

# Interplay of Spectral Efficiency, Power and Doppler Spectrum for Reference-Signal-Assisted Wireless Communication

Angel Lozano\*

June 26, 2008

## Abstract

Expressions relating spectral efficiency, power, and Doppler spectrum, are derived for Rayleigh-faded wireless channels with Gaussian signal transmission. No side information on the state of the channel is assumed at the receiver. Rather, periodic reference signals are postulated in accordance with the functioning of most wireless systems.

The analysis relies on a well-established lower bound, generally tight and asymptotically exact at low SNR. In contrast with most previous studies, which relied on block-fading channel models, a continuous-fading model is adopted. This embeds the Doppler spectrum directly in the derived expressions, imbuing them with practical significance.

Closed-form relationships are obtained for the popular Clarke-Jakes spectrum and informative expansions, valid for arbitrary spectra, are found for the low- and high-power regimes. While the paper focuses on scalar channels, the extension to multi-antenna settings is also discussed.

**Keywords:** Rayleigh Fading; Spectral Efficiency; Mutual Information; Channel Estimation; Doppler Spectrum;

---

\*Angel Lozano is with Bell Laboratories (Alcatel-Lucent), Holmdel, NJ07733, USA.

# I Introduction

Consider a Rayleigh-faded wireless channel. If we posit that a genie provides the receiver with perfect CSI (channel state information), then a Gaussian signal distribution achieves capacity and, furthermore, such capacity can be conveniently computed.

Consider now the same channel and the same signal distribution, but remove the genie and periodically insert RS (reference signals, a.k.a. pilots or training signals), to assist in the coherent detection of the data [1, 2]. This typical setting reflects the operational conditions of many wireless systems, yet its achievable spectral efficiency is not easy to evaluate (cf. Appendix A). If the channel varies slowly, then it is reasonable to expect such efficiency to be close to the perfect-CSI capacity. If the channel varies rapidly, on the other hand, the spectral efficiency might be much lower. There is thus an underlying dependence between achievable spectral efficiency and velocity (or, more generally, Doppler spectrum) and, as it turns out, this dependence is further a function of the SNR (signal-to-noise ratio). To this assortment of quantities, we may add the RS overhead. The objective of this paper is to establish relationships between these various quantities to answer, within the context of the above setting, questions that naturally arise:

- How much RS overhead is required to achieve a certain spectral efficiency, as a function of the SNR and Doppler spectrum?
- How does the spectral efficiency degrade with decreasing SNR and increasing Doppler? How much additional power is required to offset a certain increase in Doppler?
- How advantageous is it to power-boost the RS?

Treatments of the RS-assisted spectral efficiency include [3]–[8], all of which rely on the block-fading channel model (except for [3], which employs a Gauss-Markov model). Simplicity is the main attribute of the block-fading model and, at the same time, its main shortcoming. In particular, the notion of a Doppler spectrum is absent thereby precluding precise relationships involving the velocity, the symbol period, and the wavelength. Surpassing these limitations, a continuous-fading model and the associated Doppler spectrum is

adopted in [9]–[11]. This leads to relationships that encompass the aforementioned variables. Extending the excellent analysis in [9]–[11], this paper furnishes convenient closed-form expressions for the all-important Clarke-Jakes Doppler spectrum as well as new expansions, general but informative, for the low- and high-power regimes.

## II Channel Model

We shall consider a discrete-time model for a scalar frequency-flat fading channel. Let us indicate by  $\mathcal{D}$  the set of time indices corresponding to data symbols. For  $k \in \mathcal{D}$ ,

$$Y(k) = H(k)X(k) + N(k) \quad (1)$$

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

RS are inserted periodically in the transmission [12] and the fraction thereof is denoted by  $\alpha$ , i.e., one in every  $1/\alpha$  symbols is a reference while the rest are data (cf. Fig. 1). For  $k \notin \mathcal{D}$ ,

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

Notice that the same power is transmitted during both RS and data. In Section V, we shall lift this restriction allowing for power-boosted RS.

The fading process,  $H \sim \mathcal{CN}(0, 1)$ , is modeled as stationary and as having an absolutely continuous spectral distribution function whose derivative is the Doppler spectrum  $S_H(\nu)$ ,  $-1/2 \leq \nu \leq 1/2$ . It follows that  $\{H(k)\}$  is ergodic.

Let  $S(f)$  be the spectrum of the underlying continuous-time fading process. We consider bandlimited processes, i.e., such that  $S(f) = 0$  for  $|f| > f_m$ . Then, denoting by  $T$  the symbol period and by  $\Pi(\cdot)$  the Fourier transform of the transmission pulse shape,

$$S_H(\nu) = \frac{1}{T} S\left(\frac{\nu}{T}\right) \Pi^2(\nu) \quad (3)$$

which is confined to  $|\nu| \leq f_m T$ . It is thus necessary that

$$\alpha \geq 2f_m T \quad (4)$$

to ensure that the decimated channel observed through the RS has an unaliased spectrum [2, 10]. On account of its bandlimited nature,  $\{H(k)\}$  is a nonregular fading process [13].

It is worth remarking that  $S_H(\cdot)$  need not be strictly positive within  $\pm f_m T$ . It is strictly positive only if the power azimuth spectrum itself is strictly positive, i.e., if the environment scatters energy to/from the user on every azimuthal direction.

Note also that (3) implies a matched-filter front-end at the receiver. This entails no loss of optimality if  $f_m \ll 1/T$ , a premise usually satisfied. User motion, specifically, results in  $f_m = v/\lambda$  where  $v$  is the velocity and  $\lambda$  is the carrier wavelength. Such  $f_m$  is typically several orders of magnitude below the symbol rate  $1/T$  and the smooth pulse shaping  $\Pi(\cdot)$  can be thereby disregarded altogether.<sup>1</sup>

Strong emphasis shall be placed on the popular Clarke-Jakes spectrum [14]

$$S_H(\nu) = \frac{1}{\pi \sqrt{(f_m T)^2 - \nu^2}} \quad (5)$$

which has been shown to accurately describe the behavior of mobile wireless channels and which is strictly positive within  $\pm f_m T$ .

To provide examples, we shall invoke some typical situations:

- $f_m T = 0.001$ . For  $\lambda = 15$  cm and  $T \approx 100$   $\mu$ s, consistent with emerging wireless systems [15, 16], this situation is representative of pedestrian motion,  $v \approx 5$  Km/h.
- $f_m T = 0.02$ , corresponding to vehicular motion,  $v \approx 100$  Km/h.
- $f_m T = 0.05$  for high-speed trains with  $v \approx 250$  Km/h.

In terms of SNR, the range of interest to wireless communication starts a few dB below zero and extends to about 25–30 dB. Further below, synchronization mechanisms start to fail. Above, nonlinear behaviors set in.

---

<sup>1</sup>Higher values for  $f_m$  may result if the objects scattering the signal to/from the user are also in motion, or if multiple signal reflections between the user and the environment take place, but the condition  $f_m \ll 1/T$  is still very likely to be satisfied.

### III Perfect CSI

With perfect CSI at the receiver, there is no need for RS (i.e.,  $\alpha = 0$ ). The Gaussian signal  $\{X(k)\}$  is optimal and the capacity, in bits/s/Hz, is [17, 18]

$$C(\text{SNR}) = E [\log_2 (1 + \text{SNR} |H|^2)] \quad (6)$$

$$= \log_2(e) e^{1/\text{SNR}} E_1 \left( \frac{1}{\text{SNR}} \right) \quad (7)$$

with  $E_q(\cdot)$  the exponential integral of order  $q$ ,

$$E_q(\zeta) = \int_1^\infty t^{-q} e^{-\zeta t} dt. \quad (8)$$

Under the assumption of perfect CSI, the Doppler spectrum plays no role.

In the low-power regime,

$$C(\text{SNR}) = \log_2(e) [\text{SNR} - \text{SNR}^2] + o(\text{SNR}^2) \quad (9)$$

while, in the high-power regime [17],

$$C(\text{SNR}) = \log_2 \left( \frac{\text{SNR}}{e^\gamma} \right) + o(1) \quad (10)$$

where  $\gamma \approx 0.557$  is the Euler-Mascheroni constant.

### IV RS-Assisted Detection

Let us now remove the perfect CSI while inserting a reference every  $1/\alpha$  symbols. The Gaussian signal  $\{X(k)\}$ , which in general is no longer optimal, achieves a spectral efficiency

$$\lim_{K \rightarrow \infty} \frac{1}{K} I \left( \underbrace{\{X(k)\}_{k=0}^{K-1}; \{Y(k)\}_{k=0}^{K-1}}_{k \in \mathcal{D}} \mid \underbrace{\{Y(k)\}_{k=0}^{K-1}}_{k \notin \mathcal{D}} \right) \quad (11)$$

In the standard operational mode of most systems, only the RS observations  $\{Y(k)\}$ ,  $k \notin \mathcal{D}$ , assist in the detection of the data symbols. This reduces the spectral efficiency to

$$\bar{I}(\text{SNR}) = \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{\substack{k=0 \\ k \in \mathcal{D}}}^{K-1} I \left( X(k); Y(k) \mid \underbrace{\{Y(q)\}_{q=0}^{K-1}}_{q \notin \mathcal{D}} \right) \quad (12)$$

The RS observations may be used to form channel estimates  $\hat{H}(k), k \in \mathcal{D}$ . Provided the channel estimation function does not destroy information,

$$\underbrace{\{Y(q)\}_{q=0}^{K-1}}_{q \notin \mathcal{D}} \rightarrow \hat{H}(k) \rightarrow (X(k), Y(k)) \quad (13)$$

constitutes a Markov chain and (12) is equivalent to

$$\bar{\mathcal{I}}(\text{SNR}) = \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{\substack{k=0 \\ k \in \mathcal{D}}}^{K-1} I(X(k); Y(k) | \hat{H}(k)) \quad (14)$$

$$= (1 - \alpha) E \left[ I(X(k); Y(k) | \hat{H}(k)) \right] \quad (15)$$

where, in (15),  $k$  is the index of an arbitrary data symbol and the expectation is over the distribution of  $\hat{H}(k)$ .

