A Unified Treatment of Optimum Pilot Overhead in Multipath Fading Channels

Nihar Jindal, *Member, IEEE*, and Angel Lozano, *Senior Member, IEEE*

***Abstract*—The optimization of the pilot overhead in single-user wireless fading channels is investigated, and the dependence of this overhead on various system parameters of interest (e.g., fading rate, signal-to-noise ratio) is quantified. The achievable pilot-based spectral efficiency is expanded with respect to the fading rate about the no-fading point, which leads to an accurate order expansion for the pilot overhead. This expansion identifies that the pilot overhead, as well as the spectral efficiency penalty with respect to a reference system with genie-aided CSI (channel state information) at the receiver, depend on the square root of the normalized Doppler frequency. It is also shown that the widely-used block fading model is a special case of more accurate continuous fading models in terms of the achievable pilot-based spectral efficiency. Furthermore, it is established that the overhead optimization for multiantenna systems is effectively the same as for single-antenna systems with the normalized Doppler frequency multiplied by the number of transmit antennas.**

***Index Terms*—Spectral efficiency, mutual information, channel estimation, doppler spectrum, MIMO, multiantenna.**

I. Introduction

MOST wireless communication systems perform coherent data detection with the assistance of pilot signals (a.k.a. reference signals or training sequences) that are inserted periodically [2], [3]. The receiver typically performs channel estimation on the basis of the received pilot symbols, and then applies standard coherent detection while treating the channel estimate as if it were the true channel. When such an approach is taken[1] and Gaussian inputs are used, the channel estimation error effectively introduces additional Gaussian noise [5]. This leads to a non-trivial tradeoff: increasing the fraction and/or power of pilot symbols improves the channel estimation quality and thus decreases this additional noise, but also decreases the fraction and/or power of data symbols. To illustrate the importance of this tradeoff, Fig. 1 depicts the spectral efficiency as function of the pilot overhead (cf. Section III for details) for some standard channel conditions. Clearly, an incorrect overhead can greatly diminish the achievable spectral efficiency. This optimization is critical and has been extensively studied in the literature [5]–[14], on the basis of both the simplified block-fading model as well as the more accurate continuous-fading model. In most systems the pilot symbol power is fixed due to peak-to-average and interference considerations, and thus the optimization is only over the fraction of pilot symbols. This version of the optimization has been posed in prior work [7], [11]–[13], and explicit results have been established in the low- and high-power asymptotes. However, these asymptotes become accurate only for very extreme power levels [14]; as a result, this version of the optimization must generally be solved numerically. When both the power and fraction of pilot symbols can be varied, a closed-form solution for the optimal power and fraction is known for the special cases of block-fading [7] and continuous-fading with a rectangular Doppler spectrum [12], [13].

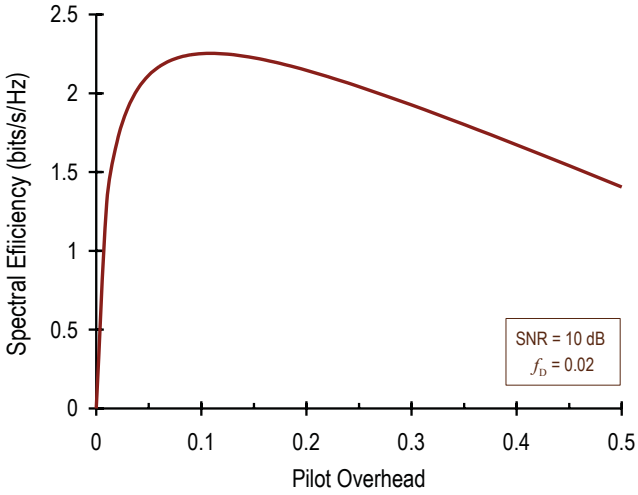

Fig. 1. Spectral efficiency as function of the pilot overhead, α, for SNR $=$ 10 dB. The Doppler spectrum is Clarke-Jakes with a maximum normalized frequency $f_D = 0.02$ corresponding, for instance, to 100 Km/h in a WiMAX subcarrier. The pilot and data symbols have equal power.

In this paper, we study the overhead optimization (with and without pilot power boosting) in the limiting regime of slow fading. More precisely, by expanding the spectral efficiency around the perfect-CSI point, i.e., for small fading rates, both versions of the optimizations can be tackled and useful expansions (in terms of the fading rate) for the optimum pilot overhead are obtained. These expansions are seen to be accurate over a wide range of relevant operating points, and provide valuable insight on the dependence of the optimal overhead and corresponding spectral efficiency on the various parameters of interest (velocity, power, etc). For non-boosted

Paper approved by E. Serpedin, the Editor for Synchronization and Sensor Networks of the IEEE Communications Society. Manuscript received November 13, 2009; revised April 1, 2010.

N. Jindal is with the University of Minnesota, Minneapolis, MN55455, USA (e-mail: nihar@umn.edu).

A. Lozano is with the Universitat Pompeu Fabra, Barcelona 08005, Spain (e-mail: angel.lozano@upf.edu).

His work is supported by the projects CONSOLIDER-INGENIO CSD2008-00010 "COMONSENS" and TEC2009-13000, and by the Marie Curie IRG-224755. An earlier version of this work appeared at the European Wireless Conference[1].

Digital Object Identifier 10.1109/TCOMM.2010.083110.090696

[1]The focus of this paper is on *separate* processing of pilots and data. Other approaches such as data-aided channel estimation or semi-blind estimation are feasible, but are not currently employed (cf. [4]) and are not investigated here.

pilots in particular, these expansions are the first explicit expressions that capture these dependencies for arbitrary power levels. Some additional insights reached in the paper are as follows:

- In terms of the spectral efficiency achievable with pilot-based communication, block-fading is a special case of continuous (symbol-by-symbol) fading.
- The optimal pilot overhead scales with the square root of the Doppler frequency; this result holds regardless of whether pilot power boosting is allowed.[2]
- The spectral efficiency penalty w.r.t. the perfect-CSI capacity also scales with the square-root of the fading rate.
- The pilot overhead optimization for multiantenna transmission is essentially the same as the optimization for single-antenna transmission except with the true Doppler frequency multiplied by the number of transmit antennas.

By showing that block-fading is a special case of the richer set of continuous fading models, the two models are unified in the context of pilot-based communication. This is important because, thus far, these models had been treated separately (block fading is used in [6]–[9] while continuous-fading is considered in [11]–[14]) and no connection had been established between them. In this paper we show that the two need not be separately treated; moreover, results obtained for one model can be applied to the other. The connection established between single-antenna and multiantenna transmission has similar benefits.

II. Preliminaries

A. Channel Model

Consider a discrete-time frequency-flat scalar fading channel $H(k)$ where k is the time index. (The extension to multiantenna channels is considered in Section VI.) Pilot symbols are inserted periodically in the transmission [4] and the fraction thereof is denoted by α, i.e., one in every $1/\alpha$ symbols is a pilot while the rest are data.[3] Moreover, $\alpha \geq \alpha_{\min}$ where $\alpha_{\min}$ is established later in this section.

Let $\mathcal{D}$ denote the set of time indices corresponding to data symbols. For $k \in \mathcal{D}$,

$$Y(k) = H(k)\sqrt{P}X(k) + N(k) \tag{1}$$

where the transmitted signal, $X(k)$, is a sequence of IID (independent identically distributed) complex Gaussian random variables with zero mean and unit variance that we indicate by $X \sim \mathcal{N}_{\mathbb{C}}(0,1)$. The additive noise is $N \sim \mathcal{N}_{\mathbb{C}}(0, N_0)$ and we define $\mathsf{SNR} = P/N_0$.

For $k \notin \mathcal{D}$, unit-amplitude pilots are transmitted and thus

$$Y(k) = H(k)\sqrt{P} + N(k). \tag{2}$$

Notice that pilot symbols and data symbols have the same average power. In Section V, we shall lift this constraint allowing for power-boosted pilots.

[2]To the best of our knowledge, this square-root dependence was first identified in the context of a different (and weaker) lower bound for the multiantenna broadcast channel in [15].

[3]Although α should be restricted to integer-reciprocals, our derivations relax this constraint. Thus, our results should be rounded to an integer-reciprocal.

1) Block Fading: In the popular block-fading model, the channel is drawn as $H \sim \mathcal{N}_{\mathbb{C}}(0,1)$ at the beginning of each block and it then remains constant for the n_{b} symbols composing the block. This process is repeated for every block in an IID fashion.

In order for the receiver to estimate the channel, at least one pilot symbol must be inserted within each block. If n_{p} represents the number of pilot symbols in every block, then

$$\alpha = \frac{n_{\mathsf{p}}}{n_{\mathsf{b}}} \tag{3}$$

and clearly $\alpha_{\min} = 1/n_{\mathsf{b}}$.

2) Continuous Fading: In this model, $H(k)$ is a discrete-time complex Gaussian stationary[4] random process, with an absolutely continuous spectral distribution function whose derivative is the Doppler spectrum $S_H(\nu)$, $-1/2 \leq \nu \leq 1/2$ [16]. It follows that the channel is ergodic.

