

The Effect of Channel Estimator Memory Mismatch on the Performance of MLSE in Wireless Data Communications

Yumin Lee Jyh-Hau Chen

Dept. of Electrical Engineering and
Graduate Institute of Communication Eng.,
National Taiwan University, Taipei 10617, Taiwan

Abstract –The maximum likelihood sequence estimator (MLSE) has been shown to be an effective measure against inter-symbol interference. An estimate of the discrete-time equivalent channel is usually required in order to perform sequence estimation. In practice, the length (or memory) of the channel is unknown to the receiver for wireless communication channels. In this paper, upper and lower bounds for the performance loss due to channel estimator memory mismatch are derived for unbiased estimators, while approximate expressions for the irreducible BER floor caused by channel estimator memory mismatch are derived for biased estimators. Simulation results are also provided for random channels to verify the correctness of the analysis.

I. INTRODUCTION

Equalization [1] has been known to be an effective measure against inter-symbol interference (ISI) in wireless data communication systems. A popular class of equalizers is known as the channel-estimation-based equalizers[2]. Various analytical[4,5,6] and simulation[2,7] studies all indicate that the performance of channel-estimation-based equalizers is very sensitive to the accuracy of the channel estimate.

For wireless communications, the actual length of the channel is generally location-dependent and unknown. Since the length of the channel estimator in a receiver is usually fixed, *channel estimator memory mismatch* inevitably occurs and is also a source of channel estimation error. When the length of the channel estimator is shorter than the actual channel, the channel estimate is always biased. This case is referred to as under-estimation in this paper. On the other hand, when the length of the channel estimator is longer than the actual channel, the channel estimate may or may not be biased. This case is referred to as over-estimation. Qualitatively, an irreducible BER floor exists if the channel estimate is biased as will be discussed later in this paper. On the other hand, if the channel estimate is unbiased (which implies over-estimation), there will be no BER floor. However, a performance loss may still be incurred due to over-estimation.

In this paper, we investigate the effect of channel estimator memory mismatch (under-estimation and over-estimation) on the BER performance of the channel-estimation-based MLSE. An analytical expression is first derived for the BER of an MLSE with biased, noisy channel estimate that may be under- or over-estimated. We next show that an unbiased over-estimated channel estimator does not cause irreducible BER floor, however it still incurs a performance loss that

increases with the extent of over-estimation. A formula and upper and lower bounds for this performance loss are also obtained. Furthermore, we also show that channel estimation bias causes an irreducible BER floor in the MLSE, and approximate formulas for the BER floor are also derived. Computer simulation results are also provided for the least-squares channel estimator to verify the correctness of the analysis.

II. CHANNEL ESTIMATION

Consider the linear modulation transmission system in Fig. 1. The discrete-time signal at the input of the MLSE is given by

$$y_k = \sum_{j=0}^v h_j x_{k-j} + z_k, \quad (1)$$

where $h_j, j=0,1,\dots,v$, is the impulse response of the baseband discrete-time equivalent channel of the system, x_k is the k -th complex modulation symbol, and z_k are the samples of complex Gaussian noise. The integer v is referred to as the channel memory. It is assumed that the modulation symbols are zero-mean, independent and identically distributed (i.i.d.) random variables with average symbol energy given by

$$E_X \equiv E[|x_k|^2] \quad (2)$$

where $E[\bullet]$ denotes statistical expectation, and that symbols in the signal constellation are equally likely. It is also assumed that the noise samples z_k are independent of x_k and are i.i.d. zero-mean circularly symmetric complex Gaussian random variables with variance given by

$$\sigma^2 \equiv E[|z_k|^2] \quad (3)$$

Finally, it is assumed that v and h_j are deterministic unless otherwise stated.

