Broad-Band Fading Channels: Signal Burstiness and Capacity

Vijay G. Subramanian, *Member, IEEE,* and Bruce Hajek, *Fellow, IEEE*

***Abstract*—Médard and Gallager recently showed that very large bandwidths on certain fading channels cannot be effectively used by direct sequence or related spread-spectrum systems. This paper complements the work of Médard and Gallager. First, it is shown that a key information-theoretic inequality of Médard and Gallager can be directly derived using the theory of capacity per unit cost, for a certain fourth-order cost function, called fourthegy. This provides insight into the tightness of the bound. Secondly, the bound is explored for a wide-sense-stationary uncorrelated scattering (WSSUS) fading channel, which entails mathematically defining such a channel. In this context, the fourthegy can be expressed using the ambiguity function of the input signal. Finally, numerical data and conclusions are presented for direct-sequence type input signals.**

***Index Terms*—Channel capacity, fading channels, spread spectrum, wide-sense-stationary uncorrelated scattering (WSSUS) fading channels.**

I. Introduction

A PROMINENT feature of wireless media is time-varying multipath fading. A fading channel is a very different entity from an additive Gaussian noise (AGN) channel. If the channel changes rapidly, then it may be better to adopt noncoherent techniques for reliable communication instead of having a structure to measure and track the channel accurately. References [15], [18], and [25] present examples of noncoherent receiver structures that achieve capacity for certain channels. Another important fact is that for pure Rayleigh-fading channels, the output signal has mean zero for any input signal. Thus, the input signal only affects the second-order statistics and higher order statistics of the output. In contrast, the input signal directly affects the mean of the output signal for AGN channels. Owing to these differences, principles of signal design used for additive Gaussian noise channels do not directly apply to fading channels [2].

Even though wireless channels have been used for a long time, they are not as well understood as the additive Gaussian noise channel. At the same time, there has been substantial work on the capacity of such channels. In this paper, we are interested specifically in the case in which neither the transmitter nor the receiver knows the channel but both know the statistics of the channel. An important aspect of our assumption is that we do not assume feedback to the transmitter. In particular, this rules out power control. Note that there can also be the case where there is no knowledge of the statistics of the channel. Lapidoth and Narayan [16] give a comprehensive treatment of such channels and Biglieri *et al.* [3] give a detailed survey of capacity-related results on fading channels.

Broad-band channels are a special case of channels with a large number of degrees of freedom. In a seminal work, Gallager [9] discussed energy-limited channels, i.e., channels where the energy per degree of freedom is very small. He showed that the reliability function per unit energy can be computed exactly for all rates if there is a finite capacity per unit energy. Telatar [26] specialized Gallager's results to the Rayleigh-fading channel and obtained the capacity divided by energy as a function of bandwidth and signal energy, concluding from this that the infinite bandwidth Rayleigh fading channel has the same capacity as the infinite bandwidth additive WGN (AWGN) channel. Verdú [27] considered capacity per unit cost for general cost functions and derived a simple expression for the capacity per unit cost for memoryless channels for certain cost functions.

Kennedy [15] considered the capacity per unit time of diffuse wide-sense-stationary uncorrelated scattering (WSSUS) fading channels. Using an M-ary frequency-shift-keying (FSK) signaling set and under the assumptions that certain bandwidth considerations are met and no intersymbol interference (ISI) exists between blocks over which this input is transmitted, Kennedy derived the reliability function using the optimum demodulator and showed that for the infinite-bandwidth WSSUS fading channel, the capacity is the same as that for the infinite-bandwidth AWGN channel with the same average signal-to-noise ratio (SNR). Jacobs [14] presented the latter result in a simpler context, and [8, Sec. 8.6] gives a nice discussion of early work on fading channels.

Viterbi [28] clearly exhibited a loss in channel capacity due to the randomness of fading. He considered M-ary orthogonal signaling normalized in such a way that the transmitted signal is not bursty in the time domain. That is, the symbol duration increases in proportion to the number of bits per symbol, to maintain constant transmit power. The received signal is due to a combination of the tone that is transmitted and the stochastic fading. The fading itself can be viewed as amplitude modulation. The capacity for the digital part is, in the large M limit, equal to the mutual information rate between the stochastic signal and sto-

Manuscript received July 17, 2000; revised June 1, 2001. This work was supported by a Motorola Fellowship, by the U.S. Army Research Office under Grant DAAH05-95-1-0246, and by the National Science Foundation under Contracts NCR 93-14253, CCR 99-79381, and ITR 00-85929. The material in this paper was presented in part at the IEEE Information Theory Workshop, Kruger National Park, South Africa, June 1999.

V. G. Subramanian was with the University of Illinois at Urbana-Champaign. He is now with Motorola Corporation, Arlington Heights, IL 60004 USA (e-mail: vijay.subramanaian@motorola.com).

B. Hajek is with the Department of Electrical and Computer Engineering and the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: b-hajek@uiuc.edu).

Communicated by M. L. Honig, Associate Editor for Communications.

Publisher Item Identifier S 0018-9448(02)01996-X.

chastic signal plus white noise, less the capacity of the amplitude modulated channel (see [28, p. 418]). In particular, the channel capacity tends to zero as the doppler spread tends to infinity, assuming a diffuse doppler spectrum. The paper explains that the uncertainty or randomness of the fading subtracts directly from the capacity of the channel. This gives intuitive appeal to the use of the relation

$$I(Y;U) = I(Y;H,U) - I(Y;H|U)$$

which is used in Appendix A of this paper to give an alternative proof of a key inequality.

Médard and Gallager [10], [19] analyzed a broad-band channel with WSSUS fading. They consider a time–frequency expansion of the input signal. A constraint is imposed that is satisfied by direct-sequence code-division multiple-access (DS-CDMA) type signals—namely, each coefficient X in the signal expansion is assumed to satisfy $E[|X|^4] \leq \alpha\epsilon^2$, where ϵ is a bound on $E[|X|^2]$ and α is a bound on the peakedness of the distribution of X. They showed that the mutual information per unit time between the input and the output for a broad-band system is upper-bounded by a constant times ϵ. The interpretation of this bound is as follows: as the spread factor increases without bound, ϵ, which is inversely proportional to the spread factor, decreases to zero, and therefore, the mutual information per unit time between the input and the output decreases to zero. The intuitive explanation given for the poor performance of DS-CDMA is that spreading the energy too thinly does not allow the channel to be measured accurately enough, which ultimately limits the performance of DS-CDMA.

Telatar and Tse [25] considered specular multipath channels with multipath components subject to time-varying delays and with no ISI, such that each channel can be approximated by a time-invariant system. They showed that the capacity of an infinite-bandwidth WSSUS channel is the same as the capacity of an infinite bandwidth AWGN channel with the same average SNR and that with DS-CDMA-type input the mutual information between the input and the output varies inversely with the number of effective diversity paths. Biglieri *et al.* [3, pp. 2636–2638] give a nice exposition on the subject of bandwidth scaling, including a unifying discussion and physical interpretations of results of [10], [18], [19], [25], [28], and other works. The paper of Ganti *et al.* [11] considers several of the concepts considered here, such as channels with memory, capacity per unit cost, wide-band limit, and spread-spectrum signaling, though the focus of that paper, namely mismatched decoding, is considerably different.

A central theme in [10], [19] is that burstiness in time–frequency is necessary to achieve capacity in broad-band fading channels. To expound on this further, we define the notion of fourthegy of an input signal, which is related to the number of diversity paths of Kennedy. A key inequality of [10] shows that the capacity per unit fourthegy of Rayleigh-fading channels is finite. We show that the inequality can be proved by using the notion of capacity per unit cost. An implication of the inequality is that if the mean fourthegy of the input signal is small, so will be the number of bits that can be transmitted reliably. The fourthegy of a signal is roughly proportional to the sum of the squares of the local energy in time–frequency bins. For fixed power, nonbursty signals, the fourthegy per unit time tends to zero, and hence, by the basic inequality, so does the information rate. We show that the fourthegy is a function of the signal ambiguity function and this aids us in evaluating the fourthegy directly for DS-CDMA-like signals. This avoids imposing constraints on the fourth moment of coordinate values for a decomposition of the continuous-time signals, as in [10], or an asymptotic analysis based on peak-value constraints in time–frequency bins, as in [24]. Numerical bounds are given on the information rates possible for DS-CDMA-type signals.

Another contribution of this paper is making the notion of the WSSUS channel model mathematically precise. For completeness, we also discuss in Section IV-C the amount of information that can be transmitted per unit energy for a WSSUS fading channel, using a similar approach.

The organization of this paper is as follows. Section II briefly reviews the notion of capacity per unit cost, and presents the capacity per unit fourthegy for a vector memoryless Rayleigh channel. Section III presents the basic definitions and a mathematical foundation for WSSUS fading channels, and is independent of Section II. Section IV points out that the definition of fourthegy and the basic bound on information per unit fourthegy identified in Section II carry over to the WSSUS channel model described in Section III. Section IV goes on to describe several properties of fourthegy, and complementary results are given. The basic inequality is applied in Section V to DS-CDMA signals over broad-band WSSUS fading channels. We conclude in Section VI with some discussion. All capacity computations are in natural units for analytical simplicity. One natural unit, nat, is $\frac{1}{\log(2)} = 1.4427$ b. It is also to be understood that $Z = Z_r + jZ_i$ has the complex normal distribution $\mathcal{CN}(\mu, \mathrm{var})$ with mean μ and variance var, if Z_r and Z_i are independent Gaussian random variables with means $\mathrm{Re}(\mu)$ and $\mathrm{Im}(\mu)$, respectively, and with variance $\mathrm{var}/2$ each.

II. Fourth Moment Information Bound for a Vector Rayleigh Channel

In this section, the theory of capacity per unit cost is applied to derive a basic information inequality for a vector Rayleigh-fading channel.

A. Background: Capacity Per Unit Cost

We briefly review the notion of capacity per unit cost in this section, following [27]. Consider a discrete-time channel without feedback and with arbitrary input and output alphabets denoted by A and B, respectively. An (N, M, β, ϵ) code is one in which the block length is equal to N; the number of codewords is equal to M; each codeword $(x_{m1}, \ldots, x_{mN})$, $m = 1, \ldots, M$, satisfies the constraint $\sum_{n=1}^{N} h(x_{mn}) \leq \beta$, where h: $A \to [0, +\infty)$ is a function that assigns a cost to each input symbol, and the average probability of decoding the message is at least $1-\epsilon$. Given $0 \leq \epsilon < 1$ and $\beta > 0$, a nonnegative number R is an ϵ-achievable rate with cost per symbol not exceeding β if for every $\gamma > 0$ there exists N_0 such that if $N \geq N_0$, then an $(N, M, N\beta, \epsilon)$ code can be found whose rate satisfies $\log M > N(R-\gamma)$. Furthermore, R is said to be achievable if it is ϵ-achievable for all $0 < \epsilon < 1$. The maximum achievable rate with cost per symbol not exceeding β is the channel capacity denoted by $C(\beta)$. Results in information theory about the capacity of an input-constrained memoryless channel imply that $C(\cdot)$ is given by

$$C(\beta) = \sup_{X:\, E[h(X)] \le \beta} I(X; Y)$$

where the supremum is taken to be zero if the set of distributions therein is empty. The capacity per unit cost is then the maximum number of bits per unit cost that can be transmitted through the channel. Verdú [27] showed the capacity per unit cost for a memoryless channel satisfies

$$C_{\text{cost}} = \sup_{\beta > 0} \frac{C(\beta)}{\beta} = \sup_X \frac{I(X; Y)}{E[h(X)]}. \tag{1}$$

Verdú [27] also showed that if there is a unique input symbol "0" with zero cost, then the capacity per unit cost is given by minimizing a ratio of the divergence between two measures and the cost function

$$C_{\text{cost}} = \sup_{x:\, h(x) \ne 0} \frac{D(P_{Y|X=x} \| P_{Y|X=0})}{h(x)}. \tag{2}$$

In particular, it follows that

$$\sup_X \frac{I(X; Y)}{E[h(X)]} = \sup_{x:\, h(x) \ne 0} \frac{D(P_{Y|X=x} \| P_{Y|X=0})}{h(x)}. \tag{3}$$

The relation (3) is interesting in itself, even though it does not involve the capacity per unit cost. Only basic measurability assumptions are required for the above results, as shown by Verdú [27]. The assumption that there is a unique, zero-cost input symbol was explored by Gallager [9] in the context of reliability functions per unit cost. The assumption greatly simplifies the computation of C_{cost}, since the supremum in (2) is over the input space, rather than over the space of probability distributions on the input space.

In many contexts, the capacity per unit cost for a given channel with constrained input signal bandwidth is equal to the limit of the capacity (in bits per second) divided by cost per second (e.g., power) for the same channel in the limit as the bandwidth of the channel tends to infinity, with the cost per second fixed. Reference [27] illustrates this with the AWGN channel. One can explain this in the following manner. Suppose that there is a discrete-time memoryless channel (DTMC) such that use of the original channel with input signals constrained to bandwidth W is equivalent to using the DTMC W times per second. In particular, suppose that the cost for an input signal for the original channel is equal to the cost of the equivalent signal for repeated use of the DTMC channel, and that there is a unique zero cost input for the DTMC. Then the original channel and the DTMC have the same capacity per unit cost. Given a code for the DTMC which achieves a given information rate per unit cost, by varying the number of uses W of the DTMC per second, we obtain a code that has a given cost per unit time, and the same ratio of information per unit cost. While the assumptions of this explanation are rarely exactly satisfied, it at least offers a heuristic explanation for why capacity per unit cost is often equal to the infinite bandwidth limit of capacity divided by cost per second.

