ASYMPTOTIC EFFICIENCY OF THE BAD ALGORITHM

J.K. Nelson, A.C. Singer

University of Illinois at Urbana-Champaign
Coordinated Science Laboratory
1308 W. Main St., Urbana, IL 61801

U. Madhow

University of California, Santa Barbara
Electrical and Computer Engineering
Santa Barbara, CA 93106

ABSTRACT

We study the high-SNR performance of the bi-directional arbitrated DFE (BAD) equalization algorithm. Using asymptotic efficiency analysis, we characterize the performance of BAD as the channel noise variance approaches zero. We consider the case in which the channel is perfectly known to the receiver, as well as the case in which the channel estimate is in error. Asymptotic efficiency analysis allows us to quantify the loss in performance due both to inter-symbol interference (ISI) and imperfect channel estimation. The asymptotic efficiency of BAD is compared to that of a matched filter receiver, a maximum likelihood sequence estimator, and a standard DFE. Our results indicate that for the channel model considered, BAD can achieve gains of up to 3 dB over the standard DFE, even when the channel estimate is in error.

1. INTRODUCTION

The concept of asymptotic efficiency has been used extensively in both single and multi-user detection to provide insight into the performance of various receivers in the high SNR regime [1]. In the following, we apply such high-SNR analysis to the bi-directional arbitrated DFE (BAD), an equalization algorithm that employs a local MAP-like criterion to arbitrate between symbol estimates generated by DFEs operating in the forward and reverse directions [2]. Asymptotic efficiency analysis allows us to characterize the performance of the BAD receiver relative to the performance when interference is not present. We derive the asymptotic efficiency of a matched filter (MF) detector, a standard DFE, maximum likelihood sequence estimation (MLSE), and the BAD receiver, all for both perfect and imperfect channel estimation. The results of this analysis give us insight into why BAD outperforms the standard DFE in simulation, as well as how we might alter the structure of BAD to increase these gains.

2. THE BAD ALGORITHM

Consider a discrete-time, symbol-spaced, time-invariant channel corrupted by additive white Gaussian noise (AWGN). Denoting the transmitted data bit by $b[n]$, the channel output at time n is given by

$$r[n] = \sum_{k=-L_1}^{L_2} h[k]b[n-k] + w[n], \tag{1}$$

This work was supported by the National Science Foundation under grants NCR96-24008 (CAREER), CCR-0092598 (CAREER), CCR99-79381, ITR00-85929, ITR02-20118, by the Army Research Office under grant DAAD19-00-1-0567, by the Office of Naval Research under Grant N00014-01-1-0117, and by the Motorola Center for Communications.

where $\mathbf{h} = \{h[k], -L_1 \le k \le L_2\}$ is the channel impulse response, and $w[n]$ is additive white Gaussian noise (AWGN). The noise variance σ_w^2 and the channel impulse response $\mathbf{h}$ are assumed to be known to the receiver.

The BAD algorithm, which is described in more detail in [3], comprises three stages: (1) bi-directional processing with MMSE DFEs, (2) data reconstruction, and (3) symbol arbitration. In the bi-directional processing stage, the received sequence is passed through a forward DFE designed for the channel $h[n]$ and is also time-reversed and passed through a reverse DFE designed for the time-reversal of $h[n]$. Hence, two estimates of the transmitted data sequence are generated. These sequences can differ substantially, which provides the diversity exploited by the BAD algorithm. In the reconstruction stage, each sequence of symbol estimates is convolved with the channel impulse response to form a noise-free estimate of the received data. When the estimates disagree, arbitration between the two values must be employed to produce the final estimate. In the arbitration stage, the Euclidean distance between each estimated received sequence and the actual received sequence is computed over a window around the desired symbol. The final estimate is taken from the DFE for which the reconstructed data sequence provides the best local fit.

3. COMPUTATION OF ASYMPTOTIC EFFICIENCY

For simplicity, we restrict out attention to BPSK signaling over the channel $H(z) = az + 1 + az^{-1}$, $0 < a < 1$. The nth received sample is given by

$$r[n] = ab[n-1] + b[n] + ab[n+1] + w[n], \tag{2}$$

where $w[n]$ has variance $\sigma^2 = N_0/2$.