In practice, v and h_j are unknown to the receiver, therefore

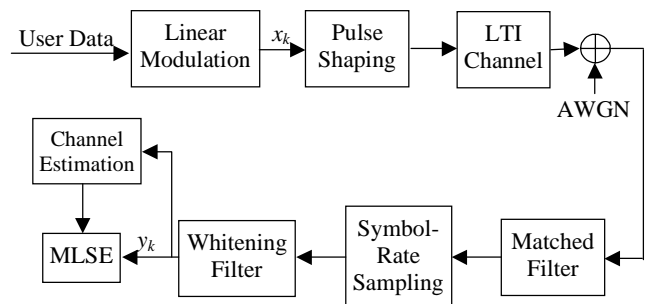


Fig. 1. Transmission system using MLSE in the receiver.

the channel estimator assumes that the channel memory is v' and computes the channel estimates $\hat{h}_j, j = 0, 1, \dots, v'$ from the received signal. In order to accommodate both over-estimation ($v' \geq v$) and under-estimation ($v' < v$), in this paper the channel vector and channel estimate vector are defined as $(\mu+1) \times 1$ vectors given by

$$\mathbf{p} = [p_0 \quad \dots \quad p_\mu]^T \quad (4)$$

and

$$\hat{\mathbf{p}} = [\hat{p}_0 \quad \dots \quad \hat{p}_\mu]^T, \quad (5)$$

respectively, where $\mu = \max(v, v')$,

$$p_j \equiv \begin{cases} h_j, & j \leq v \\ 0, & \text{else} \end{cases}, \quad (6)$$

and

$$\hat{p}_j \equiv \begin{cases} \hat{h}_j, & j \leq v' \\ 0, & \text{else} \end{cases}. \quad (7)$$

We also define the bias vector as

$$\mathbf{b} = \mathbb{E}[\hat{\mathbf{p}} - \mathbf{p}] \quad (8)$$

and zero-mean estimation error vector as

$$\mathbf{e} = \hat{\mathbf{p}} - \mathbf{p} - \mathbf{b}. \quad (9)$$

The definitions in (4) to (9) are depicted in Fig. 2 for both under-estimation and over-estimation. Note that although \mathbf{e} is a $(\mu+1) \times 1$ vector, e_j is random only for $0 \leq j \leq v'$. We assume that $\mathbf{e}(1:v'+1)$ is a random vector¹ with covariance matrix Σ_e .

One of the most well-known is the least-squares (LS) algorithm[3]. The LS channel estimate is given by [3]

$$\hat{\mathbf{h}} \equiv \begin{bmatrix} \hat{h}_0 \\ \vdots \\ \hat{h}_{v'} \end{bmatrix} = (\mathbf{X}_1^H \mathbf{X}_1)^{-1} \mathbf{X}_1^H \begin{bmatrix} y_{Q-1} \\ \vdots \\ y_0 \end{bmatrix} \quad (10)$$

where $\mathbf{X}_1 = \mathbf{X}(:, 1:v'+1)$,

$$\mathbf{X} \equiv \begin{bmatrix} X_{Q-1} & X_{Q-2} & \dots & X_{Q-\mu-1} \\ \vdots & \vdots & & \vdots \\ X_0 & X_{-1} & \dots & X_{-\mu} \end{bmatrix}, \quad (11)$$

and Q is the number of training symbols. For over-estimation ($v' \geq v$), it can be shown that $\mathbf{b} = \mathbf{0}$ and

$$\Sigma_e = \sigma^2 (\mathbf{X}_1^H \mathbf{X}_1)^{-1}. \quad (12)$$

On the other hand, for under-estimation ($v' < v$) it can be shown that Σ_e is the same as (12) and

$$\mathbf{b} = \begin{bmatrix} (\mathbf{X}_1^H \mathbf{X}_1)^{-1} \mathbf{X}_1^H \mathbf{X}_2 \\ -\mathbf{I}_{v-v'} \end{bmatrix} \begin{bmatrix} p_{v'+1} \\ \vdots \\ p_v \end{bmatrix}, \quad (13)$$

where $\mathbf{X}_2 = \mathbf{X}(:, v'+2:v+1)$. Furthermore, for both cases

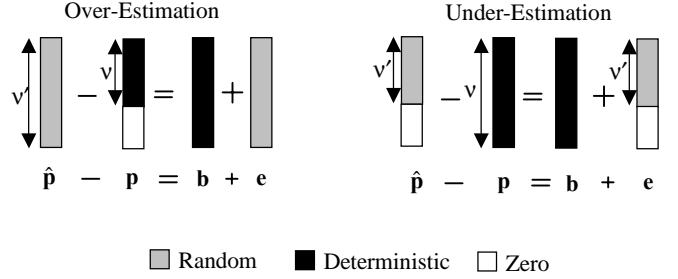


Fig. 2. The relation between the various vectors for over- and under-estimation.

