Average-Rate Optimal PSAM Transmissions Over Time-Selective Fading Channels

Shuichi Ohno, *Member, IEEE*, and Georgios B. Giannakis, *Fellow, IEEE*

***Abstract*—Enabling linear minimum-mean square error (LMMSE)-based estimation of random time-selective channels, pilot-symbol-assisted modulation (PSAM) has well-documented merits as a fading counter-measure boosting bit-error rate performance. In this paper, we design average-rate optimal PSAM transmissions by maximizing a tight lower bound of the average channel capacity. Relying on a simple closed-form expression of this bound in terms of the LMMSE channel estimator variance, we obtain PSAM transmissions with optimal spacing of pilot symbols and optimal allocation of the transmit-power budget between pilot and information symbols. Equi-powered transmitted symbols, channels with special Doppler spectra, and analytical and simulation based comparisons of possible alternatives shed more light on information-theoretic aspects of PSAM-based transmissions.**

***Index Terms*—Average channel capacity of fading channels, linear minimum mean-square error estimation, pilot symbol assisted modulation, time-selective channels.**

I. Introduction

MOST WIRELESS communication systems rely on coherent detection, which requires channel state information (CSI) to be available at the receiver. For the receiver to acquire CSI, training sequences are often sent before data transmission. For a fixed (or slowly fading) channel, it is sufficient to transmit a training sequence once (or occasionally). However, for rapidly fading wireless channels, frequent retraining is required to track the time-varying channel and this operation may consume considerable transmission bandwidth.

Instead of long training sequences sent at the beginning of each burst, inserting training symbols throughout the transmission constitutes a popular alternative that is known as pilot-symbol-assisted modulation (PSAM) [5]. Acquiring the channel based on the periodically embedded pilots, PSAM is particularly suitable for transmissions over rapidly fading time-selective environments. Being also known at the receiver end, pilot symbols are used not only for channel estimation [5], [10], but also for timing- and frequency-offset synchronization [6], [13], [14]. PSAM has also been suggested for decision-feedback (DF) equalization of block transmissions [12] (see also [1] for pilots placed to optimize DF equalization performance). The bit-error rate (BER) performance of PSAM with linear minimum mean-square error (LMMSE) estimation of time-selective channels was reported in [5]. Based on BER for BPSK, optimal pilots were designed for direct-sequence code-division multiple-access (DS-CDMA) systems in [19] and for systems with diversity in [20]. However, channel capacity issues with PSAM have not been fully addressed. Certainly, the insertion of pilot symbols reduces the information rate and, thus, reduces the utilization of the channel's capacity.

When CSI is perfectly known at the receiver, channel capacity bounds the maximum rate possible both for deterministic as well as for random channel realizations (see, e.g., [4] and references therein). For random channels in particular, averaging the maximum achievable rate over random channel realizations yields the so-termed average channel capacity that should be distinguished from Shannon's capacity defined for deterministic channels. However, when CSI is not available and has to be acquired at the receiver, a closed form expression of the average channel capacity is hard to derive, because it depends on the accuracy of the channel estimators used that are generally non-Gaussian distributed. The reduction of information rate due to channel estimation error was studied in [17], where lower and upper bounds on average channel capacity were given for a specified channel estimation accuracy, but [17] did not consider the information loss due to training.

Based on lower bounds on average channel capacity, optimal training was pursued for multiantenna wireless communications over frequency-flat block-fading channels in [9]; and for point-to-point wireless links over block-fading frequency-selective channels in [2], [18], and [22]. The effect of channel estimation accuracy on the transmission rate of a modified nearest-neighbor decoding rule was also developed in [15]. For time-selective channels, an expression for the achievable maximum transmission rate was reported in [3], which requires Monte Carlo experiments for its evaluation. Although the average channel capacity can be numerically evaluated via Monte Carlo experiments, an analytically tractable expression is preferable when it comes to understanding how the fading affects the average channel capacity.

This paper deals with point-to-point transmissions which include training (also known as pilot) symbols that are optimally designed not only for time-selective random channel estimation, but also for maximizing lower bounds on average capacity. Specifically, we develop upper and lower bounds on the average channel capacity when LMMSE channel estimation is adopted. For bandlimited time-selective channels, the lower bound is expressed as a function of: the noise and Doppler spectra, the number and spacing of pilot symbols and the power allocated to

Manuscript received January 18, 2002; revised March 7, 2002; accepted March 15, 2002. The editor coordinating the review of this paper and approving it for publication is Dr. Annamalai. The work of G. B Giannakis was supported by the National Science Foundation (NSF) Wireless Initiative under Grant 9979442. This paper was presented in part at the Wireless Communications and Networking Conference (WCNC), Orlando, FL, March 2002.

S. Ohno is with the Department of Artificial Complex Systems Engineering, Hiroshima University, Hiroshima, 739-8527, Japan (e-mail: o.shuichi@ieee.org).

G. B. Giannakis is with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: georgios@ece.umn.edu).

Digital Object Identifier 10.1109/TWC.2002.804183

pilot versus information-bearing symbols. Based on this expression, we select the number and allocate the power of pilot symbols to maximize the lower bound on the average channel capacity of time-selective channels. Our optimal training strategy for fading channels having ideal low-pass spectra is expressed in closed form and provides useful insights on the effect time-selectivity has on the average channel capacity. Simulations illustrate robustness of our theoretical findings to channels with nonideal (e.g., Jakes-like) spectra and tightness of our average capacity bounds.

This paper is organized as follows. After shortly reviewing PSAM in Section II, we introduce the maximum achievable transmission rate averaged over random channels, i.e., the average channel capacity in Section III, and develop a lower bound on the average channel capacity in Section IV. Subsequently, we maximize this lower bound with respect to the spacing of pilot symbols and determine the optimal power distribution between pilot and information-bearing symbols in Section V. Simulations are presented in Section VI before concluding the paper.

II. Background and Preliminaries

We consider point-to-point wireless transmissions over frequency-flat time-selective fading channels, where neither the transmitter nor the receiver have knowledge of the CSI. At the transmitter, the information-bearing sequence $\{s(i)\}$ is parsed into blocks $\boldsymbol{s}$ of size $N-K$, where $N>K$. Symbols in $\boldsymbol{s}$ may be linearly block precoded over the complex field (as in [23]) and are, thus, not necessarily adhering to a finite alphabet. To acquire CSI at the receiver, we rely on K training (also known as pilot) symbols $b(k)\neq 0$ for $k\in[0,K-1]$, which are known to the receiver. These pilot symbols are inserted in every block $\boldsymbol{s}$ to obtain the transmitted block $\boldsymbol{u}:=[u(0),u(1),\ldots,u(N-1)]^T$ of size N. For the simplicity of presentation, we select N so that the following condition holds.

Condition 1: The block size N is a multiple of the number K of pilot symbols such that $N=MK$.

We will see later that our design Condition 1 will not be necessary if N is sufficiently large. We denote the position of the information-bearing symbols in $\boldsymbol{u}$ by the ordered index set

$$\mathcal{I}:=\{i_k|u(i_k)=s(k),i_k<i_{k+1},k\in[0,N-K-1]\}. \tag{1}$$

Its complement, $\mathcal{I}^{\perp}$, will contain the K pilot symbol indices i_k for which $u(i_k)=b(k), i_k<i_{k+1}$ for $k\in[0,K-1]$.

Each block is parallel-to-serial (P/S) and digital-to-analog (D/A) converted, pulse-shaped and carrier-modulated for transmission through the continuous channel $h_c(t)$. We consider that timing has been acquired perfectly and sample the output of the front-end filter (that we select to have square-root Nyquist characteristics) at the symbol rate $f_s:=1/T_s$, to obtain the following discrete-time baseband equivalent model:

$$x(n)=h(n)u(n)+w(n) \tag{2}$$

where $h(n):=h_c(nT_s)$ is the sampled multiplicative fading channel and $w(n)$ is additive white Gaussian noise (AWGN) with variance σ_w^2.

