Spectral Efficiency in Reference-Signal-Assisted Low-Power Wireless Communication

Angel Lozano

Bell Labs (Alcatel-Lucent)

Holmdel, NJ 07733, USA

Email: lozano@stanfordalumni.org

Abstract—Expressions relating spectral efficiency, power and Doppler spectrum are derived for low-power Rayleigh-faded wireless channels with proper complex signaling. 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. 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 thereby imbuing them with practical significance.

I. INTRODUCTION

Consider a Rayleigh-faded wireless channel in the low-power regime. If we posit that a genie provides the receiver with perfect CSI (channel state information), the capacity-achieving signal distribution is well established and the corresponding capacity can be conveniently computed.

Consider now the same channel and the same signal, 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. This typical setting reflects the operational conditions of most wireless systems, yet its spectral efficiency is not easy to evaluate. 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. Our objective 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 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?

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

The focus of this paper is on the low-power regime, an important one in modern interference-limited systems.

Treatments of the RS-assisted spectral efficiency include [1]–[6], all of which rely on the block-fading channel model (except for [1], 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 [7]–[9]. This leads to relationships that encompass the aforementioned variables. Extending the analysis in [7]–[9], this paper furnishes new and convenient expansions for the low-power regime.

II. CHANNEL MODEL

We shall consider a discrete-time 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) proper complex random variables with zero mean and variance P , which we indicate by $X \sim \mathcal{N}_{\mathbb{C}}(0, P)$. The noise is also IID with $N \sim \mathcal{N}_{\mathbb{C}}(0, N_0)$ and we define $\text{SNR} = P/N_0$.

RS are inserted periodically in the transmission and the fraction thereof is denoted by α , i.e., one in every $1/\alpha$ symbols is a reference while the rest are data symbols (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 VI, we shall lift this restriction allowing for power-boosted RS.

The fading process, $H \sim \mathcal{CN}(0, 1)$, is stationary and ergodic with Doppler spectrum $S_H(\nu)$, $-1/2 \leq \nu \leq 1/2$. 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. Because of its bandlimited nature, $\{H(k)\}$ is a nonregular fading process [10]. User motion results in $f_m = v/\lambda$ where v is the velocity and λ the carrier wavelength. Such f_m is typically several orders of magnitude below $1/T$ and the smooth pulse shaping $\Pi(\cdot)$ can be thereby disregarded altogether.¹

Emphasis shall be placed on the popular Clarke-Jakes spectrum

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

and, to provide examples, we shall use:

- $f_m T = 0.001$. For $\lambda = 15$ cm and $T \approx 100$ μ s, as in emerging wireless systems [11], this is representative of pedestrian motion, $v \approx 5$ Km/h.
- $f_m T = 0.02$ for vehicular motion, $v \approx 100$ Km/h.
- $f_m T = 0.05$ for high-speed trains, $v \approx 250$ Km/h.

III. PERFECT CSI

With perfect CSI at the receiver, there is no need for RS (i.e., $\alpha = 0$). The capacity, in bits/s/Hz, is [12]

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

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

with $E_1(\cdot)$ the exponential integral. Although achieving $C(\text{SNR})$ requires a Gaussian signal, at low SNR any proper complex signal suffices to achieve it to second order [13]

$$C(\text{SNR}) = \log_2(e) [\text{SNR} - \text{SNR}^2] + o(\text{SNR}^2). \quad (8)$$

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

¹Higher 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.

IV. RS-ASSISTED COMMUNICATION

Let us now remove the perfect CSI while inserting a reference every $1/\alpha$ symbols. The signal $\{X(k)\}$ achieves a spectral efficiency

$$\lim_{K \rightarrow \infty} \frac{1}{K} I \left(\underbrace{\{X(k)\}_{k=0}^{K-1}}_{k \in \mathcal{D}}; \underbrace{\{Y(k)\}_{k=0}^{K-1}}_{k \notin \mathcal{D}} \right). \quad (9)$$

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. This reduces the spectral efficiency to

$$\bar{I}(\text{SNR}) = \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{\substack{k=0 \\ k \in \mathcal{D}}}^{K-1} I \left(X(k); Y(k) | \underbrace{\{Y(q)\}_{q=0}^{K-1}}_{q \notin \mathcal{D}} \right). \quad (10)$$