$\mathbf{e}(1:v'+1)$ is a circularly symmetric zero-mean Gaussian random vector.

III. ANALYSIS OF CHANNEL-ESTIMATION-BASED MLSE

The channel-estimated-based MLSE observes a sequence of N received signal samples $y_0 \dots y_{N-1}$ and finds an estimated sequence of modulated symbols $\hat{\mathbf{x}} = \hat{x}_0 \dots \hat{x}_{N-1}$ that minimizes, over all possible length- N modulation symbol sequences \mathbf{s} , the quantity

$$d(\mathbf{s}) = \sum_{j=0}^{N-1} \left| y_j - \sum_{l=0}^{v'} \hat{p}_l s_{j-l} \right|^2. \quad (14)$$

The error rate of the channel-estimation-based MLSE using channel estimate $\hat{\mathbf{p}}$ can be derived using Forney's union-bounding technique[8].

Let \mathbf{x} denote the transmitted sequence, and let $A_m(\mathbf{x})$ be the set of all paths through the trellis that diverge from and remerge for the first time with \mathbf{x} at times k and $k+m$, respectively, where $0 \leq k \leq N-1$ and $m \geq v'+1$. Define an error event of length m as the event that some member of $A_m(\mathbf{x})$ has a metric smaller than $d(\mathbf{x})$. Let $\mathbf{\epsilon}$ be an error sequence such that $\mathbf{x} + \mathbf{\epsilon}$ is a valid sequence of modulated symbols, and let

$$\mathbf{\epsilon}_m \equiv [\epsilon_{k+m-1} \quad \dots \quad \epsilon_{k-v'}]^T \quad (15)$$

be a vector representing the subsequence of $\mathbf{\epsilon}$ from time $(k-v')$ to $(k+m-1)$. The probability that $\mathbf{\epsilon}$ characterizes an error event of length m is given by

$$\Pr[\mathbf{\epsilon}_m] = \sum_{\mathbf{x}: \mathbf{x} + \mathbf{\epsilon} \in A_m(\mathbf{x})} \Pr[\mathbf{x}] \Pr[d(\mathbf{x} + \mathbf{\epsilon}) < d(\mathbf{x}) | \mathbf{x} \text{ transmitted}] \quad (16)$$

where $\Pr[\mathbf{x}]$ is the probability of transmitting \mathbf{x} . We can show that

$$\Pr[\mathbf{\epsilon}_m] \approx \sum_{\mathbf{x}: \mathbf{x} + \mathbf{\epsilon} \in A_m(\mathbf{x})} \Pr[\mathbf{x}] \times \mathcal{Q} \left(\frac{\|\hat{\mathbf{a}}_m\|}{\sqrt{2}\sigma} \left(1 + \frac{2 \operatorname{Re}[\hat{\mathbf{a}}_m^H \mathbf{B}_m \mathbf{x}_m]}{\|\hat{\mathbf{a}}_m\|^2} \right) \left(1 + \frac{(v'+1) \hat{\mathbf{a}}_m^H \mathbf{\Gamma}_m \hat{\mathbf{a}}_m}{\|\hat{\mathbf{a}}_m\|^2} \right)^{-1/2} \right) \quad (17)$$

¹ For a vector \mathbf{x} , $\mathbf{x}(a:b)$ denotes a sub-vector consisting of the a -th to the b -th elements of \mathbf{x} . Similarly, for a matrix \mathbf{A} , $\mathbf{A}(:,a:b)$ denotes a sub-matrix consisting of the a -th to the b -th columns of \mathbf{A} .

where

$$\hat{\mathbf{a}}_m \equiv \hat{\mathbf{P}}_m \mathbf{\epsilon}_m, \quad (18)$$

$$\hat{\mathbf{P}}_m = \begin{bmatrix} \hat{p}_0 & \cdots & \hat{p}_{v'} \\ & \hat{p}_0 & \cdots & \hat{p}_{v'} \\ & & \ddots & \ddots \\ & & & \hat{p}_0 & \cdots & \hat{p}_{v'} \end{bmatrix}, \quad (19)$$

