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

Shuichi Ohno[1] and Georgios B. Giannakis[2]

[1]Dept. of Mathematics and Computer Science
Shimane Univ., Shimane, 690-8504, Japan
[2]Dept. of ECE, Univ. of Minnesota,
200 Union Street SE, Minneapolis, MN 55455, USA

***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 countermeasure capable of improving bit error rate performance. In this paper, we establish average-rate optimality of PSAM by showing that the insertion of equi-powered and equi-spaced pilot symbols in PSAM transmissions maximizes 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 further design PSAM transmissions with optimal spacing of pilot symbols, and optimal allocation of the transmit-power budget between pilot and information symbols.**

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 links, frequent re-training 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 aided modulation (PSAM) [4]. Acquiring the channel based on the periodically embedded pilots, PSAM is particularly suitable for transmissions over rapidly fading time-selective environments [1,5,7]. The bit error rate (BER) performance of PSAM with linear minimum mean-square error (LMMSE) estimation of time-selective channels was reported in [4]. However, channel capacity issues with PSAM have not been fully addressed.

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., [3] 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. Actually, for a given channel estimation accuracy, only lower and upper bounds on the average channel capacity are available [9]. Although the average channel capacity can be numerically evaluated via Monte-Carlo experiments as in [2], an analytically tractable expression is preferable when it comes to understanding how the fading affects the average channel capacity.

In this paper, we address the optimality of periodically inserting pilot symbols for PSAM-based LMMSE channel estimation, and 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 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 of time-selectivity on the average channel capacity. Simulations illustrate robustness of our theoretical findings to channels with non-ideal (e.g., Jakes-like) spectra, and tightness of our average capacity bounds.

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 channel state information (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 [11]), and are thus not necessarily adhering to a finite alphabet. To acquire CSI at the receiver, we rely on K training (a.k.a. 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 correspondingly each transmitted block $\boldsymbol{u} := [u(0), u(1), \ldots, u(N-1)]^T$ of size N. For simplicity, we select N so that:

C1. *The block size N is an integer multiple of the number K of pilot symbols such that $N = MK$.*

We denote the position of the information-bearing symbols in $\boldsymbol{u}$ by the ordered index set

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

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 $k \in [0, K-1]$.

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 [4]:

$$x(n) = h(n)u(n) + w(n), \quad (2)$$

where $h(n)$ is the sampled multiplicative fading channel, and $w(n)$ is AWGN with variance σ_w^2. We assume that:

A1 *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 in blocks $\boldsymbol{x}_s$ and $\boldsymbol{x}_b$, respectively, to obtain the matrix-vector model:

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

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.

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:

A2 *The (sampled) time-selective channel's Doppler spread obeys: $\omega_d \leq \pi/2$; or, for the continuous-time channel: $f_d \leq 1/(4T_s) := f_s/4$.*

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

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

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. Moreover, this does not tell us anything about the optimum power allocation between pilot and information-bearing symbols. In this paper, we will 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.

III. Average Channel Capacity and CSI

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, i.e., the average channel capacity, will be defined as

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

where $E\{\cdot\}$ denotes the expectation operator with respect to $\boldsymbol{D}_{h,s}$; 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 [3,8]. For our model (3), it is found to be [10]

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

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

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

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. Thus, in the following, we assume that:

A3 *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.

Let now $\hat{\boldsymbol{D}}_{h,s}$ be an estimate of $\boldsymbol{D}_{h,s}$. Since the channel estimator $\hat{\boldsymbol{D}}_{h,s}$ is used for recovering information-bearing symbols, we deduce from (3) that

$$\boldsymbol{x}_s = \boldsymbol{D}_{h,s}\boldsymbol{s} + \boldsymbol{w}_s = \hat{\boldsymbol{D}}_{h,s}\boldsymbol{s} + \boldsymbol{\nu}, \quad (8)$$

where $\boldsymbol{\nu} := \Delta\boldsymbol{D}_{h,s}\boldsymbol{s} + \boldsymbol{w}_s$, and $\Delta\boldsymbol{D}_{h,s} := \boldsymbol{D}_{h,s} - \hat{\boldsymbol{D}}_{h,s}$ captures the channel estimation error.