B. The Information Bound for a Vector Rayleigh-Fading Channel

A single use of a discrete-time memoryless vector Rayleigh-fading channel is given by the following equation:

$$Y = H^\dagger U + N \tag{4}$$

where U is the channel input in $\mathcal{C}^n$, Y is the output of the channel, N is additive Gaussian noise distributed as $\mathcal{CN}(0, \sigma^2 I_{n,n})$, and H is an $n \times n$ matrix of jointly circularly symmetric mean-zero Gaussian random variables, for some $n \ge 1$. In addition, N, H, and U are assumed to be mutually independent. The columns of H are denoted by $h_1, \ldots, h_n$ (so $H = [h_1|h_2|\cdots|h_n]$), and the complex conjugate transpose of H is denoted by $H^\dagger$.

Let $S = H^\dagger U$ denote the output signal without the additive noise term N. The conditional covariance of S given $U = u$ is given by

$$\begin{aligned}\Sigma_u &= E[SS^\dagger | U = u] \\ &= \begin{bmatrix} u^\dagger E[h_1 h_1^\dagger] u & u^\dagger E[h_2 h_1^\dagger] u & \cdots & u^\dagger E[h_n h_1^\dagger] u \\ u^\dagger E[h_1 h_2^\dagger] u & u^\dagger E[h_2 h_2^\dagger] u & \cdots & u^\dagger E[h_n h_2^\dagger] u \\ \vdots & \vdots & \ddots & \vdots \\ u^\dagger E[h_1 h_n^\dagger] u & u^\dagger E[h_2 h_n^\dagger] u & \cdots & u^\dagger E[h_n h_n^\dagger] u \end{bmatrix}.\end{aligned}$$

The cost function we consider is $J_C(u) = \text{Trace}(\Sigma_u^2)$ which in this specific case is

$$J_C(u) = \sum_i \lambda_i^2 = \sum_{i,j} |u^\dagger E[h_i h_j^\dagger] u|^2$$

where $\{\lambda_i\}_{i=1}^n$ are the eigenvalues of Σ_u. We call $J_C(u)$ the fourthegy of the vector u, relative to the channel. The name is motivated by the fact that $J_C(u)$ is fourth order in u, and that it is a positive sounding name (like energy) rather than a negative sounding name, like cost.

Let C_J denote the capacity per unit cost where cost is measured by the fourthegy J_C.

Proposition II.1: The capacity per unit fourthegy for the discrete-time memoryless vector Rayleigh-fading channel is $C_J = \frac{1}{2\sigma^4}$. In particular, for any U

$$I(U; Y) \le \frac{1}{2\sigma^4} E[J_C(U)]. \tag{5}$$

Proof: We will prove that

$$\sup_{u \ne 0} \frac{D(u)}{J_C(u)} = \frac{1}{2\sigma^4} \tag{6}$$

where

$$D(u) = D(P_{Y|U=u} \| P_{Y|U=0}).$$

Once (6) is proved, (2) will imply the expression given for C_J, (3) will imply (5), and the proof will be complete. Conditioned on $U = u$, Y is a mean-zero Gaussian random vector with covariance matrix $E[YY^\dagger | U = u] = \Sigma_u + \sigma^2 I_{n,n}$ having eigenvalues $\{\lambda_i + \sigma^2\}_{i=1}^n$. This covariance matrix is Hermitian and can therefore be diagonalized by a linear transformation of space by a unitary matrix. That transformation leaves the distribution of N invariant, it transforms the conditional distribution of Y given u into a vector of n independent complex normal random variables with respective variances $\{\lambda_i + \sigma^2\}_{i=1}^n$, and it preserves the value of the divergence $D(\|)$. Therefore,

$$D(u) = \sum_{i=1}^{n} \phi(\lambda_i)$$

where ϕ is defined by

$$\begin{aligned}\phi(\lambda) &= D(\mathcal{CN}(0,\, \lambda + \sigma^2)\|\mathcal{CN}(0,\, \sigma^2)) \\ &= \frac{\lambda}{\sigma^2} - \log\left(1 + \frac{\lambda}{\sigma^2}\right).\end{aligned} \tag{7}$$

Using the fact $\log(1+x) \ge x - \frac{x^2}{2}$ for each $x \ge 0$, we have

$$D(u) = \sum_{i=1}^{n} \phi(\lambda_i) \le \sum_{i=1}^{n} \frac{\lambda_i^2}{2\sigma^4} = \frac{J_C(u)}{2\sigma^4}. \tag{8}$$

This proves (6) with "$=$" replaced by "$\le$." To complete the proof, scale u by ϵ for some $\epsilon > 0$. Note that this scales the eigenvalues by ϵ^2. Therefore,

$$\lim_{\epsilon \to 0} \frac{D(u\epsilon)}{J_C(u\epsilon)} = \frac{1}{\sum_{i=1}^{n} \lambda_i^2} \sum_{i=1}^{n} \lim_{\epsilon \to 0} \frac{\phi(\lambda_i \epsilon^2)}{\epsilon^4} = \frac{1}{2\sigma^4}.$$

Thus, (6) is established, and the proof is complete. □

Notes:

1) Inequality (5) is a key inequality of [10], [19]. Proposition II.1 shows that the inequality (5) is asymptotically tight for U an on–off signal, as the on probability tends to zero and the on signal value is scaled toward zero. Inequality (5) is applied to a WSSUS channel model in Sections IV and V, but in the spirit of Médard and Gallager, one can make, right away, for a simple channel and input scaling, the argument that capacity goes to 0 as the bandwidth goes to ∞. Suppose the channel is block-fading in frequency: there are b frequency bands that fade independently. The fourthegy for the total input is the sum of the fourthegies over the individual bands. If energy is spread evenly across the b bands, then the fourthegy per band scales as $1/b^2$ as $b \to \infty$, so the total fourthegy scales with bandwidth as $1/b$. Moreover, if the channel also decorrelates in time, then the fourthegy for a constant power input over an interval of length T is asymptotically linear in T. Hence, for channels that decorrelate sufficiently in time and input signals that are evenly distributed in time and frequency, the overall fourthegy per unit time, and hence the mutual information per unit time, is finite and tends to zero as $1/b$ as $b \to \infty$.

2) *Scalar Channel*: For the scalar Rayleigh-fading channel with $H \sim \mathcal{CN}(0, \gamma^2)$ the fourthegy is given by $J_C(u) = \gamma^4|u|^4$. In other words, the fourthegy is proportional to the fourth power of the signal magnitude. Thus, the capacity per unit fourth moment of the discrete-time scalar Rayleigh fading channel is proportional to the capacity per unit fourthegy of the same channel, and therefore, is finite. It is interesting to compare the capacity per unit fourth moment of the discrete-time scalar Rayleigh-fading channel with the capacity per unit fourth moment of the discrete-time AWGN channel. A single use of the AWGN channel is given by $Y = U + N$ where $N \sim \mathcal{N}(0,\, \sigma^2)$. For this case it is evident that

$$\frac{D(P_{Y|U=u}\|P_{Y|U=0})}{|u|^4} = \frac{1}{2\sigma^2|u|^2}$$

so the capacity per unit fourth moment for the AWGN channel is infinite.

3) *Alternative Proof*: An alternative proof of (5) along the lines of [3], [28] is given in Appendix A.

III. The WSSUS Fading Channel

A wireless channel can be reasonably modeled as a time-varying linear channel. The observed output $y(t)$ can be represented by

$$y(t) = \int h(t,\, \tau)u(t-\tau)\, d\tau + n(t) \tag{9}$$

where $u(t)$ is the input, $h(t,\, \tau)$ is the time-varying channel impulse response function, and $n(t)$ is white Gaussian noise. Owing to the high complexity of such channels, a stochastic characterization is useful. Considering a single tone transmitted to a moving receiver with isotropic scattering, Clarke [5] showed that the complex envelope of the signal at the receiver is a complex-valued wide-sense stationary (WSS) Gaussian random process with the zeroth-order Bessel function as the autocorrelation function. The magnitude at each time instance has the Rayleigh distribution. Bello [1] analyzed random time-varying linear channels and gave a statistical characterization in time and frequency variables. Usually, $h(t,\, \tau)$ for fixed τ is assumed to be a WSS process, i.e., $E[h(t,\, \tau)] = \mu(\tau)$, and $E[h(s,\, \tau)h^*(t,\, \tau)] = R_H(s-t,\, \tau)$. We can also have $h(t,\, \tau)$ uncorrelated for different values of τ. This is called uncorrelated scattering (US). Often these two simplifying features are combined (see [7]), leading to the consideration of WSSUS fading channels. For a WSSUS channel, the second moments of h have the form

$$E[h(t,\, \tau)h^*(s,\, v)] = R_H(t-s,\, \tau)\delta(\tau - v). \tag{10}$$

Finally, it is often assumed that the random process h is a complex Gaussian random process. See, for example, the urban propagation model or the GSM propagation model [7].

In this paper, we assume that h is WSSUS, Gaussian, and mean zero. The second variable, τ, indexes the path delays, and we also assume that $h(t,\, \tau) = 0$ unless $\tau \in [0,\, T_{\max}]$, where $T_{\max}$ is a bound on the maximum delay spread of the channel. Such a model, for suitable choices of R_H, fits empirical measurement data and has been used extensively in evaluating the performance of various systems like GSM, ATDMA, IS-95, etc., as mentioned in the COST and CODIT studies [7]. The channel model allows for two extreme cases, namely, specular where there is a set of distinct paths, and diffuse where there is a continuum of irresolvable paths. The general WSSUS model allows for a mixture of these two extremes [20].

A nice feature of a WSSUS channel is that the ratio of the mean output energy (excluding the additive noise) to input energy does not depend on the choice of input signal. This ratio is called the energy gain, G_H, and is given by

$$G_H = \int R_H(0, \tau)\, d\tau. \tag{11}$$

We generally assume that G_H is finite. We refer the reader to Proakis [22, Ch. 14] for a detailed treatment of WSSUS fading channels.

We have introduced the WSSUS channel model in standard engineering terminology. In the remainder of this section, we describe how the channel can be put on a firm mathematical foundation. The assumption of uncorrelated scattering means that the process $h(t, \tau)$ is white-noise-like as a function of τ, as evidenced by the delta function in (10). Also, the observed output signal y has AWGN. Despite $h(t, \tau)$ being white-noise-like as a function of τ, it can be shown that the required integrals involving $h(t, \tau)$ are ordinary square-integrable random variables, in the same way that white noise integrals yield square integrable random variables.

The following will be used instead of R_H in order to summarize the channel statistics, and then the connection back to R_H will be made. Let Γ_H be a finite measure on $\Re$ with support $[0, T_{\max}]$. This measure is the power gain distribution across different path delays. The total gain is $\Gamma_H(\Re) = G_H$. Let $r_H(t, \tau)$ be a positive-definite function for τ fixed which has $r_H(0, \tau) \equiv 1$ and which is jointly measurable in (t, τ). The function $r_H(t, \tau)$, for τ fixed, is the normalized autocorrelation function for the set of paths with delay τ. We shall give a description of a WSSUS fading channel with power gain distribution Γ_H and normalized autocorrelation function r_H.

Proposition III.1: Given Γ_H, $r_H(\cdot, \cdot)$ and $T \le +\infty$, there exists on some probability space a family of jointly Gaussian, measurable random processes $(s_{\text{out}}(u; t)\colon u \in L^2[0, T])$ with finite average energy such that for all $u, v \in L^2[0, T]$ and a.e. $s, t \in \Re$

$$\begin{aligned} &E[s_{\text{out}}(u; s) s_{\text{out}}(v; t)^*] \\ &\quad = \int_0^{T_{\max}} u(s-\tau) r_H(s-t, \tau) v(t-\tau)^* \Gamma_H(d\tau). \end{aligned} \tag{12}$$

Proof: Refer to Appendix B.

The channel is said to be diffuse if Γ_H has a density γ_H. In this case, the function R_H described at the beginning of the section is given by

$$R_H(t, \tau) = \gamma_H(\tau) r_H(t, \tau)$$

so that

$$r_H(t, \tau)\Gamma_H(d\tau) = R_H(t, \tau) d\tau.$$

The channel is said to be specular if there is a countable set of path delays $\{\tau_l\} \subset [0, T_{\max}]$ and positive constants $\{\Gamma_H^l\}$ so that for any set $A \subset \mathcal{R}$

$$\Gamma_H(A) = \sum_l \mathbf{1}_{\{\tau_l \in A\}} \Gamma_H^l.$$

In this case, the function R_H described at the beginning of the section is given by

$$R_H(t, \tau) = \sum_l \delta(\tau - \tau_l) \Gamma_H^l r_H(t, \tau).$$

In general, the measure Γ_H can have both discrete and continuous components.

The $(\Gamma_H, r_H(t, \tau))$ notation is used in Appendix A, and avoids the use of generalized functions. The $(\Gamma_H, r_H(t, \tau))$ is also used quite often in other sections of the paper for ease of exposition. In the next section of this paper, we also use the notation $R_H(t, \tau)$, primarily to maintain compatibility with the literature.

On the basis of Proposition III.1, we can write the observed output of the WSSUS channel for a finite energy input u as

$$y_s = s_{\text{out}}(u; s) + n_s, \qquad s \ge 0 \tag{13}$$

where n is complex Gaussian white noise with one-sided power spectral density σ^2. A standard mathematical interpretation of this (see, for example, [17], [21], [29]), that avoids the use of generalized random processes is that the observed signal is $(Y_t\colon t \ge 0)$ defined by

$$Y_t = \int_0^t s_{\text{out}}(u; v)\, dv + \sigma W_t$$

where $(W_t\colon t \ge 0)$ is a standard complex Wiener process. The process Y takes values in $C[0, T]$, the set of continuous complex-valued functions on the interval $[0, T]$. The signal s in the following proposition can be taken to be s_{out} for a fixed finite energy input signal u.

