A Decision-Feedback Channel Estimation Receiver for Independent Nonidentical Rayleigh Fading Channels

Yonglan Zhu, Pooi-Yuen Kam, and Yan Xin
National University of Singapore, 117576, Singapore
Email: {zhuyonglan, elekampy, elexy}@nus.edu.sg

Abstract—We derive a decision-feedback channel estimation receiver for independent and non-identically distributed (i.n.d.) Rayleigh fading channels. The receiver has memory to store signals received over the past several symbol intervals, and then use them to adaptively estimate instantaneous channel gains. The receiver adapts its decision rule based on estimates of the varying channel gains, and hence it is partially coherent. We also obtain the bit error rate (BER) of the channel estimation receiver for binary phase-shift keying (BPSK) and quadrature phase-shift keying (QPSK) modulations. The BER results are obtained in simple forms which explicitly show the effects of system parameters on the BER. In addition, our results are applicable to the cases where all the branches have arbitrary fade rates.

I. INTRODUCTION

In wireless communications, multi-path fading degrades the reliability of transmission links significantly, and becomes a bottleneck for enhancing data rate. Diversity reception is an efficient technique to mitigate the deleterious effects of fading without losing bandwidth efficiency. There are many methods for combining the received signals in diversity reception. When channel state information (CSI) of all the branches is perfectly known at the receiver, the maximal ratio combiner (MRC) is optimal [1], and this is called conventional MRC. However, channel estimate is always imperfect, and the estimation errors cannot be neglected. The effects of channel estimation errors and the bit error rate (BER) of the MRC system with weighting errors have been presented in [2], [3]. Taking account of channel estimation errors, an optimal diversity combining receiver which uses a pilot symbol based maximum likelihood (ML) channel estimation approach was proposed in [4].

To obviate channel estimation, non-coherent or differential modulations are considered. The optimum diversity combining receivers and their corresponding error performances have been presented in [5] for noncoherent frequency shift keying (FSK), and in [6] for differential phase-shift keying (DPSK) with differential detection. Both [5] and [6] deal with independent and non-identically distributed (i.n.d.) Rayleigh fading channels. The complex gain of a fading channel is usually modeled as an autorecorrelated random process. To make use of this channel autocorrelation, [7] proposed an adaptive receiver which is a decision-feedback channel estimation MRC, a generalization of the conventional MRC. The channel estimation MRC adapts its decision rule according to the varying channel conditions, and thereby improves the error performance.

In this paper, we further generalize the decision-feedback channel estimation MRC in [7] to the case of i.n.d. Rayleigh fading channels. It is assumed that the receiver has a memory to store the signals received over the past M symbol intervals. The receiver makes use of its memory to perform partially coherent detection. With decision-feedback, the receiver can estimate the instantaneous channel gains based on information retained in the memory and feedback decisions. The derived decision feedback channel estimation receiver has the structure of an estimator-detector. The past decisions and the received signals stored in the memory are used to obtain the minimum mean square error (MMSE) channel estimates. In detecting the current symbol, the signals received over the current symbol interval, the MMSE channel estimates and the MMSEs are used. Moreover, we derive the BER of the channel estimation receiver for i.n.d. Rayleigh fading channels with binary phase-shift keying (BPSK) and quadrature phase-shift keying (QPSK) modulations. The fading processes here are assumed to have an arbitrary Doppler spectrum with arbitrary Doppler bandwidth. The BER results are obtained in simple closed forms which depend on signal-to-noise ratio (SNR), the autocorrelation coefficients of the fading processes, the number of the diversity branches, the imbalance between the diversity branches, and the size of memory at the receiver.

II. SYSTEM MODEL

Consider that data is transmitted over L independent, non-selective Rayleigh fading channels with additive white Gaussian noise (AWGN). This is similar to a single input multiple output (SIMO) system. The signal transmitted during the k-th symbol interval is \( x[k] \), and the received signal at the i-th branch during the k-th symbol interval is given by

\[
y_i[k] = h_i[k]x[k] + n_i[k], \quad i = 1, 2, \ldots, L.
\]  

(1)