Assumption 1: The channel $\{h(n)\}$ is a stationary complex Gaussian process that is independent of $\{s(n)\}$ and $\{w(n)\}$ and has zero mean, variance σ_h^2 and power spectral density (psd), $S_h(e^{j\omega})$, bandlimited to $[-\omega_d,\omega_d]$ with maximum Doppler frequency $\omega_d\geq 0$ [in Hz, $f_d:=\omega_d/(2\pi T_s)$].

Based on (2), we collect the received information-bearing and pilot symbols as $\boldsymbol{x}_s$ and $\boldsymbol{x}_b$ in a matrix-vector form to obtain

$$\tilde{\boldsymbol{x}}:=\begin{bmatrix}\boldsymbol{x}_s\\ \boldsymbol{x}_b\end{bmatrix}=\begin{bmatrix}\boldsymbol{D}_{h,s} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{D}_{h,b}\end{bmatrix}\begin{bmatrix}\boldsymbol{s}\\ \boldsymbol{b}\end{bmatrix}+\begin{bmatrix}\boldsymbol{w}_s\\ \boldsymbol{w}_b\end{bmatrix} \tag{3}$$

where $\boldsymbol{D}_{h,s}(\boldsymbol{D}_{h,b})$ is a diagonal matrix with the kth entry being $h(i_k)$ for $i_k\in\mathcal{I}(\mathcal{I}^{\perp})$ and $k\in[0,N-K-1]([0,K-1])$; block $\boldsymbol{w}_s(\boldsymbol{w}_b)$ denotes the corresponding noise vector; and $\boldsymbol{b}:=[b(0),\ldots,b(K-1)]^T$ stands for the block of pilot symbols.

Intuitively thinking, since for each sample n only one equation becomes available for estimating the two unknowns $\{h(n),u(n)\}$ in (2), the problem of channel and symbol estimation is not well posed (even in the absence of noise), unless we constrain the class of time-selective channels. To solidify this claim, observe that given $x(i_k)$ and knowing $u(i_k)=b(k)$ for $i_k\in\mathcal{I}^{\perp}$, we can in principle find $h(i_k)$ from (2). Notice that although we sample the continuous received signal every T_s seconds, we sample the channel (via the pilots inserted every M information symbols) every MT_s seconds. Except for the trivial channel that has $h(n)=constant$ for all n, we need to collect $M\geq 2$ channel samples to reconstruct $h_c(t)$, which in turn will give us the channel coefficients $h(i_k)$ for $i_k\in\mathcal{I}$, that are needed to enable detection of the information-bearing symbols. According to Nyquist's Theorem, sampling the channel uniformly with period MT_s will entail no aliasing provided that $MT_s<1/(2f_d)$; or equivalently, with $\omega_d=2\pi f_d T_s$, we must have $\omega_d\leq\pi/M$. Since $M\geq 2$, we will henceforth focus on channels for which we have the following.

Assumption 2: The (sampled) time-selective channel's Doppler spread obeys: $\omega_d\leq\pi/2$; or, for the continuous-time channel: $f_d\leq 1/(4T_s):=f_s/4$.

Satisfying Assumption 2 excludes fading channels that vary very rapidly (relative to T_s), but renders our channel-symbol estimation problem well posed.

The bandwidth efficiency of our PSAM transmissions is defined as the ratio of the number of information-bearing symbols over the block size

$$\mathcal{E}(N,K):=\frac{N-K}{N}=\frac{M-1}{M}. \tag{4}$$

In terms of bandwidth efficiency, the maximum M satisfying $\omega_d\leq\pi/M$ is certainly optimal. However, it is not clear whether this choice is optimal under other communication metrics such as channel capacity, BER performance, or, symbol MMSE. Moreover, this does not tell us anything about the optimum power allocation between pilot and information-bearing symbols.

In this paper, we wish to optimize the placement of pilot symbols, as well as the power allocation between pilots and information-bearing symbols. Our performance criterion will be the maximum achievable transmission rate averaged over random

channel realizations—a quantity we earlier defined as the average channel capacity. We remark that the model (3) has the same form as an uncoded OFDM system in the frequency-domain. The design of average-rate optimal training sequences for uncoded OFDM transmissions over *frequency-selective* fading channels was dealt with in [18]. In fact, many results in the present paper for single carrier transmissions over *time-selective* fading channels can be considered as dual and, thus, complement rather nicely those developed in [18] for frequency-selective channels.

III. Average Channel Capacity and CSI

In this section, we will first review the notion of average channel capacity, when CSI acquisition is perfect. Afterwards, we will derive useful lower bounds on the average channel capacity for the practical setting, where the underlying time-selective channel needs to be estimated.

Suppose first that the channel estimation step is perfect. The mutual information between the information-bearing block and the corresponding received block is given by $I(\boldsymbol{x}_s;\boldsymbol{s}|\boldsymbol{D}_{h,s})$. For a fixed power $\mathcal{P}_s := E\{\|\boldsymbol{s}\|^2\}$, the channel capacity (normalized per transmitted symbol) averaged over random channels $\boldsymbol{D}_{h,s}$, i.e., the average channel capacity, will be defined as

$$\bar{C} := \frac{1}{N} E\left\{ \max_{p_s(\cdot),E\{\|\boldsymbol{s}\|^2\}=\mathcal{P}_s} I\left(\boldsymbol{x}_s;\boldsymbol{s}|\boldsymbol{D}_{h,s}\right) \right\} \tag{5}$$

where $E\{\cdot\}$ denotes the expectation operator with respect to the random matrix $\boldsymbol{D}_{h,s}$; and $p_s(\cdot)$ denotes the probability density function (pdf) of $\boldsymbol{s}$. For linear channel models, the average channel capacity (here expressed in nats/Hz) is attained if and only if $\boldsymbol{s}$ is Gaussian [4], [16], [21]. For our model (3), it is found to be [18]

$$\bar{C} = \frac{M-1}{M} E\left\{\log\left(1+\rho_{\text{ideal}}|h|^2\right)\right\} \tag{6}$$

where the expectation is taken with respect to the standardized complex Gaussian variate $h \sim \mathcal{CN}(0,1)$ and ρ_{ideal} is the *effective* output SNR that is defined as

$$\rho_{\text{ideal}} := \frac{\mathcal{P}_s \sigma_h^2}{(N-K)\sigma_w^2}. \tag{7}$$

Although $\boldsymbol{s}$ is generally non-Gaussian, if $N-K$ is sufficiently large and/or $\boldsymbol{s}$ is linearly precoded over the complex field, then $\boldsymbol{s}$ will be approximately Gaussian.

Assumption 3: The information-bearing symbol block $\boldsymbol{s}$ is zero-mean Gaussian with covariance matrix $\sigma_s^2\boldsymbol{I}$.

We remark that for a fixed power of the information-bearing symbols, the $\bar{C}$ given by (6) for *perfectly known* channels, can be considered as an *upper bound* on the average channel capacity with estimated channels, simply because it expresses the ideal average channel capacity without channel estimation error.