Proposition III.2: Let $T \le \infty$ and let $s(t)$ be a measurable Gaussian random process with $E[\int_0^T |s(t)|^2\, dt] < \infty$ and let $\Sigma(s, t)$ be the covariance function of $s(t)$. Let

$$Y_t = \int_0^t s(v)\, dv + \sigma W_t, \qquad \text{for } 0 \le t < T \tag{14}$$

where W_t is a standard complex Weiner process and $\sigma \ne 0$. Then $\Sigma(s, t)$ has associated nonnegative eigenvalues $\{\lambda_i\}_{i=1}^\infty$ and eigenfunctions $\{\psi_i(t)\}_{i=1}^\infty$ such that

$$\sum_{i=1}^\infty \lambda_i = \int_0^T \Sigma(t, t)\, dt$$

and μ_Y and $\mu_{\sigma W}$ are each absolutely continuous with respect to the other with the Radon–Nikodym derivative given by

$$\begin{aligned} \frac{d\mu_Y}{d\mu_{\sigma W}} &= L_\infty \\ &= \exp\left(\sum_{i=1}^\infty - \log\left(1 + \frac{\lambda_i}{\sigma^2}\right) + \frac{\lambda_i |z_i|^2}{\sigma^2(\lambda_i + \sigma^2)}\right) \end{aligned} \tag{15}$$

where the coordinates of Y are given by

$$Z_i = \int_0^T Y_t \psi_i(t)\, dt.$$

The coordinate Z_i is $\mathcal{CN}(0, \lambda)$ under μ_W and $\mathcal{CN}(0, \lambda+\sigma^2)$ under μ_Y.

Proof: Refer to Appendix C.

Since for the fading channel model $s = s_{\text{out}}$ depends on the input signal u, the eigenvalues $\{\lambda_i\}_{i=1}^{\infty}$ also depend on u. In the analysis that follows, we will be scaling the input and letting the scale factor either tend to zero or to infinity. Scaling the input does not change the eigenfunctions, but scales the eigenvalues by the square of the scale factor for u. A consequence of Proposition III.2 is that

$$D\left(P_{Y|U=u} \| P_{Y|U=0}\right) = E_Y[\log L_\infty] = \sum_{i=1}^{\infty} \phi(\lambda_i) \tag{16}$$

where $\phi(\cdot)$ is given by (7).

IV. Fourth Moment Information Bound for a WSSUS Channel

A. Definition of Fourthegy and the Information Bound

In this section, a bound analogous to (5) is proved, using essentially the same proof, for the WSSUS channel model (9) described in the previous section. The notation and assumptions of the previous section are in force. In particular, for each finite-energy input signal u, the covariance function $\Sigma(s, t)$ of the output signal $s = s_{\text{out}}$ is given by

$$\begin{aligned}\Sigma(s, t) &= \int_0^{T_{\max}} u(s-\tau)R_H(s-t, \tau)u^*(t-\tau)\,d\tau \\ &= \sum_{i=1}^{\infty} \lambda_i \psi_i(s)\psi_j^*(t). \end{aligned} \tag{17}$$

As noted in Appendix B, we can also consider Σ to be the kernel of an integral operator, also called Σ, and the eigenvalues associated with Σ are the eigenvalues of that operator, and are denoted by $\{\lambda_i\}_{i=1}^{\infty}$.

Define the fourthegy $J_C(u)$ of the input u by

$$J_C(u) = \int_0^T \int_0^T |\Sigma(s, t)|^2\, ds\, dt. \tag{18}$$

An equivalent expression for J_C is $J_C(u) = \text{Trace}(\Sigma^2)$ where Σ^2 denotes the kernel convolution of Σ with itself. Equation (17) yields a third expression: $J_C(u) = \sum_{i=1}^{\infty} \lambda_i^2$. This definition is consistent with the definition of fourthegy for a vector Rayleigh channel given in Section II-B. As before, scaling the input by a given factor scales the fourthegy by the fourth power of the factor. Some basic properties of J_C are given in the next subsection. The following theorem gives the key bound on information per unit fourthegy.

Theorem IV.1: For any measurable input random process U that has finite energy with probability one

$$I(U; Y) \le \frac{1}{2\sigma^4} E[J_C(U)] \tag{19}$$

where Y is the output random process and the expectation is carried out with respect to the measure of U.

Proof: Let

$$D(u) = D(P_{Y|U=u} \| P_{Y|U=0}).$$

In view of (16), the relations in (8) hold with "n" replaced by "∞." Thus,

$$\sup_{u \ne 0} \frac{D(u)}{J_C(u)} \le \frac{1}{2\sigma^4} \tag{20}$$

so the theorem follows from (3). □

Various complements to Theorem IV.1 are given in Section IV-C. Applications of the theorem are given in Section V.

B. Properties of Fourthegy

Let $S_H(\cdot, \tau)$ denote the Fourier transform of $R(\cdot, \tau)$ for each τ fixed. For τ fixed, $S_H(\cdot, \tau)$ is the power spectral density of the channel fading for delay τ. Using

$$R_H(s-t, \tau) = \int_{-\infty}^{\infty} S_H(f, \tau) e^{j(s-t)2\pi f}\, df$$

in (18) yields

$$\begin{aligned} J_C(u) = &\iint \left| \int_\tau u(s-\tau)R_H(s-t, \tau)u^*(t-\tau)\,d\tau\right|^2 ds\, dt \\ = &\int_{f_1}\int_{f_2}\int_{\tau_1}\int_{\tau_2} S_H(f_1, \tau_1)S_H(f_2, \tau_2) \\ &\times \left(\int u(s-\tau_1)u^*(s-\tau_2)e^{j2\pi s(f_1-f_2)}\, ds\right) \\ &\times \left(\int u(t-\tau_2)u^*(t-\tau_1)e^{-j2\pi t(f_1-f_2)}\, dt\right) \\ &\times d\tau_2\, d\tau_1\, df_2\, df_1 \\ = &\int_{f_1}\int_{f_2}\int_{\tau_1}\int_{\tau_2} S_H(f_1, \tau_1)S_H(f_2, \tau_2) \\ &\times |\chi(\tau_2-\tau_1, f_2-f_1)|^2\, d\tau_2\, d\tau_1\, df_2\, df_1 \end{aligned} \tag{21}$$

where $\chi(\nu, \tau)$ is the symmetric ambiguity function [4] of the signal $u(t)$ which is defined as

$$\chi(\nu, \tau) = \int_{-\infty}^{\infty} u(t+\tau/2)u^*(t-\tau/2)e^{-j2\pi\nu t}\, dt. \tag{22}$$

Thus, (21) can be rewritten to yield a fourth useful expression for $J_C(u)$

$$J_C(u) = \int_\nu \int_\tau |\chi(\nu, \tau)|^2 \psi_H(\nu, \tau)\, d\tau\, d\nu \tag{23}$$

where $\psi_H(\nu, \tau)$, called the channel response function, is given by

$$\psi_H(\nu, \tau) = \int_f \int_t S_H(f, t)S_H(f+\nu, t+\tau)\, dt\, df.$$

An important property of ambiguity functions is the volume invariance property [4, p. 153]

$$\iint |\chi(\nu, \tau)|^2\, d\tau\, d\nu = \chi(0, 0)^2 = \left(\int |u(t)|^2\, dt\right)^2 = E(u)^2. \tag{24}$$

Since $|\chi(\nu, \tau)| \le \chi(0, 0) = E(u)^2$ for all ν, τ, J_C is bounded above as follows:

$$\begin{aligned} J_C(u) &\le E(u)^2 \int_\nu \int_\tau \psi_H(\nu, \tau)\, d\tau\, d\nu \\ &= E(u)^2 \left(\int_f \int_t S_H(f, t)\, dt\, df\right)^2 = E(u)^2 G_H^2. \end{aligned}$$

The expression (23) shows that $J_C(u)$ captures both time and frequency aspects of the signal u. For example, it can be shown that $J_C(u) \le K_1 \int |u(t)|^4\, dt$ and $J_C(u) \le K_2 \int |U(f)|^4\, df$, where $U(f)$ is the Fourier transform of $u(t)$, as follows. By (23)

$$J_C(U) \le \int_\tau \left[\max_\nu \psi_H(\nu, \tau)\right] \left(\int_\nu |\chi(\nu, \tau)|^2\, d\nu\right) d\tau$$

and the ambiguity function has the following property [4, p. 154]:

$$\int_\nu |\chi(\nu, \tau)|^2\, d\nu = \int_\nu |\chi(\nu, 0)|^2 e^{j2\pi\nu\tau}\, d\nu.$$

Using this gives

$$\int_\nu |\chi(\nu, \tau)|^2\, d\nu \le \int_\nu |\chi(\nu, 0)|^2\, d\nu = \int |u(t)|^4\, dt.$$

Therefore,

$$J_C(u) \le K_1 \int |u(t)|^4\, dt$$

with

$$K_1 = \int_\tau \max_\nu \psi_H(\nu, \tau)\, d\tau.$$

Similarly,

$$J_C(u) \le K_2 \int |U(f)|^4\, df$$

where

$$K_2 = \int_\nu \max_\tau \psi_H(\nu, \tau)\, d\nu.$$

A drawback of the definition of fourthegy is that, unlike the definition of energy, it involves the channel. However, applying the Cauchy–Schwartz inequality to (23) yields

$$J_C(u) \le \sqrt{\int_\nu \int_\tau |\chi(\nu, \tau)|^4\, d\tau\, d\nu} \sqrt{\int_\nu \int_\tau \psi_H(\nu, \tau)^2\, d\tau\, d\nu}. \tag{25}$$

The right-hand side of (25) is the product of two terms, the first involving only the input signal, and the second involving only the channel. Perhaps the first term on the right-hand side would be a good channel-independent notion of fourthegy, but it seems too complicated to work with.

Recall that in Section III a general alternative way to describe the statistics of a WSSUS channel was given, involving a power gain distribution Γ_H and normalized autocorrelation function $r_H(t, \tau)$. If we were to follow through with that general notation in this section we would see that the channel response function is best considered as the measure given by

$$\phi_H(\nu, d\tau) = \iint s_H(f, t) s_H(f+\nu, t+\tau) \times \Gamma_H(dt)\Gamma_H(dt+d\tau)\, df$$

where

$$s_H(f, \tau) = \int e^{-j2\pi ft} r_H(t, \tau)\, dt$$

and the fourthegy $J_C(u)$ is the integral of the ambiguity function squared with respect to this measure, i.e.,

$$J_C(u) = \iint |\chi(\nu, \tau)|^2 \phi_H(\nu, d\tau)\, d\nu.$$

C. Complements

Various complements to the other results of this section are given in this subsection. First, we note that Kennedy [15] defined the number of effective diversity paths D to be the reciprocal of $J_C(u)$. In [15], u is the M-ary FSK waveform while here it is the on signal for on–off keying. Thus, Kennedy's D increasing without bound implies that $J_C(u)$ decreases to zero and the result of the error exponent for M-ary FSK going to zero in [15] is mirrored by the mutual information between the input and the output going to zero.

Second, the astute reader will note that Theorem IV.1 does not mention the notion of capacity per unit fourthegy for the WSSUS fading channel model, unlike Proposition II.1. The reason is that the bound given in Theorem IV.1, essentially a converse half of a coding theorem, has a clean proof and is all that is needed for the applications of the next section. Still, for completeness, we pursue the notion of capacity per unit fourthegy here. To begin with, we claim that equality actually holds in (20). To prove this, note that since $0 \le \frac{d^3 \log(y)}{dy^3} = \frac{2}{y^3} \le 2$ for $y \ge 1$, Taylors formula yields

$$x - \log(1+x) = \frac{x^2}{2} - \eta(x) \tag{26}$$

where

$$0 \le \eta(x) \le \frac{x^3}{3}, \qquad \text{for } x \ge 0.$$

Using (16) and (26) yields

$$D(u) = \frac{1}{2\sigma^4} J_C(u) - \alpha(u) \tag{27}$$

where α satisfies

$$0 \le \alpha(u) \le \frac{\sum_{i=1}^\infty \lambda_i^3}{3\sigma^6}. \tag{28}$$

Thus,

$$\frac{D(u\epsilon)}{J_c(u\epsilon)} = \frac{1}{2\sigma^4} - \frac{\alpha(u\epsilon)}{J_c(u\epsilon)}.$$

The eigenvalues are scaled by ϵ^2 if u is scaled by ϵ, so $\frac{\alpha(u\epsilon)}{J_c(u\epsilon)} = o(1)$ as $\epsilon \to 0$, which, in turn, shows that equality holds in (20) as claimed.

One consequence of this claim is simply that the bound of Theorem IV.1 is tight in the sense that the ratio of the right-hand side to the left-hand side tends to one for an on–off input process in which the on probability tends to zero and the on signal is scaled toward zero.

Another consequence is that we can apply (2) to conclude that the capacity per unit fourthegy of the WSSUS channel is equal to $1/2\sigma^4$. However, this result requires repeated independent use of the WSSUS channel to form a discrete-time memoryless channel. Intuitively, if the channel memory has a reasonable decay rate, one can simulate repeated independent use of

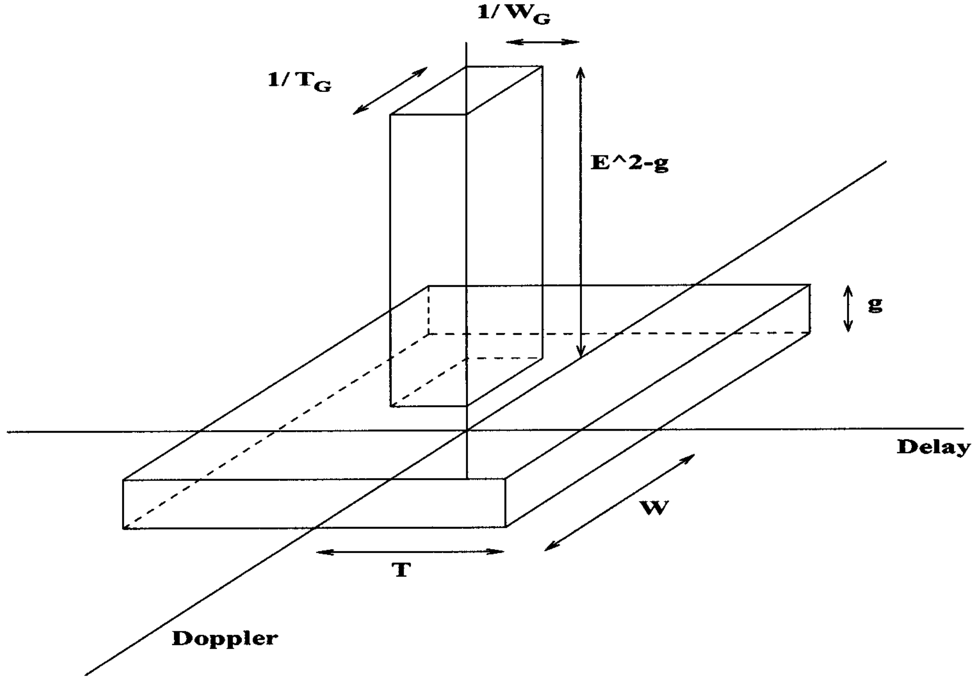

Fig. 1. Ambiguity function of a typical signal.

the WSSUS channel by using a single WSSUS channel and signaling in time intervals well separated by guard bands. Hence, it can be shown under reasonable conditions that information can be reliably sent at rates arbitrarily close to $1/2\sigma^4$ nats per unit fourthegy over a single WSSUS channel. The bound of Theorem IV.1 shows that higher rates are not possible.

