Diversity Combining With Imperfect Channel Estimation

Roy You, *Member, IEEE*, Hong Li, *Student Member, IEEE*, and Yeheskel (Zeke) Bar-Ness, *Fellow, IEEE*

***Abstract*—The optimal diversity-combining technique is investigated for a multipath Rayleigh fading channel with imperfect channel state information at the receiver. Applying minimum mean-square error channel estimation, the channel state can be decomposed into the channel estimator spanned by channel observation, and the estimation error orthogonal to channel observation. The optimal combining weight is obtained from the first principle of maximum *a posteriori* detection, taking into consideration the imperfect channel estimation. The bit-error performance using the optimal diversity combining is derived and compared with that of the suboptimal application of maximal ratio combining. Numerical results are presented for specific channel models and estimation methods to illustrate the combined effect of channel estimation and detection on bit-error rate performance.**

***Index Terms*—Diversity combining, imperfect channel estimation, Rayleigh fading.**

I. Introduction

DIVERSITY-combining techniques have often been used to combat the deleterious effect of channel fading [1], [2]. If the channel state is known perfectly at the receiver, maximum ratio combining (MRC) can be applied to minimize system bit-error rate (BER) [2]. However, in practice, since the channel estimation at receiver is often imperfect, the estimation error will degrade the BER performance. While such a problem has long been studied [3], [4], the current development of high-data-rate, multiple-input multiple-output mobile communication systems has renewed the interest in understanding the impact of imperfect channel estimation on diversity techniques [5]–[7]. In [4] and [5], the BER performance of MRC of independent and identically distributed (i.i.d.) diversity branches with Rayleigh fading is studied. In [6], the distribution of signal-to-noise ratio (SNR) is given for similar scenario. In [7], the BER performance of MRC with independent but not identically distributed (i.n.d.) branches is studied.

In this paper, we apply the techniques developed in [8] and study the impact of imperfect channel estimation on a general correlated Rayleigh fading channel. We assume that an imperfect channel observation is available at the receiver through a pilot scheme, which is jointly Gaussian distributed with the channel state. With the imperfect channel observation, minimum mean-square error (MMSE) channel estimation is performed where the channel state is decomposed into the MMSE channel estimator spanned by channel observation, and the estimation error orthogonal to channel observation. The optimal detection rule is derived based on maximum *a posteriori* (MAP) detection, given the channel observation and the decision variable. The detection error performance is analyzed, and we show that the channel-estimation error should be treated as an additional source of noise, which might not be white. Therefore, MMSE combining with reliable channel estimation should be used, instead of MRC with imperfect channel observation. Furthermore, since the combiner output SNR is quadratic, the error performance can be calculated analytically as a function of the eigenvalues of certain covariance matrices. As examples of the theory put forth, BER performance of frequency- and space-diversity systems with pilot-symbol-aided channel estimation [9] are analyzed and simulated to verify the combined effect of channel estimation and diversity combining.

Paper approved by A. Abu-Dayya, the Editor for Modulation and Diversity of the IEEE Communications Society. Manuscript received October 8, 2004; revised March 9, 2005. This paper was presented in part at the Conference on Information Sciences and Systems, Princeton, NJ, March 2004.

The authors are with the Center for Communications and Signal Processing Research, Electrical and Computer Engineering Department, New Jersey Institute of Technology, Newark, NJ 07102 USA (e-mail: you@njit.edu; hl8@njit.edu; barness@njit.edu).

Digital Object Identifier 10.1109/TCOMM.2005.857156

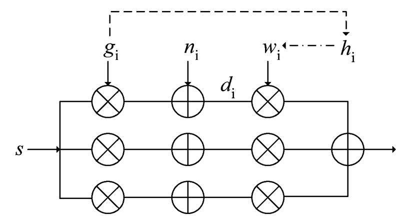


Fig. 1. Diversity combiner with imperfect channel state information.

II. System Model and Optimal Detection Rule

A. System Model

The system model is shown in Fig. 1. s is the transmitted signal and is, in general, complex-valued. In this paper, we use binary phase-shift keying (BPSK) modulation, where $s \in \{\sqrt{E_s}, -\sqrt{E_s}\}$ is drawn from an i.i.d. source with equal symbol probability. The fading channel is modeled with a correlated Rayleigh fading model, where the channel state $\mathbf{g} = [g_1, g_2, \ldots, g_N]^H$ is a proper complex Gaussian random vector [10] with zero mean and covariance matrix

$$C_{\mathbf{gg}} = E[\mathbf{gg}^H]. \tag{1}$$

Here, $\mathbf{g}^H$ denotes the Hermitian of $\mathbf{g}$, which is the complex-conjugate transpose for complex vectors, and reduces to transpose operator for real vectors. We denote the channel state as $\mathbf{g} \sim \mathcal{CN}(\mathbf{0}, C_{\mathbf{gg}})$. The channel noise $\mathbf{n} = [n_1, n_2, \ldots, n_N]^H$ is also zero-mean proper complex white Gaussian with covariance matrix

$$\Lambda_{\mathbf{n}} = E[\mathbf{nn}^H] = N_0 I \tag{2}$$

and is denoted as $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, N_0 I)$. At the receiver, a random vector $\mathbf{d}_s = [d_1, d_2, \ldots, d_N]^H$ is received, where

$$\mathbf{d}_s = \mathbf{g}s + \mathbf{n}. \tag{3}$$

Using a combiner with weight vector $\mathbf{w} = [w_1, w_2, \ldots, w_N]^H$, we create a combined decision random variable r, where

$$r = \mathbf{w}^H \mathbf{d}_s = \mathbf{w}^H(\mathbf{g}s + \mathbf{n}). \tag{4}$$

If the channel state $\mathbf{g}$ is known at the receiver, MRC can be applied to maximize the combiner output SNR, thus minimizing BER. However, the channel state is usually not known perfectly, and only an imperfect channel observation is available at the receiver through the pilot scheme. The imperfect channel observation $\mathbf{h} = [h_1, h_2, \ldots, h_N]^H$ is also assumed to be proper complex Gaussian and zero mean with covariance matrix

$$C_{\mathbf{hh}} = E[\mathbf{h}\mathbf{h}^H] \tag{5}$$

and is denoted as $\mathbf{h} \sim \mathcal{CN}(\mathbf{0}, C_{\mathbf{hh}})$. Furthermore, we assume that the channel observation $\mathbf{h}$ and the channel state $\mathbf{g}$ are jointly proper and have the cross-correlation

$$C_{\mathbf{gh}} = E[\mathbf{g}\mathbf{h}^H]. \tag{6}$$

In this paper, while we do not assume the perfect knowledge of the channel state, we do assume that perfect channel statistics are available at the receiver. Such an assumption is made since the channel statistics vary much more slowly than the channel state itself, and thus can be obtained at the initialization stage using a long training sequence, and continuously improved during the whole communication period. The channel statistics depend on the channel-state model and pilot scheme, which is kept generic here. Numerical results with specific models will be given as examples at the end of this paper.