Let now $\hat{\boldsymbol{D}}_{h,s}$ be an estimate of $\boldsymbol{D}_{h,s}$. For a given $\hat{\boldsymbol{D}}_{h,s}$, the mutual information (averaged over $\hat{\boldsymbol{D}}_{h,s}$) between the information-bearing block and the corresponding received block is defined as $I_{\text{av}} := (1/N)E\{I(\boldsymbol{x}_s;\boldsymbol{s})|\hat{\boldsymbol{D}}_{h,s})\}$; while the average channel capacity is expressed as $C := \max_{p_s(\cdot),E\{\|\boldsymbol{s}\|^2\}=\mathcal{P}_s} I_{\text{av}}$. It is not easy to evaluate the average channel capacity but is possible to evaluate its lower bound [9], [17], [18]. The lower bound on the average channel capacity for an unbiased channel estimator, is given by [18]

$$\underline{C} = \frac{1}{N}\sum_{i\in\mathcal{I}} E\left\{\log\left(1+\frac{|\hat{h}(i)|^2}{E\{|\Delta h(i)|^2\}+\sigma_w^2/\sigma_s^2}\right)\right\} \tag{8}$$

where $\hat{h}(i)$ and $\Delta h(i)$ are the channel estimator and the corresponding error at sample i.

We will let $b(k)=b$ for all $k\in[0,K-1]$ and, without loss of generality, insert the pilot symbols at time samples Mk for integers k. Formally stated, we consider the following.

Condition 2: Equi-powered pilot symbols $b(k)=b$ are inserted equi-spaced at positions Mk for all $k\in[0,K-1]$.

We will close this section by noting that in LMMSE channel estimation, the receiver has to know (or estimate) the psd of the random channel. If the second-order channel statistics are not available at the receiver, we should adopt a least-squares (LS) channel estimator that is applicable even to deterministic channels. Analytical results similar to those we will develop in the next section for LMMSE estimators are possible for LS channel estimators too, but we will omit them due to lack of space.

IV. Lower Bound on Average Channel Capacity

Using the PSAM-based LMMSE channel estimation in the maximum lower bound $\underline{C}$ on the average channel capacity, in this section we will further simplify the relationship between $\underline{C}$ and the channel estimation error variance $E\{|\Delta h(i)|^2\}$. Although it is possible to derive such a relationship even for finite block sizes, we will let the block size N go to infinity because the resulting closed form expression in this asymptotic case offers useful insights that will guide the optimization of PSAM parameters, which we will pursue in the ensuing section.

Let $S_{h,m}(e^{j\omega})$ be the under-sampled version of the channel's Doppler spectrum, $S_h(e^{j\omega})$, by a factor of M. It then holds that

$$S_{h,m}(e^{j\omega}) = \frac{1}{M}\sum_{\mu=0}^{M-1} S_h\left(e^{j(\omega+2\pi m\mu/M)}\right) e^{-j(\omega+2\pi m\mu/M)}. \tag{9}$$

Based on (9), we show in the Appendix that the channel MMSE at time $Mk+m$ is given by

$$\sigma^2_{\Delta h}(Mk+m) = \sigma_h^2 - \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{|b|^2\left|S_{h,m}\left(e^{j\omega}\right)\right|^2}{|b|^2 S_{h,0}\left(e^{j\omega}\right)+\sigma_w^2} d\omega, \tag{10}$$

where $|b|$ denotes the amplitude of the equi-spaced pilot symbols. But from the orthogonality principle, it follows that:

$$\sigma^2_{\hat{h}}(m) := E\left\{\left|\hat{h}(Mk+m)\right|^2\right\} = \sigma_h^2 - \sigma^2_{\Delta h}(Mk+m) \tag{11}$$

where we used the fact that $\sigma^2_{\Delta h}(Mk+m)$ is periodic and, thus, independent of M.

Normalizing $\hat{h}(i)$ in (8) by its variance in (11), we obtain

$$\underline{C}=\frac{1}{N}\sum_{k=0}^{K-1}\sum_{m=1}^{M}E\left\{\log\left(1+\frac{\sigma^2_{\hat{h}}(Mk+m)|h|^2}{\sigma^2_{\Delta h}(Mk+m)+\sigma^2_w/\sigma^2_s}\right)\right\}. \tag{12}$$

Recalling from Condition 1 that $N=MK$ and using the periodicity of $\sigma^2_{\Delta h}(i)$ and $\sigma^2_{\hat{h}}(i)$, we find

$$\underline{C}=\frac{1}{M}\sum_{m=1}^{M-1}E\left\{\log\left(1+\frac{\sigma^2_h-\sigma^2_{\Delta h}(m)}{\sigma^2_{\Delta h}(m)+\sigma^2_w/\sigma^2_s}|h|^2\right)\right\} \tag{13}$$

where the term $m=0$ has been omitted from the last sum because it corresponds to the pilot symbol that is inserted every M information-bearing symbols as per Condition 1.

Equation (13) is our closed form expression for the lower bound on the average channel capacity. A related expression for the average capacity bound was derived in [3]. But the latter does not impose Assumption 2 and requires Monte Carlo experiments to evaluate the lower bound. In contrast, given M, the noise variance, the channel's spectrum and the powers of pilot and information-bearing symbols, our expression (13) is easy to evaluate by simply generating standardized complex Gaussian variates h.

Let $\bar{f}_d:=\omega_d/(2\pi T_s)$ be the normalized maximum Doppler spread. When below the Nyquist rate, i.e., when $M>1/(2\bar{f}_d)$, the channel MMSE increases due to aliasing [cf. (9) and (10)]. Hence, for a fixed M, the $\underline{C}$ in (13) will be maximized, if $\sigma^2_{\Delta h}(m)$ is minimized. We are, thus, motivated to design our pilot symbols as follows.

Condition 3: The spacing between successive pilot symbols satisfies: $M\leq\lfloor 1/(2\bar{f}_d)\rfloor:=M_{\max}$, where $\lfloor\ \rfloor$ denotes integer part.

We note, however, that since $\underline{C}$ depends also on the channel psd, the condition $M\leq\lfloor 1/(2\bar{f}_d)\rfloor$, is not always necessary for the maximization of the average capacity bound.

Under Condition 3, we obtain from (9) that $|S_{h,m}(e^{j\omega})|=S_{h,0}(e^{j\omega})=S_h(e^{j\omega/M})/M$, for $m\in[1,M-1]$; hence, (10) reduces to

$$\begin{aligned}\sigma^2_{\Delta h}(Mk+m)=&\sigma^2_h-\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{|b|^2S^2_{h,0}\left(e^{j\omega}\right)}{|b|^2S_{h,0}\left(e^{j\omega}\right)+\sigma^2_w}d\omega\\ =&\sigma^2_h-\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{|b|^2S^2_h\left(e^{j\omega}\right)}{|b|^2S_h\left(e^{j\omega}\right)+M\sigma^2_w}d\omega\\ :=&\sigma^2_{\Delta h}.\end{aligned} \tag{14}$$

Equation (14) shows that the LMMSE channel estimation variance is constant, independent of the discrete-time index m. Substituting (14) into (10) we, thus, obtain

$$\underline{C}=\frac{M-1}{M}E\left\{\log\left(1+\rho|h|^2\right)\right\} \tag{15}$$

where

$$\rho:=\frac{\sigma^2_h-\sigma^2_{\Delta h}}{\sigma^2_{\Delta h}+\sigma^2_w/\sigma^2_s}. \tag{16}$$

Thanks to its simplicity, (15) will enable us to optimize PSAM parameters in closed form without Monte Carlo experiments. Specifically, we will subsequently rely on (15) to select the optimal spacing between the equi-spaced and equi-powered pilot symbols and allocate optimally a prescribed transmit-power budget between pilot and information-bearing symbols.

V. Optimizing PSAM Parameters

In this section, we will give answers to a couple of basic questions: how often should the average-rate optimal equi-spaced and equi-powered pilot symbols be inserted? and how much transmit-power should be allocated to channel estimation? Clearly, the spacing M of pilot symbols affects our bandwidth efficiency defined by (4). And among the M's minimizing the channel MMSE, the maximum, $M_{\max}=\lfloor 1/(2\bar{f}_d)\rfloor$, is optimal in terms of bandwidth efficiency. We will show that $M=M_{\max}$ is also optimal in the sense of maximizing $\underline{C}$ in (15). We will also find the optimal power distribution among pilot and information-bearing symbols.