It is not easy to evaluate the average channel capacity when there are channel estimation errors, because $\boldsymbol{\nu}$ depends on the adopted channel estimator, and $\boldsymbol{\nu}$ is non-Gaussian in general. However, it is possible to evaluate its lower bound [6,9,10]. The lower bound on the average channel capacity for an unbiased channel estimator, is given by [10]

$$\underline{C} = \frac{1}{N}\sum_{i\in\mathcal{I}} E\left\{\log\left(1 + \frac{|\hat{h}(i)|^2}{E\{|\Delta h(i)|^2\} + \sigma_w^2/\sigma_s^2}\right)\right\}, \quad (9)$$

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

To proceed, we will borrow a result that was originally derived in [10] for OFDM transmissions:

Lemma 1: *If A1-A3 and C1 hold true, then for N sufficiently large, $\underline{C}$ in (9) is maximized at high SNR if and only if the channel's MMSE, $E\{|\Delta h(i)|^2\}$, is constant for all i's.*

Assured by Lemma 1 on PSAM's average-rate optimality with equi-spaced and equi-powered pilot symbols, 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 that:

C2 *Equi-powered pilot symbols $b(k) = b$ are inserted equi-spaced at positions Mk for all $k \in [0, K-1]$.*

IV. Lower Bound on Average Channel Capacity

Having linked PSAM-based LMMSE channel estimation with the maximum lower bound $\underline{C}$, we will further simplify it. We will let the block size N go to infinity because the resulting closed form expression in this asymptotic case offers useful insights.

Normalizing $\hat{h}(i)$ in (9) by its variance, we can show that (proofs are omitted due to lack of space)

$$\underline{C} = \frac{1}{M}\sum_{m=1}^{M-1} E\left\{\log\left(1 + \frac{\sigma_h^2 - \sigma_{\Delta h}^2(m)}{\sigma_{\Delta h}^2(m) + \sigma_w^2/\sigma_s^2}|h|^2\right)\right\} , \tag{10}$$

where $\sigma_{\Delta h}^2(m) := E\{|h(m) - \hat{h}(m)|^2\}$. Eq. (10) 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 [2]. But the latter does not impose A2 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 (10) is simple to evaluate numerically.

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. For a fixed M, the $\underline{C}$ in (10) will be maximized, if $\sigma_{\Delta h}^2(m)$ is minimized. We are thus motivated to design our pilots so that:

C3 *The spacing between successive pilot symbols satisfies: $M \leq \lfloor 1/(2\bar{f}_d)\rfloor := M_{max}$, where $\lfloor\ \rfloor$ denotes integer-floor.*

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.

If C3 holds true, then the conditions of Lemma 1 are satisfied such that $\sigma_{\Delta h}^2(m) := \sigma_{\Delta h}^2$ for any m, and the lower bound reduces to

$$\underline{C} = \frac{M-1}{M} E\left\{\log\left(1 + \rho|h|^2\right)\right\} , \tag{11}$$

where

$$\rho := \frac{\sigma_h^2 - \sigma_{\Delta h}^2}{\sigma_{\Delta h}^2 + \sigma_w^2/\sigma_s^2} . \tag{12}$$

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 implied by Lemma 1 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 (11). We will also find the optimal power distribution among pilot and information-bearing symbols.

Defining the average transmit-power as

$$\bar{\mathcal{P}} := \frac{1}{M}\left[|b|^2 + (M-1)\sigma_s^2\right] , \tag{13}$$

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.,

$$\sigma_s^2 = \frac{\alpha M\bar{\mathcal{P}}}{M-1}, \quad |b|^2 = (1-\alpha)M\bar{\mathcal{P}}. \tag{14}$$

Our design parameters for the maximization of the average capacity bound are the integer M and the power ratio α. First, we fix α and optimize $\underline{C}$ in (11) with respect to M.

A. Optimal Spacing of Pilot Symbols

For a fixed α, defining $\beta := (M-1)/M$, we can re-express (11) as a function of β as follows:

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

where $\rho(\beta) = (\sigma_h^2 - \sigma_{\Delta h}^2)/(\sigma_{\Delta h}^2 + a\beta)$ with $a := \sigma_w^2/(\alpha\bar{\mathcal{P}})$.

Treating β as a continuous variable, we differentiate (15) with respect to β to obtain

$$\frac{\partial\underline{C}}{\partial\beta} = E\left\{\log\left(1 + \rho|h|^2\right) + \beta\frac{\partial\rho}{\partial\beta}\frac{|h|^2}{1+\rho|h|^2}\right\} . \tag{16}$$

Substituting $\partial\rho/\partial\beta = -\rho a/(\sigma_{\Delta h}^2 + a\beta)$ into (16), we can show that $\partial\underline{C}/\partial\beta > 0$. Thus, 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 A1-A3 and C1-C3 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 (15) 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.