B. Optimal Detection Rule

Given the channel observation $\mathbf{h}$, the MMSE estimation of the channel state $\mathbf{g}$ is [11]

$$\mathbf{m} = \mathrm{E}[\mathbf{g}|\mathbf{h}] = C_{\mathbf{gh}} C_{\mathbf{hh}}^{-1} \mathbf{h} \tag{7}$$

where $\mathbf{m} \sim \mathcal{CN}(\mathbf{0}, C_{\mathbf{mm}})$ is zero-mean Gaussian distributed with covariance matrix

$$C_{\mathbf{mm}} = C_{\mathbf{gh}} C_{\mathbf{hh}}^{-1} C_{\mathbf{gh}}^H. \tag{8}$$

The channel state is the combination of the linear estimator $\mathbf{m}$ and the estimation error $\mathbf{e}$

$$\mathbf{g} = \mathbf{m} + \mathbf{e} \tag{9}$$

where $\mathbf{h}$ (hence, $\mathbf{m}$) and $\mathbf{e} \sim \mathcal{CN}(0, C_{\mathbf{ee}})$ are orthogonal, and

$$C_{\mathbf{ee}} = C_{\mathbf{gg}} - C_{\mathbf{gh}} C_{\mathbf{hh}}^{-1} C_{\mathbf{gh}}^H = C_{\mathbf{gg}} - C_{\mathbf{mm}}. \tag{10}$$

Since the channel state $\mathbf{g}$ is not available at the receiver, the MAP detection is performed conditioning on the channel observation $\mathbf{h}$ and combiner output r

$$\hat{s} = \arg\max_s p(s|r, \mathbf{h}). \tag{11}$$

Given s and $\mathbf{h}$, r has a complex Gaussian distribution with mean and variance

$$\mathrm{E}[r|s, \mathbf{h}] = \mathbf{w}^H \mathbf{m} s \tag{12}$$

$$\mathrm{VAR}[r|s, \mathbf{h}] = \mathbf{w}^H (C_{\mathbf{ee}} E_s + N_0 I)\mathbf{w}. \tag{13}$$

We write the conditional probability density function (pdf) as

$$p(r|s, \mathbf{h}) = \frac{1}{\pi(\mathrm{VAR}[r|s, \mathbf{h}])} \exp\left(-\frac{\left|r - \mathrm{E}[r|s, \mathbf{h}]\right|^2}{\mathrm{VAR}[r|s, \mathbf{h}]}\right). \tag{14}$$

Assuming that the transmitted symbol has equal probability and the channel observation is independent of the transmitted symbols, the decision rule is simply the minimum-distance decision rule, where

$$\begin{aligned}\hat{s} &= \arg\max_s p(s|r, \mathbf{h}) \\ &= \arg\min_s |r - \mathbf{w}^H \mathbf{m} s|^2. \end{aligned} \tag{15}$$

With BPSK modulation, the decision rule reduces to a threshold test

$$\begin{aligned} |r - \mathbf{w}^H \mathbf{m} s_0|^2 &\underset{\hat{s}=s_1}{\overset{\hat{s}=s_0}{\lessgtr}} |r - \mathbf{w}^H \mathbf{m} s_1|^2 \\ \Re(r^H[\mathbf{w}^H \mathbf{m}]) &\underset{\hat{s}=s_1}{\overset{\hat{s}=s_0}{\lessgtr}} 0 \end{aligned} \tag{16}$$

where $\Re(\cdot)$ denotes the real part of a complex number. The combining weight vector $\mathbf{w}$ is determined such that the average BER can be minimized, which we derive in the next section.

III. Optimal Combining and Corresponding Error Performance

The system error performance is measured by the average BER P_e over additive noise $\mathbf{n}$, channel state $\mathbf{g}$, channel observation $\mathbf{h}$, and transmitted symbol s

$$P_e = \sum_s \Pr(s) \int\int\int \Pr(e|\mathbf{n}, \mathbf{g}, \mathbf{h}, s)\, p(\mathbf{n}, \mathbf{g}, \mathbf{h})\, d\mathbf{n}\, d\mathbf{g}\, d\mathbf{h}. \tag{17}$$

Define the average error probability conditioned on the transmitted symbol and channel observation as

$$\Pr(e|\mathbf{h}, s) = \int\int \Pr(e|\mathbf{n}, \mathbf{g}, \mathbf{h}, s) p(\mathbf{n}, \mathbf{g}|\mathbf{h}, s)\, d\mathbf{n}\, d\mathbf{g} \tag{18}$$

and assuming that the channel observation $\mathbf{h}$ is independent of the transmitted symbol s, the average bit error can be expressed as

$$P_e = \sum_s \Pr(s) \int \Pr(e|\mathbf{h}, s)\, p(\mathbf{h})\, d\mathbf{h}. \tag{19}$$

A. Optimal Combining With Perfect Channel Knowledge

When channel state information is perfect ($\mathbf{m} = \mathbf{h} = \mathbf{g}$), (19) becomes

$$P_e' = \sum_s \Pr(s) \int \Pr(e|\mathbf{g}, s)\, p(\mathbf{g})\, d\mathbf{g} \tag{20}$$

where $\Pr(e|\mathbf{g}, s)$ is the average error due to the additive Gaussian noise $\mathbf{n}$, given the perfect channel knowledge. When the MAP detection rule is applied, it can be shown that the conditional error probability $\Pr(e|\mathbf{g}, s_1)$ is

$$\begin{aligned}\Pr(e|\mathbf{g}, s_1) &= \Pr(\Re(r^H[\mathbf{w}^H\mathbf{m}]) \leq 0 \mid \mathbf{g}, s_1) \\ &= Q\left(\sqrt{\frac{2|\mathbf{w}^H\mathbf{g}|^2 E_s}{|\mathbf{w}^H\mathbf{w}|N_0}}\right). \end{aligned} \tag{21}$$

We also get $\Pr(e|\mathbf{g}, s_0) = \Pr(e|\mathbf{g}, s_1)$ due to the symmetry in BPSK. The conditional error probability is minimized when $\mathbf{w} = \mathbf{g}$. After averaging over s_i, we get

$$\min_{\mathbf{w}} \Pr(e|\mathbf{g}) = Q\left(\sqrt{\frac{2|\mathbf{g}|^2 E_s}{N_0}}\right). \tag{22}$$

This is just MRC, which is optimal when the channel state $\mathbf{g}$ is known perfectly at the receiver, and the noise $\mathbf{n}$ is white.