Defining the average transmit-power as

$$\bar{\mathcal{P}}:=\frac{1}{M}\left[|b|^2+(M-1)\sigma^2_s\right] \tag{17}$$

our transmit-power budget for M successive transmitted symbols is $M\bar{\mathcal{P}}$. With $\alpha\in(0,1)$, suppose we allocate power $\alpha M\bar{\mathcal{P}}$ to information-bearing symbols and $(1-\alpha)M\bar{\mathcal{P}}$ to each pilot symbol; i.e.,

$$\sigma^2_s=\frac{\alpha M\bar{\mathcal{P}}}{M-1},\quad |b|^2=(1-\alpha)M\bar{\mathcal{P}},\qquad \text{for } 0<\alpha<1. \tag{18}$$

Our design parameters for the maximization of the average capacity bound are the integer M and the power ratio α. First, we fix the power ratio α and optimize $\underline{C}$ in (15) with respect to M.

A. Optimal Spacing of Pilot Symbols

For a fixed α, we substitute $|b|^2=(1-\alpha)M\bar{\mathcal{P}}$ into (14) to obtain $\sigma^2_{\Delta h}=\sigma^2_h-1/(2\pi)\int(1-\alpha)\bar{\mathcal{P}}S_h(e^{j\omega})/[(1-\alpha)\bar{\mathcal{P}}S_h(e^{j\omega})+\sigma^2_w]d\omega$, which is found to be independent of M. Keeping this in mind and defining $\beta:=(M-1)/M$, we can reexpress (15) as a function of β as follows:

$$\underline{C}=\beta E\left\{\log\left(1+\rho(\beta)|h|^2\right)\right\} \tag{19}$$

where

$$\rho(\beta)=\frac{\sigma^2_{\hat{h}}}{\sigma^2_{\Delta h}+a\beta},\qquad a:=\frac{\alpha\bar{\mathcal{P}}}{\sigma^2_w}. \tag{20}$$

Treating β as a continuous variable, we differentiate (19) with respect to β to obtain

$$\frac{\partial\underline{C}}{\partial\beta}=E\left\{\log\left(1+\rho|h|^2\right)+\beta\frac{\partial\rho}{\partial\beta}\frac{|h|^2}{1+\rho|h|^2}\right\}. \tag{21}$$

It follows from (20) that

$$\frac{\partial \rho}{\partial \beta} = -\frac{a\sigma_{\hat{h}}^2}{(\sigma_{\Delta h}^2 + a\beta)^2} = -\rho \frac{a}{\sigma_{\Delta h}^2 + a\beta}. \tag{22}$$

Substituting (22) into (21), we arrive at

$$\begin{aligned}\frac{\partial \underline{C}}{\partial \beta} =& E\left\{\log\left(1+\rho|h|^2\right) - \left(\frac{\rho|h|^2}{1+\rho|h|^2}\right)\frac{a\beta}{\sigma_{\Delta h}^2 + a\beta}\right\} \\ \geq& E\left\{\frac{\rho|h|^2}{1+\rho|h|^2}\left(1 - \frac{a\beta}{\sigma_{\Delta h}^2 + a\beta}\right)\right\} \\ =& E\left\{\frac{\rho|h|^2}{1+\rho|h|^2}\frac{\sigma_{\Delta h}^2}{\sigma_{\Delta h}^2 + a\beta},\right\} > 0 \end{aligned} \tag{23}$$

where in the second step, we used the inequality $\log(1+\rho|h|^2) \geq \rho|h|^2/(1+\rho|h|^2)$. From (23), we infer that we should take β, or equivalently M, as large as possible. We can summarize our findings so far as follows.

Result 1: Suppose that Assumption 1 through Assumption 3 and Condition 1 through Condition 3 hold true. For a fixed power ratio α, the lower bound on the average channel capacity $\underline{C}$ is maximized at $M_{\max} = \lfloor 1/(2\bar{f}_d) \rfloor$.

B. Optimal Power Allocation

Let us now turn to the optimal power allocation to pilot versus information-bearing symbols. For a fixed M, $\underline{C}$ in (19) becomes a function of α only. We can certainly optimize $\underline{C}$ for α numerically. However, closed form expressions for the optimal α become available in the two special cases that we discuss next.

Ideal Low-Pass Doppler Spectrum: Let us first consider that the channel has an ideal low-pass Doppler spectrum given by

$$S_h(e^{j\omega}) = \begin{cases} \frac{\sigma_h^2}{2\bar{f}_d}, & \text{for } |\omega| < \omega_d \\ 0, & \text{otherwise.} \end{cases} \tag{24}$$

From (14) it follows that:

$$\sigma_{\Delta h}^2 = \frac{1}{2\pi}\int_{-\omega_d}^{\omega_d} \frac{M S_h(e^{j\omega})\sigma_w^2}{|b|^2 S_h(e^{j\omega}) + M\sigma_w^2} d\omega \tag{25}$$

$$= \frac{M\sigma_h^2\sigma_w^2}{\frac{|b|^2\sigma_h^2}{(2\bar{f}_d)} + M\sigma_w^2}. \tag{26}$$

Substituting the latter into (16), we obtain

$$\rho = \left(\frac{\sigma_h^2\sigma_s^2}{\sigma_w^2}\right)\frac{\frac{|b|^2}{(2\bar{f}_d)}}{\frac{|b|^2}{(2\bar{f}_d)} + M\sigma_s^2\left[1 + \frac{\sigma_w^2}{(\sigma_h^2\sigma_s^2)}\right]}. \tag{27}$$

With (18), we can also reexpress ρ as

$$\rho = \rho_{\mathrm{SNR}}\left(\frac{M}{M-1}\right)\zeta\,\frac{\alpha(1-\alpha)}{(\gamma-\alpha)} \tag{28}$$

where ρ_{SNR} is the output SNR (defined as $\rho_{\mathrm{SNR}} := \sigma_h^2\bar{\mathcal{P}}/\sigma_w^2$) and

$$\zeta := \frac{1}{1 - \frac{2M\bar{f}_d}{(M-1)}} \tag{29}$$

$$\gamma := \left(1 + \frac{2\bar{f}_d}{\rho_{\mathrm{SNR}}}\right)\zeta. \tag{30}$$

Notice that Assumption 2 implies $\bar{f}_d \leq 0.25$, from which it follows that $\zeta > 0$ and, hence, $\gamma > 0$.

For a fixed M, we deduce that maximizing $\underline{C}$ with respect to α is equivalent to maximizing ρ, since $\log(\cdot)$ is an increasing function. Differentiating $f(\alpha) := \alpha(1-\alpha)/(\gamma-\alpha)$, we obtain $f'(\alpha) = (\alpha^2 - 2\gamma\alpha + \gamma)/(\gamma-\alpha)^2$. Thus, we find that ρ is maximized at

$$\alpha_{lp} := \frac{1}{1 + \sqrt{1 - \frac{1}{\gamma}}}. \tag{31}$$

Substituting this into ρ leads to

$$\rho_{lp} := \rho_{\mathrm{SNR}}\left(\frac{M}{M-1}\right)\frac{\alpha_{lp}^2}{1 + \frac{2\bar{f}_d}{\rho_{\mathrm{SNR}}}}. \tag{32}$$

We thus arrive at the following result:

Result 2: Suppose that Assumption 1 through Assumption 3 and Condition 1 through Condition 3 hold true and that the channel psd has an ideal low-pass spectrum as in (24). For a fixed M, the lower bound on the average channel capacity $\underline{C}$ is maximized for the power ratio α_{lp} in (31) and it can be expressed in closed form as

$$\underline{C}_{lp} := \frac{M-1}{M} E\left\{\log(1+\rho_{lp}|h|^2)\right\} \tag{33}$$

where ρ_{lp} is given by (32).