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

$$S_h(e^{j\omega}) = \begin{cases} \dfrac{\sigma_h^2}{2\bar{f}_d} & \text{for } |\omega| < \omega_d, \\ 0 & \text{otherwise.} \end{cases} \quad (17)$$

In this case, we can also re-express ρ as

$$\rho = \rho_{snr} \left(\frac{M}{M-1} \right) \zeta \, \frac{\alpha(1-\alpha)}{(\gamma - \alpha)}, \quad (18)$$

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

$$\zeta := \frac{1}{1 - 2M\bar{f}_d/(M-1)}, \quad \gamma := \left(1 + \frac{2\bar{f}_d}{\rho_{snr}}\right)\zeta. \quad (19)$$

Notice that A2 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

$$\alpha_{lp} := \frac{1}{1 + \sqrt{1 - 1/\gamma}}. \quad (20)$$

Substituting this into ρ leads to

$$\rho_{lp} := \rho_{snr} \left(\frac{M}{M-1} \right) \frac{\alpha_{lp}^2}{1 + 2\bar{f}_d/\rho_{snr}}. \quad (21)$$

We thus arrive at the following result:

Result 2: *Suppose that A1-A3 and C1-C3 hold true, and that the channel psd has an ideal low-pass spectrum as in (17). For a fixed M, the lower bound on the average channel capacity $\underline{C}$ is maximized for the power ratio α_{lp} in (20), and it can be expressed in closed form as*:

$$\underline{C}_{lp} := \frac{M-1}{M} E\left\{\log(1 + \rho_{lp}|h|^2)\right\}, \quad (22)$$

where ρ_{lp} is given by (21).

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 ρ_{lp}/ρ_{ideal}. Since $\mathcal{P}_s = \alpha_{lp} N \bar{\mathcal{P}}$, we obtain from (7) that $\rho_{ideal} = \alpha_{lp}\rho_{snr} M/(M-1)$. Thus, we find that $\rho_{lp}/\rho_{ideal} = \alpha_{lp}/(1 + 2\bar{f}_d/\rho_{snr})$. Since γ is a decreasing function of SNR, α_{lp} is found from (20) to be an increasing function of SNR, and is bounded such that $\alpha_{lp} \geq 0.5$. On the other hand, for $2\bar{f}_d/\rho_{snr} \leq 1$, i.e., $\rho_{snr} \geq 2\bar{f}_d$, we have that $\rho_{lp}/\rho_{ideal} \geq \alpha_{lp}/2 \geq 0.25$. This implies that for $\rho_{snr} \geq 2\bar{f}_d$, the performance loss due to the channel estimation error is upper bounded by about 6.0dB. Since $\rho_{lp}/\rho_{ideal} = \alpha_{lp}$ at high SNR, the performance loss is at most 3.0dB at high SNR. These observations also imply that $\underline{C}_{lp}$ is tight 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 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 $\rho_{snr} \to \infty$, (20) and (21) reduce to

$$\alpha_\infty := \frac{1}{1 + \sqrt{2M\bar{f}_d/(M-1)}} \quad (23)$$

$$\rho_\infty := \rho_{snr} \left(\frac{M}{M-1} \right) \alpha_\infty^2, \quad (24)$$

respectively. Thus, we have established the following result:

Result 3: *Suppose that A1-A3 and C1-C3 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*

$$\underline{C}_\infty := \frac{M-1}{M} E\left\{\log(1 + \rho_\infty|h|^2)\right\}, \quad (25)$$

where ρ_∞ is defined as in (24).

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 $\rho_\infty/\rho_{ideal} = \alpha_\infty \cong 1/[1 + \sqrt{2\bar{f}_d/(1 - 2\bar{f}_d)}]$. 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 back-offs needed to alleviate nonlinear power amplifier distortions.

To maintain constant modulus of all transmitted symbols, we will impose the constraint $|b|^2 = E\{|s(i)|^2\} = \sigma_s^2$. Then, we can show that as long as $\mu \leq 1 + \sqrt{1 + 1/[2\bar{f}_d(1 + 1/\rho_{snr})]}$, we have $\partial\underline{C}/\partial\mu > 0$. This implies that the optimal M lies between $1 + \sqrt{1 + 1/[2\bar{f}_d(1 + 1/\rho_{snr})]}$ and M_{max}. In other words, the optimal M may not be M_{max}, unlike the unconstrained case we dealt with in subsection V-A. To obtain the optimal M, one has to resort to a numerical line search.