\mathbf{B}_m is a convolution matrices associated with the bias vector \mathbf{b} with similar definition in (19), and

$$\mathbf{\Gamma}_m = \frac{E_X}{\sigma^2} \begin{bmatrix} \langle \Sigma_{\mathbf{e}} \rangle & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad (20)$$

where $\langle \Sigma_{\mathbf{e}} \rangle$ denotes a Toeplitz matrix of the same dimensions as $\Sigma_{\mathbf{e}}$ whose elements are obtained by summing the entries of $\Sigma_{\mathbf{e}}$ on the corresponding diagonals and dividing the result by the number of elements on the main diagonal. Note that the first term in the argument of the Q function characterizes the performance of an MLSE if the channel were indeed $\hat{\mathbf{p}}$, whereas the second and third terms correspond to the additional performance loss due to bias and noise.

We now present some special cases.

A. Ideal Channel Estimator

For an ideal channel estimator, we have $v' = v = \mu$ and $\mathbf{B}_m = \mathbf{\Gamma}_m = \mathbf{0}$. In this case (17) degenerates to

$$\Pr[\mathbf{\epsilon}_m] = \Pr[\mathbf{x} + \mathbf{\epsilon} \in A_m(\mathbf{x})] Q\left(\frac{\|\mathbf{P}_m \mathbf{\epsilon}_m\|}{\sqrt{2}\sigma}\right), \quad (21)$$

where \mathbf{P}_m is the $m \times (m+v)$ convolution matrix associated with the channel vector \mathbf{p} with definition similar to (19). This result is consistent with the result in [8].

B. Unbiased Channel Estimator

For an unbiased channel estimator we have $v' \geq v$ and $\mathbf{B}_m = \mathbf{0}$. We can show that an unbiased, over-estimated channel estimator incurs an asymptotic performance loss with respect to (21). The performance loss generally increases with v' if the nonzero diagonal elements of $\mathbf{\Gamma}_m$ remain approximately constant. This is intuitively appealing because as v' increases, the channel estimate contains more unnecessary taps that are purely noise. These taps introduce additional noise into the MLSE and cause performance degradation. We can show that the performance loss can be bounded by

$$10 \log_{10}(1 + (v'+1)\lambda_{\min}) \leq \gamma \leq 10 \log_{10}(1 + (v'+1)\lambda_{\max}) \quad (22)$$

where λ_{\min} and λ_{\max} are, respectively, the smallest and largest eigenvalues of the matrix $\mathbf{\Gamma}_m$. Finally, substituting $v = v'$ into (22) gives results that were previously derived in [5].

C. Biased Channel Estimator

For a biased noisy estimator, the error event probability is given by (17). Assuming that σ^2 approaches 0, and that μ

is sufficiently large, it can be shown that the probability of bit error is approximated by

$$P_b \approx \frac{K_{\min}}{\log_2 M} Q\left(\frac{\delta_{\min}(\mathbf{b})}{\sqrt{2E_X}}\right), \quad (23)$$

where M is the size of the signal constellation,

$$\delta_{\min}(\mathbf{b}) \equiv \min_{\mathbf{\epsilon}_m: \mathbf{x} + \mathbf{\epsilon}_m \in A_m(\mathbf{x})} \delta(\hat{\mathbf{a}}_m; \mathbf{b}), \quad (24)$$

$$K_{\min} = \sum_{\mathbf{\epsilon}_m \in E_{\min}} w(\mathbf{\epsilon}_m) \Pr[\mathbf{x} + \mathbf{\epsilon} \in A_m(\mathbf{x})], \quad (25)$$

and $w(\mathbf{\epsilon}_m)$ is the number of nonzero components of $\mathbf{\epsilon}_m$. In (25) E_{\min} is the set of $\mathbf{\epsilon}_m$ for which $\delta(\hat{\mathbf{a}}_m; \mathbf{b}) = \delta_{\min}(\mathbf{b})$.

It has been observed that for most channels the minimum value of $\delta(\hat{\mathbf{a}}_m; \mathbf{b})$ over error events of length $m = v'+1$ is very close to $\delta_{\min}(\mathbf{b})$. Furthermore, $\hat{\mathbf{p}} \approx \mathbf{p} + \mathbf{b}$ when the channel SNR is high. Thus an approximate formula for the irreducible BER floor caused by channel estimation bias is

$$P_b \approx \frac{1}{\log_2 M} Q\left(\frac{d_{\min} \|\mathbf{p}(1:v'+1) + \mathbf{b}(1:v'+1)\|^2}{\sqrt{2E_X} \|\mathbf{B}_{v'+1}^H (\mathbf{p}(1:v'+1) + \mathbf{b}(1:v'+1))\|}\right), \quad (26)$$

where d_{\min} is the minimum distance between two constellation points.

