to have limited peakedness is characterized. In Section IV, the average-power-limited Rician fading with phase-noise channel, where there is phase uncertainty in the specular component, is introduced and the structure of the capacity-achieving input is investigated. In Section V, numerical results are given, and in Section VI, conclusions are presented.

II. Channel Model

In this paper, the following discrete-time memoryless Rician fading channel model is considered:

$$y_i = mx_i + a_i x_i + n_i \tag{1}$$

where $\{a_i\}$ and $\{n_i\}$ are sequences of independent identically distributed (i.i.d.) circular zero mean complex Gaussian random variables, independent of each other and of the input, with variances $E\{|a_i|^2\} = \gamma^2$ and $E\{|n_i|^2\} = N_0$, m is a deterministic complex constant, x_i is the complex channel input and y_i is the complex channel output. $\{a_i\}$ and $\{n_i\}$ represent sequences of fading coefficients and background noise samples, respectively.

The Rician fading channel model is particularly appropriate when there is a direct line-of-sight (LOS) component in addition to the faded component arising from multipath propagation. Moreover, the Rician model includes both the unfaded Gaussian channel and the Rayleigh fading channel as two special cases. Hence, results obtained for this model provide a unifying perspective.

In the channel model (1), fading is assumed to be flat, and hence, has a multiplicative effect on the channel input. This is a valid assumption if the delay spread of the channel is much smaller than the symbol duration. Moreover, frequency-selective fading channels can often be decomposed into parallel noninteracting flat fading subchannels using orthogonal multicarrier techniques. Note also that the fading coefficients assume independent realizations at every symbol period. Under such fast fading conditions, reliable estimation of the fading coefficients may be quite difficult because of the short duration between independent fades. Therefore, the noncoherent scenario where neither the receiver nor the transmitter knows the fading coefficients $\{a_i\}$ is considered.

III. Channel Capacity and Optimal Input Distribution

In this section, we elaborate on the structure of the capacity-achieving input distribution for the Rician fading channel when the input has limited peakedness, which is achieved by imposing a fourth moment or a peak limitation on the input amplitude.

A. Second and Fourth Moment Limited Input

First, it is assumed that the input amplitude is subject to second and fourth moment constraints

$$E\left\{|x_i|^2\right\} \le P_{\text{av}} \qquad \forall i \tag{2}$$

$$E\left\{|x_i|^4\right\} \le \kappa P_{\text{av}}^2 \qquad \forall i \tag{3}$$

where P_{av} is the average power constraint and $1 < \kappa < \infty$. When the average power constraint is active, the fourth moment constraint is equivalent to limiting the kurtosis, $((E\{|x_i|^4\})/(E\{|x_i|^2\})^2) \le \kappa$, which is a measure of the peakedness of the input signal. For the Rician channel model (1), the capacity is the supremum of the input–output mutual information over the set of all input distributions satisfying the constraints (2) and (3)[1]

$$C = \sup_{\substack{F_x(\cdot) \\ E\{|x|^2\} \le P_{\text{av}} \\ E\{|x|^4\} \le \kappa P_{\text{av}}^2}} \int_{\mathbb{C}} \int_{\mathbb{C}} f_{y|x}(y|x) \log \frac{f_{y|x}(y|x)}{f_y(y)} \, \mathrm{d}y \, \mathrm{d}F_x(x) \tag{4}$$

where the conditional density of the output given the input

$$f_{y|x}(y|x) = \frac{1}{\pi\left(\gamma^2|x|^2 + N_0\right)} \exp\left(-\frac{|y - mx|^2}{\gamma^2|x|^2 + N_0}\right) \tag{5}$$

is circular complex Gaussian. Moreover, note that $f_y(y) = \int_{\mathbb{C}} f_{y|x}(y|x) \, \mathrm{d}F_x(x)$ is the marginal output density, F_x is the distribution function of the input and $\mathbb{C}$ denotes the complex plane.

First, we have the following preliminary result on the optimal phase distribution.

Proposition 1: For the Rician fading channel (1) and input constraints (2) and (3), uniformly distributed phase that is independent of the input amplitude is capacity achieving.[2]

[1] Since the channel is memoryless, without loss of generality, the index i can be dropped.

[2] The result holds in wider generality: the feasible set defined by (2) and (3) can be replaced by any set of constraints that are imposed only on the input magnitude.

Proof: The result follows readily from the arguments in [18, Sec. IV.D.6] where the optimality of circularly symmetric input distributions is pointed out in a more general setting. Assume that an input random variable x generates a mutual information, I_0. Consider a new input $x_1 = x\mathrm{e}^{j\theta}$ where θ is independent of x and uniformly distributed on $(-\pi, \pi]$. Since the conditional distribution has the property $f_{y|x}(\mathrm{e}^{j\theta}y|\mathrm{e}^{j\theta}x) = f_{y|x}(y|x)$ for any θ, it can be easily seen that mutual information is invariant under deterministic rotations of the input distribution, and hence, $I(x_1; y|\theta) = I_0$. By the concavity of the mutual information over input distributions, $I(x_1; y) \geq I_0$ holds. Further, note that input constraints (2) and (3) are also invariant to rotation. Therefore, there is no loss in optimality in considering input random variables that has uniformly distributed phase independent of the amplitude. ■

Note that if $|m| = 0$, we have a Rayleigh fading channel where the phase cannot be used to convey information when the channel is unknown. However, if there is a LOS component, i.e., $|m| > 0$, phase can indeed carry information, and by Proposition 1, the uniform distribution maximizes the transmission rate.

With this characterization, the optimization problem (4) has been reduced to optimal selection of the distribution function of the input amplitude $F_{|x|}(\cdot)$ under the constraints $E\{|x|^2\} \leq P_{\mathrm{av}}$ and $E\{|x|^4\} \leq \kappa P_{\mathrm{av}}^2$. For the sake of simplification in the notation, we introduce new random variables $R = (1/N_0)|y|^2$ and $r = (\gamma/\sqrt{N_0})|x|$. Assuming a uniform input phase that is independent of the input magnitude, the mutual information in nats, after a straightforward transformation, can be expressed as

$$I(F_r) \stackrel{\mathrm{def}}{=} I(x; y) = -\int_0^\infty f_R(R; F_r) \ln f_R(R; F_r)\, \mathrm{d}R - \int_0^\infty \ln(1 + r^2)\, \mathrm{d}F_r(r) - 1 \tag{6}$$

where $f_R(R; F_r) = \int_0^\infty g(R, r)\mathrm{d}F_r(r)$ is the density function of R with the kernel given by $g(R, r) = (1/(1 + r^2)) \exp(-(R+\mathsf{K}r^2)/(1+r^2)) I_0((2\sqrt{\mathsf{K}r}\sqrt{R})/(1+r^2))$. Furthermore, F_r is the distribution function of r and $\mathsf{K} = (|m|^2/\gamma^2)$ is the Rician factor. Now, the capacity formula can be recasted as

$$C(\alpha, \kappa, \mathsf{K}) = \sup_{\substack{F_r \\ E\{r^2\} \leq \alpha \\ E\{r^4\} \leq \kappa\alpha^2}} I(F_r) \tag{7}$$

depending on three parameters, namely: $\alpha = \gamma^2(P_{\mathrm{av}}/N_0)$, the normalized SNR; κ; and K, the Rician factor. In [1], the existence of an optimal amplitude distribution achieving the supremum in (7) is shown and the following sufficient and necessary condition for an amplitude distribution to be optimal is derived employing the techniques used in [8].

Proposition 2 (Kuhn–Tucker Condition): For the Rician channel (1) with input constraints (2) and (3), F_0 is a capacity-achieving amplitude distribution if and only if there exist λ_1, $\lambda_2 \geq 0$ such that the following is satisfied

$$\int_0^\infty g(R, r) \ln f_R(R, F_0)\, \mathrm{d}R + \ln(1 + r^2) + \lambda_1(r^2 - \alpha) + \lambda_2(r^4 - \kappa\alpha^2) + C + 1 \geq 0 \qquad \forall r \geq 0 \tag{8}$$

with equality if $r \in E_0$ where E_0 is the set of points of increase[3] of F_0.

Note that in the above formulation, λ_1 and λ_2 are the Lagrange multipliers for the second and fourth moment constraints, respectively. Using Proposition 2, the following result on the optimal amplitude distribution is derived.

Theorem 1: For the Rician fading channel (1) with input amplitude constraints (2) and (3), if the fourth moment constraint (3) is active, then the capacity-achieving input amplitude distribution is discrete with a finite number of mass points.

Proof: The result is shown by contradiction. The proof can be summarized as follows:

1) First the assumption that the optimal distribution has an infinite number of points of increase on a bounded interval is contradicted.
2) Next, the assumption that the optimal distribution has an infinite number of points of increase (mass points) but only finitely many of them on any bounded interval is contradicted.
3) Ruling out the above assumptions leaves us with the only possibility that the optimal input has a finite number of mass points.

