Capacity of Underspread WSSUS Fading Channels in the Wideband Regime

Giuseppe Durisi and Helmut Bölcskei
ETH Zurich
8092 Zurich, Switzerland
E-mail: {gdurisi,boelcskei}@nari.ee.ethz.ch

Shlomo Shamai (Shitz)
Technion, Israel Institute of Technology
32000 Haifa, Israel
E-mail: sshlomo@ee.technion.ac.il

***Abstract*—We characterize the infinite bandwidth capacity behavior of the general class of underspread wide-sense stationary uncorrelated scattering (WSSUS) time-frequency selective Rayleigh fading channels. In particular, we propose a signaling scheme, termed time-frequency pulse position modulation (TF-PPM), which is shown to achieve AWGN channel capacity in the infinite bandwidth limit. As a trivial consequence of this result, the infinite bandwidth capacity of WSSUS underspread fading channels, irrespectively of the scattering function, equals the AWGN channel's infinite bandwidth capacity. The wideband slope achieved by TF-PPM is found to be zero, irrespectively of the channel's scattering function, even in the presence of perfect receive channel state information. Our proof techniques use the fact that underspread fading channels have a highly structured set of eigenfunctions and a property of orthogonal signaling schemes first presented in Butman and Klass, *Jet Propulsion Lab., Tech. Rep.*, 1973.**

I. Introduction

The wideband regime is defined as the region where the energy-per-bit and the spectral efficiency (i.e., the rate in bits/second divided by bandwidth in Hertz) are low [1]. Wireless communication in the wideband regime is attractive due to ease of coexistence with legacy systems, ease of multiple access and improved link reliability due to frequency diversity.

In an additive white Gaussian noise (AWGN) channel, we have

$$\lim_{W\to\infty} W \log_2\left(1+\frac{P}{WN_0}\right) = \frac{P}{N_0}\log_2 e$$

where P is the received power, N_0 stands for the one-sided noise spectral level and W denotes bandwidth. Although in an AWGN channel, capacity, irrespectively of the bandwidth, is achieved using Gaussian inputs, in the infinite bandwidth limit other signaling schemes such as binary antipodal modulation and on-off keying with very low duty cycle are as good as Gaussian signals [1] (in the sense of the ratio between the corresponding mutual information to capacity approaching unity). The picture changes, however, significantly in the case of fading channels[1] [1]–[4]. Jacobs [5] and Gallager [6, Sec. 8.6] showed that the infinite bandwidth capacity of a time-selective Gaussian fading channel equals the infinite bandwidth capacity of an AWGN channel (with the same received power). The capacity-achieving (in the wideband limit) codebook used in [6] employs frequency-shift-keying (FSK) with spacing between successive frequencies larger than the Doppler spread of the channel and the duty cycle going to zero as the bandwidth goes to infinity. The proof technique in [6] relies on an analysis of the error exponent. More recently, an extension of Gallager's result to the case of block-fading multipath channels with independent fading across blocks and arbitrary fading statistics was provided in [3]. Converse results, under various fading channel models, showing that spread-spectrum (i.e., white-like) signals result in the achievable rate going to zero as bandwidth approaches infinity, have been provided in [2]–[4].

A necessary and sufficient condition for a signaling scheme to achieve AWGN channel capacity over a *memoryless* fading channel (using a finite-dimensional signal model) in the infinite bandwidth limit was provided in [1]. This condition is essentially a "peakiness" condition (a.k.a. "flash-signaling").

Contributions: The results summarized above (except for the converse result in [4]) were obtained under various restrictive assumptions on the fading channel model (i.e., time-selective only, block-fading, memoryless). In this paper, we consider the general class of underspread [7] (continuous-time) wide-sense stationary uncorrelated scattering (WSSUS) [8] time-frequency selective Rayleigh fading channels without imposing the block-fading assumption. We show that a signaling scheme, which we call time-frequency PPM (TF-PPM), achieves the AWGN channel capacity in the infinite bandwidth limit, irrespectively of the channel's scattering function [8]. Even though it is not straightforward to extend the notion of flash-signaling [1] to our signal and channel model (due to the infinite dimensionality of the signal model) and draw the following conclusion directly from [1], we can still show that TF-PPM achieves zero wideband slope, even in the presence of perfect receive CSI.

The proof techniques used in this paper are built on the fact that underspread channels have a well-structured set of TF-localized (approximate) eigenfunctions [7], and on a property of the information divergence of orthogonal signaling schemes first presented in a different context in [9].

This research was supported in part by NOKIA Research Center Helsinki, Finland and by the EC FP6 Network of Excellence NEWCOM.