IV. SIMULATION RESULTS

The block diagram of the simulated system is shown in Fig. 3. Simulation results are obtained using bursts of quaternary phase shift keying (QPSK) symbols with average energy E_X . Each burst consists of training symbols followed by information symbols. The wireless channel is modeled as a random time-invariant discrete-time equivalent channel with memory v corrupted by the discrete-time complex AWGN samples with variance σ^2 . Channels are assumed to be independent from burst to burst. Each instance of the random channel is normalized so that the sum of magnitude-squares of the channel taps averages to 1. At the receiver, LS channel estimation with memory v' is performed

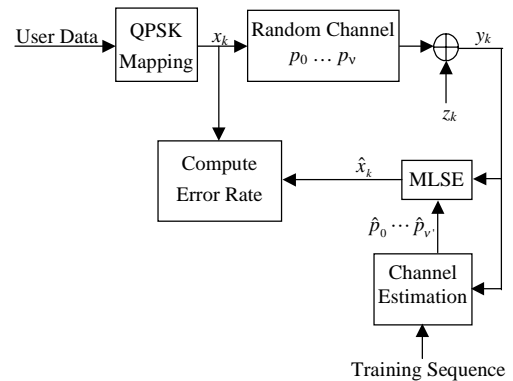


Fig. 3. Simulation system block diagram.

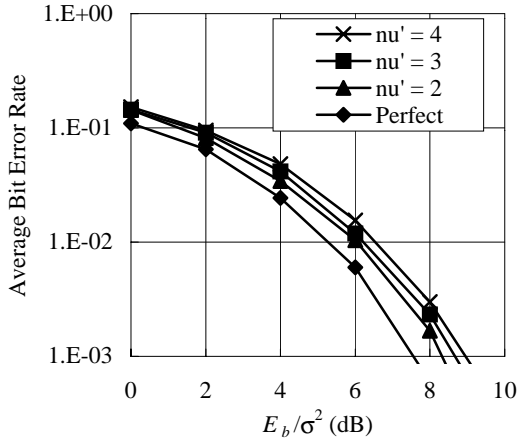


Fig. 4. Average BER performance of LS channel estimators with over-estimation ($v' \geq v$) obtained by computer simulation for random channels with $v = 2$.

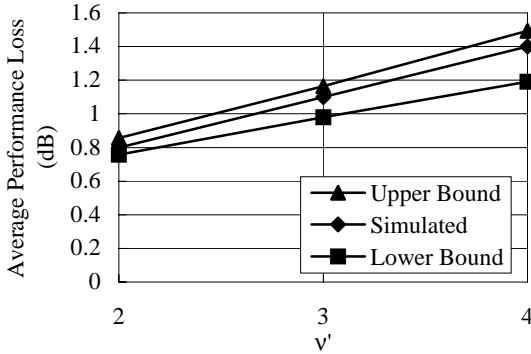


Fig. 5. Simulated and analytical bounds of the average performance loss of LS channel estimators with over-estimation ($v' \geq v$) for random channels with $v = 2$. For the simulated case the performance loss is measured at 10^{-3} average BER.

according to (10). MLSE is then performed on the remaining information symbols using the LS channel estimate, and the number of bit errors is tabulated.

The effect of over-estimation ($v' \geq v$) is shown in Fig. 4. Here the average BER of the simulated system is plotted as functions of E_b/σ^2 , where $E_b = E_x/2$ is the energy per transmitted bit. The random discrete-time equivalent channel has memory $v = 2$, with the channel taps being i.i.d. circularly symmetric zero-mean complex Gaussian random variables. LS channel estimators with $v' = 2, 3$, and 4 are simulated according to (10). There are a total of 25 training symbols consisting of repetitions of a pseudo-random (PN) sequence with period 15 mapped onto two antipodal QPSK signal points. The number of observations used for channel estimation is set to $Q = 15$. The performance of the ideal channel estimator is also plotted as a baseline for comparison.