With respect to (11), the suboptimality in (15) lies in that it precludes deriving information about  $\hat{H}(k)$  from the data symbols, i.e., it renders joint data detection and channel estimation beyond the scope of this paper. Note that simply using the channel estimates as if they represented the true channel values is in general a suboptimal approach [19, 20] and thus receivers pursuing such strategy may fall short of (15). More elaborate receiver structures that take into account the joint distribution of  $H(k)$  and  $\hat{H}(k)$  are in general necessary to attain (15).

As argued in [3]–[10] (see also [21] where a similar framework is presented for error probability analysis),  $\bar{\mathcal{I}}(\text{SNR})$  can be lower-bounded by postulating the use of a conditional-mean channel estimator

$$\hat{H}(k) = E[H(k) | \underbrace{\{Y(q)\}_{q=0}^{K-1}}_{q \notin \mathcal{D}}] \quad (16)$$

which minimizes the mean-square error  $E[|H(k) - \hat{H}(k)|^2]$ . For  $K \rightarrow \infty$ , this minimum mean-square error approaches [2, 10, 22]<sup>2</sup>

$$\text{MMSE} = 1 - \int_{-f_m T/\alpha}^{f_m T/\alpha} \frac{\text{SNR} (\alpha S_H(\alpha\nu))^2}{1 + \text{SNR} (\alpha S_H(\alpha\nu))} d\nu \quad (17)$$

---

<sup>2</sup>Although the spectral efficiency is defined asymptotically in  $K$ , in practice the blocklength is truncated to a value that results in an acceptable delay and thus the number of channel observations is curtailed. This has no significant effect on  $\hat{H}(k)$  as long as such blocklength exceeds the coherence time of the fading process.

which depends on the SNR and on the Doppler spectrum,  $S_H(\cdot)$ , spectrally expanded by  $\alpha$  to account for the RS periodicity. In turn,  $\hat{H} \sim \mathcal{N}_{\mathbb{C}}(0, 1 - \text{MMSE})$ . With a simple change of variables, (17) can be rearranged as

$$\text{MMSE} = 1 - \int_{-f_m T}^{f_m T} \frac{\text{SNR} S_H^2(\nu)}{1/\alpha + \text{SNR} S_H(\nu)} d\nu \quad (18)$$

which more directly evinces the increase in MMSE that results from a reduction in  $\alpha$ .

We can now rewrite (1) as

$$Y(k) = \hat{H}(k)X(k) + N'(k) \quad (19)$$

where  $N'(k) = (H(k) - \hat{H}(k))X(k) + N(k)$ . Thus,  $\{N'(k)\}$  is a zero-mean IID sequence but  $N'(k)$  is, in general, neither Gaussian nor independent of  $\hat{H}(k)X(k)$ . For the conditional-mean estimator, however,  $N'(k)$  is uncorrelated with  $\hat{H}(k)X(k)$ . Replacing  $\{N'(k)\}$  with a zero-mean IID complex Gaussian sequence of the same variance,  $N' \sim \mathcal{N}_{\mathbb{C}}(0, P_{\text{MMSE}} + N_0)$ , we obtain [3]–[10]

$$\bar{\mathcal{I}}(\text{SNR}) \geq (1 - \alpha) C(\text{SNR}_{\text{eff}}) \quad (20)$$

with

$$\text{SNR}_{\text{eff}} = \frac{\text{SNR} (1 - \text{MMSE})}{1 + \text{SNR} \text{MMSE}}. \quad (21)$$

The lower bound in (20), which is the starting point for our analysis, evidences two differences with respect to the capacity:

- A pre-log loss of  $(1 - \alpha)$  associated with the RS insertion.
- An effective signal-to-noise ratio,  $\text{SNR}_{\text{eff}} \leq \text{SNR}$ .

These differences reflect the tension between improving the quality of the channel estimates (by increasing  $\alpha$ ) and diminishing the RS overhead (by reducing  $\alpha$ ). This tradeoff is studied in [4, 5] using the block-fading channel model. Here it is addressed using the model described in Section II, with the Doppler spectrum centerstage.

Before proceeding, though, it is worth assessing the tightness of the bound in (20). A comparison is provided in Fig. 2, with the details on how to compute the exact  $\bar{\mathcal{I}}(\text{SNR})$  given in

Appendix A. The bound is seen to be very tight for this typical spectrum and asymptotically exact for  $\text{SNR} \rightarrow 0$  (cf. Section IV-A).

Our first result, obtained through a tedious integration of (18), is a closed-form expression for MMSE in Clarke-Jakes environments.

**Proposition 1** *For the Clarke-Jakes spectrum in (5),*

$$\text{MMSE} = 1 - \frac{\text{arctanh} \sqrt{1 - \left(\frac{\alpha \text{SNR}}{\pi f_m T}\right)^2}}{\frac{\pi}{2} \sqrt{\left(\frac{\pi f_m T}{\alpha \text{SNR}}\right)^2 - 1}} \quad (22)$$

In combination with (7), (20) and (21), the handy expression in Proposition 1 allows optimizing the RS overhead  $\alpha$  for any operational point. Precisely, the optimum RS overhead is the solution to the convex optimization

$$\alpha^* = \arg \max_{2f_m T \leq \alpha \leq 1} (1 - \alpha) C(\text{SNR}_{\text{eff}}) \quad (23)$$

which is readily solved and displayed in Fig. 3 for a wide range of SNR values, parameterized by  $f_m T$ . Plugging  $\alpha^*$  back into the corresponding expressions, the spectral efficiency lower bound for a Clarke-Jakes spectrum is as shown in Fig. 4. The optimization may be performed for other spectra as well, although possibly without the convenience of a closed-form expression for MMSE and  $\text{SNR}_{\text{eff}}$ . (The case of a uniform power spectrum is treated in [10].)

Using the block-fading channel model, it was shown in [4, 5] that:

- At low SNR,  $\text{SNR}_{\text{eff}} = \alpha b \text{SNR}^2 + o(\text{SNR}^2)$  where  $b$  is the number of symbols in a fading block.

It follows that

$$\bar{\mathcal{I}}(\text{SNR}) = (1 - \alpha) \alpha b \log_2(e) \text{SNR}^2 + o(\text{SNR}^2). \quad (24)$$

and, therefore, that  $\alpha^* = 1/2$ .

- At high SNR,

$$\bar{\mathcal{I}}(\text{SNR}) \geq (1 - \alpha) \log_2 \left( \frac{\text{SNR}}{e^\gamma} \frac{\alpha b}{1 + \alpha b} \right) + o(1) \quad (25)$$

from which  $\alpha^* = 1/b$  (single RS per fading block) and thus a 3-dB power penalty ensues.



In the remainder of this section, we reexamine the above findings using our continuous-fading model.

## A Low-Power Regime

With perfect CSI, any proper complex signal achieves (9) [23]. The class of proper complex signals includes, in particular,  $m$ -PSK for  $m > 2$ . This has welcome implications since, because of the constant amplitude of a PSK signal,  $N' \sim \mathcal{N}_{\mathbb{C}}(0, P_{\text{MMSE}} + N_0)$  in (19) [9]. Hence, the lower bound in (20) becomes an equality to second order.

A continuous-fading counterpart of (24) is then obtained through a simple series expansion (cf. Appendix B) as

$$\bar{\mathcal{I}}(\text{SNR}) = \alpha(1 - \alpha) \log_2(e) \left[ \int_{-f_m T}^{f_m T} S_H^2(\nu) d\nu \right] \text{SNR}^2 + o(\text{SNR}^2) \quad (26)$$

Drawing parallels between (24) and (26) we can establish that, at low SNR, the blocklength  $n$  in the block-fading model amounts to the integral of the squared Doppler spectrum.

Eq. (26), however, has two drawbacks:

- For important spectra, such as Clarke-Jakes, the integral does not exist. (This signifies that the spectral efficiency is superquadratic—but sublinear—in the SNR and, also, that the block-fading model is unable to represent a Clarke-Jakes spectrum at low SNR.)
- When the integral exists, (26) tends to be accurate only for very small SNR (cf. Fig. 5). This mirrors the shortcoming exhibited by (24), not enabling insights that apply within the SNR range of operational interest.

The following result resolves both drawbacks providing a more informative expression.