Recall now that the upper bound (6) on the average channel capacity is a function of ρ_{ideal}, which assumes that perfect CSI is available. To measure the performance loss incurring due to channel estimation errors, we can utilize the ratio $\rho_{lp}/\rho_{\mathrm{ideal}}$. Since $\mathcal{P}_s = \alpha_{lp}N\bar{\mathcal{P}}$, we obtain from (7) that $\rho_{\mathrm{ideal}} = \alpha_{lp}\rho_{\mathrm{SNR}}M/(M-1)$. Thus, we find that $\rho_{lp}/\rho_{\mathrm{ideal}} = \alpha_{lp}/(1 + 2\bar{f}_d/\rho_{\mathrm{SNR}})$. Since γ is a decreasing function of SNR, α_{lp} is found from (31) to be an increasing function of SNR and is bounded such that $\alpha_{lp} \geq 0.5$. On the other hand, for $2\bar{f}_d/\rho_{\mathrm{SNR}} \leq 1$, i.e., $\rho_{\mathrm{SNR}} \geq 2\bar{f}_d$, we have that $\rho_{lp}/\rho_{\mathrm{ideal}} \geq \alpha_{lp}/2 \geq 0.25$. This implies that for $\rho_{\mathrm{SNR}} \geq 2\bar{f}_d$, the performance loss due to the channel estimation error is upper bounded by about 6.0 dB. Since $\rho_{lp}/\rho_{\mathrm{ideal}} = \alpha_{lp}$ at high SNR, the performance loss is at most 3.0 dB at high SNR. These observations also imply that the lower bound is tight at high SNR.

General Doppler Spectra at High SNR: Let us now focus our attention to time-selective channels with general Doppler spectra $S_h(e^{j\omega})$. At high SNR and for any $S_h(e^{j\omega})$, we have from (25) that $\sigma_{\Delta h}^2 \cong 2M\bar{f}_d\sigma_w^2/|b|^2$. Since $\sigma_{\Delta h}^2$ does not depend on $S_h(e^{j\omega})$, we deduce that the optimal α for any $S_h(e^{j\omega})$ coincides with the optimal α for the ideal low-pass psd at high SNR. As $\rho_{\mathrm{SNR}} \to \infty$, (31) and (32) reduce to

$$\alpha_\infty := \frac{1}{1 + \sqrt{\frac{2M\bar{f}_d}{(M-1)}}} \tag{34}$$

$$\rho_\infty := \rho_{\mathrm{SNR}}\left(\frac{M}{M-1}\right)\alpha_\infty^2 \tag{35}$$

respectively. Thus, we have established the following result:

Result 3: Suppose that Assumption 1 through Assumption 3 and Condition 1 through Condition 3 hold true and that the SNR is sufficiently high. The lower bound on the average channel capacity $\underline{C}$ is maximized at α_∞ and it is given by

$$\underline{C}_\infty := \frac{M-1}{M} E\left\{\log(1+\rho_\infty|h|^2)\right\} \tag{36}$$

where ρ_∞ is defined as in (35).

Suppose now that we select equi-spaced pilots with a spacing that maximizes bandwidth efficiency; i.e., $M = M_{\max} = \lfloor 1/(2\bar{f}_d) \rfloor$. For this choice, we have that $\underline{C}_\infty \cong (1 - 2\bar{f}_d)E\{\log(1 + \rho_\infty |h|^2)\}$ and that $\rho_\infty/\rho_{\text{ideal}} = \alpha_\infty \cong 1/[1 + \sqrt{2\bar{f}_d/(1-2\bar{f}_d)}]$. As $\bar{f}_d$ increases, $1 - 2\bar{f}_d$ and $1/[1 + \sqrt{2\bar{f}_d/(1-2\bar{f}_d)}]$ decrease, causing $\underline{C}_\infty$ to decrease as well. The former degradation results from the fact that more pilot symbols are required to estimate rapidly fading channels. The latter degradation comes from the worst case channel estimation error. Since $1 + \sqrt{2\bar{f}_d/(1-2\bar{f}_d)} \leq 2$, to attain the ideal average channel capacity without estimation error, we have to consume twice the power, at worst.

C. Equi-Powered PSAM

So far, we have dealt with equi-powered pilot symbols whose power, $|b|^2$, is allowed to be different from σ_s^2, the power of information symbols. It is interesting, however, to examine what happens when $|b|^2 = \sigma_s^2$, because such constant modulus PSAM transmissions prevent severe backoffs needed to alleviate nonlinear power amplifier distortions. The latter cause inefficient power amplification that comes with nonconstant modulus transmissions.

To maintain constant modulus of all transmitted symbols, we will impose the constraint $|b|^2 = E\{|s(i)|^2\} = \sigma_s^2$. Setting $|b|^2 = \sigma_s^2$ in (27), we find

$$\begin{aligned}\rho =& \rho_{\text{SNR}} \frac{\frac{1}{(2\bar{f}_d)}}{\frac{1}{(2\bar{f}_d)} + M(1 + \frac{1}{\rho_{\text{SNR}}})} \\ =& \rho_{\text{SNR}} \frac{\tau_d}{\tau_d + M\psi} \end{aligned} \tag{37}$$

where for notational brevity, we have set $\tau_d := 1/(2\bar{f}_d)$ and $\psi := 1 + 1/\rho_{\text{SNR}}$. Let us now treat M as a continuous variable (call it μ) and differentiate the lower bound $\underline{C}$ with respect to μ. Certainly, $\underline{C}$ is discrete in its integer argument M. However, without loss of generality, we can extend its domain from the integers to the reals and consider $\underline{C}$ as a discrete-equivalent version of a continuous function in μ sampled at the integers. We can then express its derivative as

$$\frac{\partial \underline{C}}{\partial \mu} = E\frac{1}{\mu^2}\left\{\log\left(1 + \rho|h|^2\right) + \frac{\mu - 1}{\mu}\frac{\partial \rho}{\partial \mu}\frac{|h|^2}{1 + \rho|h|^2}\right\}. \tag{38}$$

From the definition (37), $\partial\rho/\partial\mu = -\rho\psi/(\tau_d + \mu\psi)$. Substituting this into (38) leads to

$$\begin{aligned}\frac{\partial \underline{C}}{\partial \mu} =& E\left\{\frac{1}{\mu^2}\log\left(1 + \rho|h|^2\right)\right. \\ & \left. - \left(\frac{\mu - 1}{\mu}\right)\left(\frac{\psi}{\tau_d + \mu\psi}\right)\left(\frac{\rho|h|^2}{1 + \rho|h|^2}\right)\right\} \end{aligned} \tag{39}$$

$$\begin{aligned}\geq& E\left\{\frac{\rho|h|^2}{1 + \rho|h|^2}\frac{1}{\mu(\tau_d + \mu\psi)}\right. \\ & \left. \times \left(-\psi(\mu - 1)^2 + \tau_d + \psi\right)\right\} \end{aligned} \tag{40}$$

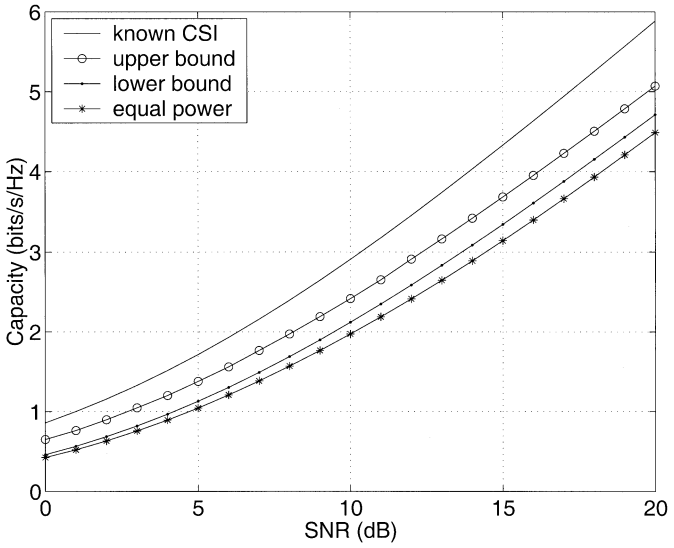


Fig. 1. Channel capacity bounds versus SNR (low-pass model).