We consider bandlimited processes such that

$$\begin{cases} S_H(\nu) > 0, & |\nu| \leq f_{\mathrm{D}} \\ S_H(\nu) = 0, & |\nu| > f_{\mathrm{D}} \end{cases} \tag{4}$$

for some Doppler frequency $f_{\mathrm{D}} \leq 1/2$. The Doppler frequency is typically given by $f_{\mathrm{D}} = Tv/\lambda$, where T is the symbol period, v is the velocity, and λ is the carrier wavelength.

To ensure that the decimated channel observed through the pilot transmissions has an unaliased spectrum, it is necessary that

$$\alpha_{\mathsf{min}} = 2f_{\mathrm{D}}. \tag{5}$$

On account of its bandlimited nature, the channel is a non-regular fading process [16]. We further consider $S_H(\cdot)$ to be strictly positive within $\pm f_{\mathrm{D}}$.[5] In order to remain consistent with earlier definitions of signal and noise power, only unit-power processes are considered.

Two important spectra are the Clarke-Jakes [17]

$$S_H(\nu) = \frac{1}{\pi\sqrt{f_{\mathrm{D}}^2 - \nu^2}} \tag{6}$$

and the rectangular

$$S_H(\nu) = \begin{cases} 1/(2f_{\mathrm{D}}) & |\nu| \leq f_{\mathrm{D}} \\ 0 & |\nu| > f_{\mathrm{D}}. \end{cases} \tag{7}$$

We will later find it useful to express the Doppler spectrum as

$$S_H(\nu) = \frac{1}{f_{\mathrm{D}}}\,\tilde{S}_H\left(\frac{\nu}{f_{\mathrm{D}}}\right) \tag{8}$$

where $\tilde{S}_H(\cdot)$ is a normalized spectral shape bandlimited to ± 1. For the Clarke-Jakes spectrum in (6), for instance, the spectral shape is

$$\tilde{S}_H(\nu) = \frac{1}{\pi\sqrt{1-\nu^2}} \tag{9}$$

while, for the rectangular spectrum in (7), the spectral shape is

$$\tilde{S}_H(\nu) = \begin{cases} 1/2 & |\nu| \leq 1 \\ 0 & |\nu| > 1. \end{cases} \tag{10}$$

[4]The block-fading model, in contrast, is not stationary but only cyclostationary. Also, note that this discrete-time process is related to the underlying continuous-time fading process as detailed in [14].

[5]This premise can be easily removed by simply restricting all the integrals in the paper to the set of frequencies where $S_H(\nu) > 0$, rather than to the entire interval $\pm f_{\mathrm{D}}$.

B. Perfect CSI

With perfect CSI at the receiver (but not at the transmitter), i.e., assuming a genie provides the receiver with $H(k)$, there is no need for pilot symbols ($\alpha = 0$). The capacity in bits/s/Hz is then [18], [19]

$$C(\mathsf{SNR}) = \mathbb{E}\left[\log_2\left(1 + \mathsf{SNR}\,|H|^2\right)\right] \tag{11}$$

$$= \log_2(e)\; e^{1/\mathsf{SNR}} E_1\left(\frac{1}{\mathsf{SNR}}\right) \tag{12}$$

where $E_1(\zeta) = \int_1^\infty t^{-1} e^{-\zeta t} dt$ is the exponential integral. The first derivative of $C(\cdot)$ can be conveniently expressed as a function of $C(\cdot)$ via

$$\dot{C}(\mathsf{SNR}) = \frac{1}{\mathsf{SNR}}\left(\log_2 e - \frac{C(\mathsf{SNR})}{\mathsf{SNR}}\right). \tag{13}$$

In turn, the second derivative can be expressed as function of $C(\cdot)$ and $\dot{C}(\cdot)$ as

$$\ddot{C}(\mathsf{SNR}) = -\frac{1}{\mathsf{SNR}^2}\left[\log_2 e + \dot{C}(\mathsf{SNR}) - 2\frac{C(\mathsf{SNR})}{\mathsf{SNR}}\right]. \tag{14}$$

III. Pilot-Assisted Detection

In pilot-assisted communication, decoding must be conducted on the basis of the channel outputs (data and pilots) alone, without the assistance of genie-provided channel realizations. In this case, the maximum spectral efficiency that can be achieved reliably is the mutual information between the data inputs and the outputs (data and pilots):

$$\lim_{K\to\infty} \frac{1}{K}\, I\left(\underbrace{\{X(k)\}_{k=0}^{K-1}; \{Y(k)\}_{k=0}^{K-1}}_{k\in\mathcal{D}} \,|\, \underbrace{\{Y(k)\}_{k=0}^{K-1}}_{k\notin\mathcal{D}}\right) \tag{15}$$

where K signifies the blocklength in symbols. Achieving (15), for which there is no known simplified expression, generally requires joint data decoding and channel estimation.

Contemporary wireless systems take the lower complexity, albeit suboptimal, approach of first estimating the channel for each data symbol—based exclusively upon all received pilot symbols—and then performing nearest-neighbor decoding using these channel estimates as if they were correct. This is an instance of mismatched decoding [20]. If we express the channel as $H(k) = \hat{H}(k) + \tilde{H}(k)$ where $\hat{H}(k)$ denotes the minimum mean-square error estimate of $H(k)$, the received symbol can be re-written as

$$Y(k) = \hat{H}(k)\sqrt{P}\,X(k) + \tilde{H}(k)\sqrt{P}\,X(k) + N(k). \tag{16}$$

Performing nearest-neighbor decoding as described above[6] has been shown to have the effect of making the term $\tilde{H}(k)\sqrt{P}\,X(k)$ appear as an additional source of independent Gaussian noise [21]. With that, the spectral efficiency becomes [5]–[7], [11], [12]

$$\bar{\mathcal{I}}(\mathsf{SNR}, \alpha) = (1-\alpha)\, C(\mathsf{SNR}_{\mathsf{eff}}) \tag{17}$$

with

$$\mathsf{SNR}_{\mathsf{eff}} = \frac{\mathsf{SNR}\,(1 - \mathsf{MMSE})}{1 + \mathsf{SNR}\cdot\mathsf{MMSE}} \tag{18}$$

[6]More specifically, the decoder finds the codeword $[X(1), \ldots, X(K)]$ that minimizes the distance metric $\sum_{k=1}^{K} |Y(k) - \sqrt{P}\hat{H}(k)X(k)|^2$.

where $\mathsf{MMSE} = \mathbb{E}[|\tilde{H}|^2]$. Although not explicitly indicated, MMSE and $\mathsf{SNR}_{\mathsf{eff}}$ are functions of SNR, α and the underlying fading model.

In addition to representing the maximum spectral efficiency achievable with Gaussian signals and channel-estimate-based nearest-neighbor decoding, $\bar{\mathcal{I}}(\cdot)$ is also a lower bound to (15). Because of this double significance, the maximization of $\bar{\mathcal{I}}(\cdot)$ over α

$$\bar{\mathcal{I}}^\star(\mathsf{SNR}) = \max_{\alpha_{\min}\le\alpha\le 1} \bar{\mathcal{I}}(\mathsf{SNR}, \alpha) \tag{19}$$

and especially the argument of such maximization, $\alpha^\star$, are the focal points of this paper.

The expressions in (17) and (19) apply to both block and continuous fading, and these settings differ only in how MMSE behaves as a function of α and SNR.

In block fading, n_{p} pilot symbols are used to estimate the channel in each block and thus [7]

$$\mathsf{MMSE} = \frac{1}{1 + \alpha\, n_{\mathsf{b}} \mathsf{SNR}}. \tag{20}$$

For continuous fading, on the other hand [3], [12]

$$\mathsf{MMSE} = 1 - \int_{-f_{\mathrm{D}}}^{+f_{\mathrm{D}}} \frac{\mathsf{SNR}\, S_H^2(\nu)}{1/\alpha + \mathsf{SNR}\, S_H(\nu)}\, d\nu \tag{21}$$

$$= 1 - \int_{-1}^{+1} \frac{\tilde{S}_H^2(\xi)}{\frac{f_{\mathrm{D}}}{\alpha\,\mathsf{SNR}} + \tilde{S}_H(\xi)}\, d\xi \tag{22}$$

where (22) is based upon the spectral shape definition in (8).