**Proposition 2** *In the low-power regime,*

$$\bar{\mathcal{I}}(\text{SNR}) = \eta_0 \log_2(e) [\text{SNR} - \text{SNR}^2 + o(\text{SNR}^2)] \quad (27)$$

with  $\eta_0 = (1 - \alpha)(1 - \text{MMSE})$ .

**Proof:** See Appendix C.

Proposition 2 is accurate over a much wider range of SNR than (26) (cf. Fig. 5). By virtue of Proposition 1, moreover,  $\eta_0$  in Proposition 2 is in closed form for a Clarke-Jakes spectrum and, in fact, further expansion of Propositions 1 and 2 leads to

$$\bar{\mathcal{I}}(\text{SNR}) = \frac{2\alpha(1-\alpha)}{\pi^2 f_m T} \log_2 \left( \frac{2\pi f_m T}{\alpha \text{SNR}} \right) \text{SNR}^2 + o(\text{SNR}^2) \quad (28)$$

revealing the precise superquadratic behavior that could not be ascertained from (26).

For  $\text{SNR} \rightarrow 0$ ,  $\alpha^* \rightarrow 1/2$  as established in [5]. This asymptote, however, must be interpreted with care as illustrated in Fig. 3, which reveals that  $\alpha^*$  remains far below  $1/2$  when the parameters are within the range of interest. A low-power approximation to the optimum RS overhead can be obtained by solving the simpler convex optimization

$$\alpha^* = \arg \max_{2f_m T \leq \alpha \leq 1} \eta_0 \quad (29)$$

in lieu of (23).

Note also, by contrasting Proposition 2 with (9), that  $\eta_0$  represents the share of the perfect-CSI capacity that can be achieved with RS-assisted detection. Furthermore,  $\eta_0$  depends on SNR and on  $f_m T$  only through their ratio indicating that, in order to maintain a given share  $\eta_0 < (1 - \alpha)$  of the perfect-CSI capacity, the Doppler need only decrease linearly with the SNR as the latter vanishes.

**Corollary 1** *A share  $\eta_0 < (1 - \alpha)$  of the low-power perfect-CSI capacity in (9) can be maintained if  $f_m T = \mathcal{O}(\text{SNR})$  for  $\text{SNR} \rightarrow 0$ .*

Interestingly then, the low-power spectral efficiency exhibits a symmetrical dependence on Doppler and SNR (i.e., a scaling in Doppler can be offset by a commensurate scaling in power) as long as  $\eta_0 < (1 - \alpha)$  is acceptable. In contrast, maintaining  $\eta_0 = (1 - \alpha)$ , i.e., fully coherent performance to second order, would require that the Doppler diminish much faster than the SNR. Specifically, it would require  $f_m T = \mathcal{O}(\text{SNR}^3)$  (cf. Appendix C).

## B High-Power Regime

A key ingredient in the high-power behavior is the measure of the set where the Doppler spectrum is nonzero. Define  $\mu = \{\nu : S_H(\nu) > 0\}$  as the Lebesgue measure of the interval

where the spectrum is strictly positive. (For any spectrum strictly positive within  $|\nu| \leq f_m T$ , e.g. Clarke-Jakes,  $\mu = 2f_m T$ .)

**Proposition 3** *In the high-power regime,*

$$\bar{\mathcal{I}}(\text{SNR}) \geq (1 - \alpha) \log_2 \left( \frac{1}{1 + \mu/\alpha} \frac{\text{SNR}}{e^\gamma} \right) + o(1) \quad (30)$$

**Proof:** *See Appendix D.*

Relating (30) with (10), the spectral efficiency lower bound is seen to asymptotically differ from the perfect-CSI capacity by a power penalty  $(1 + \mu/\alpha)|_{\text{dB}}$  in addition to the  $(1 - \alpha)$  pre-log reduction.<sup>3</sup> If the RS overhead is optimized, then such power penalty is (through  $\alpha^*$ ) a function of SNR. From Proposition 3 it also follows that, unrestricted,  $\alpha^*$  would vanish for  $\text{SNR} \rightarrow \infty$ . However, (4) limits  $\alpha^*$  from below and thus the power penalty approaches

$$\Delta_\infty = \left( 1 + \frac{\mu}{2f_m T} \right) |_{\text{dB}} \quad (31)$$

which, for spectra that are strictly positive within  $\pm f_m T$ , becomes  $\Delta_\infty = 3$  dB [11]. Thus, the insight derived using block-fading channels is upheld for this class of spectra although only for very high  $f_m T$  is it applicable within the SNR range of operational interest (cf. Fig. 6). The expression  $(1 + \mu/\alpha)|_{\text{dB}}$ , on the other hand, is accurate for all Doppler levels of interest even at modest SNR values. For spectra for which  $\mu < 2f_m T$ , the block-fading insight is not upheld and  $\Delta_\infty$  falls short of 3 dB as per (31).

For all spectra, the dependence is linear on the Doppler and logarithmic on the SNR and, thus, offsetting an increase in Doppler requires ever larger power increases as the SNR grows.

## V RS Power Boosting

An additional degree of freedom sometimes available to the designing engineer is the allocation of unequal powers for RS and data. Indeed, most emerging wireless systems feature

---

<sup>3</sup> $z|_{\text{dB}} = 10 \log_{10} z$ .

RS power boosting [15, 16]. This can be accommodated in our model by defining the signal-to-noise ratios for RS and data to be  $\rho_{rs}\text{SNR}$  and  $\rho_d\text{SNR}$ , respectively, with

$$\rho_{rs}\alpha + \rho_d(1 - \alpha) = 1. \quad (32)$$

Eq. (20) continues to hold, only with

$$\text{SNR}_{\text{eff}} = \frac{\text{SNR}(1 - \text{MMSE})}{1/\rho_d + \text{SNR MMSE}} \quad (33)$$

where

$$\text{MMSE} = 1 - \int_{-f_m T}^{f_m T} \frac{\text{SNR} S_H^2(\nu)}{1/(\rho_{rs}\alpha) + \text{SNR} S_H(\nu)} d\nu \quad (34)$$

which is sensitive only to the product  $(\rho_{rs}\alpha)$ , i.e., to the total RS power.

It can be inferred, from (20), (33) and (34), that—up to the point where (4) is violated—it is advantageous to increase  $\rho_{rs}$  while simultaneously reducing  $\alpha$ .<sup>4</sup> This leads to [11]

$$\alpha^* = 2f_m T \quad (35)$$

and the issue is then the optimization of  $\rho_{rs}$  (and consequently of  $\rho_d$ ). Once  $\alpha^*$  has been fixed, the RS power boosting that maximizes the spectral efficiency lower bound is the one that maximizes  $\text{SNR}_{\text{eff}}$ , i.e.,

$$\rho_{rs}^* = \arg \max_{\rho_{rs}\alpha^* + \rho_d(1-\alpha^*)=1} \text{SNR}_{\text{eff}} \quad (36)$$

In turn, the share of the total power allocated to RS is  $\alpha^* \rho_{rs}^* = 2f_m T \rho_{rs}^*$ . (For the case of a uniform power spectrum, (36) is solved in [10, 11].)

Invoking (35), Proposition 1 extends into the proceeding result.

**Proposition 4** *For a Clarke-Jakes spectrum with RS power boosting,*

$$\text{MMSE} = 1 - \frac{\text{arctanh} \sqrt{1 - \left(\frac{\rho_{rs}\text{SNR}}{\pi/2}\right)^2}}{\frac{\pi}{2} \sqrt{\left(\frac{\pi/2}{\rho_{rs}\text{SNR}}\right)^2 - 1}} \quad (37)$$

---

<sup>4</sup>For the block-fading model, the observation is made in [4, 5, 7] that, with RS power boosting, a single RS should be inserted on every fading block. Translated to our continuous fading model, with the constraint in (4), this can be roughly interpreted as a single RS for every coherence interval.

In order to get further insights, we can also generalize the low- and high-power analyses of Section IV and contrast them with the understanding that stems from the block-fading model. Denoting by  $b$  the number of symbols in a fading block, such understanding is [5]:

- At low SNR,  $\rho_{rs}^* = b/2$  and thus half the power is allocated to RS. The spectral efficiency is unchanged by RS power boosting and hence (24) remains valid.
- At high SNR,  $\rho_{rs}^* = b/(1 + \sqrt{b-1})$  and

$$\text{SNR}_{\text{eff}} = \frac{\text{SNR}}{1 + 2\sqrt{b-1}/b} \quad (38)$$

which amounts to a power penalty ranging between 0 dB and 3 dB depending on  $b$ . Hence, RS power boosting reduces the power penalty by

$$\frac{2}{1 + 2\sqrt{b-1}/b} \text{ dB}. \quad (39)$$

## A Low-Power Regime

**Proposition 5** *With an arbitrary Doppler spectrum and RS power boosting,*

$$\bar{\mathcal{I}}(\text{SNR}) = \eta_0 \log_2(e) \left[ \text{SNR} - \frac{1 - 2f_m T \rho_{rs}}{1 - 2f_m T} \text{SNR}^2 + o(\text{SNR}^2) \right] \quad (40)$$

where  $\eta_0 = (1 - 2f_m T)(1 - \text{MMSE})$ .