[1] Throughout the paper, whenever we speak of fading channels, we shall refer to (without explicitly stating it) the "noncoherent" case where neither transmitter nor receiver have access to channel state information (CSI) but both are aware of the channel law.

Notation: Uppercase boldface letters are used to denote matrices (both random and deterministic) and lowercase boldface letters designate vectors. The superscripts ^{T}, ^{H} and * stand for transpose, conjugate transpose and elementwise conjugation, respectively. $\det(\mathbf{A})$ and $\mathrm{Tr}(\mathbf{A})$ denote the determinant and the trace, respectively, of the matrix $\mathbf{A}$. $\mathbf{I}$ stands for the identity matrix of appropriate size. $[\mathbf{a}]_i$ is the ith element of the vector $\mathbf{a}$ and $[\mathbf{A}]_{i,j}$ is the element in row i and column j of $\mathbf{A}$. For an $M \times N$ matrix $\mathbf{A} = [\mathbf{a}_1\, \mathbf{a}_2\, \ldots\, \mathbf{a}_N]$, we define $\mathrm{vec}(\mathbf{A}) = [\mathbf{a}_1^T\ \mathbf{a}_2^T\ \ldots\ \mathbf{a}_N^T]^T$. All logarithms are to the base e unless stated otherwise. $\mathbb{E}_X\{\cdot\}$ stands for the expectation operation with respect to the random variable (RV) X. $\delta(t)$ is the Dirac delta function and $\delta_{i,j} = 1$ for $i = j$ and 0 otherwise.

II. Channel and Signal Model

Time-frequency selective underspread fading channels: We model the TF-selective fading channel as a linear random system with input-output relation

$$y(t) = (\mathbb{H}x)(t) = \int_\tau h(t, t-\tau)x(t-\tau)d\tau \qquad (1)$$

where $x(t)$ is the input signal, $y(t)$ is the output signal, $\mathbb{H}$ is the linear operator describing the effect of the ergodic channel and $h(t,t')$ is the random kernel of $\mathbb{H}$. Throughout this paper, we assume that $h(t,t')$ is a complex *Gaussian process* in t and t'. The time-varying transfer function of the channel is defined as [10]

$$L_\mathbb{H}(t,f) = \int_\tau h(t,t-\tau)e^{-j2\pi f\tau}d\tau. \qquad (2)$$

An alternative representation of (1) is

$$y(t) = \int_\tau \int_\nu S_\mathbb{H}(\tau,\nu)x(t-\tau)e^{j2\pi\nu t}d\tau d\nu$$

where $S_\mathbb{H}(\tau,\nu)$ is the channel's (delay-Doppler) spreading function, which is related to the impulse response $h(t, t-\tau)$ through a Fourier transform as

$$S_\mathbb{H}(\tau,\nu) = \int_t h(t,t-\tau)e^{-j2\pi\nu t}dt.$$

We invoke the WSSUS assumption, which is

$$\mathbb{E}_\mathbb{H}\{S_\mathbb{H}(\tau,\nu)\} = 0$$
$$\mathbb{E}_\mathbb{H}\{S_\mathbb{H}(\tau,\nu)S^*_\mathbb{H}(\tau',\nu')\} = C_\mathbb{H}(\tau,\nu)\delta(\tau-\tau')\delta(\nu-\nu')$$

where $C_\mathbb{H}(\tau,\nu) \geq 0$ denotes the scattering function [8]. Finally, we shall need the channel's correlation function defined as

$$\mathbb{E}_\mathbb{H}\{L_\mathbb{H}(t,f)L^*_\mathbb{H}(t',f')\} = R_\mathbb{H}(t-t', f-f')$$

with the Fourier correspondence

$$R_\mathbb{H}(\Delta t, \Delta f) = \int_\tau \int_\nu C_\mathbb{H}(\tau,\nu)e^{j2\pi(\nu\Delta t - \tau\Delta f)}d\tau d\nu. \qquad (3)$$

The underspread assumption and its consequences: WSSUS channels can be classified into underspread and overspread channels [7], [11]. A channel is said to be underspread if its scattering function is highly concentrated in the τ-ν plane. In order to make this statement more precise, we invoke the common assumption of a scattering function that is compactly supported within the rectangle $[-\tau_0, \tau_0] \times [-\nu_0, \nu_0]$, i.e., $C_\mathbb{H}(\tau,\nu) = 0$ for $(\tau,\nu) \notin [-\tau_0,\tau_0] \times [-\nu_0,\nu_0]$. Note that this implies that the spreading function $S_\mathbb{H}(\tau,\nu)$ is supported within this rectangle as well. Defining the spread of the channel as the area of this rectangle, $\Delta_\mathbb{H} = 4\tau_0\nu_0$, the channel is said to be underspread if $\Delta_\mathbb{H} \leq 1$ and overspread if $\Delta_\mathbb{H} > 1$ [7], [11]. The underspread assumption is relevant as most mobile radio channels are (highly) underspread.