For the Clarke-Jakes spectrum, MMSE can be computed in closed-form as [14]

$$\mathsf{MMSE} = 1 - \frac{\operatorname{arctanh}\sqrt{1 - \left(\frac{\alpha\,\mathsf{SNR}}{\pi f_{\mathrm{D}}}\right)^2}}{\frac{\pi}{2}\sqrt{\left(\frac{\pi f_{\mathrm{D}}}{\alpha\,\mathsf{SNR}}\right)^2 - 1}} \tag{23}$$

while, for the rectangular spectrum [12]

$$\mathsf{MMSE} = \frac{1}{1 + \frac{\alpha}{2 f_{\mathrm{D}}}\mathsf{SNR}}. \tag{24}$$

Proposition 1 *For any pilot overhead α, the spectral efficiency achievable with pilot-based communication, i.e., $\bar{\mathcal{I}}(\mathsf{SNR}, \alpha)$, on a block-fading channel with blocksize n_{b} equals the spectral efficiency achievable on a continuous-fading channel with a rectangular spectrum where*

$$f_{\mathrm{D}} = \frac{1}{2n_{\mathsf{b}}}. \tag{25}$$

Proof: Comparing (20) with (24), the block-fading model is seen to yield the same MMSE as a continuous fading model with a rectangular spectrum if $f_{\mathrm{D}} = 1/(2n_{\mathsf{b}})$. Because $\bar{\mathcal{I}}(\cdot)$ depends on the fading model only through MMSE, this further implies equivalence in terms of spectral efficiency. ■

Thus, for the remainder of the paper we consider only continuous fading while keeping in mind that block-fading corresponds to the special case of a rectangular spectrum with (25).

Notice by applying Jensen's inequality to the integral in (22), using the fact that the function $x^2/(a+x)$ is convex in x for any $a > 0$, that the MMSE for any spectrum (and fixed f_{D}) is

lower bounded by the right-hand side of (24). In other words, the rectangular spectrum results in the worst case estimation error; thus, block-fading analyses are also worst-case.

IV. Pilot Overhead Optimization

The optimization in (19) does not yield an explicit solution, even for the simplest of fading models, and therefore it must be computed numerically. In this section, we circumvent this difficulty by appropriately expanding the objective function $\bar{\mathcal{I}}(\cdot)$. This leads to a simple expression that cleanly illustrates the dependence of α^* and $\bar{\mathcal{I}}^*$ on the parameters of interest.

In particular, we shall expand (17) with respect to f_{D} while keeping the shape of the Doppler spectrum fixed (but arbitrary). Besides being analytically convenient, this approach correctly models different velocities within a given propagation environment.[7] We shall henceforth explicitly indicate the dependence of $\bar{\mathcal{I}}(\cdot)$ and $\bar{\mathcal{I}}^\star(\cdot)$ on f_{D}.

Proposition 2 *The optimum pilot overhead for a Rayleigh-faded channel with an arbitrary bandlimited Doppler spectrum expands as*

$$\alpha^\star = \sqrt{(1+\mathsf{SNR})\frac{\dot{C}(\mathsf{SNR})}{C(\mathsf{SNR})}\,2f_{\mathrm{D}}} - \left((1+\mathsf{SNR})\frac{\ddot{C}(\mathsf{SNR})}{\dot{C}(\mathsf{SNR})} + 2 + \frac{1}{2\,\mathsf{SNR}}\int_{-1}^{+1}\frac{d\xi}{\tilde{S}_H(\xi)}\right) f_{\mathrm{D}} + \mathcal{O}(f_{\mathrm{D}}^{3/2}) \tag{26}$$

where $\tilde{S}_H(\nu)$ is the spectral shape defined in (8). The corresponding spectral efficiency expands as

$$\bar{\mathcal{I}}^\star(\mathsf{SNR}, f_{\mathrm{D}}) = C(\mathsf{SNR}) - \sqrt{8 f_{\mathrm{D}}\,(1+\mathsf{SNR})\,C(\mathsf{SNR})\,\dot{C}(\mathsf{SNR})} + \mathcal{O}(f_{\mathrm{D}}) \tag{27}$$

Proof: *See Appendix IX-A.*

The expression for $\alpha^\star$ in Proposition 2 is a simple function involving the perfect-CSI capacity and its derivatives (cf. Section II). Furthermore, the leading term in the expansion does not depend on the particular spectral shape. Only the subsequent term begins to exhibit such dependence, through $\int_{-1}^{+1} d\nu/\tilde{S}_H(\nu)$. For a Clarke-Jakes spectrum, for instance, this integral equals $\pi^2/2$. For a rectangular spectrum, it equals 4.

Comparisons between the optimum pilot overhead expansion in Proposition 2 and the corresponding exact value obtained numerically are presented in Figs. 2–3. The agreement is excellent for a broad range of Doppler and SNR values. As with the optimum overhead, good agreement is shown in Figs. 4–5 between the spectral efficiency in (27) and its numerical counterpart as rendered by (19).

A direct insight of Proposition 2 is that the optimum pilot overhead, α^*, and the spectral efficiency penalty w.r.t. the perfect-CSI capacity, $C(\mathsf{SNR}) - \bar{\mathcal{I}}^\star(\mathsf{SNR}, f_{\mathrm{D}})$, both depend on the Doppler as $\sqrt{f_{\mathrm{D}}}$. To gain an intuitive understanding of this

[7]The propagation environment determines the spectral shape while the velocity determines f_{D}.

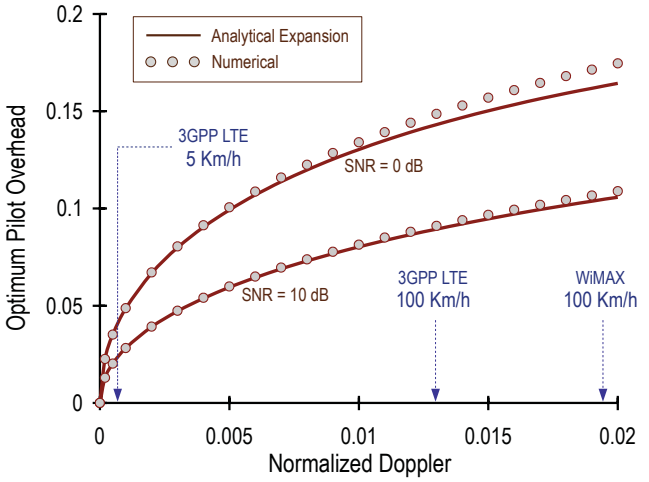

Fig. 2. Optimum pilot overhead, $\alpha^\star$, as function of f_{D} for $\mathsf{SNR} = 0$ dB and $\mathsf{SNR} = 10$ dB with a Clarke-Jakes spectrum. Relevant Doppler levels for LTE and WiMAX subcarriers are highlighted.

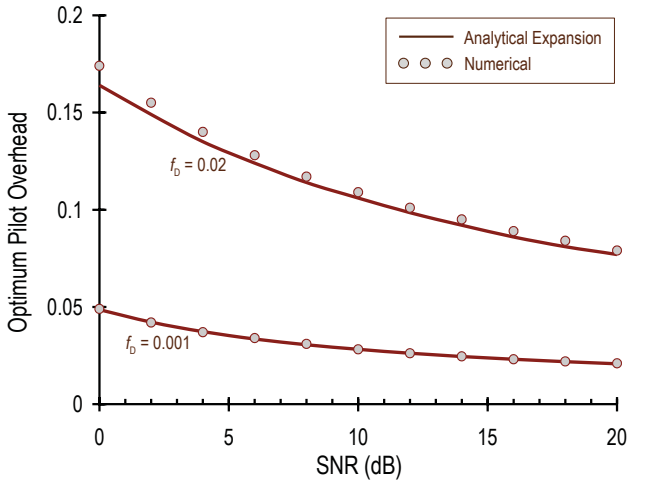

Fig. 3. Optimum pilot overhead, $\alpha^\star$, as function of SNR for $f_{\mathrm{D}} = 0.001$ and $f_{\mathrm{D}} = 0.02$ with a Clarke-Jakes spectrum.

scaling, we can express such penalty for an arbitrary α as (cf. Appendix IX-A, Eq. 62)

$$C(\mathsf{SNR}) - \bar{\mathcal{I}}(\mathsf{SNR}, \alpha, f_{\mathrm{D}}) = \alpha\, C(\mathsf{SNR}) + \frac{(1+\mathsf{SNR})\dot{C}(\mathsf{SNR})\,2f_{\mathrm{D}}}{\alpha} + \mathcal{O}(f_{\mathrm{D}}). \tag{28}$$

The first term in (28) represents the spectral efficiency loss because only a fraction $(1-\alpha)$ of the symbols contain data, while the second term is the loss on those transmitted data symbols due to the inaccurate CSI. If α is chosen to be $\mathcal{O}(f_{\mathrm{D}}^s)$ for $0 \le s \le 1$, the first and second terms in (28) are $\mathcal{O}(f_{\mathrm{D}}^s)$ and $\mathcal{O}(f_{\mathrm{D}}^{1-s})$, respectively, and thus the overall penalty is

$$\mathcal{O}\left(f_{\mathrm{D}}^{\min\{s,1-s\}}\right). \tag{29}$$

Hence, the spectral efficiency penalty is minimized by balancing the two terms and selecting $\alpha^\star = \mathcal{O}(\sqrt{f_{\mathrm{D}}})$.