For an equalizer, we define the asymptotic efficiency as the ratio of the exponent of decay of its error probability to that of the minimum error probability when there is no ISI, asymptotically as the noise variance tends to zero. The minimum error probability for a channel with energy $1 + 2a^2$ and no ISI is given by

$$P_{min} = Q\left(\sqrt{\frac{(1+2a^2)}{\sigma^2}}\right) \sim \exp\left(-\frac{1+2a^2}{2\sigma^2}\right) \tag{3}$$

using the asymptotic approximation $Q(x) \sim \exp(-x^2/2)$. Hence, the asymptotic efficiency for an equalizer with error probability $P_e = P_e(\sigma^2)$ is given by

$$\eta = \lim_{\sigma^2 \to 0} \frac{\log P_e}{\log P_{min}} = \lim_{\sigma^2 \to 0} \frac{\log P_e}{(-1+2a^2)/(2\sigma^2)}. \tag{4}$$

For an event E whose probability depends on σ^2, we define the asymptotic exponent

$$\lambda(E) = \lim_{\sigma^2 \to 0} -2\sigma^2 \log P(E). \tag{5}$$

Note the following properties of the asymptotic exponent:

(1) For a union of events, $\lambda(\cup_i E_i) = \min_i \lambda(E_i)$, i.e. the asymptotic exponent is dominated by the most likely event.

(2) For an intersection of events, $\lambda(\cap_i E_i) \geq \lambda(E_j)$ for each j, and $\lambda(\cap_i E_i) = \sum_i \lambda(E_i)$ if the events E_i are independent.

An important consequence of property (1) is that $\lambda(A)$ can be computed from the most probable subevent of A, asymptotically; hence, the strategy for computing asymptotic error exponents is to find the most likely way in which an error occurs. This is analogous to the philosophy underlying notions of minimum distance and nearest neighbors in signal and constellation design [4].

3.1. Perfect Channel Estimation

We first examine the case in which the receiver has exact knowledge of the channel. To compute the asymptotic error exponent of MF detection, we consider the decision statistic for bit $b[n]$, which is given by $\text{sgn}[r[n]*h[-n]]$, where $*$ denotes convolution. Conditioning on $b[n] = -1$, we find that the worst-case ISI combination is $b[n-k] = 1$ for $k = \pm 1, \pm 2$, which gives a probability of error of $P_{MF} = Q(\frac{1-4a}{\sigma\sqrt{1+2a^2}})$. Substituting into (4) gives the resulting asymptotic efficiency

$$\eta_{MF} = \left[\frac{\max\{0, 1-4a\}}{1+2a^2}\right]^2. \tag{6}$$

For MLSE, the asymptotic efficiency is governed by the error event that yields the minimum Euclidean distance from the true noiseless transmitted sequence. It is easy to see, using classical analysis of MLSE [5] to identify this error event, that

$$\eta_{MLSE} = \min\left\{1, \frac{2(a^2+(1-a)^2)}{1+2a^2}\right\}. \tag{7}$$

For our asymptotic analysis of the DFE and BAD, we consider zero-forcing (ZF) DFEs, since the MMSE DFE degenerates to the ZF DFE in the high SNR limit. The decision delay is chosen such that the feedforward filter degenerates to a single tap.

To compute the asymptotic exponent for a standard DFE, we consider only the most likely error event, which corresponds to the probability of a single error starting from clean feedback. Assuming such an event, the decision statistic for the forward DFE becomes $Z_f[n] = ab[n] + w[n]$, yielding probability of error $P_{DFE} = Q\left(\frac{a}{\sigma}\right)$. Hence, $\lambda_{DFE} = a^2$, and

$$\eta_{DFE} = \frac{a^2}{1+2a^2}. \tag{8}$$

Finally, we compute the asymptotic efficiency of the BAD algorithm. Errors here are a function of the error propagation events in the forward and reverse DFEs, as well as the outcome of the arbitration mechanism. There are two ways in which an error can occur in the desired bit: (A) both the forward and the reverse branches make an error in the desired bit; (B) either the forward or the reverse branch makes an error in the desired bit, and the other does not, but the arbitration mechanism chooses the branch that is in error. The error exponent for BAD is governed by the smaller asymptotic exponent between these two categories.