Approximate diagonalization of underspread channels: An important result we are going to build our developments on is the fact that underspread channels are approximately diagonalized by orthogonal Weyl-Heisenberg bases [7], which are obtained by TF-shifting of a normalized function $g(t)$ according to $g_{k,l}(t) = g(t-kT)e^{j2\pi lFt}$, with the grid parameters T and F satisfying $TF \geq 1$, and the orthonormality condition

$$\langle g_{k,l}, g_{k',l'}\rangle = \int_t g_{k,l}(t)g^*_{k',l'}(t)dt = \delta_{k,k'}\delta_{l,l'}. \qquad (4)$$

Choosing $T \leq 1/(2\nu_0)$ and $F \leq 1/(2\tau_0)$, and hence $TF \leq 1/\Delta_\mathbb{H}$, it has been shown in [7] that the impulse response $h(t,t')$ of an underspread fading channel can be well approximated by setting

$$h(t,t') = \sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} L_\mathbb{H}(kT, lF)g_{k,l}(t)g^*_{k,l}(t')$$

with suitably chosen window function $g(t)$ depending on the shape of the scattering function $C_\mathbb{H}(\tau,\nu)$. In practice $g(t)$ will be well-localized in time and frequency. Further details on the choice of $g(t)$ can be found in [7], [11]. The scalar channel coefficients $L_\mathbb{H}(kT, lF)$ are circularly symmetric complex Gaussian RVs with zero mean, variance $\sigma^2_\mathbb{H} = \int_\tau\int_\nu C_\mathbb{H}(\tau,\nu)d\tau d\nu$, and correlation function

$$\mathbb{E}_\mathbb{H}\{L_\mathbb{H}(kT,lF)L^*_\mathbb{H}(k'T,l'F)\} = R_\mathbb{H}((k-k')T, (l-l')F).$$

Canonical characterization of signaling schemes: Based on the developments in the previous paragraph, we write our transmit signal $x(t)$ as

$$x(t) = \sum_{k=-\infty}^{\infty}\sum_{l=0}^{N-1} c_{k,l}g_{k,l}(t) \qquad (5)$$

where the $c_{k,l}$ are the information bearing data symbols. This modulation scheme corresponds to pulse-shaped orthogonal frequency-division multiplexing (OFDM) with OFDM symbol duration T and subcarrier spacing F. The transmit signal bandwidth is given by $W = NF$. Furthermore, the parameters T and F have to be chosen such that $TF \geq 1$ [12]. With the received signal $r(t) = y(t) + n(t)$ and $n(t)$ circularly symmetric additive white Gaussian noise so that $\mathbb{E}\{n(t)n^*(t')\} = \delta(t-t')$, the receiver computes the inner products $\hat{c}_{k,l} = \langle r, g_{k,l}\rangle$. Exploiting the orthonormality (4) of the basis functions $g_{k,l}(t)$, we obtain the overall input-output relation

$$\hat{c}_{k,l} = L_\mathbb{H}(kT, lF)c_{k,l} + n_{k,l} \qquad (6)$$

where $\mathbb{E}\left\{n_{k,l}n^*_{k',l'}\right\} = \delta_{k,k'}\delta_{l,l'}$. To summarize, in essence we are transmitting and receiving on the channel's eigenfunctions and hence diagonalizing the channel.

We assume that each channel use takes place over N subcarriers and K OFDM symbols so that the signal transmitted in a channel use (signals are chosen independently across channel uses) is represented by the $N \times K$ matrix $[\mathbf{C}]_{l,k} = c_{k,l}$ ($l = 0, 1, \ldots, N-1$, $k = 0, 1, \ldots, K-1$) with the power constraint $\mathbb{E}\{\mathrm{Tr}(\mathbf{C}\mathbf{C}^H)\} = KP$, where P is the average power of the transmit signal. The corresponding input-output relation can be written as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{7}$$

where $\mathbf{x} = \mathrm{vec}(\mathbf{C})$ and $\mathbf{H}$ is an $NK \times NK$ diagonal matrix containing the scalar (complex-valued) channel gains $L_{\mathbb{H}}(kT, lF)$ on its main diagonal arranged to be consistent with the definition of $\mathbf{x}$. The average power constraint in terms of $\mathbf{x}$ is $\mathbb{E}\{\|\mathbf{x}\|^2\} = KP$. In the remainder of the paper, we set, for simplicity, $\sigma_{\mathbb{H}}^2 = 1$. Later, we will let $N \to \infty$ so that the dimensionality of the signal model in (7) grows linearly in N. We conclude this section by noting that since a general TF-selective fading channel exhibits correlation between the channel gains $L_{\mathbb{H}}(kT, lF)$ both in time and frequency, the channel in (6) will in general not be memoryless.