If we define the performance parameter as the combiner output SNR with perfect channel knowledge

$$\gamma_{\mathbf{g}} = |\mathbf{g}|^2 N_0^{-1} E_s \tag{23}$$

then

$$P_e' = \int Q(\sqrt{2\gamma_{\mathbf{g}}})\, p(\mathbf{g})\, d\mathbf{g}. \tag{24}$$

B. Optimal Combining With Imperfect Channel Knowledge

When the channel state is not perfectly known, but the channel statistics (i.e., $C_{\mathbf{gg}}$, $C_{\mathbf{hh}}$, $C_{\mathbf{gh}}$) are known, MMSE estimation can be applied to the channel observation to obtain the channel estimate $\mathbf{m} = C_{\mathbf{gh}} C_{\mathbf{hh}}^{-1} \mathbf{h}$. Then, the channel state can be decomposed as $\mathbf{g} = \mathbf{m} + \mathbf{e}$, where $\mathbf{e} \sim \mathcal{CN}(0, C_{\mathbf{ee}})$ is the estimation error independent of $\mathbf{m}$. Applying MAP detection as before, the conditional error probability can be expressed as

$$\begin{aligned}\Pr(e|\mathbf{h}, s_1) &= \Pr(\Re(r^H[\mathbf{w}^H\mathbf{m}]) \leq 0 | \mathbf{h}, s_1) \\ &= Q\left(\sqrt{\frac{2|\mathbf{w}^H\mathbf{m}|^2 E_s}{(\mathbf{w}^H(C_{\mathbf{ee}}E_s + N_0 I)\mathbf{w})}}\right). \end{aligned} \tag{25}$$

Again, $\Pr(e|\mathbf{h}, s_0) = \Pr(e|\mathbf{h}, s_1)$, and the $\mathbf{w}$ that minimizes $\Pr(e|\mathbf{h}, s_i)$ is

$$\mathbf{w} = (C_{\mathbf{ee}} E_s + N_0 I)^{-1} \mathbf{m}. \tag{26}$$

After averaging over s_i

$$\min_{\mathbf{w}} \Pr(e|\mathbf{h}) = Q\left(\sqrt{2\mathbf{m}^H (C_{\mathbf{ee}} E_s + N_0 I)^{-1} \mathbf{m} E_s}\right). \tag{27}$$

It can be seen that when performing MAP detection conditioned on the channel observation $\mathbf{h}$, only the MMSE estimator $\mathbf{m}$ portion of the channel state $\mathbf{g}$ is known to the receiver, and the estimation error $\mathbf{e}$ is independent of $\mathbf{h}$, thus, not observable. Therefore, we need to treat $\mathbf{e}s$ as additive noise on top of $\mathbf{n}$. The total noise, taking into account the imperfect observation, is $\mathbf{e}s + \mathbf{n}$ with zero mean and covariance

$$\Sigma = C_{\mathbf{ee}} E_s + N_0 I. \tag{28}$$

Since $\mathbf{e}s$ is often not white, MMSE combining should be applied with $\mathbf{m}$ as the observable channel state, and $\mathbf{e}s + \mathbf{n}$ as the total noise to get the appropriate weighting coefficients $\mathbf{w}$.

As before, we can define the combiner output SNR with optimal combining as

$$\gamma_o = \mathbf{m}^H (C_{\mathbf{ee}} E_s + N_0 I)^{-1} \mathbf{m} E_s = \mathbf{m}^H \Sigma^{-1} \mathbf{m} E_s \tag{29}$$

then

$$P_e = \int Q(\sqrt{2\gamma_o})\, p(\mathbf{m})\, d\mathbf{m}. \tag{30}$$

C. Evaluation of Optimal Combining Performance

Once the conditional detection error probability $\Pr(e|\mathbf{h}, s)$ is determined, the error performance P_e is evaluated by averaging the conditional error over all possible channel observations, as in (30). The approach used traditionally [4], [5] is to use (29) to derive the distribution of SNR $p(\gamma_o)$ and then calculate

$$P_e = \int Q(\sqrt{2\gamma_o})\, p(\gamma_o)\, d\gamma_o. \tag{31}$$

However, as shown in [8], when the combiner output SNR γ_o is in quadratic form, P_e can be alternatively calculated by using

$$Q(x) = \frac{1}{\pi} \int_0^{\pi/2} \exp\left(-\frac{x^2}{2\sin^2\theta}\right) d\theta. \tag{32}$$

Using this analytical expression of $Q(x)$ as in [12], for optimal combining with imperfect channel knowledge, P_e of (30) can be written as shown in (33) at the bottom of the page, where we use the fact that $\mathbf{m}$ is zero-mean proper complex Gaussian. Denote

$$\Lambda = C_{\mathbf{mm}} \Sigma^{-1} E_s \tag{34}$$

and let $\{\lambda_i\}_{i=1}^N$ denote the set of eigenvalues of Λ, then, after integration with respect to $\mathbf{m}$, we get

$$\begin{aligned}P_e &= \frac{1}{\pi} \int_0^{\pi/2} \left[\det\left(\frac{\Lambda}{\sin^2\theta} + I\right)\right]^{-1} d\theta \\ &= \frac{1}{\pi} \int_0^{\pi/2} \left[\prod_{n=1}^{N} \left(\frac{\lambda_n}{\sin^2\theta} + 1\right)^{-1}\right] d\theta \\ &= \frac{1}{\pi} \int_0^{\pi/2} \left[\sum_{k=1}^{K} \sum_{n=1}^{N_k} \mu_{k,n} \left(\frac{\lambda_k}{\sin^2\theta} + 1\right)^{-n}\right] d\theta. \end{aligned} \tag{35}$$

$$\begin{aligned}P_e &= \int_{\mathbf{m}} \frac{1}{\pi} \int_0^{\pi/2} \exp\left[-\frac{\mathbf{m}^H \Sigma^{-1} \mathbf{m} E_s}{\sin^2\theta}\right] \cdot \frac{1}{\pi^N \det(C_{\mathbf{mm}})} \exp\left[-\mathbf{m}^H C_{\mathbf{mm}}^{-1} \mathbf{m}\right] d\theta\, d\mathbf{m} \\ &= \frac{1}{\pi} \int_0^{\pi/2} \int_{\mathbf{m}} \frac{1}{\pi^N \det(C_{\mathbf{mm}})} \exp\left[-\mathbf{m}^H \left(\frac{\Sigma^{-1} E_s}{\sin^2\theta} + C_{\mathbf{mm}}^{-1}\right) \mathbf{m}\right] d\mathbf{m}\, d\theta \end{aligned} \tag{33}$$