Events in category (A), denoted by E_A, can occur only if $b[n]$ is included in error propagation events on both branches. For the channel and DFE structure under consideration, such an event must be initiated by different noise samples in the two branches. Since the two noise samples are independent, property (2) tells us that the asymptotic exponent is $\lambda(E_A) = 2a^2$, twice that of the DFE.

When computing the asymptotic exponent of events in category (B), $\lambda(E_B)$, we will consider an arbitration window of length 3 in order to maintain a simple analysis. Though a longer window would give better results, a shorter window simplifies the analysis and allows for clearer presentation of the technique. Without loss of generality, we assume that the desired bit $b[n] = -1$ and that the forward branch makes an error estimating this bit. Let $\mathbf{r}$ denote the received vector in the arbitration window, $\mathbf{b}$ denote the bits which affect $\mathbf{r}$, and $\mathbf{w}$ denote the WGN in the arbitration window. The received vector is then given by $\mathbf{r} = \mathbf{Ab} + \mathbf{w}$, where $\mathbf{A}$ is a 3×5 convolution matrix whose rows are shifted versions of the channel impulse response. Let $\hat{\mathbf{b}}_f$ and $\hat{\mathbf{b}}_r$ denote the estimates of $\mathbf{b}$ produced by the forward and reverse branches, respectively. The corresponding error vectors are defined as $\mathbf{e}_f = (\mathbf{b} - \hat{\mathbf{b}}_f)/2$ and $\mathbf{e}_r = (\mathbf{b} - \hat{\mathbf{b}}_r)/2$.

BAD error events that fall in category (B) can be viewed as the intersection of two events: (i) a pair of error vectors $(\mathbf{e}_f, \mathbf{e}_r)$ for which $e_f[n] = -1$, and $e_r[n] = 0$ occurs; and (ii) the BAD arbitration mechanism makes an error, i.e. chooses the estimate generated by the forward DFE. Both events (i) and (ii) are dependent upon the noise samples $\mathbf{w}$. However, for the purpose of computing asymptotic efficiency, we can consider the events independently by invoking property (2), i.e. by recognizing that

$$\begin{aligned} \lambda(E_b) &= \lambda((\mathbf{e}_f, \mathbf{e}_r) \cap \text{arbitration error}) \\ &\geq \max\{\lambda(\mathbf{e}_f, \mathbf{e}_r), \lambda(\text{arbitration error})\}. \end{aligned} \tag{9}$$

Since we are looking for the minimum asymptotic exponent over all error events to compute λ_{BAD}, using (9), we can ignore any events in category (B) for which either $\lambda(\mathbf{e}_f, \mathbf{e}_r) \geq 2a^2$ or $\lambda(\text{arbitration error}) \geq 2a^2$. How are these quantities computed? Consider first $\lambda(\mathbf{e}_f, \mathbf{e}_r)$. Given specific error vectors $\mathbf{e}_f$ and $\mathbf{e}_r$, we first determine the magnitude of the various noise samples required to generate the vectors and then compute the probability of those noise samples occurring. To compute $\lambda(\text{arbitration error})$, we look at the difference between the received vector and the noiseless reconstructions for the forward and reverse branches, which are given by $\mathbf{r} - \mathbf{A}\hat{\mathbf{b}}_f$ and $\mathbf{r} - \mathbf{A}\hat{\mathbf{b}}_r$, respectively. The arbitration mechanism chooses the branch for which this difference has smaller norm. We define an arbitration statistic U as follows:

$$\begin{aligned} U &= \frac{1}{4}(||\mathbf{r} - \mathbf{A}\hat{\mathbf{b}}_r||^2 - ||\mathbf{r} - \mathbf{A}\hat{\mathbf{b}}_f||^2) \\ &= \langle \mathbf{w}, \mathbf{Ae}_r - \mathbf{Ae}_f \rangle + ||\mathbf{Ae}_r||^2 - ||\mathbf{Ae}_f||^2. \end{aligned} \tag{10}$$