VI. Numerical Examples

To validate our analysis and design, we tested the applicability of (6) and (11) in evaluating average capacity bounds for

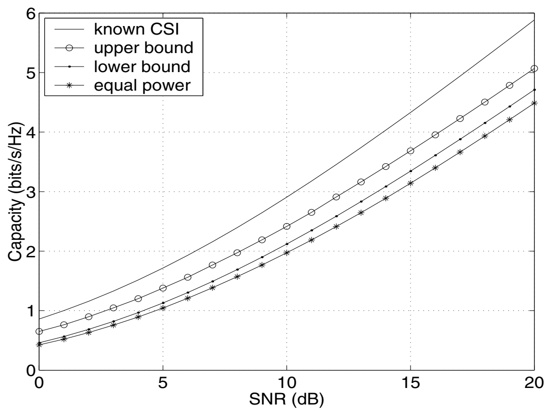

Fig. 1. Channel capacity bounds vs. SNR (Low-pass model)

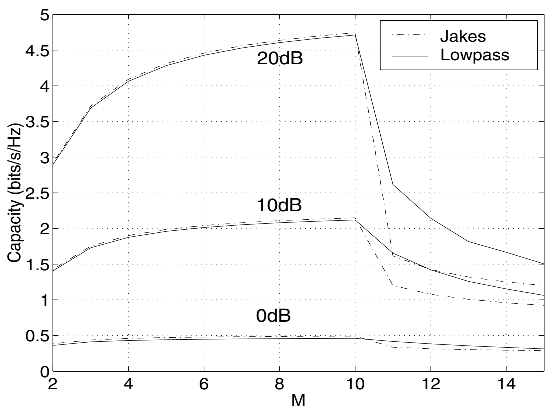

Fig. 2. Lower bounds on the average channel capacity vs. M

fading channels with: i) low-pass Doppler spectra (c.f. (17)); and ii) Jakes spectra with normalized maximum Doppler frequency 0.05. In both cases, we deduce from C3 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: i) the channel is known to the receiver and training is not required, in which case we set $\alpha = 1$, and $K = 0$ in (6); ii) the upper and lower bounds on the average capacity in (22), with the optimal α_{lp} from (20); and iii) the lower bound on the average capacity for the equi-powered PSAM with constant modulus $|b|^2 = \sigma_s^2$.

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

Test Case 2 (Optimal Spacing of Pilot Symbols): Fig. 2 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 C3.

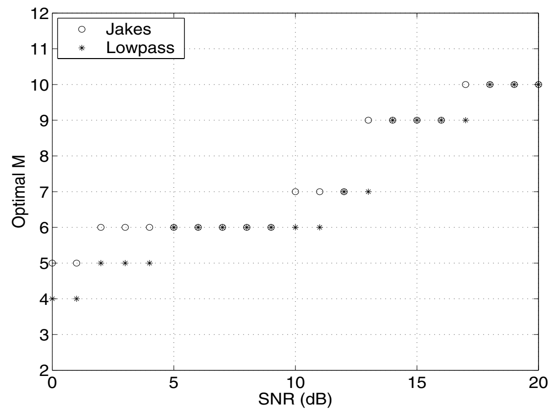

Fig. 3. Optimal pilot spacing vs. SNR (equi-powered PSAM)

Test Case 3 (Equi-Powered PSAM): Fig. 3 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 $\lfloor 1/(2\bar{f}_d) \rfloor$.

REFERENCES

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

[2] J. Baltersee, G. Fock, and H. Meyr, "An information theoretic foundation of synchronized detection," Submitted to *IEEE Trans. Com.*, 2000.

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

[4] 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.

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

[6] B. Hassibi and B. Hochwald, "How much training is needed in multiple-antenna wireless links?," Submitted to *IEEE Trans. Info. Theory*, 2000; URL: *http://mars.bell-labs.com/cm/ms/what/mars/index.html*.

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

[8] T.L. Marzetta and B.M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Transactions on Information Theory*, pp. 139–157, Jan. 1999.

[9] M. Médard, "The effect upon channel capacity in wireless communication of perfect and imperfect knowledge of the channel," *IEEE Transactions on Information Theory*, pp. 933–946, May 2000.

[10] S. Ohno and G. B. Giannakis, "Capacity maximizing pilots and precoders for wireless OFDM over rapidly fading," Submitted to *IEEE Trans. on Information Theory*, April 2001; also in *Proc. of ISSSE*, pp.246–249, Tokyo, Japan, July 24-27, 2001.

[11] Z. Wang and G. B. Giannakis, "Linearly precoded OFDM for fading wireless channels," *IEEE Trans. on Information Theory*, submitted April 2001; see also *Proc. of 3rd IEEE Work. on Signal Proc. Advances in Wireless Comm.*, pp. 267-270, 2001.