**Proof:** *Follow the steps in Appendix C, with  $\text{SNR}_{\text{eff}}$  and MMSE defined as in (33) and (34), respectively, and apply (35).*

On account of Proposition 4,  $\eta_0$  above is in closed form for a Clarke-Jakes spectrum. A further expansion then yields

$$\bar{\mathcal{I}}(\text{SNR}) = \frac{4\rho_{rs}(1 - 2f_m T \rho_{rs})}{\pi^2} \log_2\left(\frac{\pi}{\rho_{rs} \text{SNR}}\right) \text{SNR}^2 + o(\text{SNR}^2) \quad (41)$$

For  $\text{SNR} \rightarrow 0$ ,  $\rho_{rs}^* \rightarrow 1/(4f_m T)$  and the block-fading understanding is corroborated: the power allocated to RS approaches 1/2 and there is no gain from RS power boosting. Like (24), however, (41) is valid only for very small SNR. In the range of interest, the precise  $\rho_{rs}^*$  is markedly below 1/2 as can be appreciated from the table within Fig. 7. The gain from RS power boosting, however, remains scant (cf. Fig. 7, main plot).

## B High-Power Regime

**Proposition 6** *With an arbitrary Doppler spectrum and RS power boosting,*

$$\bar{\mathcal{I}}(\text{SNR}) \geq (1 - 2f_m T) \log_2 \left( \left( \frac{1 - 2f_m T}{1 - 2f_m T \rho_{\text{rs}}} + \frac{\mu}{2f_m T \rho_{\text{rs}}} \right)^{-1} \frac{\text{SNR}}{e^\gamma} \right) + o(1) \quad (42)$$

where, recall,  $\mu = \{\nu : S_H(\nu) > 0\}$ .

**Proof:** *Follow the steps in Appendix D, with  $\text{SNR}_{\text{eff}}$  and MMSE defined as in (33) and (34), respectively, and apply (35).*

The gain from RS power boosting can be appreciated in Fig. 8.

It can be verified that, with respect to the perfect-CSI capacity, the power penalty in (42) is minimized by setting the RS power boost to

$$\rho_{\text{rs}}^* = \frac{1}{2f_m T} \frac{\sqrt{\mu}}{\sqrt{1 - 2f_m T} + \sqrt{\mu}} \quad (43)$$

which leads to a power penalty

$$\Delta_\infty = \left( \sqrt{1 - 2f_m T} + \sqrt{\mu} \right)^2 \text{ dB} \quad (44)$$

ranging within 0–3 dB for spectra strictly positive within  $\pm f_m T$ . (Results equivalent to (43)–(44) were given in [10, Eqs. 34–35] and [11, Eq. 40–41].) For spectra for which  $\mu < 2f_m T$ , this is not the case and the block-fading insight is not upheld. The agreement between the exact  $(\text{SNR}|_{\text{dB}} - \text{SNR}_{\text{eff}}|_{\text{dB}})$  and  $\Delta_\infty$  in (44) is illustrated in Fig. 9. From (31) and (44), RS power boosting enables a power penalty recovery that approaches, for  $\text{SNR} \rightarrow \infty$ ,

$$\frac{1 + \mu/2f_m T}{\left( \sqrt{1 - 2f_m T} + \sqrt{\mu} \right)^2} \text{ dB} \quad (45)$$

which is the direct counterpart of the block-fading quantity in (39). For Doppler levels of interest, (45) is a very substantial part of the penalty incurred without boosting. For  $f_m T = 0.001, 0.02$  and  $0.05$ , respectively, the recovery (out of a 3-dB penalty without boosting) is 2.83 dB, 2.37 dB and 2.16 dB.

## VI Summary

The main findings can be summarized as follows. Without RS power boosting:

- The optimum RS overhead,  $\alpha^*$ , increases with diminishing SNR. Although, for  $\text{SNR} \rightarrow 0$ ,  $\alpha^* \rightarrow 1/2$ , this asymptote must be interpreted with care. For SNR and Doppler levels of operational interest,  $\alpha^*$  is well below  $1/2$ . Using the expressions provided, an accurate assessment is straightforward.
- At low SNR, the spectral efficiency is conveniently characterized via the attainable share,  $\eta_0$ , of the perfect-CSI capacity. Expressions for  $\eta_0$  have been given.
- At high SNR, the spectral efficiency lower bound is conveniently determined through two penalties on the perfect-CSI capacity. First, a pre-log factor reduction to  $(1 - \alpha)$ . Second, a power penalty  $(1 + \mu/\alpha)|_{\text{dB}}$ . For  $\text{SNR} \rightarrow \infty$ ,  $\alpha^* \rightarrow 2f_m T$  and hence the power penalty approaches 3 dB for spectra such that  $\mu = 2f_m T$ . If  $\mu < 2f_m T$ , the power penalty is bounded below 3 dB.

With RS power boosting, on the other hand,  $\alpha^*$  equals the minimum needed to preclude spectral aliasing. The issue is then the share of total power allocated to RS,  $\alpha^* \rho_{\text{rs}}^*$ .

- For a given Doppler, the power allocated to RS increases with diminishing SNR.
- At a given SNR, the power allocated to RS increases with the Doppler.
- At low SNR, there is no substantial gain from RS power boosting.
- At high SNR, conversely, boosting is modestly advantageous. For  $\text{SNR} \rightarrow \infty$ , specifically, the pre-log factor remains unchanged but most of the power penalty can be recovered.

Besides the above findings, which address the questions posed in the Introduction, it is also of obvious interest to quantify the shortcoming in spectral efficiency that results from the restriction of RS-assisted detection, i.e., the deficit of  $\bar{\mathcal{I}}(\text{SNR})$  with respect to (11). Although the computation of (11) seems, in general, unwieldy, we can observe that:

- At low SNR, for the special cases of block-fading channels and of memoryless channels, respectively, second-order expansions are given in [24, 25]. No such expressions are available for continuous fading channels other than the lower bound for PSK signaling [26]

$$\frac{\log_2(e)}{2} \left[ \int_{-f_m}^{f_m} S_H^2(\nu) d\nu - 1 \right] \text{SNR}^2 + o(\text{SNR}^2) \quad (46)$$

which can be compared directly with (26). Since the integral in both (46) and (26) is usually much greater than 1, the conclusion is that RS-assisted schemes can at most attain half the spectral efficiency of an unconstrained scheme that jointly detected the data and estimated the channel. However, since (46) and (26) are valid only at very low SNR (and only for spectra for which the integral exists), this conclusion must be taken with caution. A counterpart of Proposition 2 would be needed for a more precise assessment.

- At high SNR, no expressions are available for (11) either, but results on the capacity of continuous fading channels exist that shed light on the suboptimality of  $\bar{\mathcal{I}}(\text{SNR})$ . Specifically, the capacity behaves as [27] (see also [4, 5] for block-fading channels)

$$C(\text{SNR}) = (1 - \mu) \log \text{SNR} + O(1) \quad (47)$$

and hence, from Propositions 3 and 6, it follows that RS-assisted schemes achieve the optimum pre-log factor as long as the Doppler spectrum is strictly positive within  $\pm f_m T$ . The suboptimality is thus embodied in the power penalty characterized in (31) and (44).

## VII Extension to Multiantenna Channels

Suppose now that  $n_T$  transmit and  $n_R$  receive antennas are available. The channel at time  $k$  takes the form of a  $(n_R \times n_T)$  matrix  $\mathbf{H}(k)$ . If the entries of this matrix are IID and the  $(i,j)$ th entry is  $[\mathbf{H}]_{i,j} \sim \mathcal{N}_{\mathbb{C}}(0, 1)$ , then our analysis generalizes straightforwardly. Given a spatially isotropic Gaussian signal transmission, which is capacity-achieving with perfect CSI [28], the definition  $\text{SNR} = P/N_0$  holds with  $P$  and  $N_0$  signifying, respectively, the total power transmitted over the  $n_T$  antennas and the noise power per receive antenna.

Orthogonal RS must be used at each transmit antenna [29]. Owing to their independence, the entries of  $\mathbf{H}$  can then be separately estimated without loss of optimality and, in con-



sequence, the expressions for MMSE and  $\text{SNR}_{\text{eff}}$  remain valid. Defining  $m = \min(n_{\text{T}}, n_{\text{R}})$  and  $n = \max(n_{\text{T}}, n_{\text{R}})$ , the perfect-CSI capacity in (7) generalizes into [30]

$$C(\text{SNR}) = \log_2(e) e^{n_{\text{T}}/\text{SNR}} \sum_{i=0}^{m-1} \sum_{j=0}^i \sum_{\ell=0}^{2j} \left[ \binom{2i-2j}{i-j} \binom{2j+2n-2m}{2j-\ell} \cdot \frac{(-1)^\ell (2j)! (n-m+\ell)!}{2^{2i-\ell} j! \ell! (n-m+j)!} \sum_{q=0}^{n-m+\ell} E_{q+1} \left( \frac{n_{\text{T}}}{\text{SNR}} \right) \right] \quad (48)$$

which, for small SNR, behaves as [23]

$$C(\text{SNR}) = n_{\text{R}} \log_2(e) \left[ \text{SNR} - \frac{n_{\text{T}} + n_{\text{R}}}{2 n_{\text{T}}} \text{SNR}^2 \right] + o(\text{SNR}^2) \quad (49)$$

while, at high SNR,

$$C(\text{SNR}) = m (\log_2 \text{SNR} - \mathcal{L}_\infty) + o(1) \quad (50)$$

with [31]

$$\mathcal{L}_\infty = \log_2 n_{\text{T}} + \left( \gamma - \sum_{\ell=1}^{n-m} \frac{1}{\ell} - \frac{n}{m} \sum_{n-m+1}^n \frac{1}{\ell} + 1 \right) \log_2 e \quad (51)$$

Propositions 1 and 4 continue to apply whereas Propositions 2–3 and 5–6 generalize effortlessly. Although the spectral efficiency varies with the antenna counts, much of the insight carries over directly.