The last step follows from the partial-fraction expansion method, where K is the number of distinct eigenvalues, N_k is the eigenvalue's multiplicity, and $\mu_{k,n}$ is the kth residue associated with the nth power in the partial-fraction expansion. The term $\mu_{k,n}$ can be calculated using the residue theorem, and the bit-error probability (BEP) can be further simplified using the following closed-form integral [8]:

$$\frac{1}{\pi}\int_0^{\pi/2}\left(\frac{\lambda_i}{\sin^2\theta}+1\right)^{-n}d\theta = \sum_{k=0}^{n-1}\binom{n-1+k}{k}\left[1-G(\lambda_i)\right]^k\left[G(\lambda_i)\right]^n \quad (36)$$

where

$$G(\lambda_i)=\frac{1}{2}\left(1-\sqrt{\frac{\lambda_i}{1+\lambda_i}}\right). \quad (37)$$

In the case where all the eigenvalues are equal to λ, partial expansion is not necessary, and directly from (35), we have

$$P_e=\frac{1}{\pi}\int_0^{\pi/2}\left(\frac{\lambda}{\sin^2\theta}+1\right)^{-N}d\theta = \sum_{k=0}^{N-1}\binom{N-1+k}{k}\left[1-G(\lambda)\right]^k\left[G(\lambda)\right]^N. \quad (38)$$

In the case where the eigenvalues are distinct

$$P_e=\sum_{n=1}^{N}\mu_n\frac{1}{\pi}\int_0^{\pi/2}\left(\frac{\lambda_n}{\sin^2\theta}+1\right)^{-1}d\theta = \sum_{n=1}^{N}\mu_n G(\lambda_n), \quad (39)$$

$$\mu_n=\prod_{i\neq n}\left(1-\frac{\lambda_i}{\lambda_n}\right)^{-1}. \quad (40)$$

Other cases of eigenvalues with multiplicity can be calculated accordingly [8].

Similarly, it can be shown [12] that the error performance P_e' of (24) under perfect channel knowledge is in the same form as (35), except the parameters $\{\lambda_i\}_{i=1}^N$ should be replaced by the eigenvalues $\{\lambda_i'\}_{i=1}^N$ of

$$\Lambda'=C_{\mathbf{gg}}N_0^{-1}E_s. \quad (41)$$

IV. Suboptimal Combining and Corresponding Error Performance

Due to implementation simplicity or other constraints, sometimes suboptimal combining is used instead of the optimal weights. For example, one common case of suboptimal combining is to treat $\mathbf{h}$ as the true channel state and use it for MRC combining. In this case, the combining weight is set to $\mathbf{w}=\mathbf{h}$, and the real part of the decision variable $\tilde{r}=\mathbf{w}^H(\mathbf{g}s+\mathbf{n})$ is used for a threshold test

$$\tilde{r}=\Re[\mathbf{h}^H(\mathbf{g}s+\mathbf{n})]\underset{\hat{s}=s_1}{\overset{\hat{s}=s_0}{\lessgtr}}0. \quad (42)$$

It is interesting to compare the the suboptimal-combining error performance with the optimal case.

Following the previous development, it can be shown that the suboptimal BER is

$$\tilde{P}_e=\sum_s \Pr(s)\int \Pr(\tilde{e}|\mathbf{h},s)\,p(\mathbf{h})\,d\mathbf{h} \quad (43)$$

where

$$\Pr(\tilde{e}|\mathbf{h},s)=\Pr(\tilde{r}\geq 0|\mathbf{h},s=-\sqrt{E_s}) = Q\left(\sqrt{\frac{2(\Re[\mathbf{h}^H\mathbf{m}])^2E_s}{\mathbf{h}^H(C_{\mathbf{ee}}E_s+N_0I)\mathbf{h}}}\right). \quad (44)$$

In this case, the combiner output SNR can be defined as

$$\gamma_{\mathbf{h}}=\frac{(\Re[\mathbf{h}^H\mathbf{m}])^2E_s}{\mathbf{h}^H(C_{\mathbf{ee}}E_s+N_0I)\mathbf{h}}=\frac{(\Re[\mathbf{h}^H\mathbf{m}])^2E_s}{\mathbf{h}^H\Sigma\mathbf{h}}. \quad (45)$$

For such combining, the error performance is difficult to evaluate, since the effective combiner output SNR $\gamma_{\mathbf{h}}$ is not itself a quadratic form, but rather a quotient of quadratic forms. However, using the Cauchy–Schwartz inequality, we can find a lower bound on error performance

$$\gamma_{\mathbf{h}}=\frac{(\Re[\mathbf{h}^H\mathbf{m}])^2E_s}{\mathbf{h}^H\Sigma\mathbf{h}}\leq\frac{|\mathbf{h}^H\mathbf{m}|^2E_s}{\mathbf{h}^H\Sigma\mathbf{h}} \leq\left|\left(\sqrt{\Sigma}\right)^{-1}\mathbf{m}\right|^2E_s=\mathbf{m}^H\Sigma^{-1}\mathbf{m}E_s=\gamma_o. \quad (46)$$

Therefore, the suboptimal BEP is lower bounded by the optimal performance.

V. Examples and Numerical Results

In the previous sections, we have considered the detection error probability with imperfect channel estimation and perfect channel statistics. However, the correlation between channel state and channel observation is determined by the pilot scheme, and is, in general, complicated. In this section, a simple example of pilot-symbol-aided channel estimation [9] is given to illustrate the concept of analyzing BER while taking into consideration the channel-estimation uncertainty.

A. Pilot-Symbol-Aided Channel Estimation

The pilot-symbol-aided channel estimation refers to the general scheme where predetermined special symbols are transmitted over the same channel as the data symbols, and the received signals are used for channel estimation. For example, in frequency-diversity systems, such as multicarrier code-division multiple access (MC-CDMA), it could be sending pilot symbols over different frequency subcarriers; or in a space-diversity system, it could simply be transmitting pilot symbols over multiple antennas.