In parsing the dependence of $\alpha^\star$ upon SNR, it is worth noting that $(1+\mathsf{SNR})\,\dot{C}(\mathsf{SNR})/C(\mathsf{SNR})$ is very well approximated by $1/\log_e(1+\mathsf{SNR})$. Thus, the optimal overhead decreases

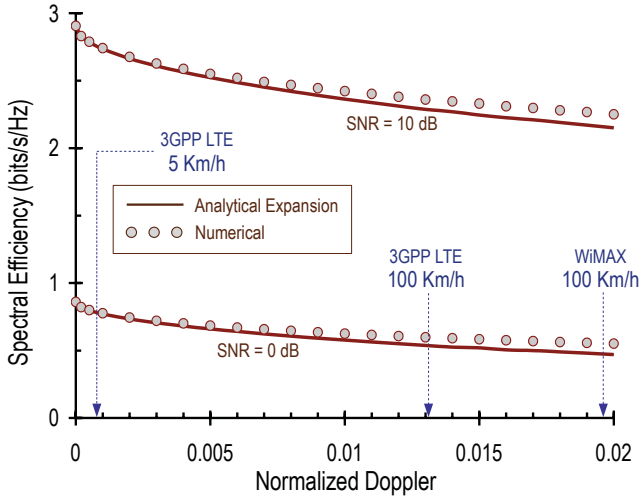

Fig. 4. Spectral efficiency with optimum pilot overhead as function of f_{D} for $\mathsf{SNR} = 10$ dB with a Clarke-Jakes spectrum. Relevant normalized Doppler levels for LTE and WiMAX subcarriers are highlighted.

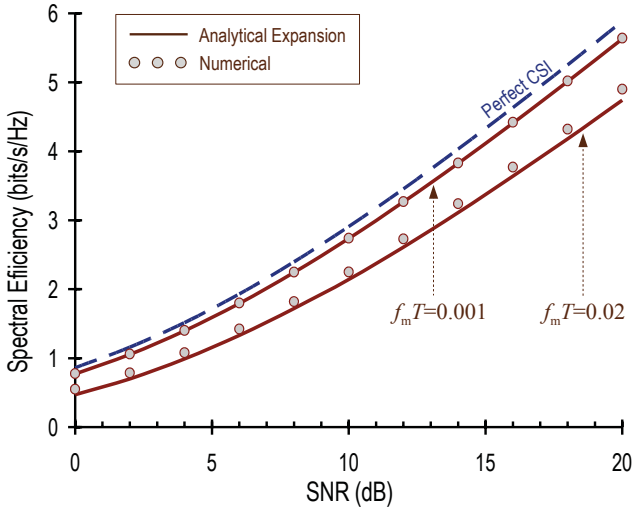

Fig. 5. Spectral efficiency with optimum pilot overhead as function of SNR for $f_{\mathrm{D}} = 0.001$ and $f_{\mathrm{D}} = 0.02$ with a Clarke-Jakes spectrum. Also shown is the capacity with perfect CSI.

with SNR approximately as $1/\sqrt{\log_e(1+\mathsf{SNR})}$. However, it is important to realize that, although the expansion in Proposition 2 is remarkably accurate for a wide range of SNR values, it becomes less accurate for $\mathsf{SNR} \to 0$ or $\mathsf{SNR} \to \infty$. In fact, in limiting SNR regimes it is possible to explicitly handle arbitrary Doppler levels [6], [7], [13], [14]. Thus, it is precisely for intermediate SNR values where the analysis here is both most accurate and most useful, thereby complementing those in the aforegiven references.

V. Pilot Power Boosting

In some systems, it is possible to allocate unequal powers for pilot and data symbols. In our models, this can be accommodated by defining the signal-to-noise ratios for pilot and data symbols to be $\rho_{\mathsf{p}}\mathsf{SNR}$ and $\rho_{\mathsf{d}}\mathsf{SNR}$, respectively, with

$$\rho_{\mathsf{p}}\alpha + \rho_{\mathsf{d}}(1-\alpha) = 1 \tag{30}$$

so that the average transmitted power is preserved. The spectral efficiency in (17) continues to hold, only with

$$\mathsf{SNR}_{\mathsf{eff}} = \frac{\mathsf{SNR}\,(1-\mathsf{MMSE})}{1/\rho_{\mathsf{d}} + \mathsf{SNR}\cdot\mathsf{MMSE}}. \tag{31}$$

The expressions for MMSE in (20) and (21) hold with SNR replaced with $\rho_{\mathsf{p}}\mathsf{SNR}$. As a result, with block fading,

$$\mathsf{MMSE} = \frac{1}{1+\alpha\, n_{\mathsf{b}}\, \rho_{\mathsf{p}}\mathsf{SNR}} \tag{32}$$

while, with continuous fading,

$$\mathsf{MMSE} = 1 - \int_{-f_{\mathrm{D}}}^{+f_{\mathrm{D}}} \frac{\mathsf{SNR}\, S_H^2(\nu)}{1/(\rho_{\mathsf{p}}\alpha) + \mathsf{SNR}\, S_H(\nu)}\, d\nu. \tag{33}$$

It is easily verified, from (32) and (33), that the identity between block fading and continuous fading with a rectangular Doppler spectrum continues to hold under condition (25).

Although we are generally interested in the maximization of spectral efficiency with respect to α and ρ_{p} (since ρ_{d} is specified by these variables according to (30)), in [12, Section V.A] and [13, Section III.B] it is shown that the spectral efficiency is maximized by using the minimum frequency of pilots, i.e.,

$$\alpha = \alpha_{\mathsf{min}} \tag{34}$$

$$= 2f_{\mathrm{D}}, \tag{35}$$

and then optimizing only with respect to the pilot power boost ρ_{p}. Because α is fixed, the power boosting that maximizes $\bar{\mathcal{I}}(\cdot)$ is directly the one that maximizes $\mathsf{SNR}_{\mathsf{eff}}$, i.e.,

$$\rho_{\mathsf{p}}^{\star} = \arg \max_{\rho_{\mathsf{p}}\alpha_{\mathsf{min}} + \rho_{\mathsf{d}}(1-\alpha_{\mathsf{min}})=1} \bar{\mathcal{I}}(\mathsf{SNR}, \alpha, \rho_{\mathsf{p}}) \tag{36}$$

$$= \arg \max_{\rho_{\mathsf{p}}\alpha_{\mathsf{min}} + \rho_{\mathsf{d}}(1-\alpha_{\mathsf{min}})=1} \mathsf{SNR}_{\mathsf{eff}}. \tag{37}$$

Although simpler than the optimization in Section IV, this nonetheless must be computed numerically, with the exception of the rectangular spectrum/block-fading [7], [12], [13].

As in Section III, we circumvent this limitation by expanding the problem in f_{D}. Again, this yields expressions that are explicit and valid for arbitrary spectral shapes.

Proposition 3 *The optimum power allocation for a Rayleigh-faded channel with an arbitrary bandlimited Doppler spectrum is given by*

$$\rho_{\mathsf{p}}^{\star} = \sqrt{\frac{1+1/\mathsf{SNR}}{2f_{\mathrm{D}}}} + \mathcal{O}(1) \tag{38}$$

$$\rho_{\mathsf{d}}^{\star} = 1 - \sqrt{\left(1+\frac{1}{\mathsf{SNR}}\right)2f_{\mathrm{D}}} + \mathcal{O}(f_{\mathrm{D}}) \tag{39}$$

and the corresponding spectral efficiency is

$$\bar{\mathcal{I}}^{\star}(\mathsf{SNR}, f_{\mathrm{D}}) = C\,(\mathsf{SNR}) - \sqrt{8f_{\mathrm{D}}\,\mathsf{SNR}\,(1+\mathsf{SNR})}\,\dot{C}(\mathsf{SNR}) + \mathcal{O}(f_{\mathrm{D}}). \tag{40}$$

Proof: *See Appendix IX-B.*

As expected, an order expansion of the closed-form solution for the rectangular spectrum [7, Theorem 2] matches the above proposition.

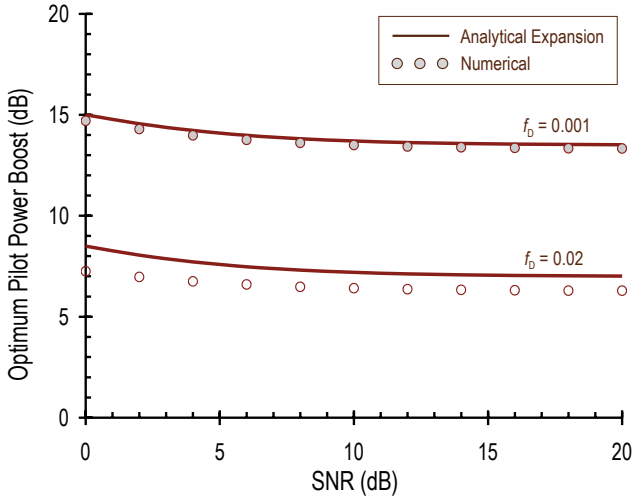

Fig. 6. Optimum pilot power boost, $\rho_{\mathsf{p}}^{\star}$, as function of SNR for $f_{\mathrm{D}} = 0.001$ and $f_{\mathrm{D}} = 0.02$ with a Clarke-Jakes spectrum.