Assume F_0 is an optimal amplitude distribution. To prove the theorem, a lower bound on the left hand side (LHS) of the Kuhn–Tucker condition (8) is first obtained. To that end, f_R is first bounded as

$$f_R(R; F_0) = \int_0^\infty g(R, r)\, \mathrm{d}F_0(r) \geq \int_0^\infty \frac{1}{1 + r^2} \exp\left(-\frac{R + \mathsf{K}r^2}{1 + r^2}\right) \mathrm{d}F_0(r) \tag{9}$$

$$\geq \exp(-R) \int_0^\infty \frac{1}{1 + r^2} \exp\left(-\frac{\mathsf{K}r^2}{1 + r^2}\right) \mathrm{d}F_0(r) \tag{10}$$

$$= D_{F_0} \exp(-R) \qquad \forall R \geq 0 \tag{11}$$

where $0 < D_{F_0} \leq 1$ is a constant depending on F_0. The first inequality is obtained from the fact that $I_0(x) \geq 1\ \forall x \geq 0$ and the second inequality comes from observing that $\mathrm{e}^{-R/(1+r^2)} \geq \mathrm{e}^{-R}\ \forall R,\ r \geq 0$. Using the lower bound (11), and noting that

[3] The set of points of increase of a distribution function F is $\{r : F(r - \epsilon) < F(r + \epsilon) \forall \epsilon > 0\}$.

$g(R, r)$ is a noncentral chi-square density function in R, the following bound on the LHS of the Kuhn–Tucker condition (8) is derived

$$\text{LHS} \geq \ln D_{F_0} - 1 - (1+\mathsf{K})r^2 + \ln(1+r^2) + \lambda_1(r^2-\alpha) + \lambda_2(r^4-\kappa\alpha^2) + 1 + C \qquad \forall r \geq 0.$$

If the fourth moment constraint is active, i.e., $\lambda_2 > 0$, then for all $D_{F_0} > 0$ and $\lambda_1 \geq 0$, the above lower bound diverges to infinity as $r \to \infty$. Now, we establish contradiction in the following assumptions.

1) Assume that the optimal distribution F_0 has an infinite number of points of increase on a bounded interval. Note that this assumption is satisfied by continuous distributions. The LHS of the Kuhn–Tucker condition (8) is first extended to the complex domain

$$\Phi(z) = \int_0^\infty g(R, z) \ln f_R(R, F_0)\, \mathrm{d}R + \ln(1+z^2) + \lambda_1(z^2-\alpha) + \lambda_2(z^4-\kappa\alpha^2) + C + 1 \quad (12)$$

where $z \in \mathbb{C}$. By the Differentiation Lemma [33, Ch. 12], it is easy to see that Φ is analytic in the region where $\mathrm{Re}\{1+z^2\} > 0$ [1, Appendix C]. This choice of region guarantees the uniform convergence of the integral expression in (12). Note that in this region, by the earlier assumption, $\Phi(z) = 0$ for an infinite number of points having a limit point.[4] By the Identity Theorem,[5] $\Phi(z) = 0$ in the region where $\mathrm{Re}\{1+z^2\} > 0$. Hence, the Kuhn–Tucker condition (8) is satisfied with equality for all $r \geq 0$. Clearly, this case is not possible from the above lower bound, which diverges to infinity as $r \to \infty$.

2) Next, assume that the optimal distribution has an infinite number of mass points but only finitely many of them on any bounded interval. Then the LHS of the Kuhn–Tucker condition should be equal to zero infinitely often as $r \to \infty$, which is again not possible by the above diverging lower bound.

Hence, the optimal distribution must be discrete with a finite number of mass points and the theorem follows. ■

The significance of Theorem 1 comes from the fact that for the Rician fading channel, any fourth moment constraint with a finite κ will eventually be active for sufficiently small SNR because, as observed in [3, Sec. V.E], if there is no such constraint, the required value of κ grows without bound as $\mathrm{SNR} \to 0$. Therefore, Theorem 1 establishes the discrete nature of the optimal input in the low-power regime. Furthermore, Theorem 1 easily specializes to the Rayleigh and the unfaded Gaussian channels. For the unfaded Gaussian channel, if the fourth moment constraint is inactive, it is well known that a Rayleigh-distributed amplitude is optimal and it has a kurtosis $\kappa = 2$. Therefore, the fourth moment constraint being active (i.e., $1 < \kappa < 2$) is also a necessary condition for the discrete nature in that case. In [8], the optimal amplitude for the average-power-limited noncoherent Rayleigh fading channel is shown to be discrete with a finite number of levels over the entire SNR range. Theorem 1 proves that this discrete character does not change when an additional fourth moment constraint is derived.

We note that the key property which leads to the proof of Theorem 1 is that we have a moment constraint higher than the second moment. Hence, the result of Theorem 1 holds in a more general setting where the fourth moment constraint (3) is replaced by a constraint in the following form: $E\{|x|^{2+\delta}\} \leq M$ for some $\delta > 0$ and $M < \infty$. In a related work, Palanki [23] has independently shown the discrete character of the optimal input for a general type of fading channels when only moment constraints strictly higher than the second moment are imposed.

B. Peak-Power-Limited Input

In this section, it is assumed that the input amplitude is subject to only a peak-power limitation

$$|x_i|^2 \overset{\text{a.s.}}{\leq} P \qquad \forall i. \quad (13)$$

Although being more stringent than the fourth moment limitation, peak-power constraint is more relevant in practical systems. For instance, efficient use of battery power in portable radio units and linear operation of RF amplifiers employed at the transmitter require peak-power-limited communication schemes.

Since the peak constraint (13) is invariant to the rotation of the input, optimality of uniform phase follows from Proposition 1. Hence, our primary focus is on obtaining a characterization for the optimal amplitude distribution. Existence of a capacity-achieving amplitude distribution readily follows from the results on the second and fourth moment limited case. Similarly, $R = (1/N_0)|y|^2$ and $r = (\gamma/\sqrt{N_0})|x|$ are defined. Specializing (8), the following sufficient and necessary condition for an amplitude distribution F_0 to be optimal over the peak-power-limited Rician channel is easily obtained

$$\int_0^\infty g(R, r) \ln f_R(R, F_0)\, \mathrm{d}R + \ln(1+r^2) + C + 1 \geq 0 \qquad \forall r \in [0, \sqrt{\alpha}] \quad (14)$$

with equality if $r \in E_0$ where E_0 is the set of points of increase of F_0. Note that $\alpha = \gamma^2(P/N_0)$. Next, the main result on the optimal amplitude distribution is stated.

Theorem 2: For the Rician fading channel (1) where the input is subject to only a peak-power constraint $|x|^2 \overset{\text{a.s.}}{\leq} P$, the capacity-achieving amplitude distribution is discrete with a finite number of mass points.

Proof: Since the input is subject to a peak constraint, the result in this case is established by contradicting the assumption

[4] The Bolzano-Weierstrass Theorem [31] states that every bounded infinite set of real numbers has a limit point.

[5] The Identity Theorem for analytic functions [32] states that if two functions are analytic in a region $\mathcal{R}$, and if they coincide in a neighborhood, however small, of a point z_0 of $\mathcal{R}$, or only along a path segment, however small, terminating in z_0, or also only for an infinite number of distinct points with the limit point z_0, then the two functions are equal everywhere in $\mathcal{R}$.

that the optimal input distribution has an infinite number of points of increase on a bounded interval.

Assume F_0 is an optimal distribution. To prove the theorem, an upper bound on the LHS of (14) is first obtained. To achieve this goal, $f_R(\cdot, F_0)$ is bounded as

$$\begin{aligned} f_R(R;F_0) &= \int_0^{\sqrt{\alpha}} \frac{1}{1+r^2} \exp\left(-\frac{R+\mathsf{K}r^2}{1+r^2}\right) \\ &\qquad \times I_0\left(\frac{2\sqrt{\mathsf{K}}r\sqrt{R}}{1+r^2}\right) \mathrm{d}F_0(r) \\ &\le \exp\left(-\frac{R}{1+\alpha} + \sqrt{\mathsf{K}R}\right) \\ &\qquad \times \int_0^{\sqrt{\alpha}} \frac{1}{1+r^2} \exp\left(-\frac{\mathsf{K}r^2}{1+r^2}\right) \mathrm{d}F_0(r) \qquad (15) \\ &= D_{F_0} \exp\left(-\frac{R}{1+\alpha} + \sqrt{\mathsf{K}R}\right) \qquad (16) \end{aligned}$$

where $0 < D_{F_0} \le 1$ is a constant depending on F_0. The upper bound in (15) is easily verified by observing $\exp(-(R/(1+r^2)) \le \exp(-(R/(1+\alpha)))$ $\forall r \le \sqrt{\alpha}$ and $I_0((2\sqrt{\mathsf{K}}r\sqrt{R})/(1+r^2)) \le I_0(\sqrt{\mathsf{K}R}) \le \exp(\sqrt{\mathsf{K}R})$ $\forall r \ge 0$. Using (16), the following upper bound is derived

$$\begin{aligned} &\int_0^{\infty} g(R,r) \ln f_R(R;F_0)\, \mathrm{d}R \\ &\le \ln D_{F_0} - \frac{1}{1+\alpha} \int_0^{\infty} g(R,r) R\, \mathrm{d}R + \sqrt{\mathsf{K}} \int_0^{\infty} g(R,r) \sqrt{R}\, \mathrm{d}R \qquad (17) \\ &\le \ln D_{F_0} - \frac{1+(\mathsf{K}+1)r^2}{1+\alpha} + \sqrt{\mathsf{K}}\sqrt{1+(1+\mathsf{K})r^2} \\ &\qquad \forall r \ge 0. \qquad (18) \end{aligned}$$