where in the second step, we used the inequality $\log(1 + \rho|h|^2) \geq \rho|h|^2/(1 + \rho|h|^2)$. Therefore, as long as $\mu \leq 1 + \sqrt{1 + 1/[2\bar{f}_d(1 + 1/\rho_{\text{SNR}})]}$, we have $\partial\underline{C}/\partial\mu > 0$. This implies that the optimal M lies between $1 + \sqrt{1 + 1/[2\bar{f}_d(1 + 1/\rho_{\text{SNR}})]}$ and $M_{\max}$. In other words, the optimal M may not be $M_{\max}$, unlike the unconstrained case we dealt with in Section V-A. To obtain the optimal M, one has to resort to a numerical line search based on (15) and (37).

VI. Numerical Examples

To validate our analysis and design, we tested the applicability of (6) and (15) in evaluating average capacity bounds for fading channels with: 1) low-pass Doppler spectra [cf. (24)] and 2) Jakes spectra with normalized maximum Doppler frequency 0.05 [11]. In both cases, we deduce from Condition 3 that $M_{\max} = \lfloor 1/(2\bar{f}_d) \rfloor = 10$.

Test Case 1 (Low-Pass Spectrum): For the low-pass channel, Fig. 1 compares the average channel capacity when: 1) the channel is known to the receiver and training is not required, in which case we set $\alpha = 1$ and $K = 0$ in (6); 2) the upper and lower bounds on the average capacity in (33), with the optimal α_{lp} from (31); and 3) the lower bound on the average capacity for the equi-powered PSAM with constant modulus $|b|^2 = \sigma_s^2$.

The difference between the average channel capacity without training and the upper bound on the average channel capacity with training is the price we have to pay to acquire CSI. The difference $\bar{C} - \underline{C}$ between the lower and the upper bounds is the penalty resulting from the worst channel estimation error. The optimally powered pilot symbols have at most $-10\log[\alpha_{lp}/(1 + 0.1/\rho_{\text{SNR}})] \approx -10\log 0.70 \approx 1.55$ dB-performance loss at 0 dB due to the channel estimation error, which converges to $-10\log\alpha_\infty = -10\log 0.75 \approx 1.25$ dB, as the SNR increases. Recall that the actual average channel capacity lies between the lower and the upper bound. Their small difference implies the tightness of the lower bound and hence validates our choice

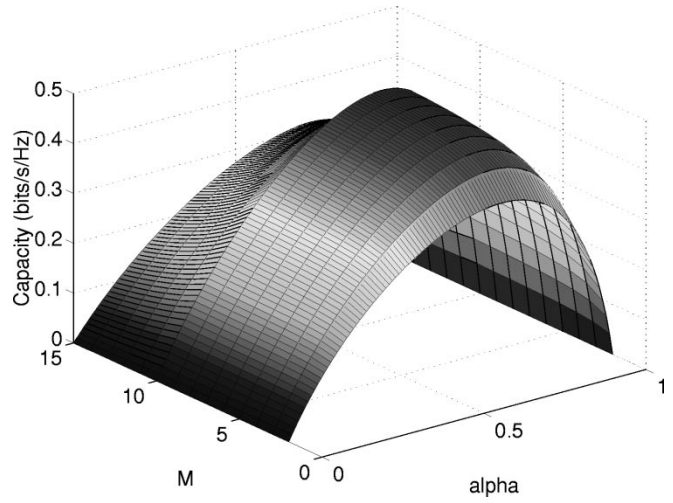


Fig. 2. Lower bounds on the average channel capacity (low-pass model, 0 dB).

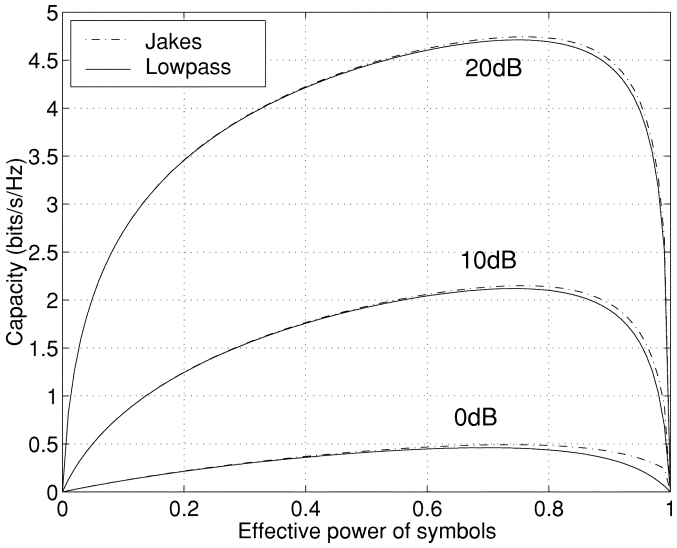


Fig. 3. Lower bounds on the average channel capacity versus α.

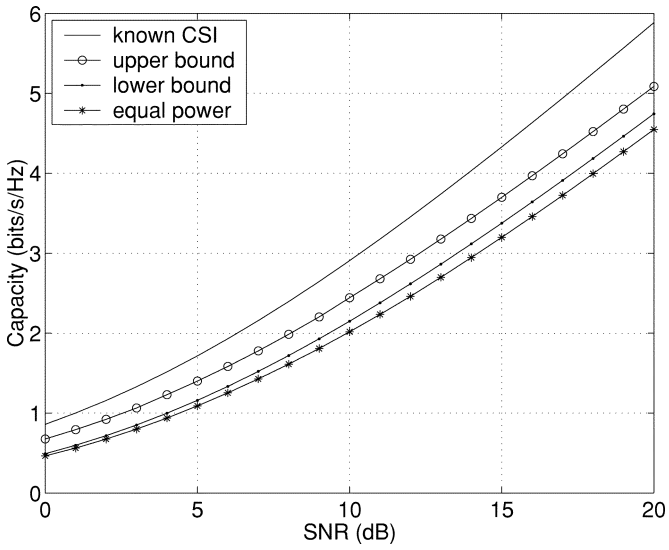


Fig. 4. Channel capacity bounds versus SNR (Jakes model).

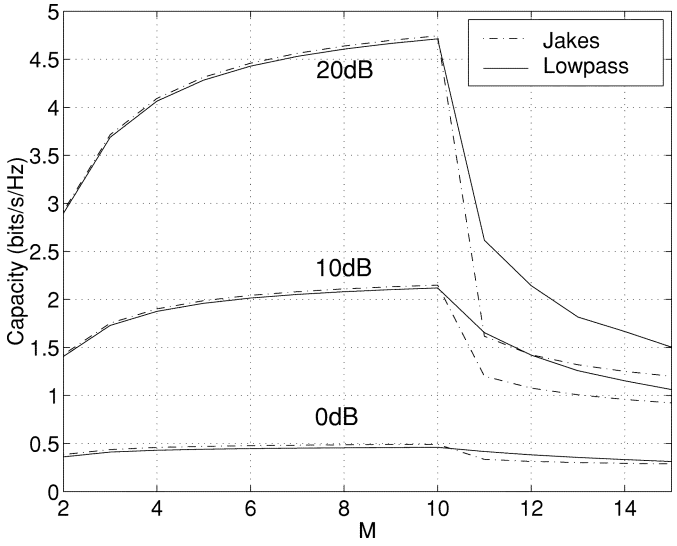


Fig. 5. Lower bounds on the average channel capacity versus M.

of $\underline{C}$ as the criterion for designing average-rate optimal PSAM transmissions.