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,

$$\underbrace{\{Y(q)\}_{q=0}^{K-1}}_{q \notin \mathcal{D}} \rightarrow \hat{H}(k) \rightarrow (X(k), Y(k)) \quad (11)$$

constitutes a Markov chain and (10) is equivalent to

$$\bar{I}(\text{SNR}) = \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{\substack{k=0 \\ k \in \mathcal{D}}}^{K-1} I(X(k); Y(k) | \hat{H}(k)) \quad (12)$$

$$= (1 - \alpha) E \left[I(X(k); Y(k) | \hat{H}(k)) \right] \quad (13)$$

where, in (13), k is the index of an arbitrary data symbol and the expectation is over the distribution of $\hat{H}(k)$. The numerical computation of $\bar{I}(\cdot)$ via (13) is feasible, but tedious and time consuming.

With respect to (9), the suboptimality in (13) 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 [14] and thus receivers pursuing such strategy may fall short of (13). More elaborate receiver structures that take into account the joint distribution of $H(k)$ and $\hat{H}(k)$ are in general necessary to attain (13).

Let us postulate a conditional-mean channel estimator

$$\hat{H}(k) = E[H(k) | \underbrace{\{Y(q)\}_{q=0}^{K-1}}_{q \notin \mathcal{D}}] \quad (14)$$

which minimizes $E[|H(k) - \hat{H}(k)|^2]$. For $K \rightarrow \infty$, this

minimum mean-square error approaches [8], [15]²

$$\text{MMSE} = 1 - \int_{-f_m T/\alpha}^{f_m T/\alpha} \frac{\text{SNR} (\alpha S_H(\alpha\nu))^2}{1 + \text{SNR} (\alpha S_H(\alpha\nu))} d\nu \quad (15)$$

which depends on the SNR and on the Doppler spectrum, $S_H(\cdot)$, spectrally expanded by α 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, (15) can be rearranged as

$$\text{MMSE} = 1 - \int_{-f_m T}^{f_m T} \frac{\text{SNR} S_H^2(\nu)}{1/\alpha + \text{SNR} S_H(\nu)} d\nu. \quad (16)$$

We can now rewrite (1) as

$$Y(k) = \hat{H}(k)X(k) + N'(k) \quad (17)$$

where $N'(k) = (H(k) - \hat{H}(k))X(k) + N(k)$ is a zero-mean IID sequence uncorrelated with $\hat{H}(k)X(k)$. Without loss of generality to second order, we can choose $\{X(k)\}$ to be drawn from a proper complex PSK constellation. This has welcome implications since, because of the constant amplitude of a PSK signal, $N' \sim \mathcal{N}_{\mathbb{C}}(0, P \cdot \text{MMSE} + N_0)$. Furthermore, $N'(k)$ is then independent of $\hat{H}(k)X(k)$ and we obtain

$$\bar{\mathcal{I}}(\text{SNR}) = (1 - \alpha) C(\text{SNR}_{\text{eff}}) + o(\text{SNR}^2) \quad (18)$$

with

$$\text{SNR}_{\text{eff}} = \frac{\text{SNR} (1 - \text{MMSE})}{1 + \text{SNR} \text{MMSE}}. \quad (19)$$

Eq. (18) evinces two differences between $\bar{\mathcal{I}}(\cdot)$ and the capacity:

- A pre-log loss of $(1 - \alpha)$ due to 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 α) and diminishing the RS overhead (by reducing α).

V. RS-ASSISTED LOW-POWER COMMUNICATION

Using the block-fading channel model, it was shown in [2]–[3] that, at low SNR and without RS power boosting,

$$\text{SNR}_{\text{eff}} = \alpha b \text{SNR}^2 + o(\text{SNR}^2) \quad (20)$$

with b the number of symbols per fading block. Thus,

$$\bar{\mathcal{I}}(\text{SNR}) = (1 - \alpha) \alpha b \log_2(e) \text{SNR}^2 + o(\text{SNR}^2) \quad (21)$$

and the optimum RS overhead is, therefore, $\alpha^* = 1/2$. In the remainder of this section, we reexamine these findings using the model described in Section II, with the Doppler spectrum centerstage.

²Although 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.

Fig. 2. 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: $\bar{\mathcal{I}}(\text{SNR})$ computed numerically. Circles: low-power expansion in (22).