The upper bound in (18) follows from the fact that $g(R,r)$ is a noncentral chi-square probability density function in R, and $\int_0^\infty g(R,r) R\ \mathrm{d}R = 1 + (1+\mathsf{K})r^2$ and $\int_0^\infty g(R,r)\sqrt{R}\ \mathrm{d}R \le \sqrt{1+(1+\mathsf{K})r^2}$, which follows from the concavity of $\sqrt{x}$ and the Jensen's inequality. From (18), the following upper bound on the LHS of (14) is obtained

$$\begin{aligned} \mathrm{LHS} \le \ln D_{F_0} - \frac{1+(\mathsf{K}+1)r^2}{1+\alpha} + \sqrt{\mathsf{K}}\sqrt{1+(1+\mathsf{K})r^2} \\ + \ln(1+r^2) + C + 1 \qquad \forall r \ge 0. \qquad (19) \end{aligned}$$

Using the above upper bound, we show that the following assumption cannot hold true.

Assume that the optimal input distribution F_0 has an infinite number of points of increase on a bounded interval. Next, the LHS of (14) is extended to the complex domain

$$\Phi(z) = \int_0^{\infty} g(R,z) \ln f_R(R,F_0)\, \mathrm{d}R + \ln(1+z^2) + C + 1 \qquad (20)$$

where $z \in \mathbb{C}$. Since the condition in (14) should be satisfied with equality at the points of increase of the optimal input distribution, by the above assumption, $\Phi(z) = 0$ for an infinite number of points having a limit point. Then, by the Identity Theorem [32], $\Phi(z) = 0$ in the whole region where it is analytic. By the Differentiation Lemma [33, Ch. 12], one can easily verify that $\Phi(z)$ is analytic in the region where $\mathrm{Re}(1+z^2) > 0$, which includes the positive real line. Therefore, it is concluded that $\Phi(r) = 0$ $\forall r \ge 0$. Clearly, this is not possible from the upper bound in (19), which diverges to $-\infty$ as $r \to \infty$ for any finite α, $\mathsf{K} \ge 0$, and $D_{F_0} > 0$.

Reaching a contradiction, it is concluded that the optimal distribution must be discrete with a finite number of mass points. ■

We note that Theorem 2 establishes the discrete structure of the optimal input distribution over the entire SNR range. The proof basically uses the observation that a bounded input induces an output probability density function that decays at least exponentially and, in turn, provides a diverging bound on the Kuhn–Tucker condition. We also note recent independent work by Huang and Meyn [14], where the discrete nature of the optimal input is proven by again showing a diverging bound on the Kuhn–Tucker condition for a general class of channels in which the input is subject only to peak amplitude constraints.

IV. Rician Fading Channel With Phase Noise

In this section, we deviate from the classical Rician channel model (1) where the specular component is assumed to be static and consider the following model

$$y_i = (a_i + m\mathrm{e}^{j\,\theta_i})x_i + n_i \qquad (21)$$

where phase noise is introduced in the specular component. Here, $\{\theta_i\}$ is assumed to be a sequence of i.i.d. uniform random variables on $[-\pi, \pi)$ and m is a deterministic complex constant. Again, the noncoherent scenario where $\{a_i\}$ and $\{\theta_i\}$ are known by neither the receiver nor the transmitter is considered. This model is relevant in mobile systems where rapid random changes in the phase of the specular component are not tracked. Moreover, such a model is suitable in cases where there is imperfect receiver side information about the fading magnitude. As another departure from the previous section, here, only an average power constraint $E\{|x|^2\} \le P_{\mathrm{av}}$ is imposed. The discrete nature of the optimal input amplitude follows immediately from the techniques of Section III when there is an additional higher moment constraint.

It is immediately realized that the channel output y is conditionally Gaussian given x and θ

$$f_{y|x,\theta}(y|x,\theta) = \frac{1}{\pi\left(\gamma^2|x|^2+N_0\right)} \exp\left(-\frac{\left|y-me^{j\theta}x\right|^2}{\gamma^2|x|^2+N_0}\right). \tag{22}$$

Integrating (22) over uniform θ, we obtain the conditional distribution of the channel output given the input

$$f_{y|x}(y|x) = \frac{1}{\pi\left(\gamma^2|x|^2+N_0\right)} \exp\left(-\frac{|y|^2+|m|^2|x|^2}{\gamma^2|x|^2+N_0}\right) \times I_0\left(\frac{2|m||y||x|}{\gamma^2|x|^2+N_0}\right). \tag{23}$$

Again the following random variables are introduced: $R = (1/N_0)|y|^2$ and $r = (\gamma/\sqrt{N_0})|x|$. Since the phase information is completely destroyed in the channel (21) and the above transformations are one to one, we have $I(x;y) = I(|x|;|y|) = I(r;R)$. Furthermore, the conditional distribution of R given r is easily obtained from (23)

$$f_{R|r}(R|r) = \frac{1}{1+r^2} \exp\left(-\frac{R+\mathsf{K}r^2}{1+r^2}\right) I_0\left(\frac{2\sqrt{\mathsf{K}}r\sqrt{R}}{1+r^2}\right) \tag{24}$$

where $\mathsf{K} = (|m|^2/\gamma^2)$ is the Rician factor. Similarly, as in the previous section, the existence of an optimal amplitude distribution is shown, and the following sufficient and necessary condition is given in [1].

Proposition 3 (Kuhn–Tucker Condition): For the Rician channel model (21) with an average power constraint $E\{|x|^2\} \le P_{\text{av}}$, F_0 is a capacity-achieving amplitude distribution if and only if there exists $\lambda \ge 0$ such that the following is satisfied

$$-D\left(f_{R|r}\|f_R\right) + \lambda(r^2-\alpha) + C \ge 0 \qquad \forall r \ge 0 \tag{25}$$

with equality if $r \in E_0$ where E_0 is the set of points of increase of F_0. In the above formulation, $D(\cdot\|\cdot)$ is the divergence (e.g. [30, Sec. 2.3]), $f_R(R;F_0) = \int_0^\infty f_{R|r}(R|r)\,\mathrm{d}F_0(r)$ is the density function of R, $\alpha = \gamma^2\text{SNR}$, and C is the capacity.

The next theorem gives the main result on the structure of the optimal input for the Rician fading channel with phase noise.

Theorem 3: For the Rician fading channel with uniform phase noise (21) and average power constraint $E\{|x|^2\} \le P$, the capacity-achieving input amplitude distribution is discrete with a finite number of mass points.

Proof: The result is shown by contradiction. Let F_0 be an optimal amplitude distribution. The proof can be summarized as follows:

1) First, it is assumed that F_0 has an infinite number of points of increase on a bounded interval. The impossibility of this case is shown by contravening the fact that under this assumption, the LHS of the Kuhn–Tucker condition, which is extended to the complex domain, is identically zero over its region of analyticity.
2) Then, it is assumed that the optimal distribution is discrete with an infinite number of mass points but only finitely many of them on any bounded interval. This assumption is also ruled out by finding a diverging lower bound on the LHS of the Kuhn–Tucker condition.
3) Having eliminated the above assumptions, what remains is the only possibility that the optimal distribution is discrete with a finite number of mass points.