The quantity \( h_i[k] \) denotes the fading coefficient in the i-th branch during the k-th symbol interval. Each sequence, \( \{h_i[k]\}_k \), consists of samples of a zero-mean complex Gaussian stochastic process with autocorrelation function \( \mathbb{E}[h_i[k]h_i^*[k-l]] = \sigma_i^2R_i(l) \), where \( \sigma_i^2 \) is the average power of the fading process, and \( R_i(l) \) is the correlation coefficient at a time difference of \( l \). We assume that the channel fading processes in different transmission links are independent, i.e., \( \forall i \neq j \), we have \( \mathbb{E}[h_i[k]h_j^*[k]] = 0 \). The quantity \( n_i[k] \) is
We also express the received signals during the all channel fading processes. The transmit signals are subject branches. Thus, the density function (PDF) of over the past memory contains the sequence $z_i[k]$, i.e., $\sum_{i=1}^{L} p_i[k] = \mathbb{E}[|x[k]|^2] = E_x$ with $\mathbb{E}[-]$ denoting the expectation operator.

III. Decision-feedback Channel Estimation Receiver

The receiver has a memory which stores the signals received over the past $M$ symbol intervals. At the $k$-th symbol interval, the memory contains the sequence

$$z[k] = [z_1[k], z_2[k], \ldots, z_L[k]],$$

where $z_i[k]$ consists of the received signals at the $i$-th branch over the past $M$ symbol intervals, i.e.,

$$z_i[k] = [y_i[k-1], y_i[k-2], \ldots, y_i[k-M]].$$

We also express the received signals during the $k$-th symbol interval over the $L$ branches as a vector

$$y[k] = [y_1[k], y_2[k], \ldots, y_L[k]].$$

Since there are autocorrelations in fading processes and receiver has memory, it is natural to determine $x[k]$ based on $y[k]$ and $z[k]$, but not the signal $y[k]$ received over the $k$-th interval alone. Without loss of generality, we assume that the first $M$ transmit signals are $E_x$ which are the initial reference signals, i.e., $x[1] = x[2] = \cdots = x[M] = E_x$.

We further assume that all the symbol points in the signal constellation, $\mathcal{X} = \{S_i\}_{i=1}^{|X|}$, are equally likely. Thus, the maximum a posteriori probability (MAP) receiver is equivalent to ML receiver. Conditioned on $x[k] = S_i$, the joint probability density function (PDF) of $y[k]$ and $z[k]$ can be computed as

$$p(y[k], z[k] | x[k] = S_i) = p(z[k]) \prod_{i=1}^{L} p(y_i[k] | x[k] = S_i, z_i[k]),$$

where the equality is due to the independence between the previous received signals $z[k]$ and current transmit signals $x[k]$, and the mutual independence of $L$ branches. Since $p(z[k])$ is irrelevant to $S_i$, we only need to evaluate the term $p(y_i[k] | x[k] = S_i, z_i[k])$, and we have

$$p(y_i[k] | x[k] = S_i, z_i[k]) = \mathbb{E}_{h_i[k] | S_i, z_i[k]} \{p(y_i[k] | x[k] = S_i, h_i[k])\},$$

which follows the facts that the fading coefficient $h_i[k]$ is independent of $x[k]$, and the noise $n_i[k]$ is independent of $z_i[k]$. It is obvious that conditioned on $x[k] = S_i$ and $h_i[k], y_i[k]$ is a complex Gaussian random variable with mean $S_i h_i[k]$ and variance $N_i$. Whereas the PDF of $h_i[k]$ conditioned on $z_i[k]$ can be computed by

$$p(h_i[k] | z_i[k]) = \sum_{x_p[k]} p(h_i[k] | z_i[k], x_p[k]) p(x_p[k] | z_i[k]),$$

where $x_p[k]$ is the sequence of $M$ previous transmit signals, i.e., $x_p[k] = [x[k-1], x[k-2], \ldots, x[k-M]]$, and the summation is over all the possible realization of $x_p[k]$. The complexity of $p(h_i[k] | z_i[k])$ in (6) increases exponentially with an increase in the memory size $M$. We thus consider decision feedback to remove unknown data modulation. Suppose that at time $k$, the prior decisions $\hat{x}_p[k] = [\hat{x}[k-1], \hat{x}[k-2], \ldots, \hat{x}[k-M]]$ on $x_p[k]$ are correct and fed back to the receiver. Thus, we have $p(x_p[k] | h_i[k], \hat{x}_p[k]) = 1$ and (6) reduces to

$$p(h_i[k] | z_i[k]) = p(h_i[k] | z_i[k], \hat{x}_p[k]) = \frac{1}{\pi V_i^2[k]} \exp \left( - \frac{|h_i[k] - \hat{h}_i[k]|^2}{V_i^2[k]} \right).$$