A continuous-fading counterpart of (21) is obtained through a simple series expansion of the RHS of (18) as

$$\bar{\mathcal{I}}(\text{SNR}) = (1 - \alpha) \alpha \log_2(e) \left[\int_{-f_m T}^{f_m T} S_H^2(\nu) d\nu \right] \text{SNR}^2 + o(\text{SNR}^2) \quad (22)$$

Drawing parallels between (21) and (22) 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. (22), however, has two drawbacks:

- For important Doppler 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, (22) tends to be accurate only for very small SNR (cf. Fig. 2). This mirrors the shortcoming exhibited by (21), not enabling insights that apply at SNR levels of operational interest.

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

Proposition 1 *Without RS power boosting,*

$$\bar{\mathcal{I}}(\text{SNR}) = \eta_0 \log_2(e) [\text{SNR} - \text{SNR}^2 + o(\text{SNR}^2)] \quad (23)$$

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

Proposition 1 is accurate over a much wider range of SNR than (22) (cf. Fig. 3). Moreover, not only does the expansion in Proposition 1 exist for a Clarke-Jakes spectrum but it is in closed form.

Fig. 3. Achievable share of the perfect-CSI capacity for a Clarke-Jakes spectrum with $f_m T = 0.001, 0.02$ and 0.05 , with $\alpha = 0.1$. In solid, exact ratio $\bar{I}(SNR)/C(SNR)$ with $\bar{I}(SNR)$ computed numerically. In circles, η_0 from Corollary 1.

Corollary 1 For a Clarke-Jakes spectrum, (23) holds with

$$\eta_0 = (1 - \alpha) \frac{\operatorname{arctanh} \sqrt{1 - \left(\frac{\alpha}{\pi} \frac{SNR}{f_m T} \right)^2}}{\frac{\pi}{2} \sqrt{\left(\frac{\pi}{\alpha} \frac{f_m T}{SNR} \right)^2 - 1}} \quad (24)$$

obtained through tedious integration of (5) in (16).

Further expansion of (23) and (24) leads to

$$\bar{I}(SNR) = \frac{2\alpha(1 - \alpha)}{\pi^2 f_m T} \log_2 \left(\frac{2\pi f_m T}{\alpha SNR} \right) SNR^2 + o(SNR^2) \quad (25)$$

revealing the precise superquadratic behavior that could not be ascertained from (22).

The optimum RS overhead can be accurately approximated by solving the convex optimization

$$\alpha^* = \arg \max_{2f_m T \leq \alpha \leq 1} \eta_0. \quad (26)$$

For $SNR \rightarrow 0$, $\alpha^* \rightarrow 1/2$ as established in [3] for the block-fading channel. This asymptote, however, must be interpreted with care since α^* remains well below $1/2$ when the parameters are within the range of interest.

Example 1 For a Clarke-Jakes spectrum with $f_m T = 0.02$, $\alpha^* = 0.17$ at $SNR = 0$ dB and $\alpha^* = 0.24$ at $SNR = -5$ dB.

Note also, by contrasting Proposition 1 with (8), that η_0 represents the share of the perfect-CSI capacity achievable with RS-assisted detection. Indeed, as shown in Fig. 3, the approximation $\bar{I}(SNR)/C(SNR) \approx \eta_0$ is excellent for a wide range of SNR and Doppler levels. Moreover, η_0 depends on SNR and on $f_m T$ only through their ratio indicating that, in order to maintain a given share $\eta_0 < (1 - \alpha)$, the Doppler need only decrease linearly with the SNR as the latter vanishes.

Corollary 2 A share $\eta_0 < (1 - \alpha)$ of the perfect-CSI capacity can be sustained to second order if $f_m T = \mathcal{O}(SNR)$ for $SNR \rightarrow 0$.

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 $\eta_0 = (1 - \alpha)$, that is, fully coherent performance to second order, would require that the Doppler diminish much faster than the SNR. Specifically, it would require $f_m T = \mathcal{O}(SNR^3)$.

VI. RS POWER BOOSTING

An additional degree of freedom sometimes available to the designing engineer is the allocation of unequal powers for RS and data. This can be accommodated in our model by defining the signal-to-noise ratios for RS and data to be $\rho_{rs} SNR$ and $\rho_d SNR$, respectively, with

$$\rho_{rs} \alpha + \rho_d (1 - \alpha) = 1. \quad (27)$$

Eq. (18) continues to hold, only with

$$SNR_{\text{eff}} = \frac{SNR(1 - \text{MMSE})}{1/\rho_d + SNR \text{MMSE}} \quad (28)$$

where

$$\text{MMSE} = 1 - \int_{-f_m T}^{f_m T} \frac{SNR S_H^2(\nu)}{1/(\rho_{rs} \alpha) + SNR S_H(\nu)} d\nu \quad (29)$$

It can be inferred, from (18), (28) and (29), that—up to the point where (4) is violated—it is advantageous to increase ρ_{rs} while simultaneously reducing α .³ This leads to $\alpha^* = 2f_m T$ and the issue is then the optimization of ρ_{rs} (and consequently of ρ_d).