Assumption 1: Assume that F_0 has an infinite number of points of increase on a bounded interval. Then, the Kuhn–Tucker condition (25) is satisfied with equality at an infinite number of points having a limit point. First, we extend the LHS of (25) to the complex domain: $\Psi(z) = -D(f_{R|r=z}\|f_R) + \lambda(z^2-\alpha) + C$, where $z \in \mathbb{C}$. Equivalently, using the fact that $f_{R|r}$ is a noncentral chi-square density function in R

$$\begin{aligned}\Psi(z) = &\int_0^\infty f_{R|r}(R|z)\ln f_R(R)\,\mathrm{d}R \\ &- \int_0^\infty f_{R|r}(R|z)\ln\left(I_0\left(\frac{2\sqrt{\mathsf{K}}z\sqrt{R}}{1+z^2}\right)\right)\mathrm{d}R \\ &+ \lambda(z^2-\alpha) + \ln(1+z^2) + \frac{2\mathsf{K}z^2}{1+z^2} + C + 1.\end{aligned} \tag{26}$$

By the Differentiation Lemma [33, Ch. 12], it is easy to see that Ψ is analytic in the region where $\text{Re}\{1+z^2\} > 0$ and $\text{Re}\{z\} > 0$. The first condition guarantees the uniform convergence of the integrals in (26) by forcing the integrands to decrease exponentially. Since I_0 has zeros on the imaginary axis and the second integral in (26) involves the logarithm of the Bessel function, with the second condition, $\text{Re}\{z\} > 0$, the imaginary axis is excluded from the region of analyticity. Since $\Psi(z) = 0$ for an infinite number of points having a limit point, by the Identity Theorem [32], $\Psi(z) = 0$ in the whole region where it is analytic. In particular, $\Psi(jb + (1/n)) = 0$ is obtained for all $|b| < 1$ and $n \in \mathbb{Z}^+$, and hence, $\lim_{n\to\infty} \Psi(jb + (1/n)) = 0\ \ \forall |b| < 1$. All the terms other than the second term in (26) are analytic also on the imaginary axis with $|z| < 1$ and the limiting expression is obtained by letting $(1/n) \to 0$ in the arguments of these functions. For the second term in (26), the Dominated Convergence Theorem [31] needs to be invoked to justify the interchange of limit and integral. An integrable upper bound on the magnitude of the integrand of the second integral is shown in [1, Appendix D]. Therefore,

$$\begin{aligned}\lim_{n\to\infty} \Psi\left(jb+\frac{1}{n}\right) &= \int_0^\infty f_{R|r}(R|jb)\ln f_R(R)\,\mathrm{d}R \\ &\quad - \int_0^\infty \lim_{n\to\infty} f_{R|r}\left(R\,\middle|\,jb+\frac{1}{n}\right) \\ &\quad \times \ln\left(I_0\left(\frac{2\sqrt{\mathsf{K}}\left(jb+\frac{1}{n}\right)\sqrt{R}}{1+\left(jb+\frac{1}{n}\right)^2}\right)\right)\mathrm{d}R \\ &\quad + \lambda(-b^2-\alpha) + \ln(1-b^2) - \frac{2\mathsf{K}b^2}{1-b^2} + C + 1 \\ &= 0 \qquad \forall |b| < 1.\end{aligned} \tag{27}$$

Note that all the terms other than the second term in (27) are real. Next, we show that the second term in (27) has a nonzero imaginary component yielding a contradiction. First, the limit in the integrand is evaluated in (28) (see equation at the bottom of the page), which is obtained easily by observing that all the terms are analytic in the entire complex plane (excluding $z = \pm j$) except the logarithm function, which is not analytic at the zeros of the Bessel function. However, as we approach the zeros of the Bessel function, $I_0 \ln I_0 \to 0$, and hence, we obtain (28). Noting that $J_0(z) = I_0(jz)$ where J_0 is the zeroth-order Bessel function of the first kind, and that the set of zeros of the I_0 function along the imaginary axis has a measure of zero, the second integral in (27) can now be expressed as

$$\int_0^\infty \frac{1}{1-b^2} \exp\left(-\frac{R - \mathsf{K}b^2}{1-b^2}\right) J_0\left(\frac{2\sqrt{\mathsf{K}}b\sqrt{R}}{1-b^2}\right) \times \ln\left(J_0\left(\frac{2\sqrt{\mathsf{K}}b\sqrt{R}}{1-b^2}\right)\right) \mathrm{d}R. \quad (29)$$

By applying a change of variables $v = (2\sqrt{\mathsf{K}}b\sqrt{R})/(1-b^2)$ and expressing $\ln z = \ln|z| + j\arg(z)$, the above integral becomes

$$\int_0^\infty \frac{1-b^2}{2\mathsf{K}b^2} v \exp\left(-\frac{1-b^2}{4\mathsf{K}b^2}v^2 + \frac{\mathsf{K}b^2}{1-b^2}\right) J_0(v) \ln|J_0(v)|\,\mathrm{d}v + j\pi \int_0^\infty \frac{1-b^2}{2\mathsf{K}b^2} v \exp\left(-\frac{1-b^2}{4\mathsf{K}b^2}v^2 + \frac{\mathsf{K}b^2}{1-b^2}\right) J_0(v) h(v)\,\mathrm{d}v. \quad (30)$$

In the above formulation, $h(v) = 0$ if $v \in (0, \alpha_1)$ and $h(v) = k$ if $v \in (\alpha_k, \alpha_{k+1})$ where $\{\alpha_k\}$ are the zeros of J_0. As noted in [13], the second term in (30) arises due to the fact that the logarithm jumps in value by $j\pi$ when a zero is passed. This fact can be observed by considering the argument of $J_0(b - j\epsilon)$ as $\epsilon \to 0$ where $b - j\epsilon$ is in a small neighborhood of α_k such that $J_0(b - j\epsilon) \backsimeq J_0'(\alpha_k)(b - j\epsilon - \alpha_k)$.

Using a similar bound obtained in [13], the following inequality is derived

$$\left|\int_0^\infty \frac{1-b^2}{2\mathsf{K}b^2} v \exp\left(-\frac{1-b^2}{4\mathsf{K}b^2}v^2\right) J_0(v) h(v)\,\mathrm{d}v\right| \geq \left(|J_0(\beta)| - \frac{2\mathsf{K}b^2}{\beta(1-b^2)}\right) \exp\left(-\frac{\beta^2(1-b^2)}{4\mathsf{K}b^2}\right) - \left(1 + \frac{4\mathsf{K}b^2}{\pi\alpha_2(1-b^2)}\right) \exp\left(-\frac{\alpha_2^2(1-b^2)}{4\mathsf{K}b^2}\right) \quad (31)$$

where α_2 is the second smallest zero of J_0 on the positive axis and β is the positive zero of J_1 less than α_2. As noted in [13], the above lower bound is positive for small enough values of b. In particular, for each $\mathsf{K} > 0$, the above lower bound is 9.16×10^{-5} when $b^2 = (1/(2\mathsf{K}+1)) < 1$.

Assumption 2: Next, it is assumed that the optimal distribution has an infinite number of mass points but only finitely many of them on any bounded interval. Following an approach similar to the one used in [8], f_R is first bounded as

$$\begin{aligned} f_R(R; F_0) &= \int_0^\infty f_{R|r}(R|r)\,\mathrm{d}F_0(r) \\ &= \sum_{i=0}^\infty p_i\, f_{R|r}(R|r_i) && (32) \\ &\geq p_i f_{R|r}(R|r_i) && (33) \\ &\geq p_i \frac{1}{1+r_i^2} \exp\left(-\frac{R + \mathsf{K}r_i^2}{1+r_i^2}\right) \quad \forall i\, \forall R \geq 0 && (34) \end{aligned}$$

where p_i and r_i are the probability and location, respectively, of the ith mass point of F_0. The last inequality is obtained by using the fact that $I_0(x) \geq 1\ \forall x \geq 0$. Noting that $f_{R|r}(R|r)$ is a noncentral chi-square density function in R, this bound leads to the following lower bound on the LHS of (25)

$$\begin{aligned} \text{LHS} \geq\ & \ln p_i - \ln\left(1 + r_i^2\right) - \frac{1 + \mathsf{K}r_i^2}{1+r_i^2} - \frac{(\mathsf{K}+1)r^2}{1+r_i^2} \\ & - \int_0^\infty f_{R|r}(R|r) \ln\left(I_0\left(\frac{2\sqrt{\mathsf{K}}r\sqrt{R}}{1+r^2}\right)\right) \mathrm{d}R \\ & + \lambda(r^2 - \alpha) + \ln(1+r^2) + \frac{2\mathsf{K}r^2}{1+r^2} \\ & + C + 1 \qquad \forall i\, \forall r \geq 0. \quad (35) \end{aligned}$$

$$\lim_{n\to\infty} f_{R|r}\left(R \,\middle|\, jb + \frac{1}{n}\right) \ln\left(I_0\left(\frac{2\sqrt{\mathsf{K}}\left(jb + \frac{1}{n}\right)\sqrt{R}}{1 + \left(jb + \frac{1}{n}\right)^2}\right)\right) = \begin{cases} 0, & \text{if } I_0\left(\frac{2\sqrt{\mathsf{K}}jb\sqrt{R}}{1-b^2}\right) = 0 \\ \frac{1}{1-b^2} \exp\left(-\frac{R - \mathsf{K}b^2}{1-b^2}\right) I_0\left(\frac{2\sqrt{\mathsf{K}}jb\sqrt{R}}{1-b^2}\right) \ln\left(I_0\left(\frac{2\sqrt{\mathsf{K}}jb\sqrt{R}}{1-b^2}\right)\right), & \text{otherwise} \end{cases} \quad (28)$$