Using pilot-symbol-aided channel estimation, we assume that the pilot symbol has energy

$$E_p=kE_s. \quad (47)$$

The corresponding received signal is

$$\mathbf{d}_p=\sqrt{E_p}\mathbf{g}+\mathbf{n} \quad (48)$$

where $\mathbf{n}$ is the additive noise at the receiver. The channel observation can be expressed as

$$\mathbf{h} = \frac{\mathbf{d}_p}{\sqrt{E_p}} = \mathbf{g} + \frac{\mathbf{n}}{\sqrt{E_p}}. \tag{49}$$

It can be seen that, under the pilot-symbol scheme, the covariance matrices $C_{\mathbf{gg}}$ and $C_{\mathbf{hh}}$ can be diagonalized by the same unitary matrix U as

$$C_{\mathbf{gg}} = U\Lambda_{\mathbf{gg}}U^H \tag{50}$$

$$C_{\mathbf{hh}} = U\Lambda_{\mathbf{hh}}U^H = U\left(\Lambda_{\mathbf{gg}} + \left(\frac{N_o}{E_p}\right)I\right)U^H. \tag{51}$$

We define

$$\bar{\mathbf{g}} = U^H\mathbf{g} \tag{52}$$

$$\bar{\mathbf{h}} = U^H\mathbf{h} \tag{53}$$

then one can show that

$$\Lambda_{\mathbf{gg}} = E[\bar{\mathbf{g}}\bar{\mathbf{g}}^H] = \mathrm{diag}(\sigma^2_{\bar{g}_1}, \sigma^2_{\bar{g}_2}, \ldots, \sigma^2_{\bar{g}_N}) \tag{54}$$

$$\Lambda_{\mathbf{hh}} = E[\bar{\mathbf{h}}\bar{\mathbf{h}}^H] = \mathrm{diag}(\sigma^2_{\bar{h}_1}, \sigma^2_{\bar{h}_2}, \ldots, \sigma^2_{\bar{h}_N}), \quad \text{where } \sigma^2_{\bar{h}_i} = \sigma^2_{\bar{g}_i} + \frac{N_0}{E_p} \tag{55}$$

$$\Lambda_{\mathbf{gh}} = E[\bar{\mathbf{g}}\bar{\mathbf{h}}^H] = \mathrm{diag}(\rho_{\bar{g}_1\bar{h}_1}, \rho_{\bar{g}_2\bar{h}_2}, \ldots, \rho_{\bar{g}_N\bar{h}_N}), \quad \text{where } \rho_{\bar{g}_i\bar{h}_i} = E[\bar{g}_i\bar{h}_i^H] = \sigma^2_{\bar{g}_i} \tag{56}$$

$$C_{\mathbf{gh}} = U\Lambda_{\mathbf{gh}}U^H \tag{57}$$

$$C_{\mathbf{ee}} = U\Lambda_{\mathbf{ee}}U^H = U(\Lambda_{\mathbf{gg}} - \Lambda_{\mathbf{gh}}\Lambda^{-1}_{\mathbf{hh}}\Lambda^H_{\mathbf{gh}})U^H. \tag{58}$$

Furthermore, we define

$$\bar{\mathbf{m}} = U^H\mathbf{m} = \Lambda_{\mathbf{gh}}\Lambda^{-1}_{\mathbf{hh}}\bar{\mathbf{h}} \tag{59}$$

then

$$\Lambda_{\mathbf{mm}} = E[\bar{\mathbf{m}}\bar{\mathbf{m}}^H] = \Lambda_{\mathbf{gh}}\Lambda^{-1}_{\mathbf{hh}}\Lambda^H_{\mathbf{gh}}. \tag{60}$$

B. Optimal Combining

In the case with imperfect channel knowledge, the combiner output SNR γ_o can be written as

$$\gamma_o = \bar{\mathbf{m}}^H(\Lambda_{\mathbf{ee}}E_s + N_0 I)^{-1}\bar{\mathbf{m}}\,E_s = \sum_{i=1}^{N} \frac{\sigma^2_{\bar{g}_i}}{\left(\sigma^2_{\bar{g}_i} + \frac{N_0}{E_p}\right)}|\bar{h}_i|^2 \cdot \frac{E_sE_p}{N_0(E_s+E_p) + \frac{N_0^2}{\sigma^2_{\bar{g}_i}}}. \tag{61}$$

Then, it can be shown that

$$P_e = \int Q(\sqrt{2\gamma_o})\,p(\mathbf{m})\,d\mathbf{m} = \int Q(\sqrt{2\gamma_o})\,p(\bar{\mathbf{h}})\,d\bar{\mathbf{h}}. \tag{62}$$

For the two channel models of frequency and space diversity that we will consider, the eigenvalues of the channel-state covariance matrix $C_{\mathbf{gg}}$ are distinct. Therefore, we can calculate P_e as in (39) and (40) with the eigenvalues of (34) as

$$\lambda_n = \frac{|\rho_{\bar{g}_n\bar{h}_n}|^2E_s}{\sigma^2_{\bar{g}_n}\sigma^2_{\bar{h}_n}E_s - |\rho_{\bar{g}_n\bar{h}_n}|^2E_s + N_0\sigma^2_{\bar{h}_n}} = \frac{\sigma^2_{\bar{g}_n}E_s}{N_0}\cdot\frac{E_p}{E_s + E_p + \frac{N_0}{\sigma^2_{\bar{g}_n}}}. \tag{63}$$

It is interesting to compare this result with combining using perfect channel knowledge, where

$$\gamma_{\mathbf{g}} = |\bar{\mathbf{g}}|^2N_0^{-1}E_s \tag{64}$$

$$P_e' = \int Q(\sqrt{2\gamma_{\mathbf{g}}})\,p(\mathbf{g})\,d\mathbf{g} = \int Q(\sqrt{2\gamma_{\mathbf{g}}})\,p(\bar{\mathbf{g}})\,d\bar{\mathbf{g}}. \tag{65}$$

Applying a similar technique, we get the eigenvalues of (41) as

$$\lambda_n' = \frac{\sigma^2_{\bar{g}_n}E_s}{N_0}. \tag{66}$$

Comparing with λ_n', we notice that λ_n can be factored in two parts. The first term $\sigma^2_{\bar{g}_n}E_s/N_0$ is the same as λ_n', indicating the SNR gain of the decorrelated subchannel. The second term $E_p/(E_s + E_p + N_0/\sigma^2_{\bar{g}_n})$ shows the loss due to channel estimation. Notice that at low SNR, this loss could be significant, due to the poor quality of channel estimation. Furthermore, when E_s/N_0 is large, we get

$$\lambda_n \approx \frac{\sigma^2_{\bar{g}_n}E_s}{N_0}\cdot\frac{k}{k+1} \tag{67}$$

which shows that for high transmission SNR, the case with imperfect channel knowledge still has $10\log(k/k+1)$ dB power loss, compared with the case with perfect channel knowledge.