The third and final item in this subsection concerns the capacity per unit energy for the WSSUS channel. This capacity, denoted by C_E, is given by

$$C_E = \sup_{u \neq \mathbf{0}} \frac{D(u)}{E(u)}$$

where $E(u) = \int_s |u(s)|^2\, ds$ is the energy of the input waveform u. The discussion of the previous paragraph applies for the positive part of the coding theorem.

Note that $E(u) = \text{Tr}(\Sigma)/G_H$ and

$$D(u) = \frac{\text{Tr}(\Sigma)}{\sigma^2} - \sum_{i=1}^{\infty} \log\left(1 + \frac{\lambda_i}{\sigma^2}\right).$$

Therefore,

$$\frac{D(u)}{E(u)} = \frac{G_H}{\sigma^2} - \frac{\sum_{i=1}^{\infty} \log\left(1 + \frac{\lambda_i}{\sigma^2}\right)}{\sum_{i=1}^{\infty} \frac{\lambda_i}{\sigma^2}} \frac{G_H}{\sigma^2} \leq \frac{G_H}{\sigma^2}$$

and fixing an arbitrary but nonzero, finite-energy signal u we have that

$$\lim_{a\to\infty} \frac{D(ua)}{E(ua)} = \frac{G_H}{\sigma^2}.$$

Therefore, we have established that $C_E = \frac{G_H}{\sigma^2}$, and for any random input signal U, $I(U; Y) \leq C_E E[\text{Energy}(U)]$. Furthermore, the capacity can be approached by using any nonzero input as the on input for the on–off keying scheme with the energy tending to infinity and the average energy tending to zero. We should note that C_E is exactly the same as the capacity per unit energy of an AWGN channel with the same gain and noise characteristics. In view of the heuristic connection between capacity per unit cost and capacity with infinitely many degrees of freedom discussed at the end of Section II-A, this is exactly as expected from the results of [14], [15], and [25].

V. DS-CDMA Signals Over Broad-Band Fading Channels

Before deriving the actual ambiguity function for DS-CDMA signals, we intuitively explain why the capacity of DS-CDMA signals over diffuse WSSUS fading channels tends to zero as the spreading increases. The ambiguity function of a typical DS-CDMA signal is shown in Fig. 1. The ambiguity function looks like a thumb-tack. From the volume invariance property stated in (24) and assuming that the energy of the signal is normalized to be 1, we can compute the dimensions of the thumb-tack. The dimensions of the stump are as follows: height, which is (normalized) energy squared, is 1, length along the delay axis is the inverse of the Gabor bandwidth W_G which is T_c for DS-CDMA-like signals, and the width along the Doppler axis is the inverse of the Gabor time width T_G which for DS-CDMA-like signals is $\frac{1}{T}$. The dimensions of the box are as follows: (normalized) height is $\frac{T_c}{4T}$, length along the delay axis is $2T$, and width along the doppler axis is $\frac{2}{T_c}$. Thus, most of the volume of the thumb-tack is contributed by the box. Heuristically, it is reasonable to expect $J_C(u)$ to decrease to 0 as $T_c \downarrow 0$, i.e., as the bandwidth of the DS-CDMA signal is increased. This happens if $\phi_H(\tau, \nu)$ is continuous with compact support. This is typical for the channel response function as illustrated in Fig. 2. For example, if $R_H(t, \tau) \equiv \text{sinc}(F_d t)$

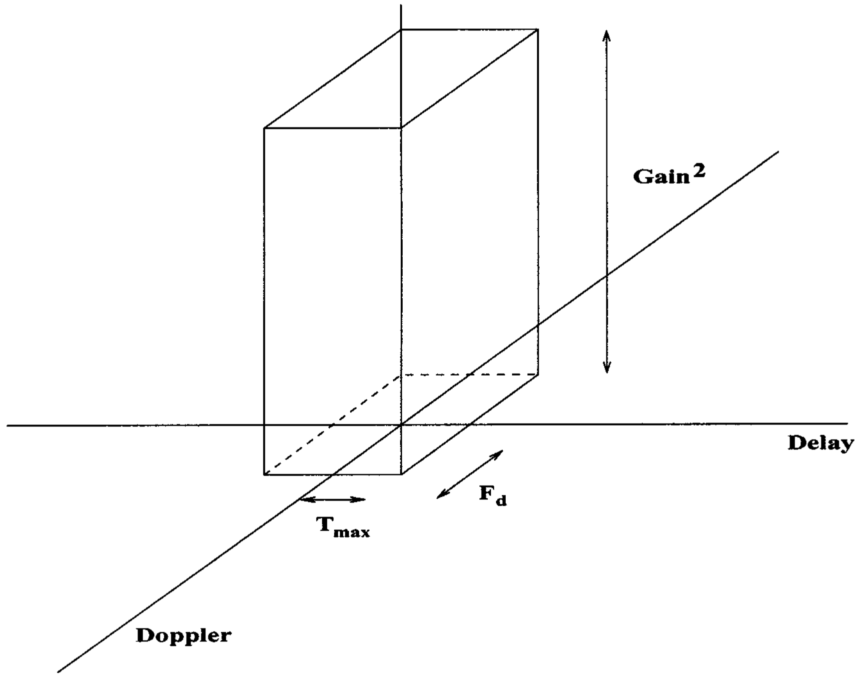

Fig. 2. A typical channel response function.

for all $\tau \in [0, T_{\max}]$, then $\phi(\tau, \nu)$ has support $[-T_{\max}, T_{\max}]$ along the delay-axis and $[-F_d, F_d]$ along the Doppler-axis. *Thus, we expect the rate of information that can be reliably transmitted with DS-CDMA signals over (nice) Gaussian WSSUS fading channels to tend to zero as the spreading factor increases.*

Fig. 1 is actually a generic picture for any signal. Specializing to frequency-hopping-like signals or to M-ary FSK signals, we find that the length of the stump along the delay axis is inversely proportional to the width of the individual frequency slots. Since that width of the frequency slots is fixed irrespective of their number, the bound does not decrease to zero for such signals. This is in conformance with [15], [10], and [25].

As another means of looking at the difference between DS-CDMA and frequency-hopping CDMA performance, we look at the distribution of the signal energy across the time–frequency grid. Roughly speaking, the fourthegy function is a sum over time and frequency bins of the local signal energy squared. Thus, the choice of the distribution of the local signal energy has a significant impact on the value of the fourthegy. It is most convenient to illustrate this for $R(t, \tau) = e^{-(t^2/2+2\tau^2)\pi}$. Using this correlation function it can be shown by expanding out in detail and using properties of the Fourier transform that

$$J_C(u) = \iint |(e^{-\pi t^2} e^{j2\pi ft}) * u(t)|^4 \, dt \, df \approx \sum (\text{local energy})^2.$$

The equivalence of $J_C(u)$ with the sum of the local energy-squared holds if we imagine the signal to assume approximately constant values in balls of unit radius in the time–frequency plane. We also have that

$$\iint |(e^{-\pi t^2} e^{j2\pi ft}) * u(t)|^2 \, dt \, df = \sqrt{\frac{1}{2}} E(u).$$

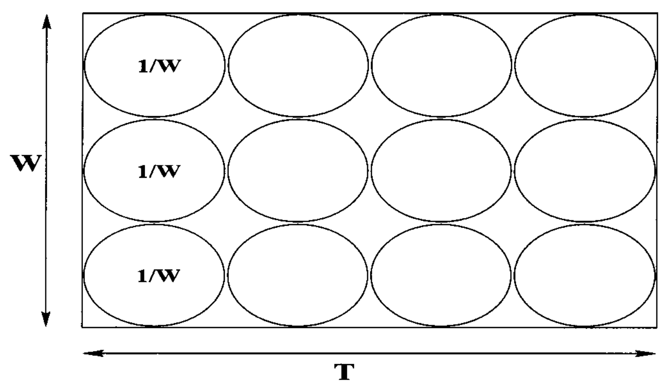

Fig. 3. A typical signal energy distribution pattern for DS-CDMA signal.

Suppose a time–frequency bin is selected at random, uniformly over the WT bins corresponding to a duration T observation of a signal of bandwidth W. This induces a probability distribution of the local energy of the transmitted signal u. The mean is the energy per unit time–frequency. The variance of the local energy of u is equal to the fourthegy per unit time–frequency (i.e., the mean square local energy) minus the square of the mean local energy. Fig. 3 illustrates the signal energy distribution of DS-CDMA signals. It is clear that DS-CDMA signals distribute the signal energy evenly, in other words, in a nonbursty manner. Assuming that the signals are fixed power signals, the energy is proportional to the duration of the signal T. Let W be the bandwidth of the signal. For DS-CDMA signals, the sum of the local energy squared is given by

$$\sum (\text{local energy})^2 \sim WT \frac{1}{W^2} = T/W.$$

Therefore, as the spreading increases, the fourthegy decreases, and so also the mutual information decreases. The variance of the local energy of these signals is zero. Fig. 4 illustrates the

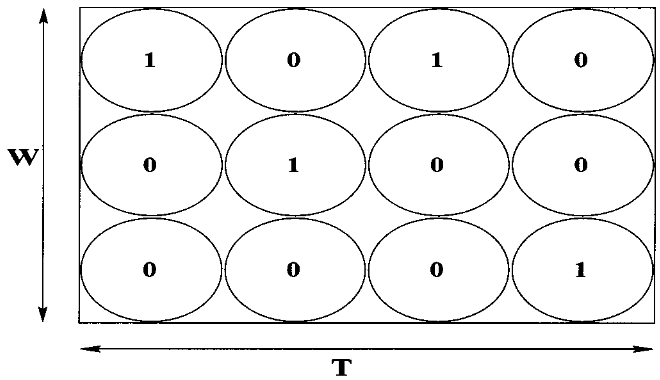

Fig. 4. A typical signal energy distribution pattern for frequency-hopping CDMA (FH-CDMA).

signal energy distribution for frequency-hopping CDMA (FH-CDMA) signals. Note that for such signals, the distribution is not even and is, in fact, bursty with large regions on the time–frequency grid having no energy. For FH-CDMA signals the sum of the local energy squared is given by

$$\sum(\text{local energy})^2 \sim T.$$

Thus, it is clear that the fourthegy of an FH-CDMA-like signal does not decrease with an increase in the bandwidth. The variance of the local energy for such signals is not zero.

A. Bound on DS-CDMA Capacity Per Unit Time

So far, we have given two qualitative arguments to explain how the capacity of DS-CDMA signals decreases as the spreading increases. From this point onwards, the objective is to justify this with quantitative/numerical results. The information rate for DS-CDMA is less than or equal to the product of the fourthegy per unit time of DS-CDMA times the maximum information per unit fourthegy for the channel. By Theorem IV.1, the second term is bounded by $\frac{1}{2\sigma^4}$, so that

$$\text{Information Rate} \le \lim_{T\to\infty} \frac{E[J_C(U)]}{T}\,\frac{1}{2\sigma^4}. \tag{29}$$

In the rest of this section we restrict our attention to diffuse WSSUS channels.

In view of (23), a good first step in calculating the mean fourthegy of DS-CDMA signals per unit time is to compute the mean magnitude squared of the ambiguity function $E[|\chi(\nu,\tau)|^2]$ of a DS-CDMA signal. The DS-CDMA signals are given by

$$u(t) = \sum_{n=0}^{N-1} a_n s(t - nT_c) \tag{30}$$

where a_n are independent and identically distributed (i.i.d.), zero-mean, complex-valued, random variables and $s(t)$, with support $[0, T_c]$ and energy T_c, is the chip waveform. All moments and integrals that appear are assumed to be finite.