It can be seen from Fig. 4 that the performance indeed

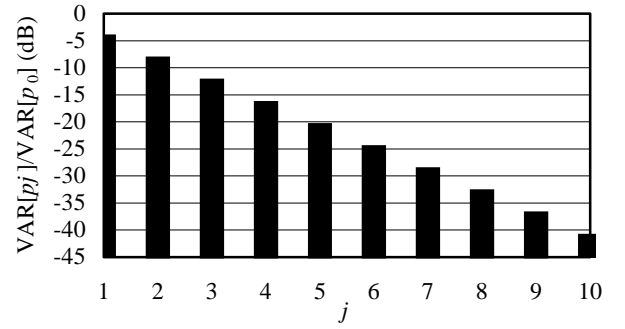


Fig. 6. The relative variance of the random channel taps for simulating under-estimation.

degrade as v' increases. The approximate average performance loss with respect to the ideal channel estimator is read off of Fig. 4 and plotted against v' in Fig. 5 together with the upper and lower bounds given by (22). It can be seen the simulated performance loss matches very well with the bounds.

The effect of under-estimation ($v' < v$) is next simulated. Here the discrete-time equivalent channel has memory $v = 10$, with each tap being independent zero-mean circularly symmetric complex Gaussian distributed, with variances that decay exponentially as shown in Fig. 6. LS channel estimators with $v' = 1, 2, 3$, and 4 are again simulated using the same training sequence as the over-estimation case. The BER for one instance of the random channel is plotted in Fig. 7 as functions of E_b/σ^2 . It can be seen that a BER floor indeed exists as predicted. It is found that for this particular instance of the random channel, (23) and (26) give the same values for all v' . These values are also plotted in Fig. 8 as a function of v' together with the simulated BER at 20 dB. Comparing the two curves, it can be seen that the analytical

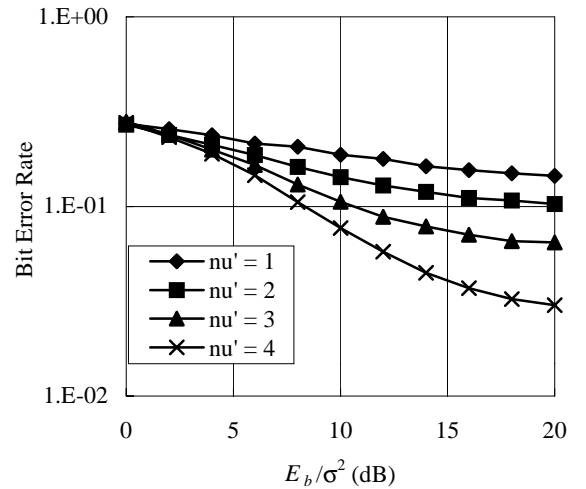


Fig. 7. Simulated BER performance of LS channel estimators with under-estimation ($v' < v$) for one particular instance of a random channel with $v = 10$.

results agree very well with the simulation results at least for this particular instance of the random channel.

Using the same channel model as in Figs. 7 and 8, the average irreducible BER floor is simulated at $E_b/\sigma^2 = \infty$. The result is plotted against v' in Fig. 9. The BER floor predicted by (23) and (26) is also evaluated for each burst, and the average over all bursts are also plotted in Fig. 9. It can be seen that (23) and (26) indeed give very close average values, with (26) being slightly lower as expected. Furthermore, it can be seen once again that the average BER floors predicted by (23) and (26) agree very well with the simulation results.

V. CONCLUSION

The effect of channel estimator memory mismatch on the performance of MLSE is investigated in this paper. It is found that as long as the channel estimate is unbiased, channel memory mismatch does not cause irreducible BER floor. However, a performance loss is still incurred. Upper and lower bounds for the performance loss with respect to an ideal channel estimator are derived in (22). Furthermore, it is also found that MLSE exhibits an irreducible BER floor when the channel estimate is biased, e.g., when the channel estimator is too short. Two approximate expressions, (23) and (26), are also obtained for the irreducible BER floor of biased channel estimators.