Actually, the conditional mean $\hat{h}_i[k]$ is the MMSE estimate of $h_i[k]$ given the information $z_i[k]$ and $x_p[k]$, and the conditional variance $V_i^2[k]$ is the MMSE of the estimates $\hat{h}_i[k]$. Because of the Gaussian statistics of $p(h_i[k] | z_i[k])$, for each branch $i$, the estimates $\hat{h}_i[k]$ are computed linearly from the past received signals $z_i[k]$ and the past decisions $\hat{x}_p[k]$, and $V_i^2[k]$ are computable at time $k$. If the channel fading processes are Markov, $\hat{h}_i[k]$ can be computed recursively by a Kalman filter [7]. If the fading processes are non-Markov, $\hat{h}_i[k]$ is given by a Wiener filter for a finite memory $M$, and the equations for computing $\hat{h}_i[k]$ and $V_i^2[k]$ in this case are given in the Appendix.

Substituting (7) into (5) and noting that $p(y_i[k] | x[k] = S_i, h_i[k])$ is Gaussian, we can obtain the conditional joint PDF, $p(y[k], z[k] | x[k] = S_i)$ in (5), as

$$p(y[k], z[k] | x[k] = S_i) = \frac{p(z[k])}{\pi L} \times \prod_{i=1}^{L} \frac{1}{\sqrt{S_i^2 V_i^2[k] + N_i}} \exp \left( - \frac{|y_i[k] - S_i \hat{h}_i[k]|^2}{S_i^2 V_i^2[k] + N_i} \right).$$

Taking natural logarithm of $p(y[k], z[k] | x[k] = S_i)$, and ignoring the terms that are independent of $S_i$, we obtain the ML decision $\hat{x}[k]$ on $x[k]$ as

$$\hat{x}[k] = \min_{x} \{ \sum_{i=1}^{L} \frac{1}{(S_i^2 V_i^2[k] + N_i)^{-1}} |y_i[k] - S_i \hat{h}_i[k]|^2 + \ln \left( (S_i^2 V_i^2[k] + N_i) \right) \}.$$
the estimator for computing the conditional mean estimates \( \{ \hat{h}_i[k] \}_{i=1}^{L} \) of the channel gains \( \{ h_i[k] \}_{i=1}^{L} \), and the MMSE \( \{ V_i^2[k] \}_{i=1}^{L} \).

For PSK modulation, \( |S_i|^2 \) is equal to \( E_s \). In this situation, after further ignoring the terms that are independent of \( S_i \), the ML decision rule in (10) reduces to

\[
\hat{x}[k] = \max_{x} \left\{ \sum_{i=1}^{L} w_i \Re (y_i[k] \hat{h}_i[k] S_i) \right\},
\]

where the weighting factors are \( w_i = (E_s V_i^2[k] + N_i)^{-1} \), and \( \hat{h}_i[k] \) and \( V_i^2[k] \) are given by (8).

**Remark:** For i.n.d. channels, the channel estimation receiver in (11) can be interpreted as a generalized MRC or a generalized differentially coherent receiver. Firstly, when the channel estimates are perfect, i.e., \( \hat{h}_i[k] = h_i[k] \) and \( V_i^2[k] = 0 \), the channel estimation receiver (11) is reduced to the conventional MRC in which the receiver has perfect knowledge of the channel gains \( \{ h_i[k] \}_{i=1}^{L} \). Secondly, when the memory size of the receiver is one, i.e., \( M = 1 \), the estimate \( \hat{h}_i[k] \) needs to be generated only from the signal \( y_i[k-1] \) received in the immediate prior interval. Applying the orthogonality principle, we can obtain

\[
\begin{align*}
\hat{h}_i[k] &= \frac{\sigma_i^2 R_i(1) \hat{x}_i[k-1]}{E_s \sigma_i^2 + N_i} y_i[k-1], \\
V_i^2[k] &= \sigma_i^2 - \sigma_i^2 E_s \sigma_i^2 |R_i(1)|^2, \\
&\quad E_s \sigma_i^2 + N_i.
\end{align*}
\]

By substituting the above \( \hat{h}_i[k] \) and \( V_i^2[k] \) into (11), the channel estimation receiver (11) is reduced to the differential receiver given in [6, Eq. (6)] and [8, Eq. (13)] for i.n.d. Rayleigh fading channels.