## Appendices

### A Numerical Computation of $\bar{\mathcal{I}}(\text{SNR})$

From (15), and dropping the time index  $k$ , we need to compute

$$I(X; Y|\hat{H}) = h(Y|\hat{H}) - h(Y|X, \hat{H}) \quad (52)$$

where  $h(\cdot)$  denotes differential entropy. The conditional distribution of the channel given its estimate is  $H|\hat{H} \sim \mathcal{N}_{\text{C}}(\hat{H}, \text{MMSE})$  and hence  $Y|X, \hat{H} \sim \mathcal{N}_{\text{C}}(\hat{H}X, |X|^2 \text{MMSE} + N_0)$ . Therefore,

$$h(Y|X, \hat{H}) = E [\log_2 (|X|^2 \text{MMSE} + N_0)] + \log_2(\pi e) \quad (53)$$

$$= \log_2(\pi e) + e^{1/(\text{SNR} \cdot \text{MMSE})} E_1 \left( \frac{1}{\text{SNR} \cdot \text{MMSE}} \right) \log_2 e \quad (54)$$

where the expectation in (53) is over the distribution of  $X$ . On the other hand,  $HX|\hat{H}$  corresponds to the product of two independent Gaussian random variables and thus the distribution of  $Y|\hat{H}$  is rather unwieldy. The computation of  $h(Y|\hat{H})$  is thus conducted via Monte Carlo. Finally, the outer expectation in (15) over  $\hat{H} \sim \mathcal{N}_{\mathbb{C}}(0, 1 - \text{MMSE})$  is also performed via Monte Carlo.

## B Proof of (26)

Differentiate both sides of (20), which recall are equal to second order. The chain rule yields

$$\frac{\partial \bar{\mathcal{I}}(\text{SNR})}{\partial \text{SNR}} \Big|_{\text{SNR}=0} = (1 - \alpha) \dot{C}(0) \frac{\partial \text{SNR}_{\text{eff}}}{\partial \text{SNR}} \Big|_{\text{SNR}=0} \quad (55)$$

$$= 0 \quad (56)$$

since, from (18) and (21),

$$\frac{\partial \text{SNR}_{\text{eff}}}{\partial \text{SNR}} \Big|_{\text{SNR}=0} = - \left( \text{SNR} \frac{\partial \text{MMSE}}{\partial \text{SNR}} \right) \Big|_{\text{SNR}=0} \quad (57)$$

$$= \lim_{\text{SNR} \rightarrow 0} \int_{-f_m T}^{f_m T} \frac{\text{SNR} S_H^2(\nu)}{(1/\alpha + \text{SNR} S_H(\nu))^2} d\nu \quad (58)$$

$$= 0 \quad (59)$$

In turn, using again (18) as well as (59),

$$\frac{1}{\log_2 e} \frac{\partial^2 \bar{\mathcal{I}}(\text{SNR})}{\partial \text{SNR}^2} \Big|_{\text{SNR}=0} = \frac{1 - \alpha}{\log_2 e} \left[ \dot{C}(0) \frac{\partial^2 \text{SNR}_{\text{eff}}}{\partial \text{SNR}^2} \Big|_{\text{SNR}=0} + \ddot{C}(0) \left( \frac{\partial \text{SNR}_{\text{eff}}}{\partial \text{SNR}} \right)^2 \Big|_{\text{SNR}=0} \right] \quad (60)$$

$$= -2(1 - \alpha) \frac{\partial \text{MMSE}}{\partial \text{SNR}} \Big|_{\text{SNR}=0} - \left( \text{SNR} \frac{\partial^2 \text{MMSE}}{\partial \text{SNR}^2} \right) \Big|_{\text{SNR}=0} \quad (61)$$

$$= 2(1 - \alpha) \alpha \int_{-f_m T}^{f_m T} S_H^2(\nu) d\nu - 2\alpha^2 \lim_{\text{SNR} \rightarrow 0} \int_{-f_m T}^{f_m T} \frac{\text{SNR} S_H^3(\nu)}{(1/\alpha + \text{SNR} S_H(\nu))^3} d\nu \quad (62)$$

$$= 2(1 - \alpha) \alpha \int_{-f_m T}^{f_m T} S_H^2(\nu) d\nu \quad (63)$$

The expansion in (26) then follows directly from (63) whenever the integral exists.

## C Proof of Proposition 2

We shall examine the behavior of  $\bar{\mathcal{I}}(\text{SNR})$  when both  $f_m T$  and SNR approach zero simultaneously. (This approach was pioneered in [7] for the block-fading model.) To that end, consider a Doppler spectrum that scales with the SNR as

$$S_H(\nu) = \frac{1}{\beta \text{SNR}} \tilde{S}_H\left(\frac{\nu}{\beta \text{SNR}}\right) \quad (64)$$

where  $\tilde{S}_H(\cdot)$  is a normalized Doppler spectrum satisfying  $\tilde{S}_H(\xi) = 0$  for  $|\xi| > 1$ . Plugging (64) into (18) and performing a change of variables,

$$\text{MMSE} = 1 - \int_{-1}^1 \frac{\tilde{S}_H^2(\xi)}{\beta/\alpha + \tilde{S}_H(\xi)} d\xi \quad (65)$$

which does not depend on the SNR. Then, differentiating both sides of (20),

$$\frac{\partial \bar{\mathcal{I}}(\text{SNR})}{\partial \text{SNR}} \Big|_{\text{SNR}=0} = (1 - \alpha) \dot{C}(0) \frac{\partial \text{SNR}_{\text{eff}}}{\partial \text{SNR}} \Big|_{\text{SNR}=0} \quad (66)$$

$$= (1 - \alpha) \log_2(e) (1 - \text{MMSE}) \quad (67)$$

and

$$\frac{\partial^2 \bar{\mathcal{I}}(\text{SNR})}{\partial \text{SNR}^2} \Big|_{\text{SNR}=0} = (1 - \alpha) \left[ \dot{C}(0) \frac{\partial^2 \text{SNR}_{\text{eff}}}{\partial \text{SNR}^2} \Big|_{\text{SNR}=0} + \ddot{C}(0) \left( \frac{\partial \text{SNR}_{\text{eff}}}{\partial \text{SNR}} \right)^2 \Big|_{\text{SNR}=0} \right] \quad (68)$$

$$= (1 - \alpha) \log_2(e) [-2(1 - \text{MMSE}) \text{MMSE} - 2(1 - \text{MMSE})^2] \quad (69)$$

$$= -2(1 - \alpha) \log_2(e) (1 - \text{MMSE}) \quad (70)$$

Defining  $\eta_0 = (1 - \alpha)(1 - \text{MMSE})$ , the series expansion of  $\bar{\mathcal{I}}(\text{SNR})$  is

$$\bar{\mathcal{I}}(\text{SNR}) = \eta_0 \log_e(e) [\text{SNR} - \text{SNR}^2 + o(\text{SNR}^2)]. \quad (71)$$

As a final step, we need to unravel  $\eta_0$  so that it applies to arbitrary operational points within the low-power regime. This simply requires ensuring that the spectrum is bandlimited to  $|\nu| \leq f_m T$ . Recognizing that, in (64), the spectrum is bandlimited to  $|\nu| \leq \beta \text{SNR}$ , (65) reverts back to (18).

If, rather than as in (64), the spectrum scales as  $S_H(\nu) = (1/\text{SNR}^3) \tilde{S}_H(\nu/\text{SNR}^3)$ , then

$$\text{MMSE} = 1 - \int_{-1}^1 \frac{\tilde{S}_H^2(\nu)}{\text{SNR}^3/\alpha + \tilde{S}_H(\nu)} d\nu \quad (72)$$

which does depend on the SNR. In this case,  $\text{MMSE}|_{\text{SNR}=0} = 0$ ,  $\partial\text{MMSE}/\partial\text{SNR}|_{\text{SNR}=0} = 0$  and

$$\left( \text{SNR} \frac{\partial^2 \text{MMSE}}{\partial \text{SNR}^2} \right) \Big|_{\text{SNR}=0} = 0 \quad (73)$$

using which (27) holds with  $\eta_0 = (1 - \alpha)$ . The same scaling requirement is found in [7] for a block-fading channel with blocklength  $b$ , i.e., maintaining a share  $(1 - \alpha)$  of the perfect-CSI capacity to second order requires  $1/b = \mathcal{O}(\text{SNR}^3)$ .