Noting that $I_0((2\sqrt{\mathsf{K}}r\sqrt{R})/(1+r^2)) \leq \exp((2\sqrt{\mathsf{K}}r\sqrt{R})/(1+r^2))$, the following is derived

$$\begin{aligned}\int_0^\infty & f_{R|r}(R|r)\ln\left(I_0\left(\frac{2\sqrt{\mathsf{K}}r\sqrt{R}}{1+r^2}\right)\right)\mathrm{d}R\\ &\leq \frac{2\sqrt{\mathsf{K}}r}{1+r^2}E\{\sqrt{R}|r\}\\ &\leq \frac{2\sqrt{\mathsf{K}}r}{1+r^2}\sqrt{E\{R|r\}}\\ &= \frac{2\sqrt{\mathsf{K}}r\sqrt{1+(1+\mathsf{K})r^2}}{1+r^2}\\ &\leq 2\sqrt{2\mathsf{K}+\mathsf{K}^2}.\end{aligned}$$

Since it has been assumed that the optimal distribution has an infinite number of mass points with finitely many of them on any bounded interval, it is seen that for any $\lambda > 0$, a sufficiently large r_i can be chosen such that the lower bound (35) diverges to ∞ as $r \to \infty$. However, by our assumption, the LHS of (35) should be equal to zero infinitely often as $r \to \infty$, which is a contradiction. $\lambda = 0$ implies that the power constraint is ineffective. The impossibility of this case is shown in [8]. Therefore, the theorem follows. ■

Recent results [8]–[10] have shown the discrete nature of the optimal distribution for the two special cases of the model (21): 1) the Rayleigh fading channel; and 2) the noncoherent AWGN channel. We have proven the discreteness of the capacity-achieving distribution in a unifying setting where there is both multipath fading and a specular component with random phase. For the noncoherent AWGN channel, Katz and Shamai [9], [10] have shown that the optimal input has an infinite number of mass points. An interesting conclusion of Theorem 3 is that the presence of an unknown multipath component induces an optimal distribution with a finite number mass points. It is also of interest to consider the classical average-power-limited Rician fading channel (1) for which tight upper and lower bounds on the capacity were derived in [18]. By Proposition 1, it is known that uniform phase is optimal for this model. Moreover, we derive the following partial result on the optimal amplitude distribution, which proves the suboptimality of input amplitude distributions with unbounded support such as the Rayleigh distribution.

Theorem 4: For the Rician fading channel (1) with only an average power limitation $E\{|x|^2\} \leq P_{\mathrm{av}}$, the optimal input amplitude distribution has bounded support.

Proof: Assume that F_0 is an optimal distribution. The proposition will be proven by contradiction. Therefore, it is further assumed that F_0 has unbounded support. With this assumption, for any finite $M \geq 0$

$$\begin{aligned}f_R(R;F_0) &= \int_0^\infty g(R,r)\,\mathrm{d}F_0(r)\\ &\geq \int_0^\infty \frac{1}{1+r^2}\exp\left(-\frac{R+\mathsf{K}r^2}{1+r^2}\right)\mathrm{d}F_0(r) \qquad (36)\\ &\geq \int_M^\infty \frac{1}{1+r^2}\exp\left(-\frac{R+\mathsf{K}r^2}{1+r^2}\right)\mathrm{d}F_0(r) \qquad (37)\\ &\geq \exp\left(-\frac{R}{1+M^2}\right)\\ &\quad\times\int_M^\infty \frac{1}{1+r^2}\exp\left(-\frac{\mathsf{K}r^2}{1+r^2}\right)\mathrm{d}F_0(r) \qquad (38)\\ &= D_{F_0,M}\exp\left(-\frac{R}{1+M^2}\right) \quad \forall R\geq 0,\ \forall M\geq 0 \qquad (39)\end{aligned}$$

where $0 < D_{F_0,M} \leq 1\ \forall M \geq 0$ and $\forall F_0$. Using (39), the following lower bound on the LHS of the Kuhn–Tucker condition is obtained[6]

$$\begin{aligned}\mathrm{LHS} \geq \ln D_{F_0,M} - \frac{1}{1+M^2} - \frac{1+\mathsf{K}}{1+M^2}r^2 + \ln(1+r^2)\\ +\lambda(r^2-\alpha)+1+C \qquad \forall r\geq 0,\ \forall M \geq 0.\end{aligned}$$

For any $\lambda > 0$[7] and $D_{F_0,M} > 0$, a sufficiently large M can be chosen such that the above lower bound diverges to infinity as $r \to \infty$. However, if the optimal input has unbounded support, the LHS of the Kuhn–Tucker condition should be zero infinitely often as $r \to \infty$. This constitutes a contradiction, and hence, the theorem follows. ■

V. Numerical Results

In general, the number of mass points of the optimal discrete distribution and their locations and probabilities depend on the SNR. Analytical expressions for the capacity and the optimal distribution as a function of SNR are unlikely to be feasible. Therefore, we resort to numerical methods to examine this behavior. The numerical algorithm used here is similar to the ones employed in [6] and [8]. In particular, we start with a sufficiently small SNR and maximize the mutual information over the set of two-mass-point discrete distributions satisfying the input constraints. Then, the maximizing two-mass-point discrete distribution is tested with the Kuhn–Tucker condition. If this distribution satisfies the necessary and sufficient Kuhn–Tucker condition, then it is optimal and the mutual information achieved by it is the capacity. As the SNR is increased, the required number of mass points monotonically increases, and therefore, to obtain the optimum distribution, the same procedure is repeated for discrete distributions with increasing numbers of mass points.[8]

For the Rician fading channel $(\mathsf{K} > 0)$ with second and fourth moment input constraints, numerical results indicate that

[6]The Kuhn–Tucker condition for the average-power-limited Rician case is essentially the same as (8) with $\lambda_2 = 0$.

[7]The impossibility of $\lambda = 0$ is shown in [8].

[8]We note that the results in this section are obtained numerically, and hence, there is no analytical claim of optimality.

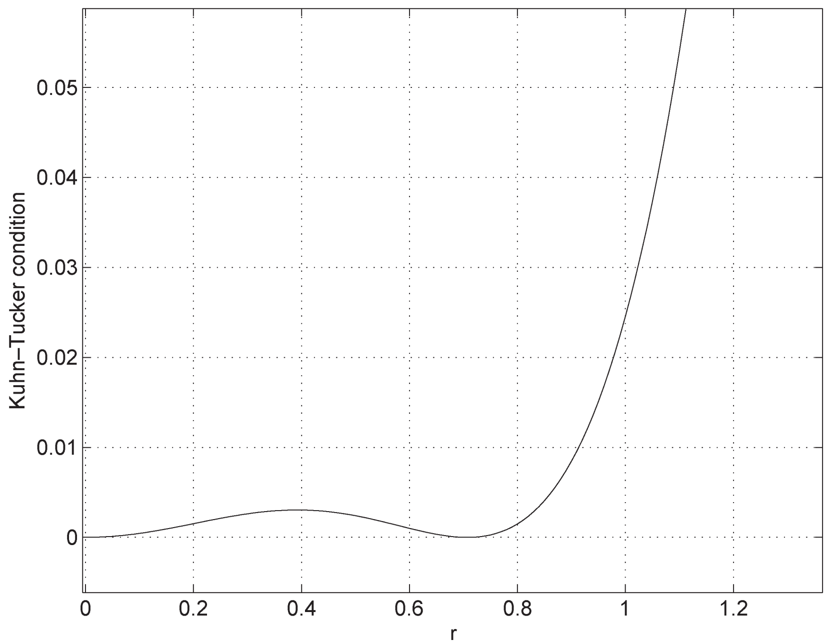


Fig. 1. Kuhn–Tucker condition for $\mathsf{K} = 1$, $\alpha = 0.05$ and $\kappa = 10$. $F(r) = 0.9u(r) + 0.1u(r - 1/\sqrt{2})$ and $C = 0.0531$. $\lambda_1 = 0.89106$, $\lambda_2 = 0.15135$.