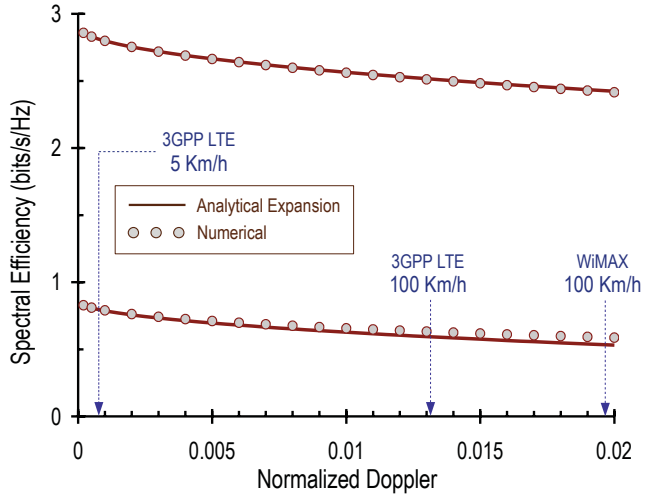

Fig. 7. Spectral efficiency with optimum pilot power boost as function of f_{D} for SNR $= 10$ dB with a Clarke-Jakes spectrum. Relevant normalized Doppler levels for LTE and WiMAX subcarriers are highlighted.

A comparison between the optimum pilot power boost given by (38) and the corresponding value obtained numerically is presented in Fig. 6. The agreement is excellent. Good agreement is further shown in Figs. 7–8 between the corresponding spectral efficiency in (40) and its exact counterpart, again obtained numerically.

While α is a direct measure of the pilot overhead in terms of bandwidth, the overhead in terms of power is measured by the product $\rho_{\mathsf{p}}\alpha$, which signifies the fraction of total transmit power devoted to pilot symbols. In light of (34) and Proposition 3, the optimum pilot power fraction when boosting is allowed equals

$$\rho_{\mathsf{p}}^{\star}\alpha = \sqrt{\left(1 + \frac{1}{\mathsf{SNR}}\right) 2f_{\mathrm{D}}} + \mathcal{O}(f_{\mathrm{D}}) \tag{41}$$

while without boosting (i.e., with $\rho_{\mathsf{p}} = 1$) the pilot power fraction is (from Proposition 2)

$$\alpha^{\star} = \sqrt{(1 + \mathsf{SNR})\,\frac{\dot{C}(\mathsf{SNR})}{C(\mathsf{SNR})}\,2f_{\mathrm{D}}} + \mathcal{O}(f_{\mathrm{D}}). \tag{42}$$

In both cases the fraction of pilot power fraction is $\mathcal{O}(\sqrt{f_{\mathrm{D}}})$. Comparing the two, the pilot power fraction with boosting is larger than the fraction without boosting by a factor

$$\sqrt{\frac{C(\mathsf{SNR})}{\mathsf{SNR}\ \dot{C}(\mathsf{SNR})}}. \tag{43}$$

This quantity is greater than unity and is increasing in SNR. Since MMSE is a decreasing function of $\rho_{\mathsf{p}}\alpha$, this implies that an optimized system with power boosting achieves a smaller MMSE than one without boosting.

Comparing (40) and (27), pilot power boosting increases the spectral efficiency by

$$\sqrt{8f_{\mathrm{D}}(1 + \mathsf{SNR})\dot{C}(\mathsf{SNR})}\left(\sqrt{C(\mathsf{SNR})} - \sqrt{\mathsf{SNR}\,\dot{C}(\mathsf{SNR})}\right) + \mathcal{O}(f_{\mathrm{D}}) \tag{44}$$

which increases monotonically with SNR and vanishes for $\mathsf{SNR} \to 0$.

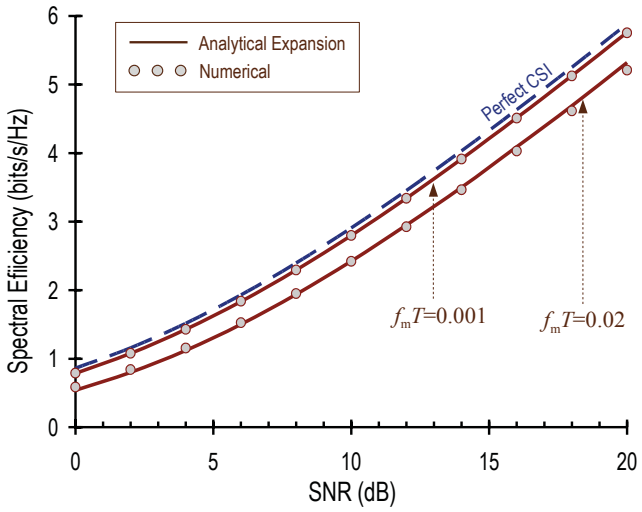

Fig. 8. Spectral efficiency with optimum pilot power boost as function of SNR for $f_{\mathrm{D}} = 0.001$ and $f_{\mathrm{D}} = 0.02$ with a Clarke-Jakes spectrum. Also shown is the capacity with perfect CSI.

VI. Multiantenna Channels

The analysis extends to multiantenna settings in a straightforward manner when there is no antenna correlation. Letting n_{T} and n_{R} denote the number of transmit and receive antennas, respectively, the channel at time k is now denoted by the $n_{\mathrm{R}} \times n_{\mathrm{T}}$ matrix $\mathbf{H}(k)$. Each of the $n_{\mathrm{T}}n_{\mathrm{R}}$ entries of the matrix varies in time according to the random processes described in Section II, for either block or continuous fading; furthermore, these processes are independent. Due to such independence, the matrix entries can be separately estimated without loss of optimality. Furthermore, the estimation error is minimized by transmitting orthogonal pilot sequences from the various transmit antennas [7], e.g., transmitting a pilot symbol from a single antenna at a time.

We denote the perfect-CSI multiantenna capacity as

$$C_{n_{\mathrm{T}},n_{\mathrm{R}}}(\mathsf{SNR}) = \mathbb{E}\left[\log_2 \det\left(\mathbf{I} + \frac{\mathsf{SNR}}{n_{\mathrm{T}}}\boldsymbol{H}\boldsymbol{H}^{\dagger}\right)\right], \qquad (45)$$

for which a closed-form expression in terms of the exponential integral can be found in [22].

The spectral efficiency with pilot-assisted detection now becomes [7]

$$\bar{\mathcal{I}}(\mathsf{SNR},\alpha) = (1-\alpha)\, C_{n_{\mathrm{T}},n_{\mathrm{R}}}(\mathsf{SNR}_{\mathsf{eff}}) \qquad (46)$$

with

$$\mathsf{SNR}_{\mathsf{eff}} = \frac{\mathsf{SNR}\,(1-\mathsf{MMSE})}{1+\mathsf{SNR}\cdot\mathsf{MMSE}} \qquad (47)$$

where MMSE is the estimation error for *each* entry of the channel matrix $\boldsymbol{H}$.

Because the definition of $\mathsf{SNR}_{\mathsf{eff}}$ carries over, the multiantenna formulation mirrors its single-antenna counterpart with $C_{n_{\mathrm{T}},n_{\mathrm{R}}}(\cdot)$ replacing $C(\cdot)$ and noting that MMSE now indicates the per-entry estimation error. The equivalence between block and continuous fading holds for each entry in terms of MMSE, and thus the equivalence in (25) extends to this multiantenna setting. Therefore, we again restrict our discussion to continuous fading.

Due to the orthogonal pilots, an overhead of α corresponds to a fraction α/n_{T} of symbols serving as pilots for a particular transmit antenna (i.e., for the n_{R} matrix entries associated with that transmit antenna). As a result, the per-entry MMSE is the same as the single-antenna expression in (22) only with α replaced by α/n_{T}, i.e.,

$$\mathsf{MMSE} = 1 - \int_{-1}^{+1} \frac{\tilde{S}_H^2(\xi)}{\frac{n_{\mathrm{T}} f_{\mathrm{D}}}{\alpha\,\mathsf{SNR}} + \tilde{S}_H(\xi)}\, d\xi. \qquad (48)$$

This equals the MMSE for a single-antenna channel with a Doppler frequency of $n_{\mathrm{T}} f_{\mathrm{D}}$. The optimization w.r.t. α in a multiantenna channel is thus the same as in a single-antenna channel, only with an effective Doppler frequency of $n_{\mathrm{T}} f_{\mathrm{D}}$ and with $C(\cdot)$ replaced by $C_{n_{\mathrm{T}},n_{\mathrm{R}}}(\cdot)$. As a result, Proposition 2 naturally extends into

$$\alpha^{\star} = \sqrt{(1+\mathsf{SNR})\,\frac{\dot{C}_{n_{\mathrm{T}},n_{\mathrm{R}}}(\mathsf{SNR})}{C_{n_{\mathrm{T}},n_{\mathrm{R}}}(\mathsf{SNR})}\,2n_{\mathrm{T}}f_{\mathrm{D}}} - \left((1+\mathsf{SNR})\,\frac{\ddot{C}_{n_{\mathrm{T}},n_{\mathrm{R}}}(\mathsf{SNR})}{\dot{C}_{n_{\mathrm{T}},n_{\mathrm{R}}}(\mathsf{SNR})} + 2 + \frac{1}{2\,\mathsf{SNR}}\int_{-1}^{+1}\frac{d\xi}{\tilde{S}_H(\xi)}\right) n_{\mathrm{T}} f_{\mathrm{D}} + \mathcal{O}(f_{\mathrm{D}}^{3/2}). \qquad (49)$$

Notice here the dependence on $\sqrt{n_{\mathrm{T}}}$ in the leading term.