## D Proof of Proposition 3

An expansion of (18) for large SNR yields

$$\text{MMSE} = \left[ \int_{\mu}^1 \frac{1}{\alpha} d\nu \right] \frac{1}{\text{SNR}} + \mathcal{O}(1/\text{SNR}^2) \quad (74)$$

$$= \frac{\mu}{\alpha} \frac{1}{\text{SNR}} + \mathcal{O}(1/\text{SNR}^2). \quad (75)$$

It follows, using (21), that

$$\text{SNR}_{\text{eff}} = \frac{\text{SNR}}{1 + \mu/\alpha} - \frac{1}{1 + \alpha/\mu} + \mathcal{O}(1/\text{SNR}) \quad (76)$$

which, via (10) and (20), yields the claimed expansion.

## References

- [1] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rauleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 686–693, Nov. 1991.
- [2] L. Tong, B. M. Sadler, and M. Dong, "Pilot-assisted wireless transmissions: general model, design criteria, and signal processing," *IEEE Signal Proc. Magazine*, vol. 21, no. 6, pp. 12–25, Nov. 2004.
- [3] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Inform. Theory*, vol. 46, no. 3, pp. 933–946, May 2000.

- [4] L. Zheng and D. N. C. Tse, "Communication on the Grassman manifold: A geometric approach to the non-coherent multiple-antenna channel," *IEEE Trans. Inform. Theory*, vol. 48, no. 2, pp. 359–383, Feb. 2002.
- [5] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?," *IEEE Trans. Inform. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.
- [6] X. Ma, L. Yang, and G. B. Giannakis, "Optimal training for MIMO frequency-selective fading channels," *IEEE Trans. Wireless Communications*, vol. 4, no. 2, pp. 453–466, Mar. 2005.
- [7] L. Zheng, D. N. C. Tse, and M. Medard, "Channel coherence in the low-SNR regime," *IEEE Trans. Inform. Theory*, vol. 53, no. 3, pp. 976–997, Mar. 2007.
- [8] S. Furrer and D. Dahlhaus, "Multiple-antenna signaling over fading channels with estimated channel state information: Capacity analysis," *IEEE Trans. Inform. Theory*, vol. 53, no. 6, pp. 2028–2043, June 2007.
- [9] J. Baltersee, G. Fock, and H. Meyr, "An information theoretic foundation of synchronized detection," *IEEE Trans. Communications*, vol. 49, no. 12, pp. 2115–2123, Dec. 2001.
- [10] S. Ohno and G. B. Giannakis, "Average-rate optimal PSAM transmissions over time-selective fading channels," *IEEE Trans. Wireless Communications*, vol. 1, no. 4, pp. 712–720, Oct. 2002.
- [11] X. Deng and A. M. Haimovich, "Achievable rates over time-varying Rayleigh fading channels," *IEEE Trans. Communications*, vol. 55, no. 7, pp. 1397–1406, July 2007.
- [12] M. Dong, L. Tong, and B. M. Sadler, "Optimal insertion of pilot symbols for transmissions over time-varying flat fading channels," *IEEE Trans. Signal Processing*, vol. 52, no. 5, pp. 1403–1418, May 2004.
- [13] J. Doob, *Stochastic Processes*, New York, Wiley, 1990.
- [14] William C. Jakes, *Microwave Mobile Communications*, New York, IEEE Press, 1974.
- [15] 3rd Generation Partnership Project (3GPP), *UTRA-UTRAN Long Term Evolution (LTE)*, Nov. 2004.
- [16] IEEE 802 LAN/MAN Standards Committee, *IEEE Standard for Local and Metropolitan Area Networks—Part 16: Air Interface for Fixed Broadband Wireless Access Systems*, Nov. 2005.

- [17] W. C. Y. Lee, "Estimate of channel capacity in Rayleigh fading environments," *IEEE Trans. Veh. Technol.*, vol. 39, pp. 187–189, Aug. 1990.
- [18] 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.
- [19] N. Merhav, G. Kaplan, A. Lapidoth, and S. Shamai, "On information rates for mismatched decoders," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1953–1967, Nov. 1994.
- [20] A. Lapidoth and S. Shamai, "Fading channels: How perfect need 'perfect side information' be?," *IEEE Trans. Inform. Theory*, vol. 48, no. 5, pp. 1118–1134, May 2002.
- [21] A. Aghamohammadi and H. Meyr, "On the error probability of linearly modulated signals on Rayleigh frequency-flat fading channels," *IEEE Trans. Communications*, vol. 38, no. 11, pp. 1966–1970, Nov. 1990.
- [22] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*, Prentice Hall, 2000.
- [23] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1319–1343, June 2002.
- [24] C. Rao and B. Hassibi, "Analysis of multiple-antenna wireless links at low SNR," *IEEE Trans. Inform. Theory*, vol. 50, no. 9, pp. 2123–2130, Sept. 2004.
- [25] V. V. Prelov and S. Verdú, "Second-order asymptotics of mutual information," *IEEE Trans. Inform. Theory*, vol. 50, no. 8, pp. 1567–1580, Aug. 2004.
- [26] W. Zhang and J. Nicholas Laneman, "How good is PSK for peak-limited fading channels in the low-SNR regime?," *IEEE Trans. Inform. Theory*, vol. 53, no. 1, pp. 236–251, Jan. 2007.
- [27] A. Lapidoth, "On the asymptotic capacity of stationary Gaussian fading channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 2, pp. 437–446, Feb. 2005.
- [28] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecom*, vol. 10, pp. 585–595, Nov. 1999.
- [29] T. L. Marzetta, "BLAST training: Estimating channel characteristics for high capacity space-time wireless," *Proc. of 37th Annual Allerton Conf. on Communication, Control and Computing*, Monticello, IL, Sept. 1999.

- [30] H. Shin and J. H. Lee, "Capacity of multiple-antenna fading channels: Spatial fading correlation, double scattering and keyhole," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2636–2647, Oct. 2003.
- [31] A. Lozano, A. M. Tulino, and S. Verdu, "High-SNR power offset in multiantenna communication," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4134–4151, Dec. 2005.

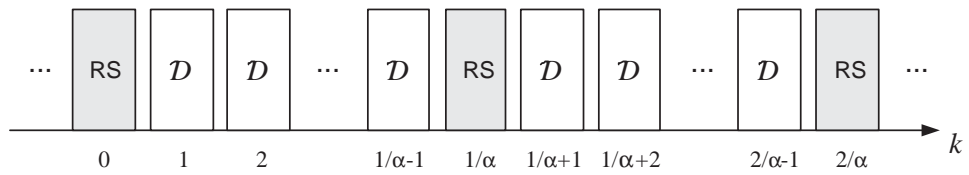


Figure 1: Transmit signal: RS and data. (The rectangular pulse shape is for illustration purposes only, a general pulse shape is allowed.)



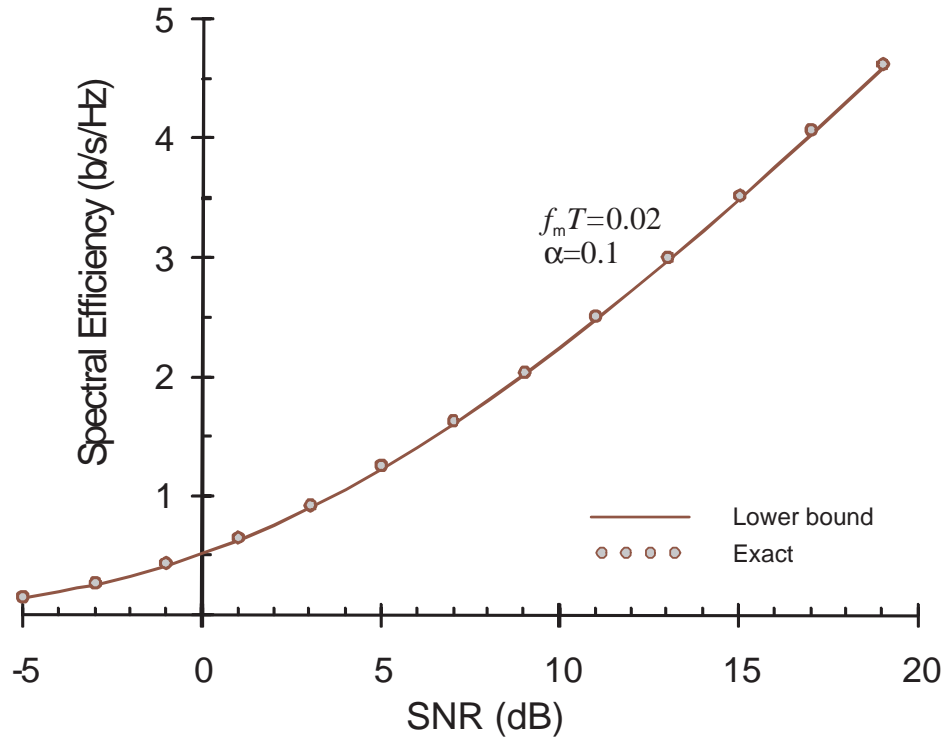


Figure 2: Comparison between the exact RS-assisted spectral efficiency,  $\bar{\mathcal{I}}(\text{SNR})$ , and the lower bound in (20) for a Clarke-Jakes spectrum with  $f_m T = 0.02$  and  $\alpha = 0.1$ .