III. TF-PPM Achieves Capacity

A lower bound on the capacity of the channel in (6) is obtained by assuming that different channel uses yield independent channel realizations. For a given bandwidth $W = NF$, the corresponding capacity lower bound in nats per time interval T, is given by

$$C = \frac{1}{K} \sup_{f_{\mathbf{x}}(\cdot)} I(\mathbf{x}; \mathbf{y}) \tag{8}$$

where the supremum is taken over the joint distribution $f_{\mathbf{x}}(\cdot)$ of the NK-dimensional transmit signal vector $\mathbf{x}$ satisfying $\mathbb{E}\{\|\mathbf{x}\|^2\} = KP$. The aim of this section is to show that TF-PPM (defined below) results in

$$\lim_{N\to\infty} \frac{1}{K} I(\mathbf{x}; \mathbf{y}) = P \tag{9}$$

and hence achieves AWGN channel capacity over the general WSSUS underspread channel (6) in the infinite bandwidth limit, irrespectively of $C_{\mathbb{H}}(\tau, \nu)$. In the following, since F is fixed, we use N (instead of $W = NF$) to denote bandwidth.

Before presenting the main result, a few comments putting our findings into perspective with respect to existing results are in order. To the best of our knowledge the infinite bandwidth capacity of the general channel in (6) as well as corresponding capacity-achieving (in the infinite bandwidth limit) signaling schemes have not been reported in the literature. In fact, previous results are often based on further simplifications of the channel model underlying (6), such as assuming that the fading is time-selective only [5] or frequency-selective only [3] where both assumptions typically come in combination with the block-fading assumption. If the correlation function $R_{\mathbb{H}}(\Delta t, \Delta f)$ of the general underspread WSSUS channel $\mathbb{H}$ has compact support in time or frequency or both, independent channel uses, and hence an effective block-fading channel, can indeed be obtained through the insertion of guard intervals. A compactly supported correlation function is, however, not compatible with the underspread assumption [see the Fourier correspondence in (3)], and also unlikely to occur in practice. The results in [1] are not directly applicable to our setup since [1] considers a memoryless channel and employs a finite-dimensional signal model. The main difficulty in extending the results in [1] to the general channel model considered in this paper lies in the fact that the signal model in the present paper is an infinite-dimensional one (recall that we allow $N \to \infty$) and that it is not clear what an appropriate definition of flash-signaling for infinite-dimensional signals is. Time-frequency selective WSSUS channels are considered in [4]; the corresponding result is, however, a converse one, showing that spread-spectrum signals lead to zero capacity in the infinite bandwidth limit. In summary, it is therefore interesting to investigate the wideband behavior of the general WSSUS underspread fading channel in (6).

TF-PPM: We consider signals with bandwidth N and (effective) time duration K. The corresponding $N \times K$ rectangle in the time-frequency plane is divided into contiguous rectangular blocks of size $N' \times K'$, where $K = K'r$ and $N = N'q$ with $r, q \in \mathbb{N}$ and K' and N' fixed. The collection of all qr blocks hence covers the entire $N \times K$ block (i.e., all signal space dimensions are occupied) and q grows with increasing N. We denote the $N'K'$-dimensional random vector transmitted in a given $N' \times K'$ block by $\widetilde{\mathbf{x}}$.

Definition 1: The TF-PPM signaling scheme consists of selecting one of the qr $N' \times K'$ blocks with uniform probability and transmitting $\widetilde{\mathbf{x}}$ satisfying $\mathbb{E}\{\|\widetilde{\mathbf{x}}\|^2\} = KP$ within that block.

It is easy to verify that TF-PPM according to Definition 1 satisfies the power constraint $\mathbb{E}\{\|\mathbf{x}\|^2\} = KP$. We are now ready to state the main result of this section.

Theorem 2: Let $\mathbf{x}$ be chosen according to Definition 1 (TF-PPM). Then, for a WSSUS underspread channel $\mathbb{H}$ with scattering function $C_{\mathbb{H}}(\tau, \nu)$, we have

$$\lim_{N\to\infty} \frac{1}{K} I(\mathbf{x}; \mathbf{y}) = P, \tag{10}$$

irrespectively of $C_{\mathbb{H}}(\tau, \nu)$.