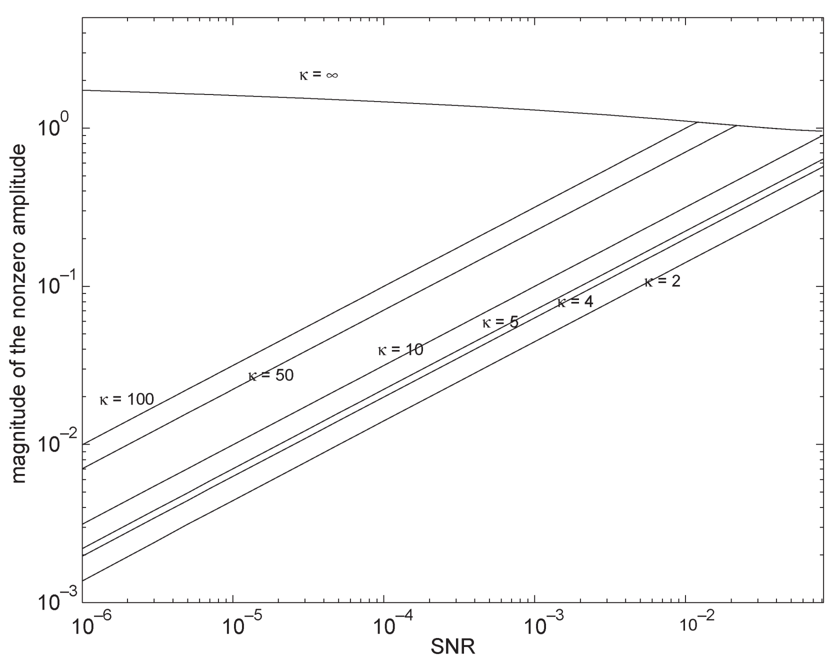


Fig. 2. Location of the second mass point versus normalized $\mathrm{SNR} = \gamma^2(P_{\mathrm{av}}/N_0)$ in the Rician channel $\mathsf{K} = 1$.

for sufficiently small SNR values, the two-mass-point discrete amplitude distribution

$$F(|x|) = \left(1 - \frac{1}{\kappa}\right) u(|x|) + \frac{1}{\kappa} u\left(|x| - \sqrt{\kappa N_0 \mathrm{SNR}}\right) \quad (40)$$

is optimal. Note that this distribution does not depend on the Rician factor K. Fig. 1 plots the LHS of the Kuhn–Tucker condition (8) as a function of r for the distribution $F(r) = 0.9u(r) + 0.1u(r - 1/\sqrt{2})$ for the Rician fading channel $(\mathsf{K} = 1)$ with $\alpha = 0.05$ and $\kappa = 10$. From the figure, it is seen that the Kuhn–Tucker condition is satisfied and the optimal distribution is in the form given by (40). Figs. 2 and 3 plot the magnitude and the probability of the nonzero amplitude, respectively, as a function of SNR $(N_0 = 1)$ for various values of κ. The significant impact of imposing a fourth moment constraint is noticed immediately. When there is no such constraint, the nonzero amplitude migrates away from the origin as $\mathrm{SNR} \to 0$ while its probability decreases sufficiently fast to satisfy the average power constraint. This type of input is called flash signaling in [3]. However, as we observe in Figs. 2 and 3, if there is a fourth moment constraint with a finite κ, then the behavior is quite different. The nonzero amplitude approaches

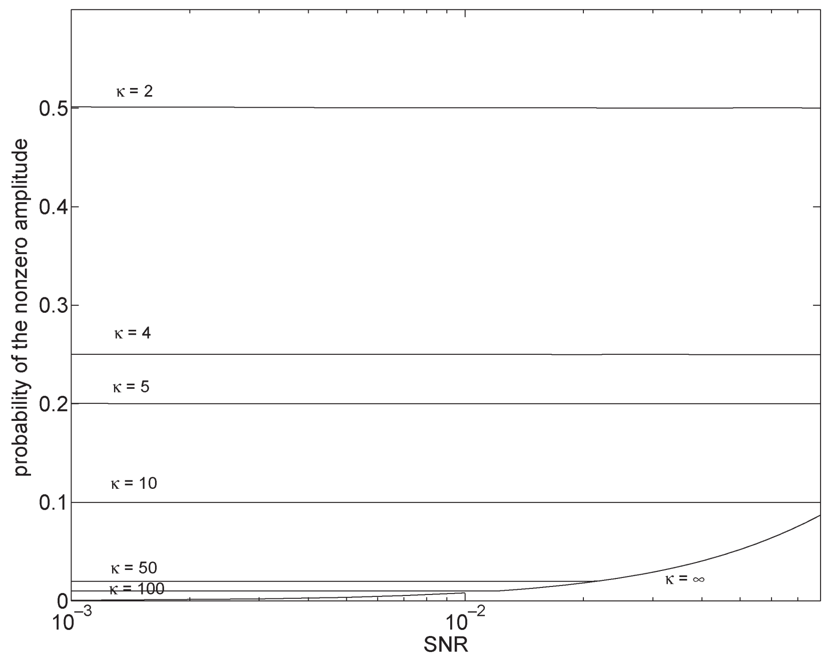


Fig. 3. Probability of the second mass point versus normalized $\mathrm{SNR} = \gamma^2(P_{\mathrm{av}}/N_0)$ in the Rician channel $\mathsf{K} = 1$.

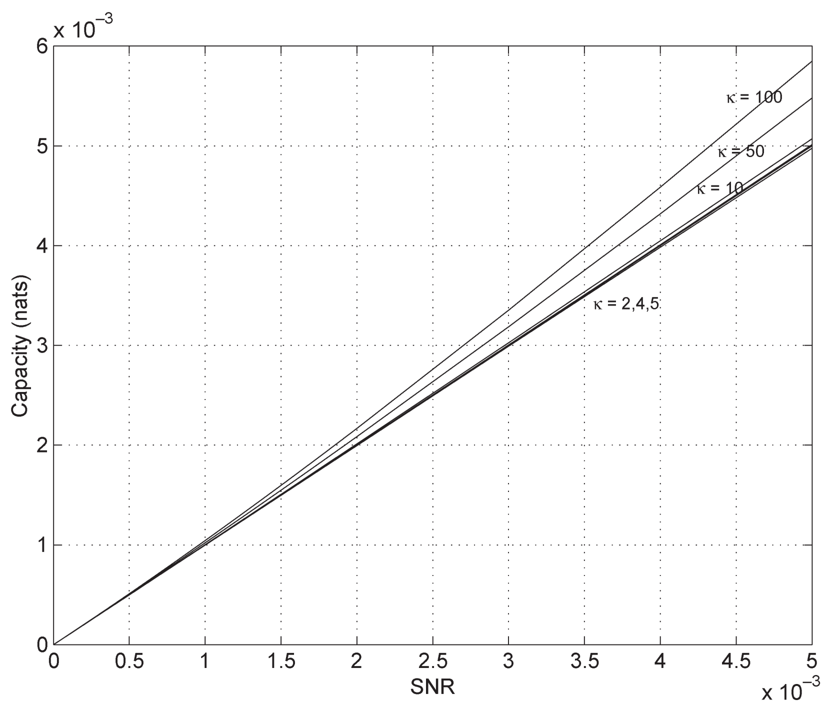


Fig. 4. Capacity (in nats) versus normalized $\mathrm{SNR} = \gamma^2(P_{\mathrm{av}}/N_0)$ for the Rician channel $(\mathsf{K} = 1)$ subject to fourth moment constraints with $\kappa = 2$, 4, 5, 10, 50, 100.

the origin as $\mathrm{SNR} \rightarrow 0$, while its probability is kept constant. In the Rayleigh channel $(\mathsf{K} = 0)$, (40) is still optimal at low SNR up to a point after which, as SNR is further lowered, the second moment constraint becomes inactive, and it is observed that the nonzero mass point approaches the origin more slowly while its probability decreases. From Fig. 4, which plots the capacity curves as a function SNR for various values of κ in the low-power regime, it is seen that all the curves have the same first derivative at zero SNR. This may suggest that performance in the low-power regime is similar for any finite value of κ. However, as shall be seen in [2], the picture radically changes when the spectral-efficiency/bit-energy tradeoff is investigated.

For the peak-power-limited Rician fading channel $(\mathsf{K} > 0)$, numerical results indicate that for sufficiently low SNR values, the optimal amplitude distribution has a single mass at the peak level $\sqrt{P}$, and hence, all the information is carried on the uniform phase. For the Rayleigh channel $(\mathsf{K} = 0)$, an equiprobable two-mass-point distribution where one mass is at the origin and the other mass at the peak level is capacity achieving in

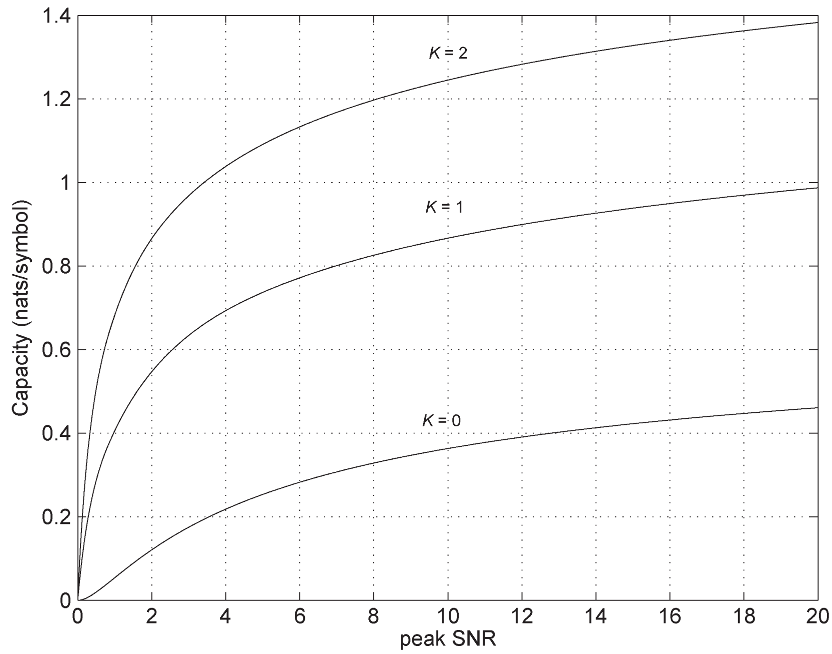


Fig. 5. Capacity curves as a function of the normalized peak $\mathrm{SNR} = \gamma^2(P/N_0)$ for the peak-power-limited Rayleigh channel $(\mathsf{K} = 0)$ and Rician channels with $\mathsf{K} = 1, 2$.