C. Suboptimal Combining

For suboptimal combining with pilot-symbol-aided channel estimation, it can be shown that

$$\gamma_{\mathbf{h}} = \frac{(\Re[\mathbf{h}^H\mathbf{m}])^2E_s}{\mathbf{h}^H\Sigma\mathbf{h}} = \frac{|\bar{\mathbf{h}}^H\Lambda_{\mathbf{gh}}\Lambda^{-1}_{\mathbf{hh}}\bar{\mathbf{h}}|^2E_s}{\bar{\mathbf{h}}^H(\Lambda_{\mathbf{ee}}E_s + N_0I)\bar{\mathbf{h}}}$$
$$= \frac{\left|\sum_{i=1}^{N}\frac{\sigma^2_{\bar{g}_i}}{\left(\sigma^2_{\bar{g}_i}+\frac{N_0}{E_p}\right)}|\bar{h}_i|^2\right|^2}{\sum_{i=1}^{N}\frac{\sigma^2_{\bar{g}_i}\left(\frac{N_0}{E_s}+\frac{N_0}{E_p}\right)+\left(\frac{N_0}{E_s}\right)\left(\frac{N_0}{E_p}\right)}{\left(\sigma^2_{\bar{g}_i}+\frac{N_0}{E_p}\right)}|\bar{h}_i|^2}$$
$$= \sum_{i=1}^{N}\frac{\sigma^2_{\bar{g}_i}}{\left(\sigma^2_{\bar{g}_i}+\frac{N_0}{E_p}\right)}|\bar{h}_i|^2 \cdot \frac{E_sE_p\sum_{i=1}^{N}\eta_i|\bar{h}_i|^2}{N_0(E_s+E_p)\sum_{i=1}^{N}\eta_i|\bar{h}_i|^2 + N_0^2\sum_{i=1}^{N}\frac{1}{\sigma^2_{\bar{g}_i}}\eta_i|\bar{h}_i|^2},$$
$$\text{where } \eta_i = \frac{\sigma^2_{\bar{g}_i}}{\left(\sigma^2_{\bar{g}_i}+\frac{N_0}{E_p}\right)}. \tag{68}$$

Although the above expression of $\gamma_{\mathbf{h}}$ is complicated, it is interesting to compare it with the expression of γ_o in (61).

It can be seen that when the eigenvalues of $C_{\mathbf{gg}}$ are identical, the suboptimal combining SNR is the same as γ_o. This implies that if the multipath channel can be diagonalized into parallel subchannels with i.i.d. (as in rich scattering environment), then there is no difference between optimal and suboptimal combining. Furthermore, as the transmission SNR (E_s/N_0) rises, $\gamma_{\mathbf{h}}$ also approaches the optimal combining SNR γ_o. This should not be surprising, since at high SNR, the channel observation is very close to the channel state and the MMSE channel estimation, thus making the optimal and suboptimal combining weights similar to each other. Although the optimal and suboptimal combining performances are similar under high SNR or identical eigenvalue conditions, the performances are indeed different for practical situations. Compared with combining with perfect channel knowledge, there are also performance losses due to pilot-symbol power.

D. Channel Models and Simulation Results

Aside from detection and estimation methods, the system BER is fundamentally determined by the underlying channels, where correlation between channel state can be over time, frequency, or space, or their combinations. In this paper, we consider two examples of correlated Rayleigh fading channels, where the correlation is either between subcarriers at different frequencies or between antennas at different space locations.

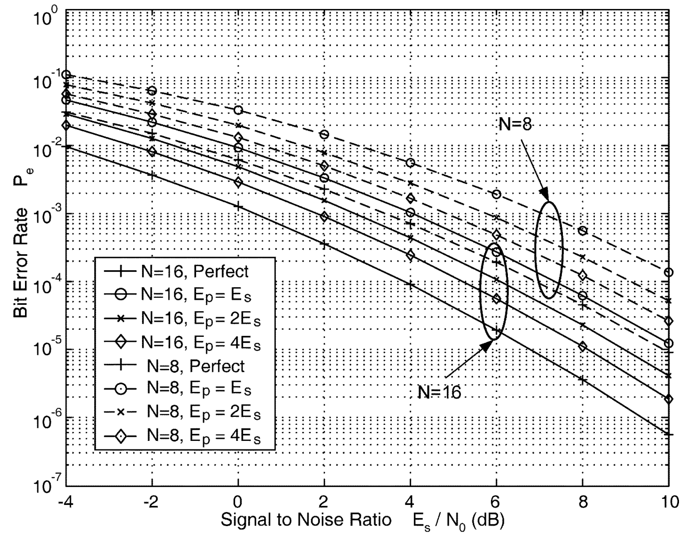
The first example we consider is to combine over frequency diversity, such as in MC-CDMA systems. The correlation for the channel state is modeled as [13]

$$C_{\mathbf{gg}}(m,n) = \frac{\sigma_g^2}{1 + j2\pi f_d \tau_d (m-n)}. \tag{69}$$

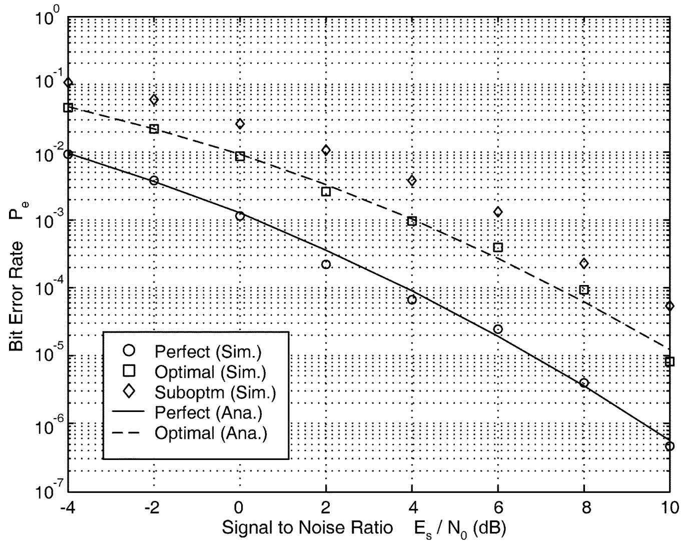
τ_d is the channel delay spread and is set to 25 ns. f_d is the frequency separation between two adjacent subcarriers. We allocate a total bandwidth of 20 MHz and divide the bandwidth into either 8 or 16 subcarriers. Thus, $f_d = 1.25$ MHz for $N = 16$ and $f_d = 2.5$ MHz for $N = 8$. From Fig. 2(a), we can see that when perfect channel knowledge is available, using more subcarriers increases BER performance because of the increased frequency diversity. However, when the channel knowledge is not perfect, the system designer can choose to improve BER performance by either improving the channel observation quality (e.g., increasing the ratio of pilot power to symbol power (E_p/E_s) in this case) or increasing system diversity (e.g., increasing the number of subcarriers from 8 to 16). When SNR is low or the subcarriers are highly correlated, the diversity gain is not as effective as improved channel observation. Therefore, even with more combining branches, the BER performance could be worse than system with fewer combining branches but better channel observation. In this situation, to improve system BER, it would be more cost effective and less complex to implement systems with better channel observation and fewer combining branches. As shown by Fig. 2(b), compared with the case with perfect channel knowledge, imperfect channel knowledge degrades system performance, even with optimal combining, and BER is further reduced when suboptimal combining is used. In this case, the loss due to suboptimal combining is significant, and can be attributed to inaccurate estimation of channel state.