Proof: Using the decomposition [13, Eq. (10)] of mutual information, we get

$$\frac{1}{K} I(\mathbf{x}; \mathbf{y}) = P - \frac{1}{K}\,\mathbb{E}_{\widetilde{\mathbf{x}}}\Big\{\log\det\Big(\mathbf{I} + \widetilde{\mathbf{X}}\mathbf{R}_{\widetilde{\mathbf{h}}}\widetilde{\mathbf{X}}^H\Big)\Big\} - \frac{1}{K} D\big(\mathcal{P}_{\mathbf{y}} \| \mathcal{P}_{\mathbf{y}|\mathbf{x}=\mathbf{0}}\big) \tag{11}$$

where $\mathbf{0}$ denotes the all zeros symbol, $[\widetilde{\mathbf{H}}]_{i,j} = L_{\mathbb{H}}(iT, jF)$ ($i = 0, 1, \ldots, K'-1$, $j = 0, 1, \ldots, N'-1$), $\widetilde{\mathbf{h}} = \mathrm{vec}\Big(\widetilde{\mathbf{H}}\Big)$, $\mathbf{R}_{\widetilde{\mathbf{h}}} = \mathbb{E}\Big\{\widetilde{\mathbf{h}}\widetilde{\mathbf{h}}^H\Big\}$, $\widetilde{\mathbf{X}}$ is an $N'K' \times N'K'$ diagonal matrix with the elements of $\widetilde{\mathbf{x}}$ on its main diagonal arranged to be consistent with the entries in $\widetilde{\mathbf{h}}$ and $D(\mathcal{P}\|\mathcal{Q})$ is the Kullback-Leibler information divergence between the probability density functions $\mathcal{P}$ and $\mathcal{Q}$. Using the fact that the individual $N' \times K'$

blocks in the TF-PPM codebook are simply TF-translates of any given $N' \times K'$ block and that individual blocks are selected according to a uniform distribution, it can be shown that

$$\lim_{N\to\infty} \frac{1}{K} D\big(\mathcal{P}_{\mathbf{y}} \| \mathcal{P}_{\mathbf{y}|\mathbf{x}=0}\big) = 0 \tag{12}$$

for finite K. The proof of (12) is lengthy and can be found in [14, App. I]. The technique used to establish (12) is based on a property of orthogonal signals first presented in [9] (to prove that, in the infinite bandwidth limit, M-ary FSK achieves capacity on the AWGN channel with phase noise). The second term on the right-hand-side (RHS) of (11) can be upper-bounded in a straightforward fashion according to

$$\frac{1}{K}\,\mathbb{E}_{\widetilde{\mathbf{x}}}\Big\{\log\det\Big(\mathbf{I}+\widetilde{\mathbf{X}}\mathbf{R}_{\widetilde{\mathbf{h}}}\widetilde{\mathbf{X}}^H\Big)\Big\} \leq$$

$$\leq \frac{1}{K}\,\mathbb{E}_{\widetilde{\mathbf{x}}}\bigg\{N'K'\log\bigg(1+\frac{\|\widetilde{\mathbf{x}}\|^2}{N'K'}\bigg)\bigg\} \tag{13}$$

$$\leq \frac{N'K'}{K}\log\bigg(1+\frac{PK}{N'K'}\bigg) \tag{14}$$

where we used Hadamard's and Jensen's inequality in (13) and again Jensen's inequality in (14). As (14) can be made as small as desired by choosing the block length K large enough (recall that K' is fixed), the proof is complete. ∎

Theorem 2 shows that in the infinite bandwidth limit TF-PPM according to Definition 1 achieves the AWGN channel capacity over the general WSSUS underspread channel (6), irrespectively of $C_{\mathbb{H}}(\tau,\nu)$. Note that since N' and K' are fixed and we take $N \to \infty$ and K large to make (14) small, TF-PPM can be seen as a signaling scheme that is peaky in the channel's eigenspace. We furthermore note that Theorem 2 can be shown to hold for a wider set of signaling schemes. In fact, our proof technique generalizes to the case where the individual blocks in the signaling scheme are obtained from TF-shifts of a finite number of differently shaped (possibly noncontiguous) blocks. A signaling scheme that neither fits Definition 1 nor falls into this more general class, but still achieves AWGN channel capacity (on a multipath block-fading channel) in the infinite bandwidth limit, is multi-tone FSK as introduced in [15]. Finally, we note that Theorem 2 generalizes Corollary 1 in [16] both in terms of the channel model and the signaling scheme used in [16].

We conclude this section by noting that Theorem 2 trivially implies that the infinite bandwidth capacity of WSSUS underspread channels $\mathbb{H}$ equals the infinite bandwidth capacity of an AWGN channel with the same received power. Even though this result may not come as a surprise, the contribution lies in its generality with the main technical difficulty stemming from the fact that we consider channels with memory in time and frequency and our signal model is infinite dimensional.