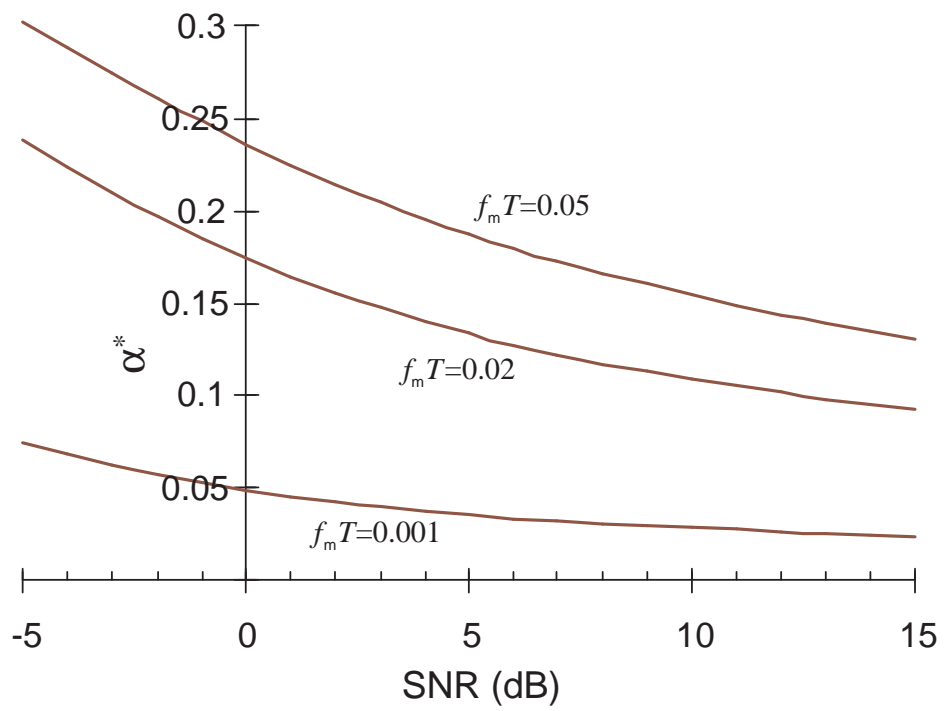


Figure 3: Optimum RS overhead  $\alpha^*$  as function of the SNR for a Clarke-Jakes spectrum with  $f_m T = 0.001, 0.02$  and  $0.05$ .

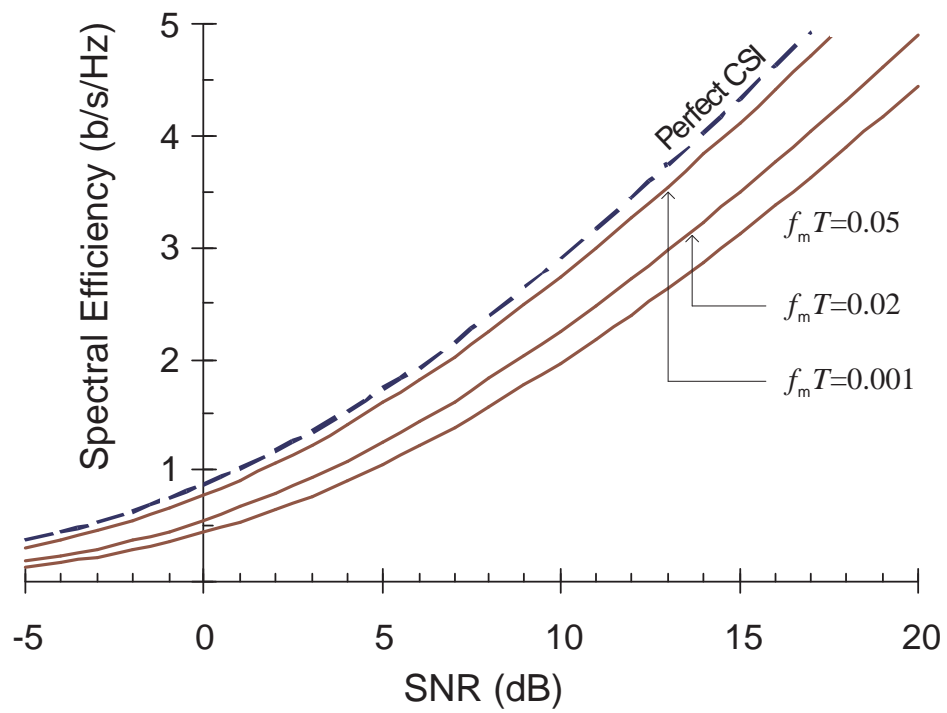


Figure 4: Solid lines: spectral efficiency lower bound with the optimum RS overhead (cf. Fig. 3) for a Clarke-Jakes spectrum. Dashed line: capacity with perfect CSI.

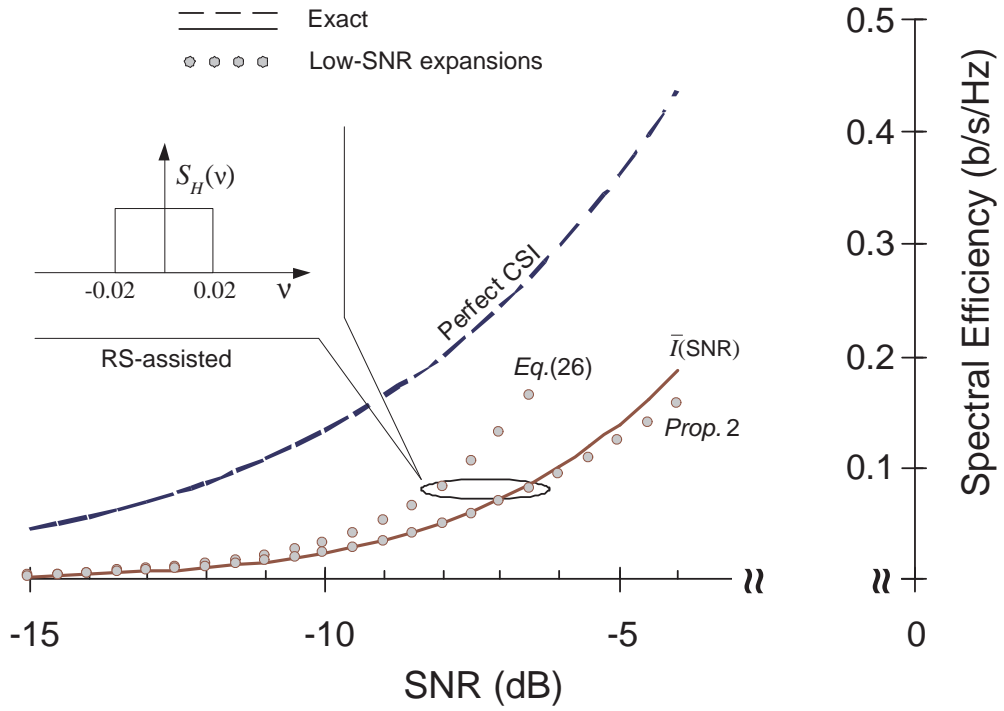


Figure 5: Low-power spectral efficiency both with perfect CSI and with RS-assisted detection ( $\alpha = 0.1$ , uniform Doppler spectrum  $S_H(\nu) = 1/(2f_m T)$  with  $f_m T = 0.02$ ). Dashed line:  $C(\text{SNR})$ . Solid line:  $\bar{I}(\text{SNR})$ . Circles: low-power expansions in (26) and in Proposition 2.

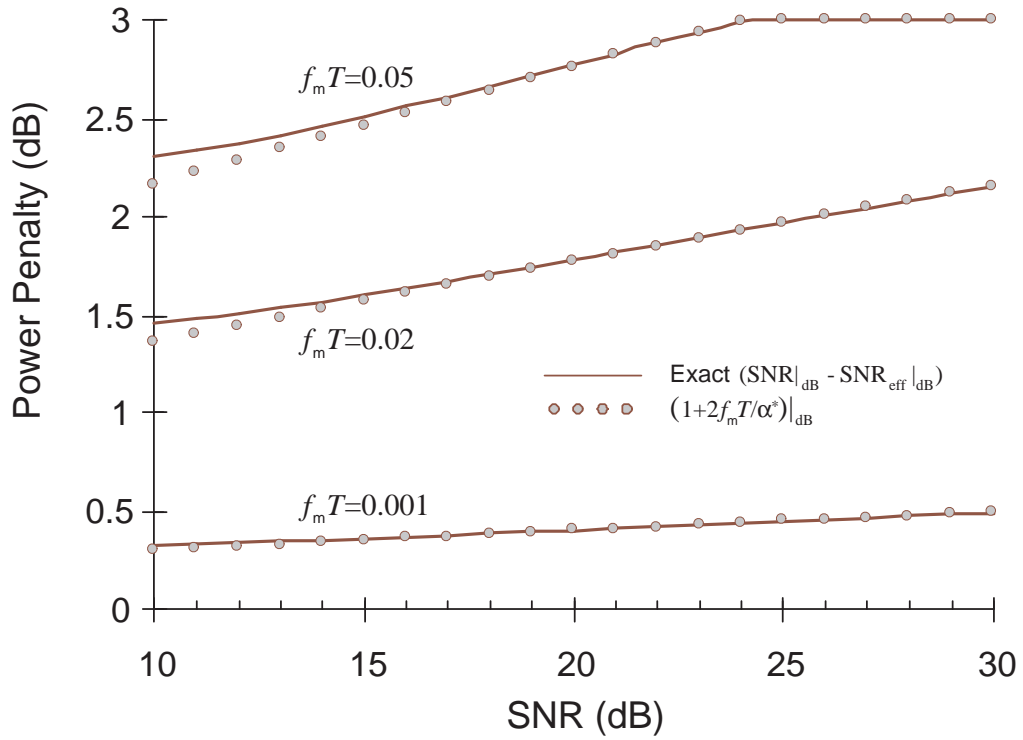


Figure 6: High-power dB-penalty as function of SNR for Clarke-Jakes spectra with  $f_m T = 0.001, 0.02$  and  $0.05$  with the optimum RS overhead  $\alpha^*$ . Solid lines: exact values. Circles:  $(1 + 2f_m T / \alpha^*)|_{\text{dB}}$ .