Expanding $\chi(\nu, \tau)$ yields that

$$\chi(\nu,\tau) = \sum_{n=0}^{N-1}\sum_{m=0}^{N-1} a_n a_m^* e^{-j2\pi\nu\frac{m+n}{2}T_c}\chi_s(\nu, \tau+(m-n)T_c)$$

where $\chi_s(\nu, \tau)$ is the ambiguity function of $s(t)$. The support of $s(t)$ is $[0, T_c]$, so the support of $\chi_s(\tau, \nu)$ along the τ-axis is $[-T_c, T_c]$, and, therefore,

$$\chi_s(\nu, \tau+(m-n)T_c)\chi_s^*(\nu, \tau+(n-m)T_c) = 0, \qquad \text{if } n \ne m.$$

This observation, and the independence and zero-mean assumptions on the a_n's yields that

$$\begin{aligned} E[|\chi(\nu,\tau)|^2] &= \sum_{n,\,i:\,n\ne i} E[|a_n|^2]E[|a_i|^2] \\ &\quad\times \exp(-j2\pi\nu(n-i)T_c)|\chi_s(\nu,\tau)|^2 \\ &\quad+ \sum_{n,\,m:\,n\ne m} E[|a_n|^2]E[|a_m|^2]|\chi_s(\nu, \tau+(m-n)T_c)|^2 \\ &\quad+ \sum_n E[|a_n|^4]|\chi_s(\nu,\tau)|^2. \end{aligned}$$

Therefore,

$$\begin{aligned} E[|\chi(\nu,\tau)|^2] &= \Bigg[E[|a_n|^2]^2 \sum_{m=-N+1}^{N-1}(N-|m|)e^{-j2\pi m\nu T_c} \\ &\qquad + N(E[|a_n|^4] - E[|a_n|^2]^2)\Bigg]|\chi_s(\nu,\tau)|^2 \\ &\quad + E[|a_n|^2]^2 \sum_{m:\,1\le|m|\le N-1}(N-|m|)|\chi_s(\nu, \tau+mT_c)|^2. \end{aligned} \tag{31}$$

By (23), the mean fourthegy $E[J_C(U)]$ is the integral of $E[|\chi(\nu,\tau)|^2]$ times the channel response function. The next step is to use this fact and the expression (31) to bound $E[J_C(U)]$ above. For simplicity, take a separable channel, i.e., a channel for which each path fades similarly. Thus, assume that

$$r_H(t,\tau) = r_H(t), \qquad \forall\,\tau \ge 0,$$

or, equivalently, that $s_H(f,\tau) = s_H(f)$, where $s_H(f,\tau)$ is the Fourier transform of $r_H(t,\tau)$ and $s_H(f)$ is the Fourier transform of $r_H(t)$. Therefore,

$$R_H(t,\tau) = r_H(t)\gamma_H(\tau) \quad \text{and} \quad \psi_H(\nu,\tau) = \psi_F(\nu)\psi_T(\tau)$$

where

$$\psi_F(\nu) = \int s_H(f)s_H(f+\nu)\,df,$$

$$\text{and} \quad \psi_T(\tau) = \int \gamma_H(t)\gamma_H(t+\tau)\,dt.$$

Assume without loss of generality (since σ^2 can be varied) that $\Gamma_H(\Re) = \int_t \gamma_H(t)\,dt = 1$; in other words, $G_H = 1$. Finally, assume that constant modulus symbols are used, meaning that $|a_n|$ is constant. Note that

$$\int_\tau \psi_T(\tau)\,d\tau = G_H^2 = 1$$

and

$$\int_\nu \psi_F(\nu)\,d\nu = 1.$$

(from $r_H(0) = 1$). Using (23), (31), and the fact $|\chi_s(\nu, \tau)| \le T_c 1_{\{|\tau| \le T_c\}}$ yields the following upper bound:

$$\begin{aligned} E[J_C(U)] \le\ & T_c^2 E[|a_n|^2]^2 \left(\int_{|\tau| \le T_c} \psi_T(\tau)\, d\tau \right) \\ & \times \int_f s_H(f) \int_\nu \sum_{m=-N+1}^{N-1} (N - |m|) \\ & \times e^{-j2\pi m \nu T_c} s_H(f+\nu)\, d\nu\, df \\ & + 2NT_c^2 E[|a_n|^2]^2 \int_\nu \int_\tau \psi_H(\nu, \tau)\, d\tau\, d\nu. \end{aligned} \quad (32)$$

Note that

$$\begin{aligned} \sum_{m=-N+1}^{N-1} & (N - |m|) \int_\nu e^{-j2\pi m\nu T_c} s_H(f+\nu)\, d\nu \\ & = \sum_{m=-N+1}^{N-1} (N - |m|) r_H(-mT_c) e^{j2\pi m f T_c} \end{aligned} \quad (33)$$

and

$$\begin{aligned} \int_f s_H(f) \int_\nu & \sum_{m=-N+1}^{N-1} (N - |m|) e^{-j2\pi m\nu T_c} s_H(f+\nu)\, d\nu\, df \\ & = \sum_{m=-N+1}^{N-1} (N - |m|) r_H(-mT_c) r_H(mT_c). \end{aligned}$$

Since $r_H(-t) = r_H(t)^*$ and $(N - |m|) \le N$, it follows that

$$\begin{aligned} \frac{E[J_C(U)]}{T} & \\ \le & \frac{NT_c E[|a_n|^2]^2}{T} \sum_{m=-N+1}^{N-1} |r_H(mT_c)|^2 T_c \int_{|\tau| \le T_c} \psi_T(\tau)\, d\tau \\ & + 2 \frac{NT_c}{T} T_c E[|a_n|^2]^2 \int_\nu \psi_F(\nu)\, d\nu \int_\tau \psi_T(\tau)\, d\tau. \end{aligned}$$

Therefore,

$$\begin{aligned} \lim_{T \to \infty} & \frac{E[J_C(U)]}{T} \le E[|a_n|^2]^2 \\ & \times \left(2T_c + \sum_{m=-\infty}^{+\infty} |r_H(mT_c)|^2 T_c \int_{|\tau| \le T_c} \psi_T(\tau)\, d\tau \right). \end{aligned} \quad (34)$$

The ratio of received signal power to noise spectral density is $\frac{P}{N_0}$, where P is the received signal power given by $P = E[|a_n|^2]$, and $N_0 = \sigma^2$. Of course, $\frac{P}{N_0} = \frac{E_b}{N_0} \times$ (data rate), where E_b is the received energy per bit. Combining (29) and (34) yields the following corollary to Theorem IV.1.

Corollary V.1: The information rate for DS-CDMA signaling with constant modulus symbols transmitted over a separable WSSUS fading channel satisfies the following:

$$\begin{aligned} \text{Information Rate} \le & \left(\frac{P}{N_0} \right)^2 \\ \times & \left(T_c + \frac{1}{2} \left(\sum_{m=\infty}^{+\infty} |r_H(mT_c)|^2 T_c \right) \int_{|\tau| \le T_c} \psi_T(\tau)\, d\tau \right). \end{aligned} \quad (35)$$

If $s_H(f)$ is band-limited with bandwidth less than $\frac{1}{2T_c}$, i.e., if the maximum doppler frequency is less than $\frac{1}{2T_c}$ (which would be common in practice) then the sampling theorem yields

$$\frac{1}{2} \sum_{m=-\infty}^{+\infty} |r_H(mT_c)|^2 T_c = T_{\mathrm{coh}}$$

where T_{coh} is the coherence time of the channel defined somewhat arbitrarily by

$$T_{\mathrm{coh}} = \frac{1}{2} \int_t |r_H(t)|^2\, dt = \frac{1}{2} \int_f |s_H(f)|^2\, df.$$

The delay power density is said to be uniform if

$$\gamma_H = \frac{1}{T_{\max}} 1_{[0,\, T_{\max}]}.$$

Corollary V.2: Suppose that the maximum doppler frequency is finite and less than $\frac{1}{2T_c}$, and suppose the delay power density is uniform. Then the information rate for DS-CDMA signaling with constant modulus symbols transmitted over a separable WSSUS fading channel satisfies the following:

$$\text{Information Rate} \le \left(\frac{P}{N_0} \right)^2 T_c \left(1 + \frac{2T_{\mathrm{coh}}}{T_{\max}} \right). \quad (36)$$

The corollaries imply that for fixed power, DS-CDMA signals convey less information per unit time, as the spreading increases (i.e., as $T_c \to 0$). In fact, the rate is proportional to T_c, and hence inversely proportional to the bandwidth over which the signal is spread.

The bounds in Corollaries V.1 and V.2 hold for any time-limited chip waveform (time-limited to the chip duration). Tighter bounds can be obtained for specific chip waveforms. In the rest of this section, we will specialize to the case of a rectangular chip waveform for which

$$|\chi_s(\nu, \tau)| = (T_c - |\tau|)^+ \mathrm{sinc}(\nu(T_c - |\tau|))$$

where $x^+ = \max(x, 0)$. For a rectangular chip waveform it is clear that $|\chi_s(\nu, \tau)| \le (T_c - |\tau|)^+$. Using this, rather than the weaker but more general bound $|\chi_s(\nu, \tau)| \le T_c 1_{\{|\tau| \le T_c\}}$ in (31) yields the following modification to (34):

$$\begin{aligned} \lim_{T \to \infty} \frac{E[J_C(U)]}{TE[|a_n|^2]^2} \le & \left(2T_c + \sum_{m=-\infty}^{+\infty} |r_H(mT_c)|^2 \right. \\ & \left. \times T_c \int_{|\tau| \le T_c} \left(1 - \frac{|\tau|}{T_c} \right)^2 \psi_T(\tau)\, d\tau \right). \end{aligned} \quad (37)$$

With the uniform power density assumption, Corollaries V.1 and V.2 can be modified as follows.

Corollary V.3: The information rate for DS-CDMA signaling with constant modulus symbols and a rectangular chip waveform transmitted over a separable WSSUS fading channel with a uniform power density satisfies the following:

$$\begin{aligned} & \mathit{Information\ Rate} \\ & \le \left(\frac{P}{N_0} \right)^2 \left(T_c + \frac{1}{2} \left(\sum_{m=\infty}^{+\infty} |r_H(mT_c)|^2 T_c \right) \right. \\ & \qquad \left. \times \int_{|\tau| \le T_c} \left(1 - \frac{|\tau|}{T_c} \right)^2 \psi_T(\tau)\, d\tau \right). \end{aligned} \quad (38)$$

Corollary V.4: Suppose that the maximum doppler frequency is finite and less than $\frac{1}{2T_c}$, and suppose the delay power

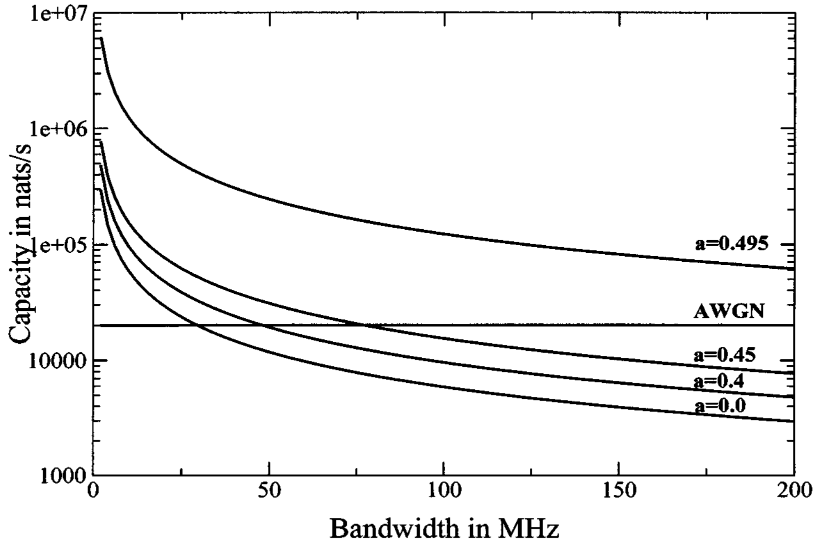

Fig. 5. Upper bound (39) for various a with $\frac{P}{N_0} = 2 \times 10^4$ Hz.

density is uniform. Then the information rate for DS-CDMA signaling with constant modulus symbols and a rectangular chip waveform transmitted over a separable WSSUS fading channel satisfies the following:

$$\text{Information Rate} \le \left(\frac{P}{N_0}\right)^2 T_c \left(1 + \frac{2T_{\text{coh}}}{3T_{\max}}\right). \quad (39)$$

B. Numerical Results

The bounds on information rate for DS-CDMA signals given in Corollaries V.1–V.4 depend strongly on the correlation function, or equivalently the doppler spectrum, of the channel. In particular, $\sum_{m=-\infty}^{+\infty} |r_H(mT_c)|^2$ needs to be finite for the bound in (35) and (38) to be finite. Frequently in the literature, the doppler spectrum is assumed to be the Clarke spectrum, which corresponds to a uniform distribution of received power over all angles of arrival in two dimensions. For such spectrum, the correlation function tends to zero as $1/\sqrt{t}$, so that T_{coh} is infinite. Therefore, the bounds given in Corollaries V.1–V.4 are infinite for the Clarke spectrum. Moreover, it is shown in Appendix D that $\frac{E[J_C[U]]}{T} \to \infty$ for DS-CDMA signaling over a channel with the Clarke spectrum. Thus, the approach of considering fourthegy per unit time is not fruitful for the case of the Clarke spectrum. The numerical results reported in this section are thus for channels for which the correlation decays more quickly than for the Clarke spectrum.

For the first set of channels it is assumed that $T_{\max} = 1\ \mu$s and the maximum doppler frequency is $F_d = 200$ Hz. The family of channel correlation functions considered is given by

$$s_H(f) \propto \frac{1}{\left(1 - \left(\frac{f}{F_d}\right)^2\right)^a}, \qquad \text{for } 0 \le a < 0.5.$$

If a converges to 0.5, this spectrum converges to the Clarke spectrum, whereas if a is near 0, the spectrum has much milder singularities, so that the correlation function decays much more quickly. The value for $\frac{P}{N_0}$ is assumed to be 2×10^4 Hz. This numerical value arises, for example, for a system with a bit rate of 10 kbit/s operating with $\frac{E_b}{N_0} = 3$ dB. The data rate 10 kbits/s is roughly the minimum data rate, and the value 3 dB is roughly the value of $\frac{E_b}{N_0}$, targeted for third-generation cellular systems such as the emerging wide-band CDMA systems proposed for UMTS. We also take bandwidth $= \frac{1}{T_c}$. Fig. 5 displays the upper bound (39) for different values of a. A region of interest in the figure is the set of bandwidths such that the upper bound falls below the capacity of an AWGN channel with the same $\frac{P}{N_0}$. As expected, the upper bound converges to infinity as a tends to 0.5.