Proposition 2 With RS power boosting,

$$\bar{I}(SNR) = \eta_0 \log_2(e) \left[SNR - \frac{1 - 2f_m T \rho_{rs}}{1 - 2f_m T} SNR^2 + o(SNR^2) \right] \quad (30)$$

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

Corollary 3 For a Clarke-Jakes spectrum, (30) holds with

$$\eta_0 = (1 - 2f_m T) \frac{\operatorname{arctanh} \sqrt{1 - \left(\frac{\rho_{rs} SNR}{\pi/2} \right)^2}}{\frac{\pi}{2} \sqrt{\left(\frac{\pi/2}{\rho_{rs} SNR} \right)^2 - 1}} \quad (31)$$

The optimum ρ_{rs}^* can be found by maximizing the expression in Proposition 2 (or, when appropriate, Corollary 3) with the constraint that the share of power allocated to RS not exceed unity, i.e., $\rho_{rs} \leq 1/(2f_m T)$. Comparing the resulting spectral efficiency with the

³For the block-fading model, it is observed in [2], [3], [5] 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 per coherence interval.

Fig. 4. 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: $C(\text{SNR})$. Circles: Proposition 1 with optimum α^* and Proposition 2 with optimum ρ_{rs}^* (and with $\alpha^* = 0.04$).

one obtained from Proposition 1 (or, when appropriate, Corollary 1) in conjunction with (26), the gain from RS power boosting can be assessed. As it turns out, such gain is scant (cf. Fig. 4).

VII. SUMMARY

We can now answer, within the context of the low-SNR regime, the questions posed in the Introduction:

- The optimum RS overhead, α^* , 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, α^* is well below $1/2$. Using (26), an accurate assessment is straightforward.
- The spectral efficiency is conveniently characterized via the attainable share, η_0 , of the perfect-CSI capacity. Expressions for η_0 have been given, in closed form for a Clarke-Jakes spectrum and in integral form for arbitrary spectra. For $\eta_0 < (1 - \alpha)$, a scaling in Doppler can be offset by a commensurate scaling in power.
- There is no substantial gain from RS power boosting and thus, given that it increases the peak-to-average power ratio, it seems preferable not to power boost.

(Note that the above findings need not apply beyond the low-SNR regime.)

Also of obvious interest would be to quantify the shortcoming in spectral efficiency that results from the restriction of RS-assisted detection, i.e., the deficit of $\bar{I}(\text{SNR})$ with respect to (9). Although the computation of (9) seems, in general, unwieldy, for the special cases of block-fading channels and of memoryless channels, second-order low-power expansions are given in [16] and [17], respectively. No such expressions are available for continuous fading channels other than the lower

bound for PSK signaling [18]

$$\frac{\log_2(e)}{2} \left[\int_{-f_m}^{f_m} S_H^2(\nu) d\nu - 1 \right] \text{SNR}^2 + o(\text{SNR}^2) \quad (32)$$

which can be compared directly with (22). Since the integral in both (32) and (22) 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 (32) and (22) 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 1 would be needed for a more precise assessment of whether the enormous additional complexity required to attain (9) is warranted.

REFERENCES

- [1] 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.
- [2] 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.
- [3] 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.
- [4] 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.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] 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.
- [9] 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.
- [10] J. Doob, *Stochastic Processes*. New York, Wiley, 1990.
- [11] UTRA-UTRAN Long Term Evolution (LTE), 3rd Generation Partnership Project (3GPP), Nov. 2004.
- [12] 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.
- [13] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1319–1343, June 2002.
- [14] N. Merhav, G. Kaplan, A. Lapidoth, and S. Shamai, "On information rates for mismatched decoders," *IEEE Trans. Inform. Theory*, vol. 40, no. 6, pp. 1953–1967, Nov. 1994.
- [15] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Prentice Hall, 2000.
- [16] 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.
- [17] 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.
- [18] W. Zhang and J. N. 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.