Optimally powered pilot symbols gain about 1 dB at high SNR relative to the equi-powered case of Section V-C. At low SNR, the difference between them is small. This is because the lower bound is a flat function around its maximum as illustrated in Fig. 2 that depicts the lower bound $\underline{C}$ as a function of M and α at 0 dB; and in Fig. 3 as a function of α for $M = M_{\max}$.

Test Case 2 (Jakes Spectrum): Fig. 4 is the counterpart of Fig. 1 for the Jakes' channel spectrum. Although the Jakes' spectrum does not exactly obey the conditions under which $\underline{C}$ was derived in (33), the resulting average capacity curves of Figs. 1 and 4 are almost the same. This similarity can be also observed in Figs. 3 and 5. This implies that our analytical results on fading channels with low-pass Doppler spectra have broader applicability to at least account for time-selective channels with Jakes' spectra.

Test Case 3 (Optimal Spacing of Pilot Symbols): Fig. 5 depicts the lower bounds on the average channel capacity as a function of the spacing M between consecutive pilot symbols. For both cases, the optimal M is found to be $M = M_{\max}(= 10)$. This suggests that the aliasing due to under-sampling severely decreases the average channel capacity, which justifies our design Condition 3.

Test Case 4 (Equi-Powered PSAM): Fig. 6 depicts the optimal M that maximizes the lower bound for constant modulus PSAM transmissions. It is observed that at low SNR, pilot symbols should be inserted more frequently than what suggested by the Nyquist frequency in order to compensate for the equal power constraint. However, as a function of SNR, the pilot symbol spacing M converges to $\lfloor 1/(2\bar{f}_d) \rfloor$ which is implied by Nyquist's Sampling Theorem and is dictated by the maximum Doppler shift that is normalized by the sampling period.

VII. Conclusion and Discussion

We analyzed point-to-point PSAM systems with channel estimation over time-selective random channels from an

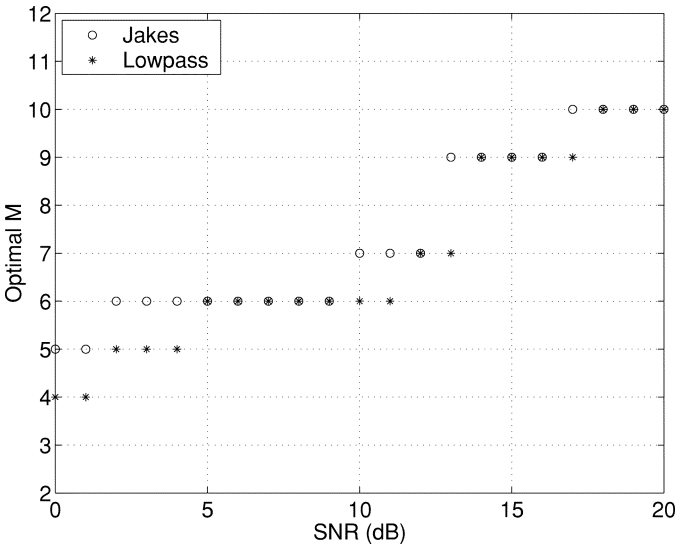


Fig. 6. Optimal pilot spacing versus SNR (equi-powered PSAM).

information-theoretic viewpoint. We presented practical lower bounds on the average channel capacity of PSAM with LMMSE channel estimation. In its simplest form, the lower bound was expressed as a function of the channel's Doppler spectrum, the noise variance, the number and spacing of pilot symbols and the powers of pilot and information-bearing symbols. For a given transmit power budget, we derived the optimal percentages of power to be distributed between information and pilot symbols. We found that with optimally powered pilot symbols, at high SNR, channel estimation incurs at most 3-dB loss and no more than 6 dB for channels with ideal low-pass spectra.

Our results in this paper relied on a fading channel model characterized by its maximum Doppler spread only. It will be interesting to explore similar information-theoretic issues for more general fading models, such as the basis expansion model overviewed in [8]. This model offers a parsimonious parameterization of time- and frequency-selective channels and may come particularly handy for addressing the challenging problem of designing average-rate optimal transmissions through such doubly-selective fading channels.

APPENDIX
DERIVATION OF LMMSE FILTERS

From our design conditions, Condition 1 and Condition 2, it follows that (A.1) at the bottom of the page, which shows that $x(n)$ is not a stationary process but a cyclostationary one with period M. The cyclostationarity necessitates an LMMSE filter that is periodic in time with the same period (see, e.g., [7]). Taking the periodicity into account, the LMMSE channel estimate at time $Mk+m$ for $m \in [0, M-1]$ can be expressed as

$$\hat{h}(Mk+m) = \sum_{l=-\infty}^{\infty} \tilde{g}^*(l;m)x(Mk+m-l) \tag{A.2}$$

where $\tilde{g}^*(l;m)$ denotes the optimal periodically time-varying Wiener filter at time $Mk+m$. It follows from (A.1) and the independence of $h(n)$ and $s(n)$ that $E\{h(Mk+m)x^*(n)\} = 0$, if n is not a multiple of M. From the orthogonality principle, it follows that $E\{[h(Mk+m)-\hat{h}(Mk+m)]x^*(n)\} = 0$. Hence, when n is not a multiple of M, we have

$$E\left\{\left[h(Mk+m) - \hat{h}(Mk+m)\right]x^*(n)\right\} \tag{A.3}$$

$$= -\sum_{l=-\infty}^{\infty} \tilde{g}^*(l;m)E\left\{x(Mk+m-l)x^*(n)\right\} \tag{A.4}$$

$$= -\tilde{g}^*(Mk+m-n;m)E\left\{|x(n)|^2\right\} = 0. \tag{A.5}$$

Since $E\{|x(n)|^2\} \neq 0$, (A.5) implies that $\tilde{g}(Mk+m-n;m) = 0$, if n is not a multiple of M; or equivalently, $\tilde{g}(Ml'+m;m) \neq 0$, only if l' is an integer. Setting $g(l';m) := \tilde{g}(Ml'+m;m)$, (A.2) can be reexpressed as

$$\hat{h}(Mk+m) = \sum_{l'=-\infty}^{\infty} \tilde{g}^*\left(Ml'+m;m\right)x\left(M\left(k-l'\right)\right)$$

$$= \sum_{l'=-\infty}^{\infty} g^*\left(l';m\right)x\left(M\left(k-l'\right)\right). \tag{A.6}$$

Because x is Gaussian, it follows readily from (A.6) that $\hat{h}$ is also Gaussian.

Next, we wish to express the optimal Wiener filter in terms of the channel's Doppler spectrum. Taking expectation of the squared error at time $Mk+m$ followed by discrete-time Fourier transform, we find from (A.6) that:

$$\sigma^2_{\Delta h}(m) := E\left\{|\Delta h(Mk+m)|^2\right\}$$

$$= \sigma_h^2 - \frac{1}{2\pi}\int_{-\pi}^{\pi}\Big[2\mathrm{Re}\left\{G_m^*\left(e^{j\omega}\right)bS_{h,m}\left(e^{j\omega}\right)\right\}$$

$$+ \left|G_m\left(e^{j\omega}\right)\right|^2\left(|b|^2S_{h,0}\left(e^{j\omega}\right)+\sigma_w^2\right)\Big]d\omega. \tag{A.7}$$

By minimizing $\sigma^2_{\Delta h}(m)$ with respect to G_m, the optimal Wiener filter is found to be

$$G_m\left(e^{j\omega}\right) = \frac{bS_{h,m}\left(e^{j\omega}\right)}{|b|^2S_{h,0}\left(e^{j\omega}\right)+\sigma_w^2}. \tag{A.8}$$

Substituting (A.8) to (A.7) yields the expression for $\sigma^2_{\Delta h}(m)$ in (10) and completes the proof.

$$x(Mk+m) = \begin{cases} h(Mk)b + w(Mk), & m = 0 \\ h(Mk+m)s[(M-1)k+m-1] + w(Mk+m), & m \neq 0 \end{cases} \tag{A.1}$$

REFERENCES

[1] S. Adireddy and L. Tong, "Detection with embedded known symbols: Optimal symbol placement and equalization," in *Proc. Int. Conf. ASSP*, vol. 5, Istanbul, Turkey, June 2000, pp. 2541–2543.