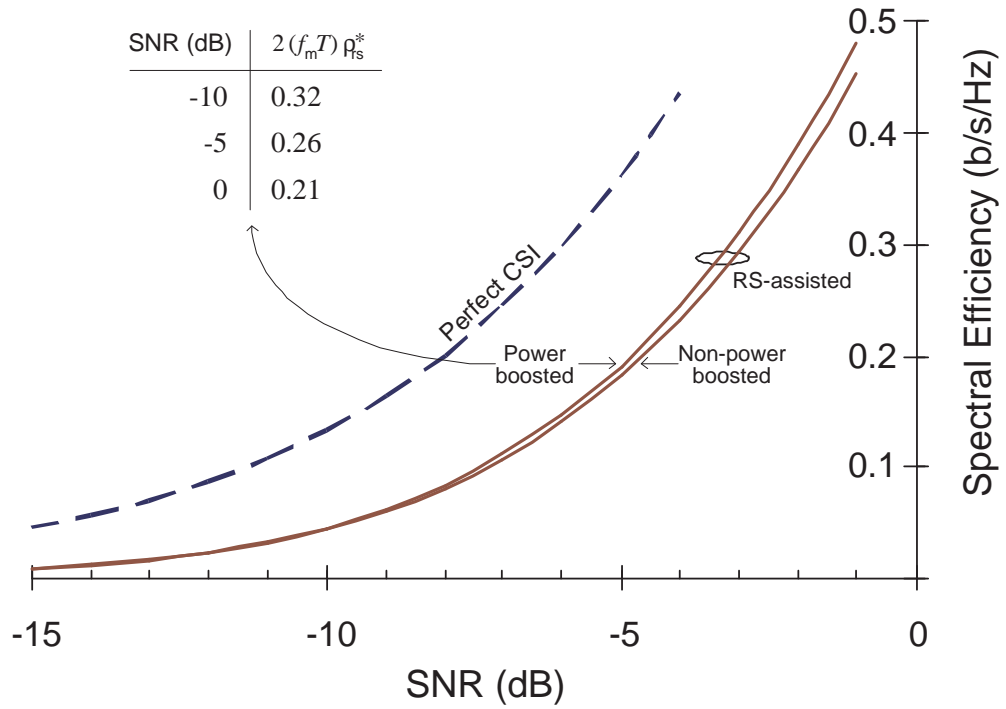


Figure 7: Low-power spectral efficiency both with perfect CSI and with RS-assisted detection for a Clarke-Jakes spectrum with  $f_m T = 0.02$ . Dashed line:  $C(\text{SNR})$ . Solid lines:  $\bar{I}(\text{SNR})$  without RS power boosting (optimum  $\alpha^*$ ) and with RS power boosting ( $\alpha^* = 0.04$ , optimum  $\rho_{rs}^*$ ). Also listed is the share of power allocated to RS, for various SNR, when boosting is used.

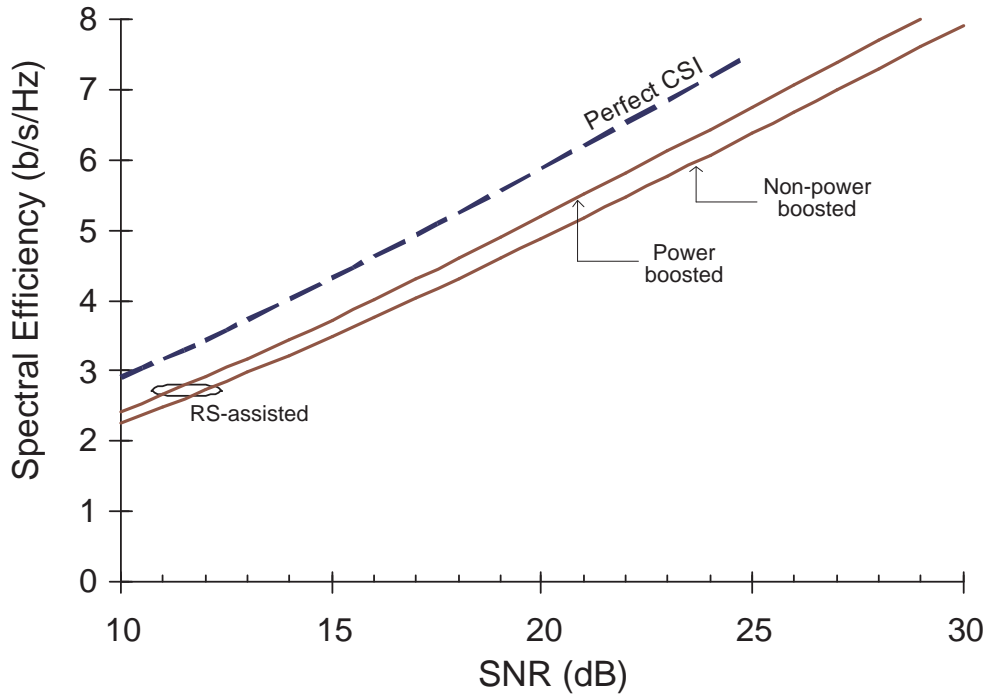


Figure 8: High-power spectral efficiency lower bound both with perfect CSI and with RS-assisted detection for a Clarke-Jakes spectrum with  $f_m T = 0.02$ . Dashed line:  $C(\text{SNR})$ . Solid lines:  $\bar{I}(\text{SNR})$  without RS power boosting (optimum  $\alpha^*$ ) and with RS power boosting ( $\alpha^* = 0.04$ , optimum  $\rho_{rs}^*$ ).

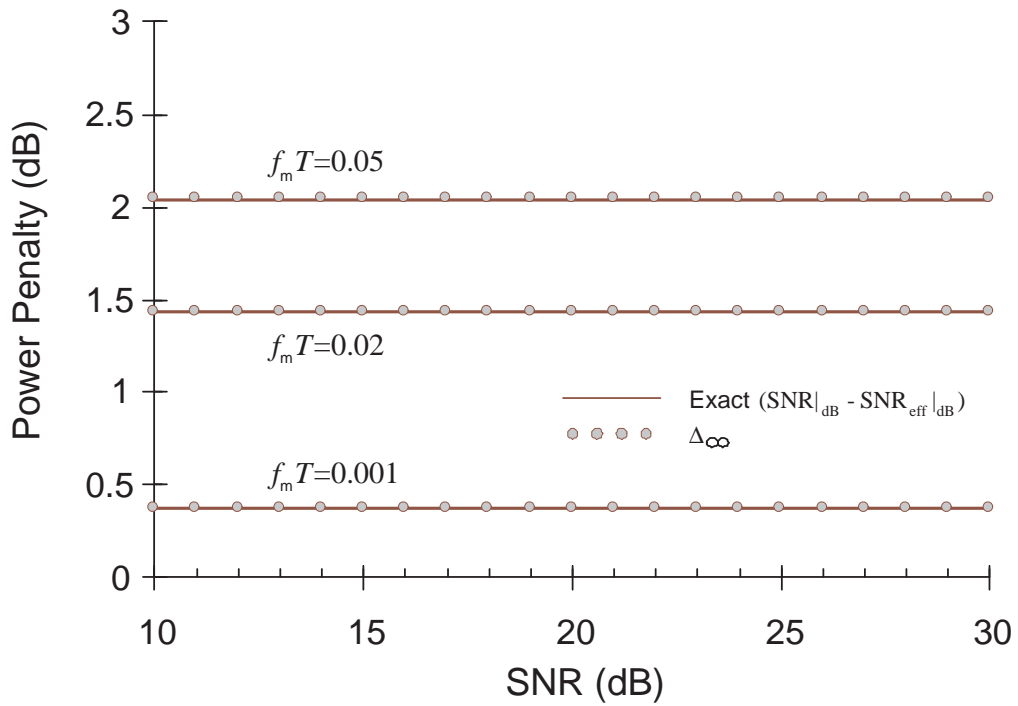


Figure 9: High-power dB-penalty for Clarke-Jakes spectra with  $f_m T = 0.001, 0.02$  and  $0.05$  with  $\alpha^* = 2f_m T$  and with the optimum RS power boost,  $\rho_{rs}^*$ . Solid lines: exact values. Circles:  $\Delta_\infty$  from (44).