IV. Wideband Slope Analysis

Achieving capacity in the infinite bandwidth limit per se does not give an indication of how capacity is approached as a function of bandwidth. In order to gain insight on this question, following the approach in [1], we characterize the first-order behavior of the spectral efficiency $\eta\,(1/N) = I(\mathbf{x};\mathbf{y})/(KN)$ as a function of the normalized energy per bit E_b/N_0 (expressed in dB). The first-order behavior of η is succinctly captured by the wideband slope (measured in bits/s/Hz/(3 dB)) defined as[2] [1]

$$S_0 = \frac{2P^2}{-\ddot{\eta}(0)}\frac{1}{10\log_{10} e}. \tag{15}$$

In Theorem 3 below, we will show that the perfect receive CSI wideband slope of TF-PPM equals 0 (for $K = \infty$). Since the perfect receive CSI mutual information constitutes an upper bound on the mutual information without receive CSI, this implies that the wideband slope of TF-PPM over the general channel in (6) satisfies $S_0 = 0$. We are now ready to state the second result of the paper.

Theorem 3: TF-PPM according to Definition 1 with $K = \infty$ achieves $S_0 = 0$ on a WSSUS underspread channel $\mathbb{H}$, independently of the scattering function $C_{\mathbb{H}}(\tau,\nu)$, even when perfect CSI is available at the receiver.

Proof: For the sake of brevity of exposition, we restrict the proof to the case $N' = K' = 1$. The proof for general N' and K' follows the same approach with minor changes. In the following, $\mathbf{x}_i$ denotes the random vector having the single nonzero entry $\widetilde{x}$, with $\mathbb{E}\{|\widetilde{x}|\}^2 = KP$, in position i. Using the decomposition [13, Eq. (10)] of mutual information for the case when the channel is known perfectly at the receiver, we obtain

$$\frac{1}{K}I(\mathbf{x};\mathbf{y}|\mathbf{H}) = P - \frac{1}{K}\,\mathbb{E}_{\mathbf{H}}\big\{D\big(\mathcal{P}_{\mathbf{y}}\|\mathcal{P}_{\mathbf{y}|\mathbf{x}=\mathbf{0}}\big)\big\}\,. \tag{16}$$

In order to show that $S_0 = 0$, it suffices to show that $\ddot{\eta}(0) = -\infty$, or, equivalently (see [14, Eq. (21)]), that

$$\lim_{K\to\infty}\lim_{N\to\infty}\frac{N}{K}\,\mathbb{E}_{\mathbf{H}}\big\{D\big(\mathcal{P}_{\mathbf{y}}\|\mathcal{P}_{\mathbf{y}|\mathbf{x}=\mathbf{0}}\big)\big\} = \infty. \tag{17}$$

The Kullback-Leibler distance in (17) can be lower bounded as follows

$$\begin{aligned}\mathbb{E}_{\mathbf{H}}\big\{D\big(\mathcal{P}_{\mathbf{y}}\|\mathcal{P}_{\mathbf{y}|\mathbf{x}=\mathbf{0}}\big)\big\} &= \mathbb{E}_{\mathbf{H}}\{\mathbb{E}_{\mathbf{y}|\mathbf{x}=\mathbf{0}}\{q(\mathbf{y})\log q(\mathbf{y})\}\}\\ &\geq \mathbb{E}_{\mathbf{y}|\mathbf{x}=\mathbf{0}}\{\mathbb{E}_{\mathbf{H}}\{q(\mathbf{y})\}\log\mathbb{E}_{\mathbf{H}}\{q(\mathbf{y})\}\}\end{aligned}$$

where

$$q(\mathbf{y}) = \frac{1}{NK}\sum_{i=1}^{NK}\frac{\mathbb{E}_{\widetilde{x}}\{\mathcal{P}_{\mathbf{y}|\mathbf{x}=\mathbf{x}_i}\}}{\mathcal{P}_{\mathbf{y}|\mathbf{x}=\mathbf{0}}}.$$