[2] ——, "Optimal embedding of known symbols for OFDM," in *Proc. Int. Conf. ASSP*, vol. 4, Salt Lake City, UT, May 2001, pp. 2393–2396.

[3] J. Baltersee, G. Fock, and H. Meyr, "An information theoretic foundation of synchronized detection," *IEEE Trans. Select. Areas Commun.*, vol. 19, pp. 2358–2368, Dec. 2001, submitted for publication.

[4] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inform. Theory*, pp. 2619–2692, Oct. 1998.

[5] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Vehicular Tech.*, vol. 40, pp. 686–693, Nov. 1991.

[6] J. A. Gansman, M. P. Fitz, and J. V. Krogmeier, "Optimum and suboptimum frame synchronization for pilot-symbol-assisted modulation," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1327–1337, Oct. 1997.

[7] G. B. Giannakis, "Cyclostationary signal analysis," in *Digital Signal Processing Handbook*, V. K. Madisetti and D. Williams, Eds. Boca Raton, FL: CRC, 1998, ch. 17.

[8] G. B. Giannakis and C. Tepedelenlioglu, "Basis expansion models and diversity techniques for blind equalization of time-varying channels," *Proc. IEEE*, vol. 86, pp. 1969–1986, Oct. 1998.

[9] B. Hassibi and B. Hochwald, "How much training is needed in multiple-antenna wireless links?," *IEEE Trans. Inform. Theory*, 2000, submitted for publication.

[10] P. Höher and F. Tufvesson, "Channel estimation with super imposed pilot sequence," in *Proc. GLOBECOM Conf.*, Brazil, Dec. 1999, pp. 2162–2166.

[11] W. A. Jakes, *Microwave Mobile Communications*. New York: Wiley, 1974.

[12] G. K. Kaleh, "Channel equalization for block transmission systems," *IEEE J. Select. Areas Commun.*, vol. 13, pp. 1728–1736, Jan. 1995.

[13] T. Keller and L. Hanzo, "Adaptive multicarrier modulation: A convenient framework for time-frequency processing in wireless communications," *IEEE Commun. Mag.*, vol. 88, pp. 611–640, May 2000.

[14] W.-Y. Kuo and M. P. Fitz, "Frequency offset compensation of pilot symbol assisted modulation in frequency flat fading," *IEEE Trans. Commun.*, vol. 45, pp. 1412–1416, Nov. 1997.

[15] A. Lapidoth and S. Shamai (Shitz), "Fading channels: How perfect need "perfect side information" be?," in *Proc. IEEE Information Theory Communications Workshop*, June 1999, pp. 36–38.

[16] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 45, pp. 139–157, Jan. 1999.

[17] M. Médard, "The effect upon channel capacity in wireless communication of perfect and imperfect knowledge of the channel," *IEEE Trans. Inform. Theory*, vol. 46, pp. 933–946, May 2000.

[18] S. Ohno and G. B. Giannakis, "Capacity maximizing pilots and precoders for wireless OFDM over rapidly fading," *IEEE Trans. Inform. Theory*, Apr. 2003, to be published.

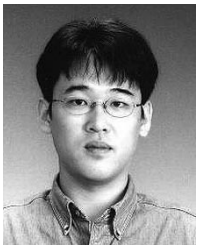
[19] P. Schramm, "Analysis and optimization of pilot-channel-assisted bpsk for DS-CDMA systems," *IEEE Trans. Commun.*, vol. 46, pp. 1122–1124, Sept. 1998.

[20] P. Schramm and R. R. Muller, "Pilot symbol assisted BPSK on Rayleigh fading channels with diversity: Performance analysis and parameter optimization," *IEEE Trans. Commun.*, vol. 46, pp. 1560–1563, Dec. 1998.

[21] I. E. Telatar, "Capacity of multiple-antenna Gaussian channels," *European Trans. Telecommun.*, vol. 10, pp. 585–596, Nov.-Dec. 1999.

[22] H. Vikalo, B. Hassibi, B. Hochwald, and T. Kailath, "Optimal training for frequency-selective fading channels," in *Proc. Int. Conf. ASSP*, vol. 4, Salt Lake City, UT, May 2001, pp. 2105–2108.

[23] Z. Wang and G. B. Giannakis, "Linearly precoded OFDM for fading wireless channels," *IEEE Trans. Inform. Theory*, Apr. 2003, to be published.



Shuichi Ohno (M'95) received the B.E., M.E., and Dr. Eng. degrees in applied mathematics and physics from Kyoto University, Kyoto, Japan, in 1990, 1992, and 1995, respectively.

From 1995 to 1999, he was a Research Associate in the Department of Mathematics and Computer Science at Shimane University, Shimane, Japan, where he became an Assistant Professor. He spent 14 months in 2000 and 2001 at the University of Minnesota, Minneapolis, MN, as a Visiting Researcher. Since 2002, he has been an Associate Professor in the Department of Artificial Complex Systems Engineering, Hiroshima University, Hiroshima, Japan. His current interests are in the areas of signal processing in communication, wireless communications, adaptive signal processing, and multirate signal processing.

Dr. Ohno is a member of the Institute of Electronics and Communication Engineers (IECE). He has been serving as an Associate Editor for IEEE SIGNAL PROCESSING LETTERS since 2001.

Georgios B. Giannakis (S'84–M'86–SM'91–F'97) received the Diploma in electrical engineering from the National Technical University of Athens, Greece, 1981. From September 1982 to July 1986, he was with the University of Southern California (USC), Los Angeles, where he received the M.Sc. degree in electrical engineering, in 1983, the M.Sc. in mathematics, in 1986, and the Ph.D. degree in electrical engineering, in 1986.

After lecturing for one year at USC, he joined the University of Virginia, Charlottesville, in 1987, where he became a Professor of electrical engineering in 1997. Since 1999, he has been a Professor in the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, where he now holds an ADC Chair in Wireless Telecommunications. His general interests span the areas of communications and signal processing, estimation and detection theory, time-series analysis, and system identification—subjects on which he has published more than 150 journal papers, 290 conference papers, and two edited books. His current research topics focus on transmitter and receiver diversity techniques for single- and multiuser fading communication channels, precoding and space–time coding for block transmissions, multicarrier, and wide-brand wireless communication systems.

Dr. Giannakis is the (co-) recipient of four best paper awards from the IEEE Signal Processing Society (1992, 1998, 2000, 2001). He also received the Society's Technical Achievement Award in 2000. He co-organized three IEEE Signal Processing Workshops, and Guest Edited (co-edited) four special issues. He has served as Editor in Chief for the IEEE SIGNAL PROCESSING LETTERS, Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, and the IEEE SIGNAL PROCESSING LETTERS, Secretary of the Signal Processing Conference Board, Member of the Signal Processing Publications Board, Member and Vice-Chair of the Statistical Signal and Array Processing Technical Committee, and Chair of the Signal Processing for Communications Technical Committee. He is a Member of the Editorial Board for the PROCEEDINGS OF THE IEEE, and the steering committee of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. He is a Member of the IEEE Fellows Election Committee, the IEEE-Signal Processing Society's Board of Governors, and a frequent consultant for the telecommunications industry.