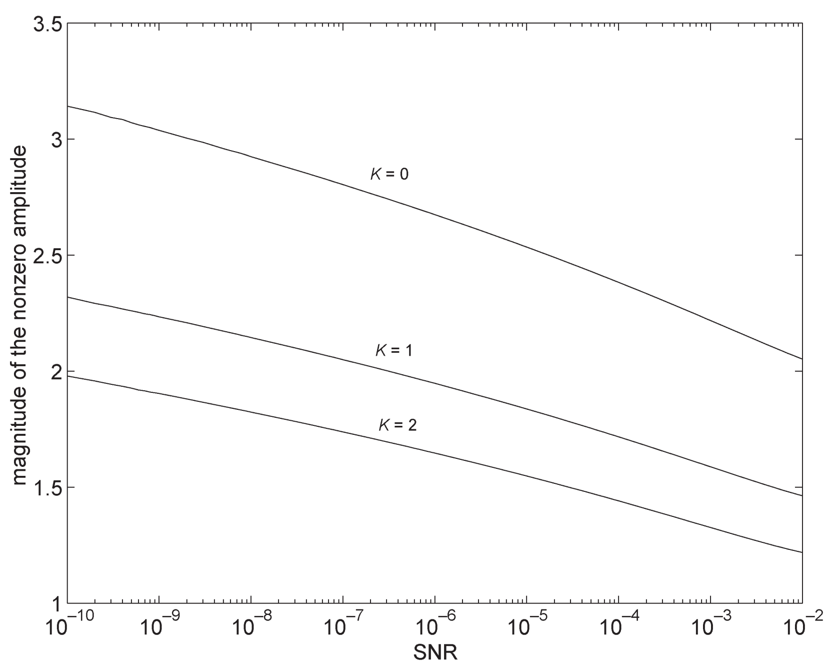


Fig. 6. Magnitude of the nonzero amplitude versus normalized $\mathrm{SNR} = \gamma^2(P_{\mathrm{av}}/N_0)$ for the Rician channel with phase noise $\mathsf{K} = 0, 1, 2$.

the low-power regime. Fig. 5 plots the capacity curves for the peak-power-limited Rayleigh channel and Rician channels with $\mathsf{K} = 1, 2$ as a function of the peak SNR. Note that the Rayleigh channel capacity curve has a zero slope at zero SNR.

For the Rician fading channel with phase noise (21), numerical results illustrate again that a two-mass-point discrete distribution is optimal for sufficiently small SNR values. Figs. 6 and 7 plot the magnitude and probability, respectively, of the optimal nonzero amplitude for this channel with Rician factors $\mathsf{K} = 0, 1, 2$. Note that only an average power constraint is imposed here. It is observed that flash-signaling-type optimal input, where the nonzero amplitude migrates away from the origin as $\mathrm{SNR} \to 0$ while its probability is decreasing, is required in the low-power regime. For a fixed SNR, it is also seen that the nonzero amplitude is closer to the origin for higher Rician factors K. Finally, Fig. 8 provides the capacity curves as a function of SNR for Rician factors $\mathsf{K} = 0, 1, 2$.

VI. Conclusion

In this paper, we have analyzed the structure of the capacity-achieving input for the noncoherent Rician fading

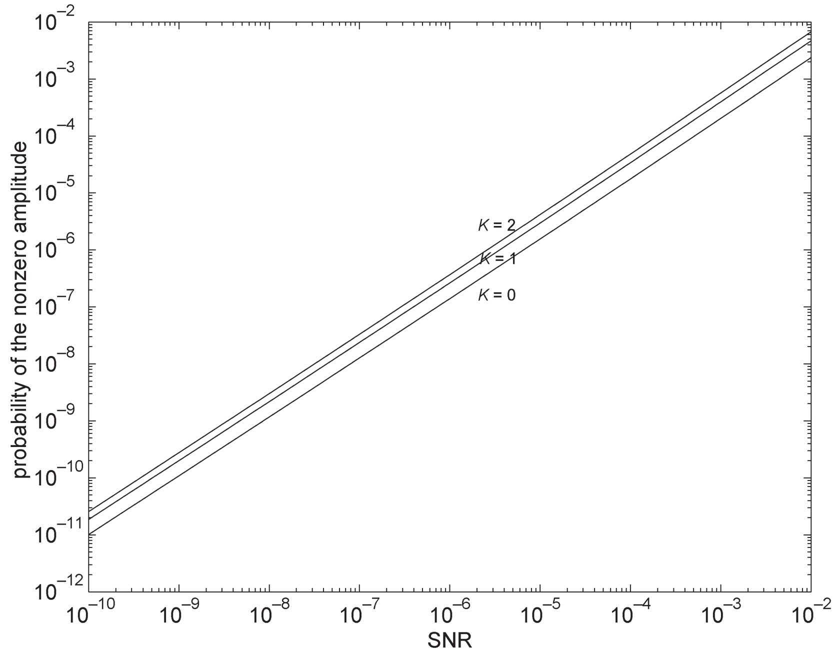


Fig. 7. Probability of the nonzero amplitude versus normalized $\mathrm{SNR} = \gamma^2(P_{\mathrm{av}}/N_0)$ for the Rician channel with phase noise $\mathsf{K} = 0, 1, 2$.

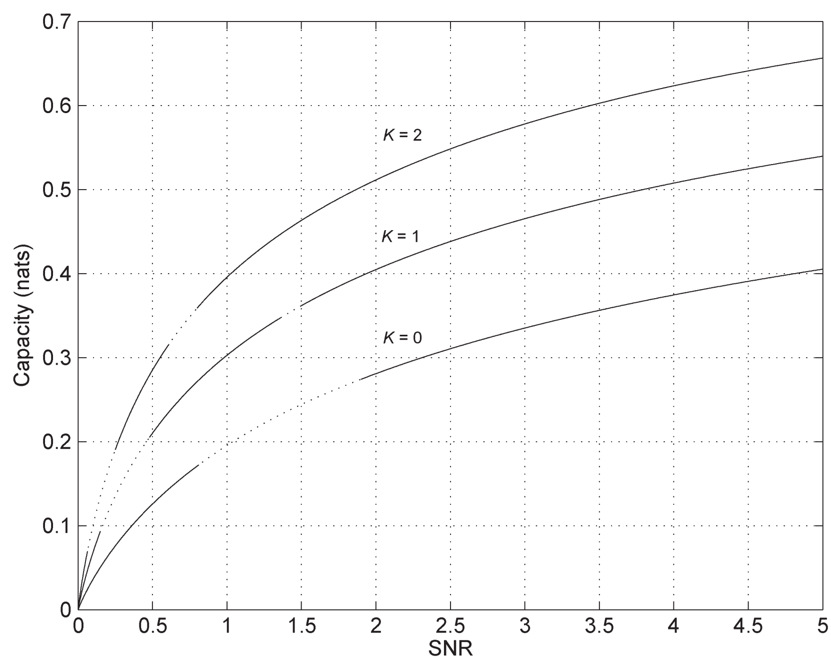


Fig. 8. Capacity curves as a function of the normalized $\mathrm{SNR} = \gamma^2(P_{\mathrm{av}}/N_0)$ for the Rician channel with phase noise with Rician factors $\mathsf{K} = 0, 1, 2$. The dashed segments are interpolated capacity curves. Numerical optimization methods do not provide stable results in these regions where a new mass point is emerging with a very small probability.

channel. The peakedness of the input has been limited by imposing a fourth moment or a peak constraint. Using a sufficient and necessary condition, we have proven that when the input is subject to second and fourth moment limitations, the optimal input amplitude is discrete with a finite number of levels in the low-power regime. It turns out that a particular two-mass-point distribution that depends only on the signal-to-noise ratio (SNR) and κ is asymptotically optimal as $\mathrm{SNR} \to 0$. Discreteness of the optimal input amplitude distribution has also been shown for the peak-power-limited Rician channel over the entire SNR range. This time, the amplitude distribution with a single mass at the peak level is optimal in the low-power regime for the Rician channel with $\mathsf{K} > 0$.