Note that $\mathbb{E}_{\mathbf{H}}\{q(\mathbf{y})\}$ is the average of NK independent and identically distributed RVs with unit mean. Hence, by the strong law of large numbers, $\lim_{K\to\infty}\lim_{N\to\infty}\mathbb{E}_{\mathbf{H}}\{q(\mathbf{y})\} = 1$. Furthermore, $\mathbb{E}_{\mathbf{H}}\{q(\mathbf{y})\}\log\mathbb{E}_{\mathbf{H}}\{q(\mathbf{y})\}$ forms a submartingale (see [14, App. I]). Therefore, by the backward submartingale convergence theorem [9], the following Taylor expansion holds for large N and K

$$\mathbb{E}_{\mathbf{y}|\mathbf{x}=\mathbf{0}}\{\mathbb{E}_{\mathbf{H}}\{q(\mathbf{y})\}\log\mathbb{E}_{\mathbf{H}}\{q(\mathbf{y})\}\} \simeq$$

$$\simeq \frac{1}{2}\mathbb{E}_{\mathbf{y}|\mathbf{x}=\mathbf{0}}\big\{|\mathbb{E}_{\mathbf{H}}\{q(\mathbf{y})\}|^2 - 1\big\} \tag{18}$$

[2]Strictly speaking the definition in (15) assumes that the signaling scheme is capacity achieving in the infinite bandwidth limit, i.e., $K = \infty$ in Theorem 2.

where we used $x \log x \simeq (x-1) + (x-1)^2/2$ for $x \to 1$ and $\mathbb{E}_{\mathbf{y}|\mathbf{x}=\mathbf{0}}\{\mathbb{E}_{\mathbf{H}}\{q(\mathbf{y})\}\} = 1$. Applying Lemma 4 (see Appendix), we get

$$\frac{N}{K}\left[\mathbb{E}_{\mathbf{y}|\mathbf{x}=\mathbf{0}}\left\{|\mathbb{E}_{\mathbf{H}}\{q(\mathbf{y})\}|^2\right\} - 1\right] = = P^2 \frac{\mathbb{E}_{z_1,\widetilde{z}_1}\{\exp(z_1\widetilde{z}_1^* + z_1^*\widetilde{z}_1)\}}{\left[\mathbb{E}\{|z_1|^2\}\right]^2} - \frac{1}{K^2} \tag{19}$$

where $z_1 = h_1\widetilde{x}$, $h_1 = [\mathbf{H}]_{1,1}$, $\mathbb{E}\left\{|z_1|^2\right\} = KP$, and $\widetilde{z}_1$ is independent of and identically distributed to z_1 (see Appendix). The proof is complete as the first term in (19) diverges for $K \to \infty$ (see the proof of [1, Th. 16]). ■

Theorem 3 essentially extends Theorem 16 in [1] to infinite-dimensional signaling schemes. We note that, since an appropriate definition of flash-signaling for the infinite-dimensional signal model considered in this paper does not seem to be available, the result in [1] cannot be used to infer directly that $S_0 = 0$ for TF-PPM. We conclude by noting that Theorem 3 is not as general as Theorem 16 in [1] as it makes a statement about a specific signaling scheme, namely TF-PPM, only.

V. Concluding Remarks

Our analysis applies to a wider class of channels than those considered in [1] and to a signaling scheme that has infinite dimensionality. Our results are, however, weaker than those in [1], in the sense that we do not provide a necessary and sufficient condition for a signaling scheme to achieve AWGN channel capacity (irrespectively of $C_{\mathbb{H}}(\tau,\nu)$) in the infinite bandwidth limit. Finding such a condition for the setup considered in this paper, in fact, constitutes an interesting open problem. The key to such a general result is probably an appropriate definition of flash-signaling for infinite-dimensional signal models. At this point the only converse results we are aware of state that full-band OFDM in combination with PSK [17] and spread-spectrum signals [4] yield achievable rates that approach zero as bandwidth goes to infinity.

Appendix

Consider the following multivariate L-dimensional probability distributions

$$f(\mathbf{y}) = \frac{1}{\pi^L}\exp(-\|\mathbf{y}\|^2)$$

$$g_l(\mathbf{y}) = \frac{1}{\pi^L}\mathbb{E}_{z_l}\left\{\exp\left(-|y_l - z_l|^2\right)\right\}\prod_{\substack{j=1\\ j\neq l}}^{L}\exp(-|y_j|^2)$$

where the z_l, $l = 1, 2, \ldots, L$, are identically distributed RVs, with marginal distribution $f_z(\cdot)$ and $y_l = [\mathbf{y}]_l$.

Lemma 4: If $\mathbf{y}$ is distributed according to $f(\mathbf{y})$, we have

$$\mathbb{E}_{\mathbf{y}}\left\{\left|\frac{1}{L}\sum_{l=1}^{L}\frac{g_l(\mathbf{y})}{f(\mathbf{y})}\right|^2\right\} = \frac{1}{L}\mathbb{E}\{\exp(z_1\widetilde{z}_1^* + z_1^*\widetilde{z}_1)\} + 1 - \frac{1}{L} \tag{20}$$

where z_1 and $\widetilde{z}_1$ are independent and have identical distribution $f_z(\cdot)$.