The remainder of the numerical results are for the channel correlation given by a two-sided exponential

$$R_H(t, \tau) = \exp(-F_d|t|)\frac{1}{T_{\max}} 1_{[0, T_{\max}]}.$$

This correlation function decays more quickly than the variations of the Clarke spectrum considered above. Again, assume a uniform power density. The upper bound in (38) and the inequality $1 + x \le \exp(x)$ yield

Information Rate

$$\le \left(\frac{P}{N_0}\right)^2 T_c \left(1 + \frac{T_c}{3T_{\max}} + \frac{1}{3F_d T_{\max}}\right). \quad (40)$$

Fig. 6 displays this bound for several different values of F_d, with $\frac{P}{N_0} = 2 \times 10^4$ Hz and $T_{\max} = 1\ \mu$s as before.

In future years, even more sensitive transmission systems will be sought, so that smaller values of $\frac{P}{N_0}$ may be relevant. As an example of how this changes the bounds, Fig. 7 shows the same upper bound for the same correlation function as in Fig. 6, except that $\frac{P}{N_0}$ is halved to 10^4 Hz. Here we find that the bandwidth at which the upper bound falls below the AWGN channel capacity is approximately half of the same value for the larger $\frac{P}{N_0}$. Detrimental effects of overspreading are indicated in Fig. 6 for a bandwidth of 8 MHz, and are indicated in Fig. 7 for a bandwidth of 4 MHz. These bandwidths are in the range of currently emerging third-generation commercial systems.

The bound (40) can also be used to produce a lower bound on $\frac{E_b}{N_0}$ for a given bandwidth and data rate, as illustrated in Fig. 8. The figure is based on a 20-MHz DS-CDMA system using $F_d = 1000$ Hz and $T_{\max} = 1\ \mu$s. Data rates from 8 to

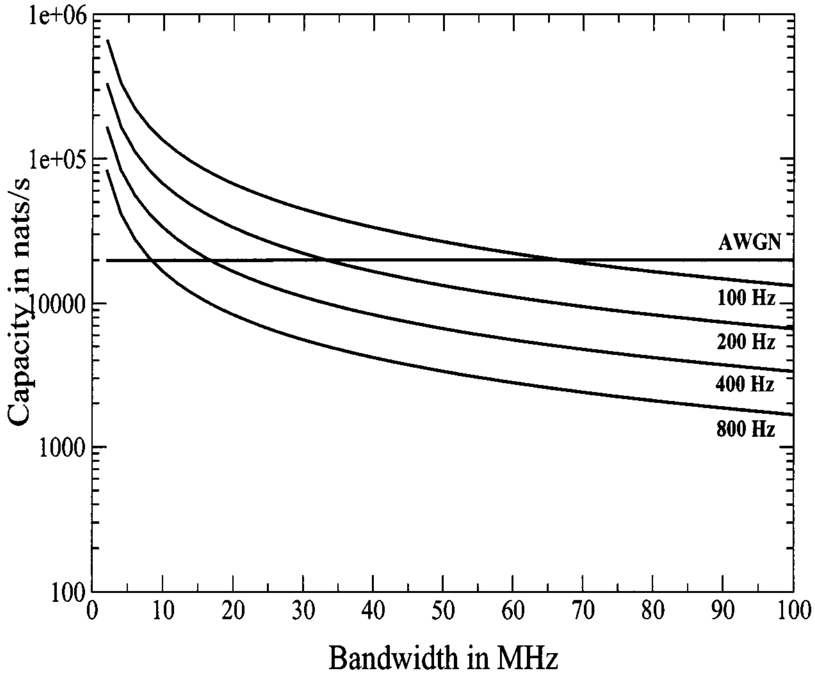

Fig. 6. Upper bound (40) for various F_d with $\frac{P}{N_0} = 2 \times 10^4$ Hz.

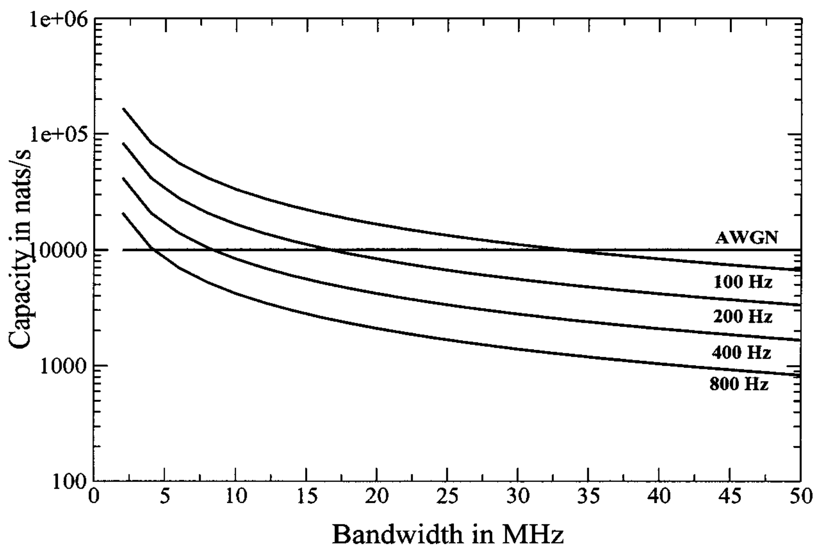

Fig. 7. Upper bound (40) for various F_d with $\frac{P}{N_0} = 10^4$ Hz.

256 kbits/s are considered. For each data rate, the bound (40) implies a lower bound on $\frac{E_b}{N_0}$. As noted in Section IV-C, it is known that $\frac{E_b}{N_0}$ must be at least -1.6 dB (same as for a nonfading AWGN channel). The larger of these two lower bounds is pictured for each data rate. Note that the required $\frac{E_b}{N_0}$ is considerably larger for the smaller data rates. Fig. 8 is qualitatively the same as a figure based on extensive system engineering and simulation for the emerging WCDMA standard for UMTS [13, Fig. 10.4].

The focus of this section is on upper bounds on the information rate of DS-CDMA as the bandwidth is increased, for fixed power. Another interesting limit is the case that the doppler spread tends to infinity for fixed bandwidth. Viterbi [28] showed that for FSK that is not bursty in the time domain, the information rate converges to zero in this limit. This fact is reflected in Figs. 6 and 7 because for practical systems the dominant term in (40) is the last one, which is inversely proportional to F_d. The remaining terms in the bound (40) do not converge to zero as $F_d \to \infty$, but an alternative analysis applied to the expressions for fourthegy per unit time with $s(t)$ assumed to be a rectangular pulse can be used to show that the information rate indeed converges to zero as $F_d \to \infty$. Since the remaining terms are very small for practical systems, the details are omitted.

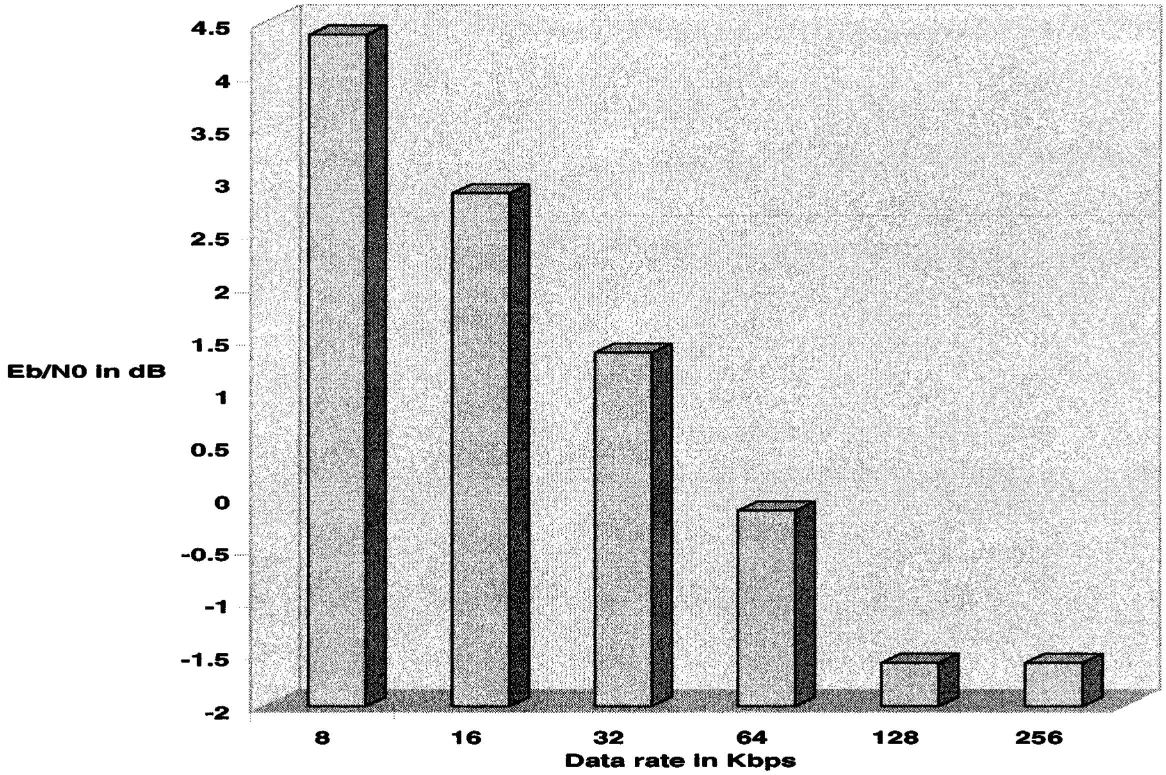

Fig. 8. Lower bound on $\frac{E_b}{N_0}$ requirement for different data rates for DS-CDMA system with bandwidth 20 MHz, $F_d = 1000$ Hz, and $T_{\max} = 1$ μs for two-sided exponential correlation function.

C. Specular Multipath Channels

In this subsection, we concentrate on specular WSSUS multipath fading channels. Since we are considering a Gaussian channel it is sufficient to specify the correlation function $R_H(t, \tau)$. For an L-path specular WSSUS multipath channel the following form for $R_H(t, \tau)$ holds:

$$R_H(t, \tau) = \sum_{l=1}^{L} \Gamma_H^l r_H(t, \tau_l)\delta(\tau - \tau_l) \tag{41}$$

where $\{\tau_1, \tau_2, \ldots, \tau_L\}$ are the time offsets of the various multipath components. Thus, $S_H(f, \tau)$ is given by

$$S_H(f, \tau) = \sum_{l=1}^{L} \Gamma_H^l s_H(f, \tau_l)\delta(\tau - \tau_l)$$

where $s_H(f, \tau)$ is the Fourier transform of $r_H(t, \tau)$. Therefore, $\psi_H(\tau, \nu)$ is given by

$$\begin{aligned}\psi_H(\nu, \tau) &= \iint S_H(f, t) S_H(f + \nu, t + \tau)\, dt\, df \\ &= \sum_{l=1}^{L}\sum_{k=1}^{L} \Gamma_H^l \Gamma_H^k \int s_H(f, \tau_l) \\ &\quad \times s_H(f + \nu, \tau_k)\, df\, \delta(\tau + \tau_l - \tau_k).\end{aligned}$$

Finally, the following expression for $J_C(u)$ holds:

$$J_C(u) = \iint |\chi(\nu, \tau)|^2 \psi(\nu, \tau)\, d\tau\, d\nu = \sum_{l=1}^{L}\sum_{k=1}^{L} J_C^{lk}(u) \tag{42}$$

where

$$J_C^{lk}(u) = \int |\chi(\nu, \tau_k - \tau_l)|^2 \psi_H^{lk}(\nu)\, d\nu \tag{43}$$

with

$$\psi_H^{lk}(\nu) = \Gamma_H^l \Gamma_H^k \int s_H(f, \tau_l) s_H(f + \nu, \tau_k)\, df.$$

Before going into detail, let us pause briefly to summarize how we will proceed to bound $J_C(u)$ from above. Roughly speaking, if there are many paths each with approximately the same energy, and if the total average received energy is fixed, then ψ_H^{lk} scales as $\frac{1}{L^2}$. Considering DS-CDMA-type signals for small enough T_c, we can expect the diagonal terms $J_C^{ll}(u)$ to dominate in the right-hand side of (42). Since there are only L dominant terms, we can expect the mutual information between the input and the output to be small for large spreading factors. In the rest of the subsection we make this statement precise.