The forward branch is chosen if $U > 0$; this condition can be rewritten as $\langle \mathbf{w}, \mathbf{A}(\mathbf{e}_r - \mathbf{e}_f) \rangle > ||\mathbf{Ae}_f||^2 - ||\mathbf{Ae}_r||^2$. Denote by $F(\mathbf{e}_f, \mathbf{e}_r)$ the set of noise vectors $\mathbf{w}$ satisfying this condition. Regarding the vectors $\mathbf{e}_f$ and $\mathbf{e}_r$ as given, it is easy to see that

$$\begin{aligned} \lambda(\text{arbitration error}) &= \lambda(\mathbf{w} \in F(\mathbf{e}_f, \mathbf{e}_r)) \\ &= \left[\frac{||\mathbf{Ae}_f||^2 - ||\mathbf{Ae}_r||^2}{||\mathbf{A}(\mathbf{e}_r - \mathbf{e}_f)||}\right]^2. \end{aligned} \tag{11}$$

For error vector pairs $(\mathbf{e}_f, \mathbf{e}_r)$ such that both $\lambda(\mathbf{e}_f, \mathbf{e}_r)$ and λ(arbitration error) are less than $2a^2$, we must compute the overall asymptotic exponent $\lambda((\mathbf{e}_f, \mathbf{e}_r) \cap$ arbitration error). This computation can be simplified by noting that

$$\lambda((\mathbf{e}_f, \mathbf{e}_r) \cap \text{arbitration error}) = \lambda(\text{arbitration error} \mid (\mathbf{e}_f, \mathbf{e}_r)) + \lambda(\mathbf{e}_f, \mathbf{e}_r). \tag{12}$$

Using these computation techniques for our three-tap symmetric channel, we find that λ_{BAD} is partitioned into 3 regions for $a \in (0, 1)$, each of which is dominated by different error events. The resulting asymptotic efficiency for the BAD algorithm is given by

$$\eta_{BAD} = \begin{cases} \frac{2a^2}{1+2a^2}, & a \in (0, \frac{1}{2}) \\ \frac{6a^2-6a+2}{1+2a^2}, & a \in (\frac{1}{2}, \frac{2}{3}) \\ \frac{3a^2-4a+2}{1+2a^2}, & a \in (\frac{2}{3}, 1) \end{cases} . \tag{13}$$

Figure 1 shows the asymptotic efficiency for the MF receiver, MLSE detection, the DFE, and BAD plotted as a function of a for $a \in (0, 1)$. Note that the asymptotic efficiency of BAD exceeds that of the DFE over the full range of a considered. In fact, BAD maintains a constant 3 dB gain over the DFE for $a \le \frac{1}{2}$ and a gain of at least 1.5 dB for $\frac{1}{2} < a < \frac{2}{3}$. As expected, the asymptotic efficiency of the MF receiver is high for small a but quickly drops to 0 for larger a. The asymptotic efficiencies of BAD and the DFE are quite low for small a, since, for the single feedforward tap DFE implementation used in our analysis, bit decisions are made based on a received element with energy a. In practice, however, a different decision delay would be used in the DFE for very small values of a, resulting in better performance relative to the MF in this regime.

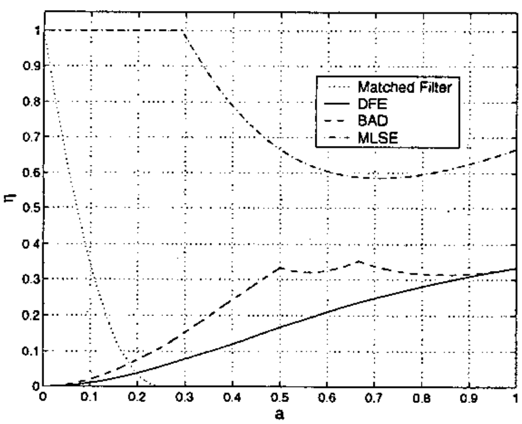

Fig. 1. Asymptotic Efficiency under Perfect Channel Knowledge

The asymptotic gain of the BAD algorithm over a standard DFE is given by

$$10 \log \frac{\lambda_{BAD}}{\lambda_{DFE}} \tag{14}$$

and is plotted as a function of a in Figure 2. The BAD algorithm achieves a full 3 dB gain over a standard DFE for $a < \frac{1}{2}$, and a gain of at least 1.75 dB for a between $\frac{1}{2}$ and $\frac{2}{3}$.