Proof: Expanding the LHS of (20), we get

$$\mathbb{E}_{\mathbf{y}}\left\{\left|\frac{1}{L}\sum_{l=1}^{L}\frac{g_l(\mathbf{y})}{f(\mathbf{y})}\right|^2\right\} = \frac{1}{L^2}\sum_{l=1}^{L}\mathbb{E}_{\mathbf{y}}\left\{\left|\frac{g_l(\mathbf{y})}{f(\mathbf{y})}\right|^2\right\} + \frac{1}{L^2}\sum_{l=1}^{L}\sum_{\substack{m=1\\ m\neq l}}^{L}\mathbb{E}_{\mathbf{y}}\left\{\frac{g_l(\mathbf{y})}{f(\mathbf{y})}\frac{g_m(\mathbf{y})}{f(\mathbf{y})}\right\}. \tag{21}$$

Since $g_l(\mathbf{y})/f(\mathbf{y}) = \exp(|y_l|^2)\mathbb{E}_{z_l}\left\{\exp(-|y_l - z_l|^2)\right\}$, the second term on the RHS of (21) is equal to $1 - 1/L$. The first term on the RHS of (21) can be simplified as follows

$$\mathbb{E}_{\mathbf{y}}\left\{\left|\frac{g_l(\mathbf{y})}{f(\mathbf{y})}\right|^2\right\} = \mathbb{E}_{\mathbf{y}}\left\{\exp(2|y_l|^2)\left[\mathbb{E}_{z_l}\left\{\exp(-|y_l - z_l|^2)\right\}\right]^2\right\} = \mathbb{E}_{z_l,\widetilde{z}_l}\{\exp(z_l\widetilde{z}_l^* + z_l^*\widetilde{z}_l)\} = \mathbb{E}_{z_1,\widetilde{z}_1}\{\exp(z_1\widetilde{z}_1^* + z_1^*\widetilde{z}_1)\}.$$

■

References

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

[2] 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.

[3] 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, July 2000.

[4] 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.

[5] I. Jacobs, "The asymptotic behavior of ideal *m*-ary systems," *Proc.* IEEE, vol. 51, pp. 251–252, Jan. 1963.

[6] R. G. Gallager, *Information Theory and Reliable Communication*. New York, NY, USA: John Wiley & Sons, Inc., 1968.

[7] W. Kozek, "Matched Weyl-Heisenberg expansions of nonstationary environments," Ph.D. dissertation, Vienna University of Technology, 1997.

[8] R. S. Kennedy, *Fading Dispersive Communication Channels*. New York, NY, USA: John Wiley & Sons, Inc., 1969.

[9] S. Butman and M. J. Klass, "Capacity of noncoherent channels," *Jet Propulsion Lab., Pasadena, CA, Tech. Rep. 32-1526*, vol. XVIII, pp. 85–93, Sept. 1973.

[10] L. Zadeh, "Time varying networks – I," *Proc. of IRE*, pp. 1488–1503, 1961.

[11] A. M. Sayeed and B. Aazhang, "Joint multipath-Doppler diversity in mobile wireless communications," *IEEE Trans. Commun.*, vol. 47, no. 1, pp. 123–132, Jan. 1999.

[12] W. Kozek and A. F. Molisch, "Nonorthogonal pulseshapes for multicarrier communications in doubly dispersive channels," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1579–1589, Oct. 1998.

[13] S. Verdú, "On channel capacity per unit cost," *IEEE Trans. Inf. Theory*, vol. 36, no. 5, pp. 1019–1029, Sept. 1990.

[14] G. Durisi, H. Bölcskei, and S. Shamai, "Capacity of underspread WSSUS fading channels in the wideband regime," ETH Zürich, Switzerland, Internal Tech. Rep., May 2006, Available upon request.

[15] C. Luo, M. Medard, and L. Zheng, "On approaching wideband capacity using multi-tone FSK," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 9, pp. 1830–1838, Sept. 2005.

[16] M. C. Gursoy, H. V. Poor, and S. Verdú, "Power efficiency of joint frequency-phase modulation in the low-SNR regime over noncoherent Rician channels," in *Proc. IEEE Vehicular Technology conference (VTC 2004-Fall)*, Los Angeles, CA, USA, Sept. 2004, pp. 2068–2072.

[17] D. Schafhuber, H. Bölcskei, and G. Matz, "System capacity of wideband OFDM communications over fading channels without channel knowledge," in *Proc. IEEE Int. Symposium on Information Theory (ISIT 2004)*, Chicago, IL, Jun. / Jul. 2004, p. 389 (with corrections available at http://www.nari.ee.ethz.ch/commth/pubs/p/ofdm04).