When pilot power boosting is allowed, it is again advantageous to reduce α to its minimum value, now given by $\alpha_{\mathsf{min}} = 2n_{\mathrm{T}}f_{\mathrm{D}}$, and to increase ρ_{p}.[8] In this case the achievable spectral efficiency becomes [7]

$$(1-2n_{\mathrm{T}}f_{\mathrm{D}})\, C_{n_{\mathrm{T}},n_{\mathrm{R}}}(\mathsf{SNR}_{\mathsf{eff}}) \qquad (50)$$

[8]The proofs of the optimality of $\alpha = \alpha_{\mathsf{min}}$ in [12], [13] are derived for single-antenna channels. The extension to the multiantenna setting follows if one notes that the function $C_{n_{\mathrm{T}},n_{\mathrm{R}}}(\cdot)$ is concave and thus $C_{n_{\mathrm{T}},n_{\mathrm{R}}}(x) \geq x\,\dot{C}_{n_{\mathrm{T}},n_{\mathrm{R}}}(x)$ for any $x \geq 0$.

with $\mathsf{SNR}_{\mathsf{eff}}$ as defined in (18) and with

$$\mathsf{MMSE} = 1 - \int_{-1}^{+1} \frac{\tilde{S}_H^2(\xi)}{\frac{n_{\mathrm{T}} f_{\mathrm{D}}}{\alpha\,\rho_{\mathsf{p}}\,\mathsf{SNR}} + \tilde{S}_H(\xi)}\, d\xi. \qquad (51)$$

The optimization of the power boost again corresponds to the maximization of $\mathsf{SNR}_{\mathsf{eff}}$ with respect to ρ_{p}. Since MMSE is the same as for a single-antenna channel with effective Doppler $n_{\mathrm{T}} f_{\mathrm{D}}$, the optimum pilot power boost for a multiantenna channel with Doppler frequency f_{D} is exactly the same as the optimum pilot power boost for a single-antenna channel with the same spectral shape and with Doppler frequency $n_{\mathrm{T}} f_{\mathrm{D}}$. As a result, the expressions in Section V apply verbatim if f_{D} is replaced by $n_{\mathrm{T}} f_{\mathrm{D}}$.

Applying (41), the fraction of power devoted to pilots is given by

$$\rho_{\mathsf{p}}^{\star}\alpha = \sqrt{\left(1+\frac{1}{\mathsf{SNR}}\right)2n_{\mathrm{T}}f_{\mathrm{D}}} + \mathcal{O}(f_{\mathrm{D}}) \qquad (52)$$

which increases with $\sqrt{n_{\mathrm{T}}}$.

Based upon these results, when power-boosting is allowed the pilot overhead optimization on a multiantenna channel with Doppler frequency f_{D} and a particular spectral shape is *exactly* equivalent to the optimization on a single-antenna channel with the same spectral shape and with Doppler frequency $n_{\mathrm{T}} f_{\mathrm{D}}$. When pilot power boosting is not allowed, this equivalence is *approximate* because the perfect-CSI capacity functions $C(\cdot)$ and $C_{n_{\mathrm{T}},n_{\mathrm{R}}}(\cdot)$ differ. Roughly speaking, multiple antennas increase the perfect-CSI capacity by a factor of $\min(n_{\mathrm{T}},n_{\mathrm{R}})$ and thus $C_{n_{\mathrm{T}},n_{\mathrm{R}}}(\mathsf{SNR}) \approx \min(n_{\mathrm{T}},n_{\mathrm{R}})C(\mathsf{SNR})$. If this approximation were exact, then the aforementioned equivalence would also be exact. Although not exact, the perfect-CSI capacities are sufficiently similar, particularly for symmetric ($n_{\mathrm{T}} = n_{\mathrm{R}}$) channels, to render the equivalence very accurate also for the case of non-boosted pilots. To illustrate this accuracy, the optimal pilot overhead for a symmetric channel at an SNR of 10 dB is plotted versus the number of antennas along with the optimal overhead for the single-antenna equivalent (with Doppler $n_{\mathrm{T}} f_{\mathrm{D}}$) in Fig. 9. Excellent agreement is seen between the two.

The main implication of the single/multiantenna equivalence is that, based upon our earlier results quantifying the dependence of the pilot overhead on the Doppler frequency, the optimal overhead (with or without power boosting) scales with the number of antennas proportional to $\sqrt{n_{\mathrm{T}}}$.

VII. Extension to Discrete Signal Constellations

The optimization of pilot overheads with discrete constellations, rather than ideal Gaussian signals, would be a natural extension of the work in this paper. For m-PSK signals, in particular, this extension appears feasible. As recognized in [23] in a related content (computation of the low-SNR spectral efficiency without pilot symbols), the constant-amplitude property of m-PSK signals ensures that the additional noise caused by imperfect channel estimation in (16) remains both Gaussian and uncorrelated with the useful signal. Hence, the optimization problem described by (17) and (18) remains

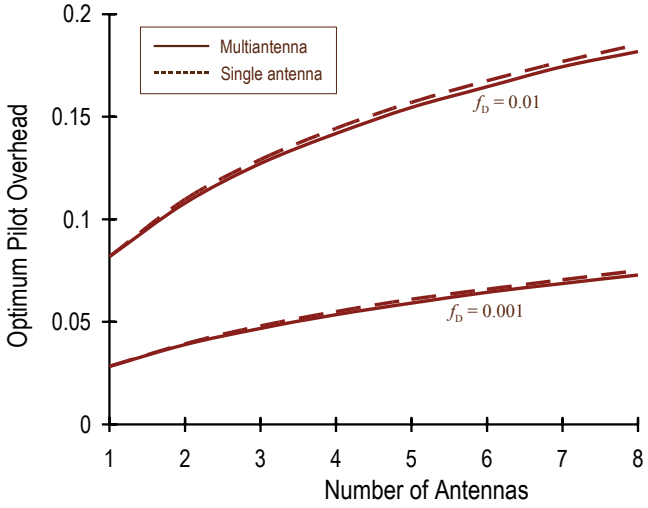

Fig. 9. Optimum pilot overhead, $\alpha^\star$, as function of number of antennas ($n_{\mathrm{T}} = n_{\mathrm{R}}$) for $f_{\mathrm{D}} = 0.001$ and $f_{\mathrm{D}} = 0.01$ for a rectangular spectrum with $\mathsf{SNR} = 10$ dB. Also shown is the optimal pilot overhead for the single-antenna equivalent with a normalized Doppler of $n_{\mathrm{T}} f_{\mathrm{D}}$.

valid, only with $C(\mathsf{SNR})$ replaced by the ergodic mutual information associated with an m-PSK signal, namely

$$C_{\mathsf{PSK}}(\mathsf{SNR}) = -\log_2(\pi e) - \mathbb{E}\left[\int f_m\left(y, \mathsf{SNR}|H|^2\right) \log_2 f_m\left(y, \mathsf{SNR}|H|^2\right)\, dy\right] \tag{53}$$

where the expectation is over the exponential distribution of $|H|^2$ and the integration is over the complex plane, with

$$f_m(y,\rho) = \frac{1}{m\pi}\sum_{\ell=1}^{m} \exp\left\{-\left|y - \sqrt{\rho}\, e^{j2\pi\ell/m}\right|^2\right\}. \tag{54}$$

Note that the computation of MMSE also remains as described in Section III because the pilots are not affected by the signal format.

Given the modified optimization obtained when the ergodic mutual information is given by $C_{\mathsf{PSK}}(\cdot)$, Propositions 2 and 3 continue to apply except with $C(\cdot)$, $\dot{C}(\cdot)$ and $\ddot{C}(\cdot)$ suitably replaced by $C_{\mathsf{PSK}}(\cdot)$ and its derivatives, respectively. Clearly, (53) cannot be integrated into a closed form, but it can be precomputed numerically and tabulated for subsequent use. The derivative of (53) can be expressed by virtue of the relationship between mutual information and nonlinear estimation [24] as in (55) at the bottom of the page, where the expectation is again over the exponential distribution of $|H|^2$ and the integration is over the complex plane. As with $C_{\mathsf{PSK}}(\cdot)$, $\dot{C}_{\mathsf{PSK}}(\cdot)$ can be precomputed numerically and tabulated.

The functions $C_{\mathsf{PSK}}(\cdot)$ and $\dot{C}_{\mathsf{PSK}}(\cdot)$ suffice for Proposition 3 to be applicable to m-PSK signals. Application of Proposition 2 further requires $\ddot{C}_{\mathsf{PSK}}(\cdot)$, which could be obtained by specializing the derivations in [25].