The second example we consider is to combine over space diversity. In this example, we allocate a fixed length of 0.05 m and assign this space to either two ($N = 2$) or three ($N = 3$) antennas. The antennas are assumed to be placed in line and spaced equally. The normalized covariance matrix for the channel state is modeled as [14]

$$C_{\mathbf{gg}}(m,n) = \sigma_g^2 J_0\left(2\pi \frac{|m-n|d \cdot f_c}{C}\right) \tag{70}$$



(a)



(b)

Fig. 2. Example of frequency-diversity model ($\tau_d = 25$ ns). (a) Comparing the analytical error performance of optimal combining with perfect channel knowledge and imperfect channel knowledge, using different estimation power for $N = 8$ and $N = 16$. (b) Analytical error performance and Monte Carlo simulation results for the perfect channel knowledge, optimal combining with imperfect channel knowledge and the suboptimal combining performance with $N = 16$ and $E_p = E_s$.

where d is the distance between two adjacent antennas, $f_c = 2.4$ GHz is the carrier frequency, and C is the light speed. $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind. In the case of $N = 2$, $d = 5$ cm, and the two antennas are uncorrelated. In the case of $N = 3$, $d = 2.5$ cm, and the three antennas are correlated. From Fig. 3(a), we can see a similar tradeoff between combining diversity and channel observation quality. More antennas provide more space diversity, which improves system BER performance. However, such improvement can be achieved by better channel observation, which could be more cost-effective to implement. In Fig. 3(b), the analytical error performance of optimal combining with perfect and imperfect channel knowledge are compared with Monte Carlo simulation and are shown to be in good agreement. It is somewhat surprising that simulation also shows that the suboptimal combining in this example has only slightly worse performance than

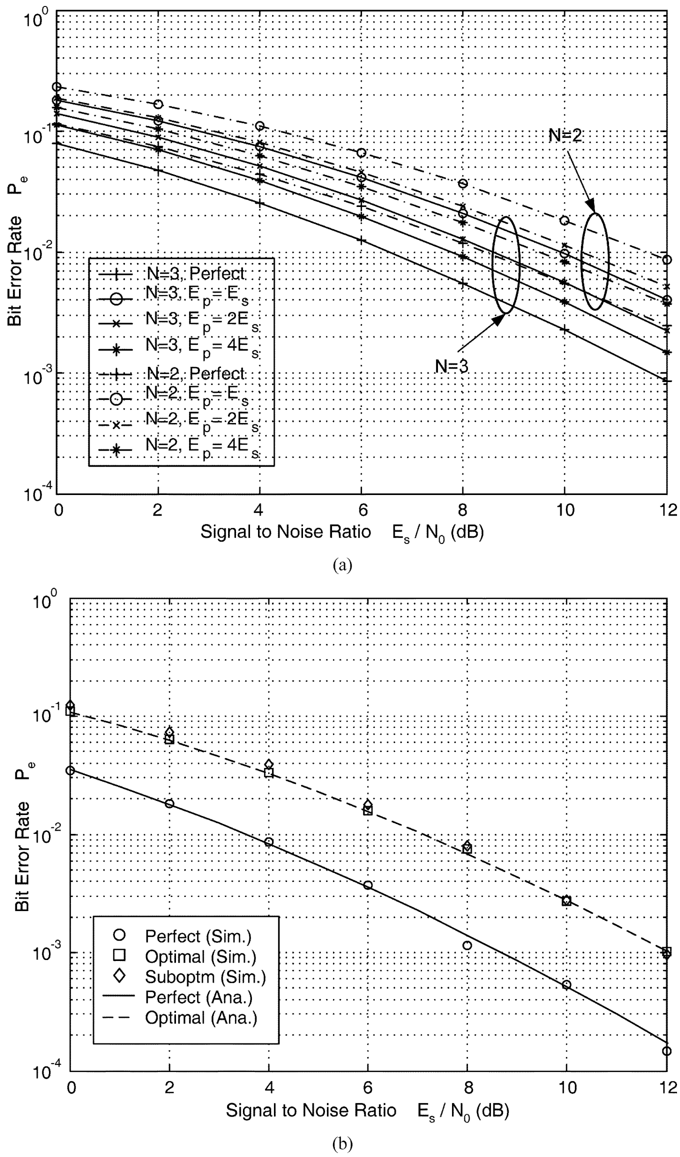


Fig. 3. Example of spatial diversity model ($f_c = 2.4$ GHz). (a) Comparing the analytical error performance of optimal combining with perfect channel knowledge and imperfect channel knowledge using different estimation power for $N = 2$ and $N = 3$. (b) Analytical error performance and the Monte Carlo simulation results for the perfect channel knowledge, optimal combining with imperfect channel knowledge, and the suboptimal combining performance with $N = 3$ and $E_p = E_s$.

optimal combining. However, by applying the corresponding parameters into (70), we get the eigenvalues $\sigma^2_{\bar{g}_1} = 0.0635$, $\sigma^2_{\bar{g}_2} = 1.0550$, $\sigma^2_{\bar{g}_3} = 1.8816$. It can be seen that the eigenvalue $\sigma^2_{\bar{g}_i}$ is either "off" ($\sigma^2_{\bar{g}_1} \approx 0$) or "on." For the eigenvalues that are "on," their values are quite close ($\sigma^2_{\bar{g}_2} \approx \sigma^2_{\bar{g}_3}$). Applying these eigenvalues in (61) and (68), we can see that the combiner output SNR γ_o and $\gamma_{\mathbf{h}}$ are very close. Hence, the suboptimal performance is very close to that of the optimal combining. In this case, the channel can be approximated by two parallel subchannels with i.i.d. distributions of channel states. However, in general, the optimal combining has better performance and should be applied.