Let

$$T_c < \min_{k, l \in \{1, 2, \ldots, L\}: k \neq l} |\tau_k - \tau_l|$$

then the terms in the right-hand side of (42) fall into two groups.

i) $l = k$. From (31) we have that

$$\begin{aligned}&E[|\chi(\nu, 0)|^2] \\ &\quad = E[|a_n|^2]^2 \left(\sum_{m=-N+1}^{N-1} (N - |m|) \exp(-j2\pi m\nu T_c) \right) \\ &\qquad \times |\chi_s(\nu, 0)|^2 + N(E[|a_n|^4] - E[|a_n|^2]^2)|\chi_s(\nu, 0)|^2.\end{aligned}$$

Let

$$r(t, 0) = \mathcal{F}^{-1}(|\chi_s(\nu, 0)|^2)$$

be the inverse Fourier transform of $|\chi_s(\nu, 0)|^2$. Note that $r(t) = |s(t)|^2 * |s(-t)|^2$. Considering individual terms gives

$$\begin{aligned}
&\int_\nu \int_f e^{-j2\pi m\nu T_c} |\chi_s(\nu, 0)|^2 s_H(f+\nu, \tau_l) s_H(f, \tau_l)\, df\, d\nu \\
&= \int_f s_H(f, \tau_l) \int_\nu e^{j2\pi m\nu T_c} |\chi_s(\nu, 0)|^2 \\
&\quad \times s_H(f-\nu, \tau_l)\, d\nu\, df \\
&= \int_t r_H(t, \tau_l) r(t - mT_c, 0)(r_H(t, \tau_l))^*\, dt \quad \text{(Parseval)} \\
&= \int |r_H(t, \tau_l)|^2 r(t - mT_c, 0)\, dt.
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
&\sum_{l=1}^{L} \int_\nu E[|\chi(\nu, 0)|^2] \psi_H^{ll}(\nu)\, d\nu \\
&\quad = \int_t \left(\sum_{l=1}^{L} \left(\Gamma_H^l\right)^2 |r_H(t, \tau_l)|^2 \right) \\
&\quad \times \left[E[|a_n|^2]^2 \sum_{m=-N+1}^{N-1} (N - |m|) r(t - mT_c, 0) \right] dt \\
&\quad + N(E[|a_n|^4] - E[|a_n|^2]^2) \\
&\quad \times \int_t \left(\sum_{l=1}^{L} \left(\Gamma_H^l\right)^2 |r_H(t, \tau_l)|^2 \right) r(t, 0)\, dt.
\end{aligned}$$

ii) $l \neq k$. Then we have $\tau_l - \tau_k = \tilde{m}T_c + \delta$ where $\tilde{m}$ is an integer not equal to 0 or -1, and $\delta \geq 0$. Define

$$r(t, \tau) = \mathcal{F}^{-1}(|\chi_s(\nu, \tau)|^2).$$

Then the contribution of such terms to $E[J_C(U)]$ is given by

$$\begin{aligned}
&\int_\nu |\chi(\nu, \tau_k - \tau_l)|^2 \psi_H^{lk}(\nu)\, d\nu \\
&\quad = E[|a_n|^2]^2 \sum_{m=1}^{N-1} (N - m) \Gamma_H^l \Gamma_H^k \\
&\quad\quad \times \int_t r_H(t, \tau_l)(r_H(t, \tau_k) \tilde{r}(t, mT_c, \tilde{m}T_c + \delta))^*\, dt
\end{aligned}$$

where $\tilde{r}(t, s, \tau)$ is given by

$$\tilde{r}(t, s, \tau) = r(t - s - \tau, s + \tau) + r(t + s - \tau, -s + \tau).$$

For simplicity, consider the case in which the chip waveform is a rectangular pulse, the channel is separable, and the gain is 1. Then

$$r(t, \tau) = ((T_c - |\tau|)^+ - |t|)^+ \quad \text{and} \quad \sum_{l=1}^{L} \Gamma_H^l = 1.$$

In this case, we can upper-bound $E[J_C(U)]$ as

$$\begin{aligned}
&E[J_C(U)] \\
&\leq NT_c E[|a_n|^2]^2 \sum_{l=1}^{L} \left(\Gamma_H^l\right)^2 \int_{-NT_c}^{NT_c} |r_H(t)|^2\, dt \\
&\quad + NT_c(E[|a_n|^4] - E[|a_n|^2]^2) \sum_{l=1}^{L} \left(\Gamma_H^l\right)^2 \int_{-T_c}^{T_c} |r_H(t)|^2\, dt \\
&\quad + 2E[|a_n|^2]^2 NT_c \left(1 - \sum_{l=1}^{L} \left(\Gamma_H^l\right)^2\right) \int_{-T_c}^{T_c} |r_H(t)|^2\, dt.
\end{aligned}$$

Thus,

$$\begin{aligned}
&\lim_{T\to\infty} \frac{E[J_C(U)]}{T} \\
&\quad \leq E[|a_n|^2]^2 \sum_{l=1}^{L} \left(\Gamma_H^l\right)^2 \int |r_H(t)|^2\, dt \\
&\quad\quad + (E[|a_n|^4] - E[|a_n|^2]^2) \sum_{l=1}^{L} \left(\Gamma_H^l\right)^2 \int_{-T_c}^{T_c} |r_H(t)|^2\, dt \\
&\quad\quad + 2E[|a_n|^2]^2 \left(1 - \sum_{l=1}^{L} \left(\Gamma_H^l\right)^2\right) \int_{-T_c}^{T_c} |r_H(t)|^2\, dt.
\end{aligned}$$

Now, letting T_c tend to 0, we get

$$\lim_{T_c\to 0} \lim_{T\to\infty} \frac{E[J_C(U)]}{T} \leq E[|a_n|^2]^2 \sum_{l=1}^{L} \left(\Gamma_H^l\right)^2 \int |r_H(t)|^2\, dt. \tag{44}$$

If it is now assumed that all paths have equal energy, then

$$\Gamma_H^l = \frac{1}{L} \quad \text{and} \quad \sum_{l=1}^{L} (\Gamma_H^l)^2 = \frac{1}{L}.$$

Therefore, the capacity per unit time is inversely proportional to the number of paths. Specializing to the case of [25] with Gaussian fading and realizing that

$$E[|a_n|^2] = P \quad \text{and} \quad r_H(t) = \max\left(1 - \frac{|t|}{T_{\text{coherence}}}, 0\right)$$

where $T_{\text{coherence}}$ is the coherence time of the channel as defined in [25], we can extend their upper bound on the capacity per unit time for very large spreading factors, namely, $\frac{2P^2 T_{\text{coherence}}}{3N_0^2 L}$, to channels with ISI. Hence, we can conclude that if there are many multipath components, then the information rate that can be transmitted reliably with DS-CDMA-like signals is small.

VI. Discussion

This paper reinforces the conclusions of Médard and Gallager [19] that signals need to be bursty in time and/or frequency to be able to achieve constant information rates per unit power over very-wide-band WSSUS fading channels. Smooth signals like those used in direct-sequence spread-spectrum systems do not have enough fourthegy per unit energy to achieve significant values of reliably communicated bits per unit energy for a WSSUS fading channel. In particular, detrimental effects of overspreading on the required energy to interference ratio are observed in Section V-B for a channel and modulation scheme not far from currently emerging CDMA systems operating at their lowest data rates. This loss in capacity for DS-CDMA signals has also been observed in practice [13, p. 245] where it is stated that, "The main reason why the E_b/N_0 depends on the bit rate is that the [control channel] is needed to keep the physical layer connection running and it contains reference symbols for channel estimation and power control signaling bits. The E_b/N_0 performance depends on the accuracy of the channel … estimation algorithms."

Numerical evaluation of the upper bounds on the information rate for direct-sequence spread-spectrum-like signals shows that these bounds are informative for large bandwidths which are close to the bandwidths for future broad-band systems. The numerical bounds suggest that for ultra-wide-band systems (20–50 MHz or more and for data rates in the tens of kilobits per second) DS-CDMA-type signaling is inefficient. This may well explain why most proposals for ultra-wide-band systems call for pulse-position modulation or on–off modulation with long off periods, which are highly bursty in the time domain.

A caveat to these conclusions is that they are based on numerical examples for a few specific channel correlation functions. For some correlation functions, such as that for the two-dimensional isotropic scattering (Clarke's spectrum), the upper bounds on capacity are infinite.

Appendix A
Alternative Proof of (5)

The following alternative proof of the basic inequality (5) was suggested by a reviewer. The proof uses the equation

$$I(Y;U) = I(Y;H,U) - I(Y;H|U)$$

which was exploited by [28], as discussed in the Introduction, and highlighted by [3]. The notation Σ_u, $\phi(u)$, and $J_C(u)$ used in this appendix is the same as in Section II-B. Since Y depends on H and U only through the product $S = H^\dagger U$, it follows that $I(Y;H,U) = I(Y;S)$. Since S is a mean zero vector, its covariance matrix is given by $E[SS^\dagger] = E[\Sigma_U]$. Here, Σ_U is the matrix Σ_u evaluated at $u = U$, and the expectation in $E[\Sigma_U]$ is with respect to U. Since Y is obtained from S by the addition of Gaussian noise, the mutual information $I(Y;S)$ is less than or equal to what it would be if S were Gaussian with the same covariance. This and the inequality $\log(1+\mu) \le \mu$ applied to the eigenvalues of $E\Sigma_U$ yield

$$\begin{aligned} I(Y;H,U) &\le \log\left(\det\left(I + \frac{1}{\sigma^2} E\Sigma_U\right)\right) \\ &\le \frac{1}{\sigma^2}\operatorname{Tr}(E\Sigma_U) = E\left[\frac{1}{\sigma^2}\operatorname{Tr}(\Sigma_U)\right]. \end{aligned}$$

On the other hand, for a given U, Y is the output of a Gaussian additive noise channel with input H, so that

$$I(Y;H|U) = E\left[\log\left(\det\left(I + \frac{1}{\sigma^2}\Sigma_U\right)\right)\right].$$

Therefore, writing $\lambda_i(U)$ for the eigenvalues of Σ_U and using the inequality $\log(1+x) \ge x - \frac{x^2}{2}$,

$$\begin{aligned} I(Y;U) &\le E\left[\frac{1}{\sigma^2}\operatorname{Tr}(\Sigma_U) - \log\left(\det\left(I + \frac{1}{\sigma^2}\Sigma_U\right)\right)\right] \\ &= E\left[\sum_i \phi(\lambda_i(U))\right] \\ &\le \frac{1}{2\sigma^4} E\left[\sum_i \lambda_i(U)^2\right] = \frac{1}{2\sigma^4} E[J_C(U)] \end{aligned}$$

so that (5) is proved.

Appendix B
Proof of Proposition III.1

First we define a family of random processes h_n, all on the same probability space, and we will define h by letting $n \to \infty$. Define

$$h_n(t,\tau) = n\, Z_{n,i}(t), \qquad \text{if } \frac{i}{n} < \tau \le \frac{i+1}{n}$$

where for each n $Z_{n,i}$ is a mean-zero Gaussian random process with autocorrelation function

$$\int_{(\frac{i}{n}, \frac{i+1}{n}]} r(t,\tau)\Gamma_H(d\tau).$$

Suppose that the $Z_{n,i}$ are independent for distinct values of i, and that whenever $n = 2^k$ for some $k \ge 0$

$$Z_{n,i}(t) = (Z_{2n,2i}(t) + Z_{2n,2i+1}(t)). \tag{45}$$

The above requirements are consistent since (45) implies

$$\begin{aligned} &E[Z_{n,i}(t+s)Z^*_{n,i}(s)] \\ &\quad = E[Z_{2n,2i}(t+s)Z^*_{2n,2i}(s) \\ &\qquad + Z_{2n,2i+1}(t+s)Z^*_{2n,2i+1}(t+s)] \\ &\quad = \int_{(\frac{2i}{2n}, \frac{2i+1}{2n}]\cup(\frac{2i+1}{2n}, \frac{2i+2}{2n}]} r_H(t,\tau)\Gamma_H(d\tau) \\ &\quad = \int_{(\frac{i}{n}, \frac{i+1}{n}]} r_H(t,\tau)\Gamma_H(d\tau) \end{aligned}$$

as required.

Let $C_0(\Re)$ be the collection of all continuous functions on $\Re$ with compact support. Let $\phi \in C_0(\Re)$ and let $t \in \Re$. We show that

$$Z_n \triangleq \int h_n(t,\tau)\phi(\tau)\,d\tau$$

converges in L^2 as $n = 2^k \to +\infty$. It suffices to show that (Z_n) is a Cauchy sequence or equivalently that $\lim_{n,m\to\infty} E[Z_n Z_m^*]$ exists and is finite. Now

$$\begin{aligned} E[Z_n Z_m^*] &= E\left[\int h_n(t,\tau)\phi(\tau)\,d\tau \int h_m^*(t,\sigma)\phi^*(\sigma)\,d\sigma\right] \\ &= \iint \phi(\tau)R_{n,m}(t,\tau,\sigma)\phi^*(\sigma)\,d\tau\,d\sigma \end{aligned}$$

where

$$R_{n,m}(t,\tau,\sigma) = mn\int_{(\frac{i}{m}, \frac{i+1}{m}]} r_H(t,\tau)\Gamma_H(d\tau)$$

if $\tau \in (\frac{j}{n}, \frac{j+1}{n}]$ and

$$\sigma \in \left(\frac{i}{m}, \frac{i+1}{m}\right] \subset \left(\frac{j}{n}, \frac{j+1}{n}\right]$$

and, otherwise, $R_{n,m}(t, \tau, \sigma) = 0$. Without loss of generality we have assumed that $n \leq m$. So the measure given by $R_{n,m}(t, \tau, \sigma)\, d\tau d\sigma$ converges weakly to the measure $r(t, \tau)\delta(\tau - \sigma)\Gamma_H(d\tau)$ and Z_n converges in L^2.

If $\phi, \psi \in C_0(\Re)$ and $s, t \in \Re$

$$\begin{aligned} V_n(\phi, t, \psi, s) &\triangleq E\left[\int h_n(t, \tau)\phi(\tau)\, d\tau \int h_n^*(s, \psi)\psi^*(\sigma)\, d\sigma\right] \\ &= \int \phi(\tau)R_n(t - s, \tau, \sigma)\psi(\sigma)\, d\tau\, d\sigma \end{aligned}$$

where

$$R_n(t, \tau, \sigma) = n^2 \int_{(\frac{i}{n}, \frac{i+1}{n}]} r_H(t, \tau)\Gamma_H(d\tau)$$

if $\tau, \sigma \in (\frac{i}{n}, \frac{i+1}{n}]$ for some i, and, otherwise, $R_n(t, \tau, \sigma) = 0$. So

$$V_n(\phi, t, \psi, s) = \int \phi_n(\tau)r_H(t - s, \tau)\psi_n^*(\tau)\Gamma_H(d\tau)$$

where

$$\phi_n(\tau) = n \int_{(\frac{i}{n}, \frac{i+1}{n}]} \phi(\tau)\, d\tau, \qquad \text{for } \tau \in \left(\frac{i}{n}, \frac{i+1}{n}\right]$$

and ψ_n is defined similarly. But $\phi_n \to \phi$ uniformly and $\psi_n \to \psi$ uniformly, so

$$V_n \to \int \phi(\tau)r_H(t - s, \tau)\psi^*(\tau)\Gamma_H(d\tau).$$

Thus, we have described a limiting procedure allowing us to construct a random variable $\Theta(\phi, t)$ for each (ϕ, t) with $\phi \in C_0(\Re)$ and $t \in \Re$ so that

$$E[\Theta(\phi, t)\Theta^*(\psi, s)] = \int \phi(\tau)r_H(t - s, \tau)\psi^*(\tau)\Gamma_H(d\tau). \tag{46}$$

We thus take $\Theta(\phi, t)$ to be the definition of $\int \phi(\tau)h(t, \tau)\, d\tau$.