Discrete constellations other than m-PSK are less amenable, chiefly because the additional noise caused by imperfect channel estimation in (16) is no longer Gaussian if $X(k)$ has varying amplitude and a nearest neighbor decoder will not perform as if this noise were Gaussian. Moreover, replacing it with Gaussian noise of the same variance need not yield a lower bound on the spectral efficiency as the Gaussian noise distribution is generally not the worst when the signal is discrete. Nonetheless, an approximation would be obtained by rendering the noise Gaussian, under which Propositions 2 and 3 would hold with suitable functions for the ergodic mutual information of the corresponding constellation, along with its derivatives.

VIII. Summary

This paper has investigated the problem of pilot overhead optimization in single-user wireless channels. In the context of earlier work, our primary contributions are two-fold.

First, we were able to unify prior work on continuous- and block-fading channels and on single- and multiantenna channels: the commonly used block-fading model was shown to be a special case of the richer set of continuous-fading models in terms of the achievable pilot-based spectral efficiency, and the pilot overhead optimization for multiantenna chanels is seen to essentially be equivalent to the same optimization for a single-antenna channel in which the normalized Doppler frequency is multiplied by the number of transmit antennas.

Second, by finding an expansion for the overhead optimization in terms of the fading rate, the square root dependence of both the overhead and the spectral efficiency penalty was cleanly identified.

IX. Acknowledgments

The authors gratefully acknowledge comments received from the editor and anonymous reviewers, as well as from Prof. Albert Guillen from the University of Cambridge.

Appendix

A. Proof of Proposition 2

We set out to expand $\bar{\mathcal{I}}(\mathsf{SNR}, \alpha)$ w.r.t. f_{D} about the point $f_{\mathrm{D}} = 0$ while holding SNR and α fixed. We need

$$\frac{\partial \bar{\mathcal{I}}(\mathsf{SNR}, \alpha, f_{\mathrm{D}})}{\partial f_{\mathrm{D}}}|_{f_{\mathrm{D}}=0} = (1-\alpha)\,\dot{C}(\mathsf{SNR})\frac{\partial \mathsf{SNR}_{\mathsf{eff}}}{\partial f_{\mathrm{D}}}|_{f_{\mathrm{D}}=0} \tag{56}$$

$$= -(1-\alpha)\,\mathsf{SNR}\,(1+\mathsf{SNR})\,\dot{C}(\mathsf{SNR}) \times \frac{\partial \mathsf{MMSE}}{\partial f_{\mathrm{D}}}|_{f_{\mathrm{D}}=0} \tag{57}$$

$$\dot{C}_{\mathsf{PSK}}(\mathsf{SNR}) = \mathbb{E}\left[|H|^2\left(1 - \frac{1}{m^2\pi^2}\int \frac{\left|\sum_{\ell=1}^{m} e^{j2\pi\ell/m} \exp\{-|y - \sqrt{\mathsf{SNR}|H|^2}\, e^{j2\pi\ell/m}|^2\}\right|^2}{f_m(y, \mathsf{SNR}|H|^2)} dy\right)\right] \tag{55}$$

and

$$\frac{\partial^2 \bar{\mathcal{I}}(\mathsf{SNR},\alpha)}{\partial f_{\mathrm{D}}^2}|_{f_{\mathrm{D}}=0} = (1-\alpha)\left[\dot{C}(\mathsf{SNR})\,\frac{\partial^2 \mathsf{SNR}_{\mathsf{eff}}}{\partial f_{\mathrm{D}}^2} + \ddot{C}(\mathsf{SNR})\left(\frac{\partial \mathsf{SNR}_{\mathsf{eff}}}{\partial f_{\mathrm{D}}}\right)^2\right]|_{f_{\mathrm{D}}=0} \quad (58)$$

$$= -(1-\alpha)\left[\dot{C}(\mathsf{SNR})\,\mathsf{SNR}\,(1+\mathsf{SNR})\right. \quad (59)$$
$$\times\left(\frac{\partial^2 \mathsf{MMSE}}{\partial f_{\mathrm{D}}^2} - 2\,\mathsf{SNR}\left(\frac{\partial \mathsf{MMSE}}{\partial f_{\mathrm{D}}}\right)^2\right)$$
$$\left.+\ddot{C}(\mathsf{SNR})\,\mathsf{SNR}^2(1+\mathsf{SNR})^2\left(\frac{\partial \mathsf{MMSE}}{\partial f_{\mathrm{D}}}\right)^2\right]|_{f_{\mathrm{D}}=0}.$$

Based upon (22), regardless of the shape of the Doppler spectrum we have

$$\frac{\partial \mathsf{MMSE}}{\partial f_{\mathrm{D}}}|_{f_{\mathrm{D}}=0} = \frac{2}{\alpha\,\mathsf{SNR}} \quad (60)$$

where we have used the fact that $\tilde{S}_H(\cdot)$ is bandlimited to ± 1 and strictly positive within. In turn,

$$\frac{\partial^2 \mathsf{MMSE}}{\partial f_{\mathrm{D}}^2}|_{f_{\mathrm{D}}=0} = -\frac{2}{(\alpha\,\mathsf{SNR})^2}\int_{-1}^{+1}\frac{1}{\tilde{S}_H(\nu)}d\nu. \quad (61)$$

Combining (57), (59), (60) and (61),

$$\bar{\mathcal{I}}(\mathsf{SNR},\alpha,f_{\mathrm{D}}) = (1-\alpha)(1+\mathsf{SNR})\left[\frac{C(\mathsf{SNR})}{1+\mathsf{SNR}} - \dot{C}(\mathsf{SNR})\frac{2f_{\mathrm{D}}}{\alpha}\right.$$
$$\left.+\left(2\,(1+\mathsf{SNR})\,\ddot{C}(\mathsf{SNR})+\dot{C}(\mathsf{SNR})\left(\frac{1}{\mathsf{SNR}}\int_{-1}^{+1}\frac{d\xi}{\tilde{S}_H(\xi)}+4\right)\right)\frac{f_{\mathrm{D}}^2}{\alpha^2}\right] + \mathcal{O}(f_{\mathrm{D}}^3). \quad (62)$$

As justified below, the extreme points $\alpha = 2f_{\mathrm{D}}$ and $\alpha = 1$ can be ruled out as optimizers of (62). Thus, (62) is maximized by equating its derivative w.r.t. α to zero. This yields a cubic equation in α, whose relevant solution expands as

$$\alpha^\star = \sqrt{2f_{\mathrm{D}}(1+\mathsf{SNR})\frac{\dot{C}(\mathsf{SNR})}{C(\mathsf{SNR})}} - \left((1+\mathsf{SNR})\frac{\ddot{C}(\mathsf{SNR})}{\dot{C}(\mathsf{SNR})} + 2 + \frac{1}{2\,\mathsf{SNR}}\int_{-1}^{+1}\frac{d\xi}{\tilde{S}_H(\xi)}\right)f_{\mathrm{D}} + \mathcal{O}(f_{\mathrm{D}}^{3/2}). \quad (63)$$

The spectral efficiency in (27) is obtained by substituting (63) into (62).

We now justify ruling out the extreme points. The case $\alpha = 1$ is trivially suboptimal. In turn, if we plug $\alpha = 2f_{\mathrm{D}}$ into (62) the spectral efficiency expands as $C(\mathsf{SNR}) - \mathcal{O}(f_{\mathrm{D}})$, which is smaller (orderwise) than (27).

Since this is an order result, $\alpha^\star$ in (63) may be smaller than $\alpha_{\min}$ if f_{D} is not sufficiently small for the SNR being considered. In such a case, the value obtained should be adjusted to $\alpha_{\min}$.