**IV. ERROR PROBABILITY ANALYSIS**

In this section, we derive the exact BER and tight Chernoff bounds on the BER for BPSK and QPSK modulations for the channel estimation receiver (11).

**A. Exact BER**

For equally likely transmit signals, we can assume \( x[k] = E_s^{1/2} \) without loss of generality. The BER for detecting \( x[k] \) involves the following probability for some angle \( \alpha \)

\[
F(\alpha) = P\left( \sum_{i=1}^{L} w_i \Re (y_i[k] \hat{h}_i[k] e^{-j\alpha}) < 0 \mid x[k] = E_s^{1/2} \right).
\]

With Gray mapping, the BER performance of BPSK and QPSK modulations is given, respectively, by [8]

\[
P_b, \text{BPSK} = F(0), \quad P_b, \text{QPSK} = F(\pi/4).
\]

Conditioned on \( x[k] = E_s^{1/2} \) and \( \hat{h}_i[k] e^{j\alpha} = u_i \), the quantity \( y_i[k] \) is conditionally complex Gaussian distributed with mean \( E_s^{1/2} u_i e^{-j\alpha} \) and variance \( (E_s V_i^2[k] + N_i) \), and thus the quantity \( \sum_{i=1}^{L} w_i \Re (y_i[k] \hat{h}_i[k] e^{-j\alpha}) \) is also Gaussian distributed with mean \( E_s^{1/2} \cos \alpha \sum_{i=1}^{L} w_i |u_i|^2 \) and variance \( 0.5 \sum_{i=1}^{L} w_i^2 (E_s V_i^2[k] + N_i) |u_i|^2 = 0.5 \sum_{i=1}^{L} w_i |u_i|^2 \). Therefore, the probability \( F(\alpha) \) conditioned on \( \hat{h}_i[k] e^{j\alpha} = u_i \) is now readily given by

\[
F(\alpha \mid \hat{h}_i[k] e^{j\alpha} = u_i) = \frac{1}{2} \text{erfc} \left( E_s \cos^2 \alpha \sum_{i=1}^{L} w_i |u_i|^2 \right)^{1/2}.
\]

Let \( g_i = w_i |u_i|^2 = w_i |\hat{h}_i[k]^2|, \) and \( g = \sum_{i=1}^{L} g_i \). Since \( \hat{h}_i[k] \) is the estimate of \( h_i[k] \), \( \hat{h}_i[k] \) is a linear combination of the past received signals, and hence \( \hat{h}_i[k] \) is a complex Gaussian random variable with mean zero and variance \( (\sigma_i^2 - V_i^2[k]) \).

It is obvious that \( g_i \) is exponential distributed, and its PDF function is given by

\[
f_{g_i}(g_i) = \frac{1}{\lambda_i} e^{-g_i/\lambda_i},
\]

where \( \lambda_i = w_i (\sigma_i^2 - V_i^2[k]) \). Since the diversity branches are independent, \( \{ g_i \}_{i=1}^{L} \) are independent exponential random variables. With the assumption that \( \lambda_i \) are distinct, the PDF of \( g \) can be obtained as [8]

\[
f_g(g) = \sum_{i=1}^{L} \frac{A_i}{\lambda_i} e^{-g/\lambda_i},
\]

where \( A_i \) is given by

\[
A_i = \prod_{j=1}^{L, j \neq i} \frac{\lambda_j}{\lambda_i - \lambda_j}.
\]

By averaging the conditional probability (14) over \( g \) using its PDF (15), the probability \( F(\alpha) \) is obtained as

\[
F(\alpha) = E_{u_i} \left[ F(\alpha \mid \hat{h}_i[k] e^{j\alpha} = u_i) \right] = \frac{1}{2} E_g [\text{erfc}(g E_s \cos^2 \alpha)^{1/2}]
\]

\[
= \sum_{i=1}^{L} \frac{A_i}{2} \left[ 1 - \left( \frac{\lambda_i E_s \cos^2 \alpha}{\lambda_i E_s \cos^2 \alpha + 1} \right)^{1/2} \right].
\]

With the expression of \( F(\alpha) \), we can easily compute the BER according to (13).

**Remark:** The expression of \( F(\alpha) \) in (17) has the same form as [8, Eq. (29)] except for the expression of \( \lambda_i \). When the memory of the receiver is one, i.e., \( M = 1 \), is reduced to [8, Eq. (29)].