For the specific case of $\phi(\tau) = u(t - \tau)$ for $u \in C_0(\Re)$ and t fixed, we can define

$$s_{\text{out}}(u; t) \triangleq \int u(t - \tau)h(t, \tau)\, d\tau.$$

Relation (12) is a consequence of (46). If $u \in C_0(\Re)$, then $s_{\text{out}}(u; t)$ is mean-square continuous, so by results of [6, pp. 61–62] there exists a separable and measurable version of $s_{\text{out}}(u; t)$. Moreover

$$E\left[\left(\int |s_{\text{out}}(u; t)|^2\, dt\right)\right] = G_H \int |u(t)|^2\, dt. \tag{47}$$

Let $\mathcal{H}$ be the Hilbert space of measurable mean-zero Gaussian random processes on the underlying probability space with norm $(E[\int |s(t)|^2 dt])^{\frac{1}{2}}$. The mapping $\Phi(\cdot)$ defined by

$$\Phi(u) = \frac{1}{\sqrt{G_H}}\, s_{\text{out}}(u; .)$$

is an isomorphism from $C_0(\Re)$, which is a dense subset of $L^2([0, T])$, to $\mathcal{H}$. The mapping Φ can, therefore, be extended to an isomorphism for all of $L^2([0, T])$ into $\mathcal{H}$, which we again call Φ. For any u in $L^2([0, T])$, we define $s_{\text{out}}(u; .)$ to be $\sqrt{G_H}\Phi(u)$. Note that $s_{\text{out}}(u; .)$ is a measurable Gaussian random process, (47) holds, and (12) continues to hold.

APPENDIX C
PROOF OF PROPOSITION III.2

Since s is a measurable Gaussian random process with finite mean energy, the measure induced by Y on the Borel subsets of $C[0, T]$ is absolutely continuous with respect to the measure induced by σW [17, Theorem 7.16]. This result does not require that s be mean-square-continuous. We shall now present a proof of this fact, and at the same time identify the Radon–Nikodym derivative.

Let $L^2[0, T]$ denote the space of complex-valued square integrable functions on $[0, T]$ with inner product given by

$$(\phi, \xi) = \int_0^T \phi(t)\xi^*(t)\, dt.$$

The autocorrelation function Σ of s is the kernel of a linear operator on $L^2[0, T]$, which we again call Σ, defined by

$$\Sigma\phi(s) = \int_0^T \Sigma(s, t)\phi(t)\, dt.$$

The operator Σ is symmetric (i.e., $(\psi, \Sigma\xi) = (\xi, \Sigma\psi)^*$) and nonnegative (i.e., $(\psi, \Sigma\psi) \geq 0$). Also, for an arbitrary complete, orthonormal basis (ψ_n) of $L^2[0, T]$

$$\begin{aligned} \sum_n (\psi_n, \Sigma\psi_n) &= \sum_n E[|(\psi_n, s)|^2] = E\left[\sum_n |(\psi_n, s)|^2\right] \\ &= E\left[\int_0^T |s(t)|^2\right]\, dt < \infty \end{aligned}$$

so that Σ has finite trace given by

$$\text{Trace}(\Sigma) = \sum_n (\psi_n, \Sigma\psi_n) = \int_0^T \Sigma(t, t)\, dt.$$

Hence, Σ is also a compact operator and, by the Hilbert–Schmidt theorem, it has a complete orthonormal basis of eigenfunctions (ψ_n) and associated eigenvalues (λ_n) [23]. Therefore, $\Sigma\psi_n = \lambda_n\psi_n$, and also, $\text{Trace}(\Sigma) = \sum_n \lambda_n$. The observations $(Y_t\colon 0 \leq t \leq T)$ have the same information content up to sets of measure zero as $(Z_n\colon n \geq 1)$, where Z_n is defined by mean-square integration

$$Z_n = \int_0^T Y_t\psi_n(t)\, dt.$$

To see this, start with the fact that for each t,

$$Y_t = \lim_{n\to\infty} \sum_{i=1}^n (\psi_n, p^t)Z_n \text{ (mean square sense)}$$

where p^t denotes the indicator function of the interval $[0, t]$. Mean square convergence implies almost sure convergence along a subsequence, and Y_t only needs to be recovered for rational values of t since it is sample continuous. Thus, up to sets of measure zero, the information in (Z_n) is indeed the same as that of Y. Under Y, the Z_n's are independent with Z_n distributed as $\mathcal{CN}(0, \lambda_n + \sigma^2)$ for each n, and under σW, the Z_n's are i.i.d. with distribution $\mathcal{CN}(0, \sigma^2)$.

The Radon–Nikodym derivative for n observations $Z_1^n = (Z_1, \ldots, Z_n)$ [29] is given by

$$L_n = \frac{d\mu_{Z_1^n}}{d\mu_{\sigma W}} = \exp\left(\sum_{1=1}^{n} -\log\left(1+\frac{\lambda_i}{\sigma^2}\right) + \frac{\lambda_i |z_i|^2}{\sigma^2(\lambda_i+\sigma^2)}\right).$$

By general theory, the sequence $(L_n\colon n \geq 1)$ is a martingale under the measure of σW. Direct computation shows that if p is a number with $1 < p < \min(\lambda_i + \sigma^2)/\lambda_i$, then $E[L_n^p]$ is uniformly bounded in n, where the expectation is taken under the measure of σW. Hence, the sequence $(L_n\colon n \geq 1)$ is a uniformly integrable martingale. It therefore converges in the L^1 sense with its limit L_∞ being given by (15). Moreover, by general theory, L_∞ is equal to the Radon–Nikodym derivative of the measure of Y with respect to that of σW [29, Proposition 1.4, p. 212 and Proposition 7.6, p. 33]. Finally, since L_∞ is strictly positive with probability one (in fact, it is bounded below), it follows that the two measures are equivalent. See [12, Ch. VII, Sec. 4] for more references and information related to the representation (15).

APPENDIX D
CLARKE SPECTRUM

In this appendix, we concentrate on the behavior of the capacity per unit fourthegy bound for the Clarke spectrum. The Clarke spectrum is commonly used for the design and analysis of systems. From the discussion in Section V-A, we expect the bound to be infinite for the Clarke spectrum, and our objective here is to show this. As a specific example, consider a separable channel with a uniform distribution of power among the multipath elements. The channel considered has F_d as the maximum doppler spread and $T_{\max}$ is the multipath delay spread. The power spectral density indexed by path delay $S_H(f, \tau)$ can be written as

$$S_H(f, \tau) = \frac{1}{F_d \pi \sqrt{1 - \left(\frac{f}{F_d}\right)^2}} \frac{1}{T_{\max}} \mathbf{1}_{[0, T_{\max}]}(\tau).$$

Consulting Fig. 1, it is clear that the *stump* region contributes the most to the fourth moment cost. Moreover, $\phi_H(\tau, \nu)$ also grows without bound in that region. Therefore, it suffices to consider the following integral:

$$\int_{|\tau| \leq T_c/2} \int_{|\nu| \leq \frac{1}{2T}} E[|\chi(\nu, \tau)|^2] \psi_H(\nu, \tau)\, d\nu\, d\tau.$$

In the region considered we can approximate $\psi_H(\nu, \tau)$ as

$$\psi_H(\nu, \tau) = \Theta\left(\frac{1}{T_{\max}}(-\log|\nu|)\right), \qquad \text{as } |\nu| \to 0$$

where it is implicitly assumed that $T \gg 1$ and $T_c \ll T_{\max}$. By (31)

$$\begin{aligned} E[|\chi(\nu, \tau)|^2] &= E[|a_n|^2]^2 \left(\frac{\sin(\pi\nu T_c N)}{\sin(\pi \nu T_c)}\right)^2 |\chi_s(\nu, \tau)|^2 \\ &\quad + N(E[|a_n|^4] - E[|a_n|^2]^2)|\chi_s(\nu, \tau)|^2 \\ &\quad + E[|a_n|^2]^2 \sum_{m\colon 1 \leq |m| \leq N-1} (N - |m|)|\chi_s(\nu, \tau + mT_c)|^2. \end{aligned}$$

Thus,

$$\begin{aligned} E[J_C(U)] &= \Theta\left(E[|a_n|^2]^2 \left(\frac{T_c}{T_{\max}}\right) T^2 \int_{-\frac{1}{2T}}^{\frac{1}{2T}} (-\log|\nu|)\mathrm{sinc}^2(\nu T)\, d\nu\right) \\ &= \Theta\left(E[|a_n|^2]^2 \left(\frac{T_c}{T_{\max}}\right) T \log(T)\right), \qquad \text{as } T \to \infty. \end{aligned}$$

Therefore,

$$\frac{E[J_C(U)]}{T} = \Theta\left(E[|a_n|^2]^2 \left(\frac{T_c}{T_{\max}}\right) \log(T)\right) \tag{48}$$

as $T \to \infty$. Thus, $\frac{E[J_C(U)]}{T}$ tends to infinity as $T \to \infty$ but only as $\log(T)$. Therefore, the bound on the information rate given by the capacity per fourthegy result is infinite.

ACKNOWLEDGMENT

The authors are grateful to pointers from E. Telatar early in this investigation and to C. Frank for providing the exact numbers for the IS-95 standards.

REFERENCES

[1] P. A. Bello, "Characterization of randomly time-varying linear channels," *IEEE Trans. Commun. Syst.*, vol. COM-11, pp. 360–393, Dec. 1963.
[2] E. Biglieri, G. Caire, and G. Taricco, "Coding and modulation for the fading channel," in *Proc. IEEE Vehicular Technology Conf.*, May 1997.
[3] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: Information—Theoretic and communication aspects," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2619–2692, Oct. 1998.
[4] R. E. Blahut, *Theory of Remote Surveillance*. Book manuscript.
[5] G. C. Clarke, "A statistical theory of mobile radio reception," *Bell Syst. Tech. J.*, vol. 47, pp. 957–1000, 1968.
[6] J. L. Doob, *Stochastic Processes*. New York: Wiley, 1953.
[7] B. H. Fleury and P. E. Leuthold, "Radiowave propagation in mobile communications: An overview of European research," *IEEE Commun. Mag.*, vol. 34, pp. 70–81, Feb. 1996.
[8] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.
[9] ——, "Energy limited channels: Coding, multiaccess, and spread spectrum," MIT, Cambridge, MA, Tech. Rep. LIDS-P-1714, LIDS, Nov. 1987.
[10] R. Gallager and M. Médard, "Bandwidth scaling for fading channels," in *Proc. Int. Symp. Information Theory (ISIT'97)*, Ulm, Germany, June/July 1997, p. 471.
[11] A. Ganti, A. Lapidath, and İ. E. Telatar, "Mismatched decoding revisited: General alphabets, channels with memory, and the wide-band limit," *IEEE Trans. Inform. Theory*, vol. 46, pp. 2315–2328, Nov. 2000.
[12] I. I. Gihman and A. V. Skorohod, *The Theory of Stochastic Processes I*. New York: Springer-Verlag, 1980.
[13] H. Holma and A. Toskala, *WCDMA for UMTS: Radio Access for Third Generation Mobile Communications*. New York: Wiley, 2000. Revised edition 2001.

[14] I. Jacobs, "The asymptotic behavior of incoherent M-ary communication systems," *Proc. IEEE*, vol. 51, pp. 251–252, Jan. 1963.
[15] R. S. Kennedy, *Fading Dispersive Communication Channels*. New York: Wiley-Interscience, 1969.
[16] A. Lapidoth and P. Narayan, "Reliable communication under channel uncertainty," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2148–2177, Dec. 1997.
[17] R. S. Lipster and A. N. Shiryayev, *Statistics of Random Processes I, General Theory*. New York: Springer-Verlag, 1977.
[18] T. Marzetta and B. Hochwald, "Capacity of a mobile multiple-antenna communication link in a Raleigh flat-fading environment," *IEEE Trans. Inform. Theory*, vol. 45, pp. 128–138, Jan. 1999.
[19] M. Médard and R. G. Gallager, "Bandwidth scaling for fading multipath channels," *IEEE Trans. Inform. Theory*, vol. 48, pp. 840–852, Apr. 2002.
[20] D. L. Noneaker and M. B. Pursley, "On the chip rate of cdma systems with doubly selective fading and rake reception," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 853–861, June 1994.
[21] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1988.
[22] J. G. Proakis, *Digital Communications*, 3rd ed. New York: McGraw-Hill, 1995.
[23] M. Reed and B. Simon, *Functional Analysis*. New York: Academic, 1972.
[24] B. Hajek and V. G. Subramanian, "Capacity and reliability function for small peak signal constraints," *IEEE Trans. Inform. Theory*, vol. 48, pp. 828–839, Apr. 2002.
[25] E. Telatar and D. Tse, "Capacity and mutual information of wide-band multipath fading channels," *IEEE Trans. Inform. Theory*, vol. 46, pp. 2315–2328, Nov. 2000.
[26] İ. E. Telatar, "Coding and multiaccess for the energy limited Raleigh fading channel," Master's thesis, Dept. Elec. Eng. Comput. Sci., MIT, Cambridge, MA, 1986.
[27] S. Verdú, "On channel capacity per unit cost," *IEEE Trans. Inform. Theory*, vol. 36, pp. 1019–1030, Sept. 1990.
[28] A. J. Viterbi, "Performance of an M-ary orthogonal communication systems using stationary stochastic signals," *IEEE Trans. Inform. Theory*, vol. IT-13, pp. 414–421, July 1967.
[29] E. Wong and B. Hajek, *Stochastic Processes in Engineering Systems*. New York: Springer-Verlag, 1985.