B. Proof of Proposition 3

The derivation closely parallels that in Appendix IX-A. The spectral efficiency equals

$$\bar{\mathcal{I}}(\mathsf{SNR},f_{\mathrm{D}}) = (1-2f_{\mathrm{D}})\,C\,(\mathsf{SNR}_{\mathsf{eff}}) \quad (64)$$

where the dependence on ρ_{p} and ρ_{d} is concentrated on $\mathsf{SNR}_{\mathsf{eff}}$. To expand $\mathsf{SNR}_{\mathsf{eff}}$ w.r.t. f_{D}, we need

$$\frac{\partial \mathsf{SNR}_{\mathsf{eff}}}{\partial f_{\mathrm{D}}}|_{f_{\mathrm{D}}=0} = -\rho_{\mathsf{d}}\mathsf{SNR}\,(1+\rho_{\mathsf{d}}\mathsf{SNR})\,\frac{\partial \mathsf{MMSE}}{\partial f_{\mathrm{D}}}|_{f_{\mathrm{D}}=0}. \quad (65)$$

In order to compute $\partial \mathsf{MMSE}/\partial f_{\mathrm{D}}$, we invoke again the normalized spectral shape in (8) and further use (30) to rewrite (33) as

$$\mathsf{MMSE} = 1 - \int_{-1}^{+1}\frac{\tilde{S}_H^2(\xi)}{\frac{f_{\mathrm{D}}}{\mathsf{SNR}\,(1-\rho_{\mathsf{d}}(1-2f_{\mathrm{D}}))}+\tilde{S}_H(\xi)}\,d\xi. \quad (66)$$

Then,

$$\frac{\partial \mathsf{MMSE}}{\partial f_{\mathrm{D}}}|_{f_{\mathrm{D}}=0} = \frac{2}{\mathsf{SNR}\,(1-\rho_{\mathsf{d}})}. \quad (67)$$

Combining (65) and (67), and using the fact that, for $f_{\mathrm{D}} \to 0$, $\mathsf{SNR}_{\mathsf{eff}}$ approaches $\rho_{\mathsf{d}}\mathsf{SNR}$, we have

$$\mathsf{SNR}_{\mathsf{eff}} = \rho_{\mathsf{d}}\mathsf{SNR} - \rho_{\mathsf{d}}\frac{1+\rho_{\mathsf{d}}\mathsf{SNR}}{1-\rho_{\mathsf{d}}}\,2f_{\mathrm{D}} + \mathcal{O}(f_{\mathrm{D}}^2). \quad (68)$$

Equating the derivative of (68) w.r.t. ρ_{d} to zero yields a quadratic equation in terms of ρ_{d}, which has solutions above and below 1. A value of $\rho_{\mathsf{d}} > 1$ leads to worse performance than non-boosted pilots, and thus the optimizer is the solution below 1. Such solution expands as

$$\rho_{\mathsf{d}}^\star = 1 - \sqrt{2f_{\mathrm{D}}\,(1+1/\mathsf{SNR})} + \mathcal{O}(f_{\mathrm{D}}). \quad (69)$$

Analogously, combining (30) and (69), and with the constraint that $\rho_{\mathsf{p}} > 1$,

$$\rho_{\mathsf{p}}^\star = \sqrt{\frac{1+1/\mathsf{SNR}}{2f_{\mathrm{D}}}} + \mathcal{O}(1). \quad (70)$$

The spectral efficiency in (40) is obtained by substituting (69) into (68) and then into (17), and expanding w.r.t. f_{D}.

REFERENCES

[1] A. Lozano and N. Jindal, "Optimum pilot overhead in wireless communication: a unified treatment of continuous and bock-fading channels," in *Proc. European Wireless Conf.*, pp. 725-732, Apr. 2010.

[2] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 686-693, Nov. 1991.

[3] L. Tong, B. M. Sadler, and M. Dong, "Pilot-assisted wireless transmissions: general model, design criteria, and signal processing," *IEEE Signal Proc. Mag.*, vol. 21, no. 6, pp. 12-25, Nov. 2004.

[4] M. Dong, L. Tong, and B. M. Sadler, "Optimal insertion of pilot symbols for transmissions over time-varying flat fading channels," *IEEE Trans. Signal Process.*, vol. 52, no. 5, pp. 1403-1418, May 2004.

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

[6] L. Zheng and D. N. C. Tse, "Communication on the Grassman manifold: a geometric approach to the non-coherent multiple-antenna channel," *IEEE Trans. Inf. Theory*, vol. 48, no. 2, pp. 359-383, Feb. 2002.

[7] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951-963, Apr. 2003.

[8] X. Ma, L. Yang, and G. B. Giannakis, "Optimal training for MIMO frequency-selective fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 453-466, Mar. 2005.

[9] S. Furrer and D. Dahlhaus, "Multiple-antenna signaling over fading channels with estimated channel state information: capacity analysis," *IEEE Trans. Inf. Theory*, vol. 53, no. 6, pp. 2028-2043, June 2007.

[10] L. Zheng, D. N. C. Tse, and M. Medard, "Channel coherence in the low-SNR regime," *IEEE Trans. Inf. Theory*, vol. 53, no. 3, pp. 976-997, Mar. 2007.

[11] J. Baltersee, G. Fock, and H. Meyr, "An information theoretic foundation of synchronized detection," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2115-2123, Dec. 2001.
[12] S. Ohno and G. B. Giannakis, "Average-rate optimal PSAM transmissions over time-selective fading channels," *IEEE Trans. Wireless Commun.*, vol. 1, no. 4, pp. 712-720, Oct. 2002.
[13] X. Deng and A. M. Haimovich, "Achievable rates over time-varying Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 55, no. 7, pp. 1397-1406, July 2007.
[14] A. Lozano, "Interplay of spectral efficiency, power and Doppler spectrum for reference-signal-assisted wireless communication," *IEEE Trans. Commun.*, vol. 56, no. 12, Dec. 2008.
[15] M. Kobayashi, N. Jindal, and G. Caire, "How much training and feedback is optimal for MIMO broadcast channels?" in *Proc. Int'l Symp. Inf. Theory (ISIT'08)*, July 2008.
[16] J. Doob, *Stochastic Processes*. New York: Wiley, 1990.
[17] W. C. Jakes, *Microwave Mobile Communications*. New York: IEEE Press, 1974.
[18] W. C. Y. Lee, "Estimate of channel capacity in Rayleigh fading environments," *IEEE Trans. Veh. Technol.*, vol. 39, pp. 187-189, Aug. 1990.
[19] L. Ozarow, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Veh. Technol.*, vol. 43, pp. 359-378, May 1994.
[20] I. Czisar and J. Koerner, *Information Theory*. Budapest: Akademia Kiado, 1981.
[21] A. Lapidoth and S. Shamai, "Fading channels: how perfect need 'perfect side information' be?" *IEEE Trans. Inf. Theory*, vol. 48, no. 5, pp. 1118-1134, May 2002.
[22] H. Shin and J. H. Lee, "Capacity of multiple-antenna fading channels: spatial fading correlation, double scattering and keyhole," *IEEE Trans. Inf. Theory*, vol. 49, pp. 2636-2647, Oct. 2003.
[23] W. Zhang and J. N. Laneman, "How good is PSK for peak-limited fading channels in the low-SNR regime?" *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 236-251, Jan. 2007.
[24] A. Lozano, A. M. Tulino, and S. Verdú, "Power allocation over parallel Gaussian channels with arbitrary input distributions," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3033-3051, July 2006.
[25] M. Payaro and D. P. Palomar, "Hessian and concavity of mutual information, differential entropy, and entropy power in linear vector Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3613-3628, Aug. 2009.

Nihar Jindal (S'99-M'04) received the B.S. degree in electrical engineering and computer science from the University of California at Berkeley in 1999 and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 2001 and 2004, respectively. He is an Assistant Professor at the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis. His industry experience includes internships at Intel Corporation, Santa Clara, CA, in 2000 and at Lucent Bell Labs, Holmdel, NJ, in 2002. His research spans the fields of information theory and wireless communication, with specific interests in multiple-antenna/multiuser channels, dynamic resource allocation, and sensor and ad hoc networks. Dr. Jindal currently serves as an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, and was a Guest Editor for a special issue of the *EURASIP Journal on Wireless Communications and Networking* on the topic of multiuser communication. He was the recipient of the 2005 IEEE Communications Society and Information Theory Society Joint Paper Award, the University of Minnesota McKnight Land-Grant Professorship Award in 2007, the NSF CAREER award in 2008, and the best paper award for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS in 2009.

Angel Lozano is a Professor of Information and Communication Technologies at UPF (Universitat Pompeu Fabra) in Barcelona, Spain, where he teaches and conducts research on wireless communications as head of the WiCom (Wireless Communications) Group.

Born in Manresa, Spain, Angel Lozano received the Telecommunications Engineering degree from UPC (Universitat Politecnica de Catalunya), Spain, in 1992 and Master of Science and Ph.D. degrees in Electrical Engineering from Stanford University in 1994 and 1998, respectively. Contemporarily, between 1996 and 1998, he also worked for Rockwell Communication Systems (now Conexant Systems) in San Diego, USA.

In 1999 he joined Bell Labs (Lucent Technologies, now Alcatel-Lucent) in Holmdel, USA, where he was a member of the Wireless Communications Research Department until 2008. Between 2005 and 2008 he was also an Adjunct Associate Professor of Electrical Engineering at Columbia University.

Dr. Lozano is a senior member of the IEEE since 1999. He served as associate editor for the IEEE TRANSACTIONS ON COMMUNICATIONS between 1999 and 2009, has guest-edited various other IEEE and non-IEEE journal special issues, and is actively involved in committees and conference organization tasks for the IEEE Communications Society. Since 2010, he is an associate editor for the *Journal of Communications & Networks*. He has further participated in standardization activities for 3GPP, 3GPP2, IEEE 802.20 and the IETF.

Dr. Lozano has authored over 80 technical journal and conference papers, holds 15 patents, and has contributed to several books. His papers have received two awards: the best paper at the 2006 IEEE Int'l Symposium on Spread Spectrum Techniques & Applications, and the Stephen O. Rice prize to the best paper published in the IEEE TRANSACTIONS ON COMMUNICATIONS in 2008.