The analytical expressions are verified by computer simulation using the LS channel estimator for random channels. For over-estimation, simulation results show that the performance loss with respect to an ideal channel estimator agree very well with the analytical bounds in (22). On the other hand, for under-estimation, simulation results also show that the irreducible BER floors agree very well with the approximate analytical expressions (23) and (26).

REFERENCES

- [1] S. Qureshi, "Adaptive Equalization," *Proceedings of IEEE*, Vol. 73, No. 9, pp. 1349 – 1387, September 1985.
- [2] Y. Lee and D. C. Cox, "Adaptive Equalization with Regularization for Indoor Wireless Data Communications," *IEEE Global Telecommunications Conference*, November 1997.
- [3] Simon Haykin, *Adaptive Filter Theory*, 3rd Ed., Prentice Hall, 1996.
- [4] D. Dzung, "Error Probability of MLSE Equalization Using Imperfect Channel Measurements," 1991 *IEEE Int. Conf. on Commun.*, pp. 558 – 562.
- [5] A. Gorokhov, "On the Performance of the Viterbi Equalizer in the Presence of Channel Estimation Errors," *IEEE Signal Processing Letters*, Vol. 5, No. 12, pp. 321 – 324, December 1998.
- [6] M.-C. Chiu and C.-c. Chao, "Analysis of LMS-Adaptive MLSE Equalization on Multipath Fading Channels," *IEEE Trans. On Communications*, Vol. 44, No. 12, pp. 1684 – 1692, December 1996.
- [7] H. -N. Lee and G. J. Pottie, "Channel Estimation Based Adaptive Equalization/Diversity Combining for Time-Varying Dispersive Channels," 1997 *IEEE Vehicular Tech. Conference*, pp. 884 – 888.
- [8] G. D. Forney, Jr., "Maximum-Likelihood Sequence Estimation of Digital Sequences in the Presence of Intersymbol Interference," *IEEE Trans. On Inf. Th.*, Vol. 18, pp. 363 – 378, May 1972.

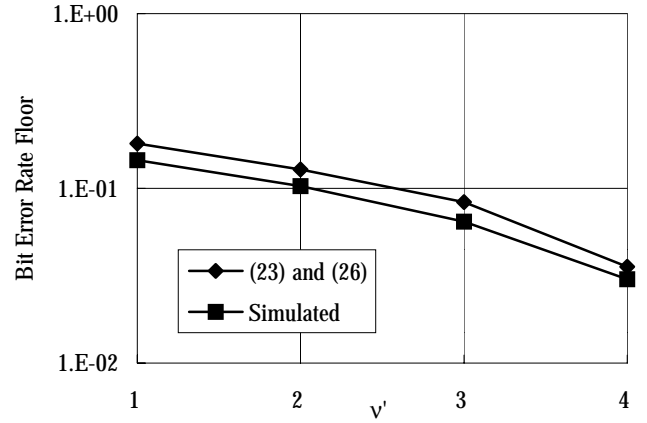


Fig. 8. Simulated and analytical BER floor of LS channel estimators with under-estimation ($v' < v$) for the channel in Figure 6. For the simulated case the BER floor is measured at E_b/σ^2 of 20 dB.

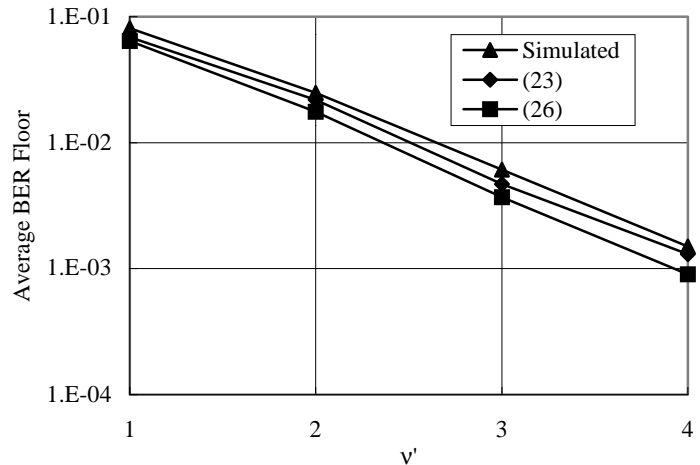


Fig. 9. Simulated and analytical average BER floors of LS channel estimators with under-estimation ($v' < v$) for one particular instance of the random channel.