VI. Conclusion

In this paper, we investigate the effect of imperfect channel estimation on diversity combining. For Rayleigh fading channel, when the channel observation at the receiver is imperfect, MMSE channel estimation should be performed. The channel state is the sum of the channel estimator and estimation error, where the estimation error mixed with transmitted signal contributes an additional source of noise. Since this noise is usually not white, MMSE combining should be used for optimal diversity combining, with the adjusted channel state and noise model. The optimal combining error performance is calculated and compared with MRC with perfect channel knowledge and suboptimal combining with imperfect channel knowledge cases. Numerical results with specific models show good agreement with Monte Carlo simulations.

References

[1] W. C. Jakes, *Microwave Mobile Communications*. New York: Wiley, 1974.

[2] J. G. Proakis, *Digital Communications*, 4 ed. New York: McGraw-Hill, 2000.

[3] ——, "Probabilities of error for adaptive reception of M-phase signals," *IEEE Trans. Commun.*, vol. COM-16, no. 2, pp. 71–81, Feb. 1968.

[4] M. Gans, "The effect of Gaussian error in maximal ratio combiners," *IEEE Trans. Commun.*, vol. COM-19, no. 8, pp. 492–500, Aug. 1971.

[5] B. R. Tomiuk, N. C. Beaulieu, and A. A. Abu-Dayya, "General forms for maximal ratio diversity with weighting errors," *IEEE Trans. Commun.*, vol. 47, no. 4, pp. 488–492, Apr. 1999.

[6] S. Roy and P. Portier, "Maximal-ratio combining architectures and performance with channel estimation based on a training sequence," *IEEE Trans. Wireless Commun.*, vol. 3, no. 7, pp. 1154–1164, Jul. 2004.

[7] J. S. Thompson, "Antenna array performance with channel estimation errors," in *Proc. ITG Workshop Smart Antennas*, Munich, Germany, Mar. 2004, pp. 75–78.

[8] M. K. Simon and M.-S. Alouini, *Digital Communication Over Fading Channel: A Unified Approach to Performance Analysis*. New York: Wiley, 2000.

[9] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 40, no. 6, pp. 686–693, Nov. 1991.

[10] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inf. Theory*, vol. 39, no. 7, pp. 1293–1302, Jul. 1993.

[11] S. M. Kay, *Fundamentals of Statistical Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall PTR, 1993, vol. I, Estimation Theory.

[12] V. V. Veeravalli, "On performance analysis of signaling on correlated fading channels," *IEEE Trans. Commun.*, vol. 49, no. 11, pp. 1879–1883, Nov. 2001.

[13] M. D. Yacoub, *Foundations of Mobile Radio Engineering*. Boca Raton, FL: CRC, 1993.

[14] G. Stuber, *Principles of Mobile Communication*. Norwell, MA: Kluwer, 1996.

Roy You (S'98–M'03) received the B.S. and Ph.D. degrees in electrical engineering from the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, in 1997 and 2003, respectively.

Since 2003, he has been an Assistant Professor of Electrical and Computer Engineering with the New Jersey Institute of Technology, Newark. In his previous research, he applied communication techniques to analyze an optical wireless channel's information capacity, developed new techniques to reduce power requirements of optical wireless systems, and investigated the application of trellis-coded modulation for fiber-optic systems with optical amplifiers. His current research includes analysis and design of multicarrier CDMA systems, study of diversity combining performance under channel uncertainty, and application of hybrid-ARQ to fading channels.

Dr. You received the UC Regents Fellowship in 1997 and NASA Faculty Fellowship in 2004. He serves as faculty advisor for the IEEE student branch at NJIT.

Hong Li (S'04) received the B.S. degree in electrical engineering from Shanghai Jiao Tong University, Shanghai, China, in 1997, and the M.S. degree in communication and electronic systems from the Second Academy of China Aerospace Science and Industry Corp., Beijing, China, in 2000. He is currently working toward the Ph.D. degree at the New Jersey Institute of Technology (NJIT), Newark.

From 2000 to 2001, he was a R&D Engineer with Alcatel Shanghai Bell, Shanghai, China. Since 2001, he has been a Research Assistant with the Center for Communication and Signal Processing Research at the Department of Electrical and Computer Engineering, NJIT. His research interests include channel estimation, signal detection, and modulation classification.

Yeheskel Bar-Ness (M'69–SM'78–F'89) received the B.Sc. and M.Sc. degrees in electrical engineering from the Technion, Haifa, Israel, and the Ph.D. degree in applied mathematics from Brown University, Providence, RI.

Currently, he is a Distinguished Professor of Electrical and Computer Engineering, and the Foundation Chair of the Center for Communication and Signal Processing Research at the New Jersey Institute of Technology (NJIT), Newark. Previously, he was with the Rafael Armament Development Authority, Israel, working in the field of communications and control; and with the Nuclear Medicine Department, Elscint Ltd., Haifa, Israel, as a Chief Engineer in the field of control, and image and data processing. In 1973, he joined the School of Engineering, Tel-Aviv University, Tel Aviv, Israel, where he was an Associate Professor of Control and Communications. Between September 1978 and September 1979, he was a Visiting Professor with the Department of Applied Mathematics, Brown University. He was on leave with the University of Pennsylvania and Drexel University, both in Philadelphia, PA. He came to NJIT from AT&T Bell Laboratories in 1985. Between September 1993 and August 1994, he was on sabbatical with the Telecommunications and Traffic Control Systems Group, Faculty of Electrical Engineering, Delft University of Technology, Delft, The Netherlands. Between September 2000 and August 2001, he was on sabbatical at Stanford University, Stanford, CA. His current research interests include adaptive multiuser detection, array processing and interference cancellation, and wireless mobile and personal communications.

Dr. Bar-Ness was an Area Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS (Transmission Systems), and Editor for Adaptive Processing Systems. He is the Founder and Editor-in Chief for IEEE COMMUNICATIONS LETTERS. He is also Editor for the journal *Wireless Personal Communications*. He was Chairman of the Communication Systems Committee, and currently is the Vice Chair of the Communications Theory Committee of the IEEE Communication Society. He served as the General Chair of the 1994 and 1999 Communication Theory Mini-Conference. He was also the Technical Chair for the IEEE Sixth International Symposium on Spread Spectrum Techniques and Applications. He is a recipient of the Kaplan Prize (1973), which is awarded annually by the government of Israel to the ten best technical contributors.