A Rician fading channel model where there is phase noise in the specular component has also been analyzed. We have shown that under an average power limitation, the optimal amplitude is discrete with a finite number of levels. For this model, numerical results for the capacity and the optimal input distribution have been provided, where it was observed that a flash-signaling-type input is required in the low-power regime. It has also been proven that the optimal input for the average-power-limited classical Rician channel has bounded support.

REFERENCES

[1] M. C. Gursoy, H. V. Poor, and S. Verdú. (2002). "The capacity of the noncoherent Rician fading channel," Princeton Univ., Princeton, NJ, Tech. Rep. 70. [Online]. Available: http://www.ee.unl.edu/people/gursoy.html

[2] ——, "The noncoherent Rician fading channel—Part II: Spectral efficiency in the low-power regime," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2207–2221, Sep. 2005.

[3] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1319–1343, Jun. 2002.

[4] R. S. Kennedy, *Fading Dispersive Communication Channels*. New York: Wiley-Interscience, 1969.

[5] J. S. Richters, "Communication over fading dispersive channels," MIT Res. Lab. Electron., Cambridge, MA, Tech. Rep. 464, Nov. 1967.

[6] J. G. Smith, "The information capacity of peak and average power constrained Gaussian channels," *Inf. Control*, vol. 18, pp. 203–219, 1971.

[7] S. Shamai and I. Bar-David, "The capacity of average and peak-power-limited quadrature Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 41, no. 4, pp. 1060–1071, Jul. 1995.

[8] I. Abou-Faycal, M. D. Trott, and S. Shamai, "The capacity of discrete-time memoryless Rayleigh fading channels," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1290–1301, May 2001.

[9] M. Katz and S. Shamai, "On the capacity-achieving distribution of the discrete-time non-coherent and partially-coherent AWGN channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2257–2270, 2004.

[10] ——, "On the capacity-achieving distribution of the discrete-time non-coherent additive white Gaussian noise channel," in *Proc. IEEE Int. Symp. Information Theory*, Lausanne, Switzerland, Jun. 30–Jul. 5, 2002, p. 165.

[11] S. Shamai, "On the capacity of a pulse amplitude modulated direct detection photon channel," *Proc. Inst. Elect. Eng.*, vol. 137, pt. I, pp. 424–430, Dec. 1990.

[12] A. Das, "Capacity-achieving distributions for non-Gaussian additive noise channels," in *Proc. IEEE Int. Symp. Information Theory*, Sorrento, Italy, Jun. 25–30, 2000, p. 432.

[13] R. Nuriyev and A. Anastasopoulos, "Capacity characterization for the noncoherent block-independent AWGN channel," in *Proc. IEEE Int. Symp. Information Theory*, Yokohama, Japan, Jun. 29–Jul. 4, 2003, p. 373.

[14] J. Huang and S. P. Meyn, "Characterization and computation of optimal distributions for channel coding," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2336–2351, Jul. 2005.

[15] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Inf. Theory*, vol. 45, no. 1, pp. 139–157, Jan. 1999.

[16] A. Lapidoth, "Capacity bounds via duality: A phase noise example," in *Workshop Concepts Information Theory*, Breisach, Germany, Jun. 26–28, 2002, pp. 58–61.

[17] ——, "On phase noise channels at high SNR," in *IEEE Information Theory Workshop*, Bangalore, India, Oct. 20–25, 2002, pp. 1–4.

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

[19] M. Médard and R. G. Gallager, "Bandwidth scaling for fading multipath channels," *IEEE Trans. Inf. Theory*, vol. 48, no. 4, pp. 840–852, Apr. 2002.

[20] V. G. Subramanian and B. Hajek, "Broad-band fading channels: Signal burstiness and capacity," *IEEE Trans. Inf. Theory*, vol. 48, no. 4, pp. 809–827, Apr. 2002.

[21] I. E. Telatar and D. N. C. Tse, "Capacity and mutual information of wideband multipath fading channels," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1384–1400, Jul. 2000.

[22] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.

[23] R. Palanki, "On the capacity achieving distributions of some fading channels," in *40th Annu. Allerton Conf. Communication, Control, and Computing*, Monticello, IL, Oct. 2–4, 2002, pp. 337–346.

[24] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1986–1992, Nov. 1997.

[25] G. Caire and S. Shamai, "On the capacity of some channels with channel state information," *IEEE Trans. Inf. Theory*, vol. 45, no. 6, pp. 2007–2019, Sep. 1999.

[26] M. Médard and A. J. Goldsmith, "Capacity of time-varying channels with channel side information," in *Proc. IEEE Int. Symp. Information Theory*, Ulm, Germany, Jun. 29–Jul. 4 1997, p. 372.

[27] M. Médard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of channel," *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 933–946, May 2000.

[28] R. B. Ash, *Information Theory*. New York: Dover, 1990.

[29] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

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

[31] W. Rudin, *Principles of Mathematical Analysis*. New York: McGraw-Hill, 1964.

[32] K. Knopp, *Theory of Functions*. New York: Dover, 1945.

[33] S. Lang, *Complex Analysis*, 2nd ed. New York: Springer-Verlag, 1985.

[34] G. N. Watson, *A Treatise on the Theory of Bessel Functions*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1958.

Mustafa Cenk Gursoy (S'01–M'04) received the B.S. degree in electrical and electronic engineering from Bogazici University, Istanbul, Turkey, in 1999, and the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, in 2004.

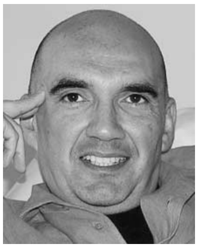
In the summer of 2000, he worked at Lucent Technologies, Holmdel, NJ, where he conducted performance analysis of digital subscriber line (DSL) modems. Since September 2004, he has been an Assistant Professor in the Department of Electrical Engineering at the University of Nebraska-Lincoln. His research interests are in the general areas of wireless communications, information theory, communication networks, and signal processing.

Dr. Gursoy is a recipient of the Gordon Wu Graduate Fellowship from Princeton University.

H. Vincent Poor (S'72–M'77–SM'82–F'87) received the Ph.D. degree in EECS from Princeton University, Princeton, NJ, in 1977.

From 1977 until 1990, he was a member of the faculty of the University of Illinois at Urbana-Champaign. Since 1990, he has been a member of the faculty of Princeton University, where he is the George Van Ness Lothrop Professor in Engineering. His research interests are in the areas of wireless networks, advanced signal processing, and related fields. Among his publications in these areas is the recent book Wireless Communication Systems: Advanced Techniques for Signal Reception (Upper Saddle River, NJ: Prentice-Hall, 2004).

Dr. Poor is a member of the U.S. National Academy of Engineering, and is a Fellow of the Institute of Mathematical Statistics, the Optical Society of America, and other organizations. In 1990, he served as President of the IEEE Information Theory Society, and in 1991–1992, he was a member of the IEEE Board of Directors. He is currently serving as the Editor-in-Chief of the IEEE TRANSACTIONS ON INFORMATION THEORY. Among his recent honors are the Joint Paper Award of the IEEE Communications and Information Theory Societies (2001), the NSF Director's Award for Distinguished Teaching Scholars (2002), a Guggenheim Fellowship (2002–2003), and the IEEE EAB Major Educational Innovation Award (2004).



Sergio Verdú (S'80–M'84–SM'88–F'93) received the Telecommunications Engineering degree from the Polytechnic University of Catalonia, Barcelona, Spain, in 1980, and the Ph.D. degree in electrical engineering from the University of Illinois at Urbana-Champaign in 1984.

He is a Professor of Electrical Engineering at Princeton University, Princeton, NJ. He has held visiting appointments at the Australian National University, the Technion-Israel Institute of Technology, the University of Tokyo, the University of California, Berkeley, and the Mathematical Sciences Research Institute, Berkeley. He served on the Board of Governors of the Information Theory Society in 1989–1999 and was President of the Society in 1997. He was Guest Editor of the 1998 Special Commemorative Issue of the IEEE TRANSACTIONS ON INFORMATION THEORY, reprinted as the IEEE Press volume *Information Theory: Fifty Years of Discovery*. He is the author of the textbook *Multiuser Detection* (Cambridge, U.K.: Cambridge University Press, 1998). He served as Associate Editor for Shannon Theory of the IEEE TRANSACTIONS ON INFORMATION THEORY. He currently serves as Editor-in-Chief of the Foundations and Trends in Communications and Information Theory and as Associate Editor for Book Reviews of the IEEE TRANSACTIONS ON INFORMATION THEORY.

Prof. Verdú is a recipient of the NSF Presidential Young Investigator Award, the IEEE Donald Fink Paper Award, a Golden Jubilee Paper Award from the Information Theory Society, the 1998 Information Theory Society Paper Award, the 2000 Paper Award from the Telecommunications Advancement Foundation of Japan, and the 2002 Leonard G. Abraham Prize Award from the IEEE Communications Society. He also received a Millennium Medal from the IEEE and the 2000 Frederick E. Terman Award from the American Society for Engineering Education.