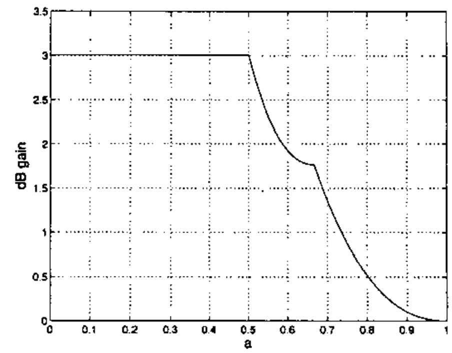

Fig. 2. Asymptotic Gain of BAD over DFE under Perfect Channel Knowledge

3.2. Effect of Channel Estimation Errors

High-SNR analysis can also be used to quantify the effect of channel estimation errors. We now consider the case in which the channel estimate used at the receiver varies slightly from the true channel impulse response. We assume that the error in channel estimation is small enough such that, though the probability of the most likely error event may change, the specific error event(s) that dominate(s) the asymptotic efficiency for each algorithm will remain the same as in the case of perfect channel estimation.

We consider the true channel given by

$$H_{act}(z) = \alpha z^{-1} + 1 + \alpha z$$

and the estimated channel given by

$$H_{est}(z) = a z^{-1} + 1 + a z.$$

To study the effects of underestimating ISI, we let $\alpha = a + \delta$, $\delta > 0$, $\delta \ll a$; to study the effects of overestimating ISI, we let $\alpha = a - \gamma$, $\gamma > 0$, $\gamma \ll a$. Since the actual channel now has energy $1 + 2\alpha^2$, we have $\lambda_{noISI} = 1 + 2\alpha^2$.

Using the same techniques developed for the analysis with perfect channel estimation, we have computed asymptotic exponents for each of the algorithms under consideration. The asymptotic exponents of the matched filter and the MLSE receiver are the same for under- or over-estimation of the channel ISI and are given by:

$$\lambda_{MF} = \max\{0, (1 - 2(a + \alpha))^2\} \tag{15}$$

$$\lambda_{MLSE} = \min\left\{ \frac{(1 + 2a^2 + 2(a - \alpha))^2}{1 + 2a^2}, \frac{(1 - 2a^2 + 2\alpha - 2a + 4a\alpha)^2}{1 + 2a^2}, \frac{(2 - 2\alpha - 2a + 4a\alpha)^2}{2a^2 + 2(1 - a)^2} \right\}. \tag{16}$$

The asymptotic exponents of the DFE and BAD algorithms vary according to whether the ISI is over- or under-estimated. For the standard DFE, the asymptotic exponents are given by

$$\lambda_{DFE} = a^2 \tag{17}$$

and

$$\lambda_{DFE} = (a - 2\gamma)^2 \tag{18}$$

$$\lambda_{BAD} = \begin{cases} \min\{2a^2, a^2 + 2(1-a)^2 - 8\delta(1-a) + 8\delta^2, a^2 + \frac{(a-1)^2+(1-2a)^2-2\delta(2-3a)^2}{(a-1)^2+(1-2a)^2}\} & , a < \frac{2-2\delta}{3} \\ \min\{2a^2, a^2 + 2(1-a)^2 - 8\delta(1-a) + 8\delta^2, 2a^2 + (3a+2\delta-2)^2 - 4\delta^2 - 4a\delta\}, & a > \frac{2-2\delta}{3} \end{cases} \tag{19}$$

$$\lambda_{BAD} = \begin{cases} \min\{2(a-2\gamma)^2, (a-2\gamma)^2 + 2(1-a)^2, (a-2\gamma)^2 + \frac{(a-1)^2+(1-2a)^2-2\gamma(2a-1)^2}{(1-a)^2+(2a-1)^2}\}, & a < \frac{2+2\gamma}{3} \\ \min\{2(a-2\gamma)^2, (a-2\gamma)^2 + 2(1-a)^2, a^2 + (a-2\gamma)^2 + (3a-2\gamma-2)^2\} & , a > \frac{2+2\gamma}{3} \end{cases} \tag{20}$$

for under- and over-estimation, respectively, of the channel ISI. The asymptotic exponent of the BAD algorithm is dependent not only on whether the channel ISI is over- or under-estimated, but also on the value of a. The value of λ_{BAD} for under- and over-estimation of the channel ISI is given in (19) and (20), respectively. The asymptotic efficiency for each algorithm is simply its asymptotic exponent divided by that of the no ISI case.