**B. Chernoff Bound on BER**

By using the Chernoff bound \( \text{erfc}(x)^{1/2} < e^{-x} \), the probability \( F(\alpha) \) conditioned on \( \hat{h}_i[k] e^{j\alpha} = u_i \) in (14) is upper bounded as

\[
F(\alpha \mid \hat{h}_i[k] e^{j\alpha} = u_i) < \frac{1}{2} e^{-g_i E_s \cos^2 \alpha} = \frac{1}{2} \prod_{i=1}^{L} e^{-g_i E_s \cos^2 \alpha}.
\]

Averaging the above upper bound over the PDF of \( g_i \), and noting that \( \{ g_i \}_{i=1}^{L} \) are independent, the upper bound on \( F(\alpha) \) can be obtained as

\[
F(\alpha) < \frac{1}{2} \prod_{i=1}^{L} \left[ 1 - \frac{1}{\lambda_i E_s \cos^2 \alpha + 1} \right].
\]
Remark: This result is consistent with [8, Eq. (40)] except for the expression of $\lambda$. When the memory of the receiver is one, i.e., $M = 1$, (18) reduces to [8, Eq. (40)].

V. NUMERICAL RESULTS

In this section, we present simulation results for the i.n.d. systems with BPSK and QPSK modulations. The autocorrelation coefficients of the fading processes are assigned according to Jake’s model, i.e., $R_i(t) = J_0(2\pi f DT_s t)$, $\forall i = 1, 2, \ldots, L$, where $f_D$ is the maximum Doppler frequency and $T_s$ is the symbol duration. We assume $f_D = 100$ Hz, $T_s = 5 \times 10^{-3}$ s, and all the branches have the same noise power, i.e., $N_0 = N_i$. We compare the BER performance of the channel estimation receiver with the ones of conventional MRC.

In Fig. 1, we show the BER of an i.n.d. system for various channel statistics. It can be observed that the nonidentical distribution of channel gains among diversity branches degrades the BER performance from the case of independent and identically distributed (i.i.d.) branches. Moreover, the BER of the channel estimation receiver has an error floor at high SNR. This is due to an irreducible MMSE in channel estimates which results from the fluctuations in the fading process. In Fig. 2, we illustrate the effects of memory size $M$ on the BER. We can see that with an increase in the receiver memory size, the BER of the adaptive receiver decreases at the expense of an increase in the receiver complexity.

VI. CONCLUSIONS

In this paper, we proposed the adaptive diversity combining receiver with memory for i.n.d. Rayleigh fading channels system. The adaptive receiver makes use of the channel continuity, and has an estimator-detector structure. Furthermore, we derived the BER of the adaptive receiver for BPSK and QPSK modulations.

APPENDIX

In this appendix, we derive the MMSE estimate of $\hat{h}_i[k]$ based on $z_i[k]$ and the assumption that $\hat{a}_i[k]$ are correct and fed back to receiver at time $k$. We also compute the MMSE $V_i^2[k]$. Since $h_i[k]$ is Gaussian distributed, by applying the orthogonality principle, the MMSE estimate $\hat{h}_i[k]$ is given by

$$\hat{h}_i[k] = c_i[k](G_i[k])^{-1} z_i^T[k],$$

where $c_i[k](G_i[k])^{-1}$ represents a Wiener filter with $c_i[k] = \mathbb{E}\left\{h_i[k]z_i^*[k]\right\}$ and $G_i[k] = \mathbb{E}\{z_i^T[k]z_i^*[k]\}$. The $m$-th entry of $c_i[k]$ is

$$c_i[k] = \sigma_i^2 R_i(m) \hat{\alpha}^*[k-m],$$

and the $(m, n)$-th entry of $G_i[k]$ is

$$G_i[k] = \left\{\begin{array}{ll}
\sigma_i^2 R_i(n-m) \hat{\alpha}[k-m] \hat{\alpha}^*[k-n], & m \neq n; \\
\sigma_i^2 R_i(0) \hat{\alpha}^*[k-m]^2 + N_i, & m = n.
\end{array}\right.$$  

Since the channel estimation error, $(\hat{h}_i[k] - h_i[k])$, is independent of $\hat{h}_i[k]$, we can compute the MSE as

$$V_i^2[k] = \sigma_i^2 R_i(0) - c_i[k](G_i[k])^{-1} c_i^H[k].$$

REFERENCES