Figure 3 shows the asymptotic efficency of the matched filter, DFE, BAD, and MLSE for perfect channel estimation, as well as for under- and over-estimation of the ISI by 10%, i.e. $\delta = \gamma = 0.1\ \alpha$. The plot reveals that only the matched filter receiver is relatively unaffected by incorrect channel estimation. The asymptotic efficiency of MLSE degrades significantly for small a, where it was equal to 1 when the channel was perfectly estimated. Both the DFE and BAD suffer from reductions in asymptotic efficiency, especially for larger values of a, but BAD still attains significant gains over the DFE for most of the range of a considered.

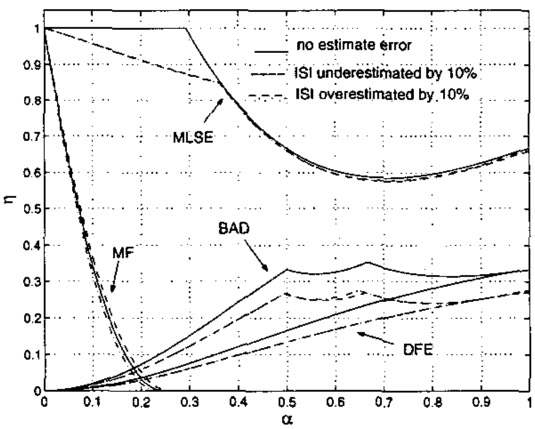

Fig. 3. Asymptotic Efficiency under Imperfect Channel Estimation

Figure 4 shows the asymptotic performance gain (in dB) of BAD over the DFE when the channel is correctly estimated, as well as when it is under- and over-estimated by 10%. The BAD algorithm maintains a constant 3 dB gain for $a < 0.5$, but its gain over the DFE is slightly lessened by imperfect channel estimation for larger values of a. For the scenario considered here, underestimating the channel ISI has a more detrimental effect on the BAD algorithm than overestimating the ISI.

4. CONCLUSION

Despite the simplicity of the setting considered, the results of our asymptotic analysis provide insight into why the BAD algorithm outperforms the DFE. By identifying the most likely failure mode,

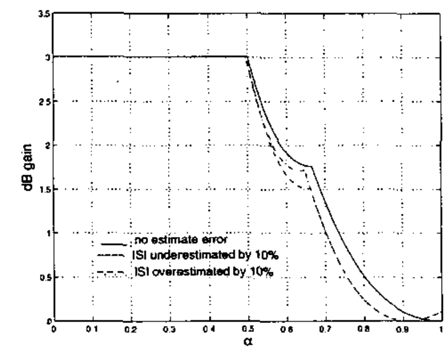

Fig. 4. Asymptotic Gain of BAD over DFE under Imperfect Channel Estimation

asymptotic efficiency analysis allows us to isolate the "weakest link" of the BAD algorithm and focus efforts on improving its performance. In addition, for a symmetric channel with identical forward and reverse DFEs, our analysis reveals that the asymptotic gain of BAD over the DFE is bounded by 3 dB, corresponding to errors in the desired bits in both branches. To surpass this bound, we must generate more candidates for arbitration (see [6] for some early attempts), possibly by iterating on the initial estimates from the two DFEs.

5. REFERENCES

[1] S. Verdu, *Multiuser Detection*, Cambridge University Press, 1998.

[2] J. K. Nelson, A. C. Singer, U. Madhow, and C. S. McGahey, "BAD: Bidirectional arbitrated decision feedback equalization," February 2003, Submitted to *IEEE Transactions on Communications*.

[3] C.S. McGahey, "BAD: Bi-directional arbitrated decision-feedback equalization," M.S. thesis, University of Illinois at Urbana-Champaign, August 2000.

[4] E. Lee and D. Messerschmitt, *Digital Communication*, Kluwer Academic Publishers, second edition, 1994.

[5] G. D. Forney, "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Transactions on Information Theory*, vol. 18, pp. 363–378, May 1972.

[6] G. Barriac and U. Madhow, "Beyond BAD: a parallel arbitration framework for low-complexity equalization," in *Proc. of the 2001 Allerton Conference on Communication, Control, and Computing*, Monticello, Illinois, October 2001.