OPTIMAL EMPIRICAL DETECTOR FUSION

Amitabh Dixit and Geoffrey C. Orsak

Department of Electrical Engineering
Southern Methodist University
Dallas, Texas 75275
amitabh@seas.smu.edu, gorsak@seas.smu.edu

ABSTRACT

In cellular communications, one often wishes to merge data from various base stations to improve overall link performance. There are two well known alternatives for doing this: (1) combine sufficient statistics from each of the various receivers using standard signal processing approaches, or (2) fuse raw bit decisions from each of the receivers into a final bit decision. For its utility and simplicity, we focus on the later approach herein.

In this paper we have derived the optimal empirical rule for fusing bit estimates from a bank of receivers when the performance of each receiver is unknown. As an integral part of this algorithm, we "blindly" estimate the performance of each receiver and use these estimates in the final fusion algorithm. We show that the resulting algorithm converges to the globally optimal algorithm as the number of bits observed grows without bound. More importantly, we show that in practical regimes, this new fusion rule outperforms the best individual detector in the bank and outperforms the standard "majority rule" fusion detector.

1. INTRODUCTION

In wireless cellular communications, fusing information from the various base stations can improve link performance. This information can be the individual bit decisions from the base stations or can be other (sufficient) statistics derived from the base stations. In order to optimally fuse bit estimates from multiple receivers, it is well know that the fusion rule must have precise information regarding the probability of error (BER) of the individual receivers. Clearly, BER estimates can be obtained by using training. However, reliance on training data consumes valuable bandwidth and thus reduces overall link performance.

In this paper, we derive a new blind algorithm which allows cellular systems to optimally fuse bit estimates from individual cell sites without any training data or any a priori knowledge of the individual link performances. Within we show that this algorithm results in a fusion rule with the following important and practical properties:

- fusion BER converges to the globally optimal bit fusion rule as the number of observed bits grows without bound.
- for small numbers of observed bits, this blind algorithm outperforms the best detector in the bank of detectors and the standard "majority rule" fusion detector.

The approach used in deriving this new algorithm is threefold: Firstly, we develop a simple algorithm for determining the best detector from a bank of detectors by only observing individual bit estimates from each of the detectors. Secondly, having identified the best detector, we are able to accurately estimate the BER of each of the individual detectors. And finally, using a standard likelihood based algorithm, we are able to implement the optimal empirical fusion rule.

The paper is organized as follows: In section 2 we derive the blind algorithm for blindly determining the best performing detector from a bank of detectors. In this section, we show that our identification algorithm converges to the best detector exponentially fast. In section 3 we derive a new algorithm for jointly estimating the BERs associated with each of the individual detectors. In section 4 we derive the fusion algorithm based on the estimated BERs. And finally in section 5, we derive an upper bound on the performance of this new fusion algorithm.

2. IDENTIFYING THE BEST DETECTOR

In figure (1) we show a bank of detectors each operating on the same input data observed through individual channels. In this paper, we demonstrate a new blind algorithm for optimally fusing the bit estimates $\widehat{b_i}$ without any a priori knowledge of the separate probability of errors $P_e(i)$ or without any training data. Herein we assume the following:

- the detectors make independent decisions.
- each channel is a BSC.
- the transmitted bits have equal priors.

In this section, we attempt to identify the detector with the minimum $P_e(i)$ from the bank of detectors when there is no training information available. To achieve this goal, we rely upon the following fact.

Fact 1 *Let G be any detector not in the bank of detectors with probability of error given by $P_e(G)$. $P_e(1) \leq P_e(2) \ldots \leq P_e(N)$ if and only if Pr(detector 1 agrees with G) $\geq$ Pr(detector 2 agrees with G) $\ldots \geq$ Pr(detector N agrees with G).*

Using this fact, we can identify the best detector by identifying the detector which agrees with some other arbi-

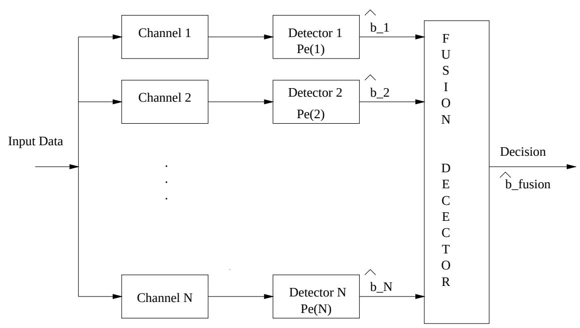

Figure 1: Blind Detector Fusion

trary detector G most frequently. In the case of this paper, we set detector G to be the "majority rule" detector since it is easily constructed and generally has good performance.

Proceeding, let M be the number of bits passed through the bank. An estimate for the probability that the i^{th} detector agrees with the majority rule is

$$\alpha_{im} = \frac{\sum_{j=1}^{M} x_j^{im}}{M}$$

where
$x_j^{im} = 1$ if detector i agrees with the majority and
$x_j^{im} = 0$ if detector i disagrees with the majority.

Hence x_j^{im} is a Bernoulli random variable that takes on a value 1 with the probability that detector i agrees with majority (say p_{im}) and a value 0 with the probability that detector i disagrees with majority (= 1-p_{im}). Therefore α_{im} is a binomial random variable with parameters $(p_{im}/M, M)$.

Thus, using Fact 1 above, we identify the best detector as the one for which α_{im} is the maximum. As M increases, the estimates α_{im}, i=1,2,...,N converge to their true values, and as such our identification of the best detector becomes more reliable.

As an example, consider a bank of three detectors. In figure 2, we plot the probability of correctly identifying the best detector amongst the three versus the number of bits observed. It should be noted that we require 97 bits to identify the best detector with a confidence of 95%.

2.1. Upper Bound on the Probability of error in Identifying the Best Detector

Computing the probability of error in identifying the best detector ($P_e^{iden\_bst}$) requires that we have the joint statistics between all the estimates α_{im} which is unrealistic. Therefore, in this subsection, we derive an upper bound on $P_e^{iden\_bst}$ that is computationally simple and is informative. Without loss of generality, assume that detector 1 has the minimum P_e. Let $P_c^{iden_bst} \triangleq$Probability of identifying the best detector correctly then,

$$P_c^{iden\_bst} = Pr(\alpha_{1m} > \alpha_{2m}; \alpha_{1m} > \alpha_{3m}; \ldots; \alpha_{1m} > \alpha_{Nm})$$

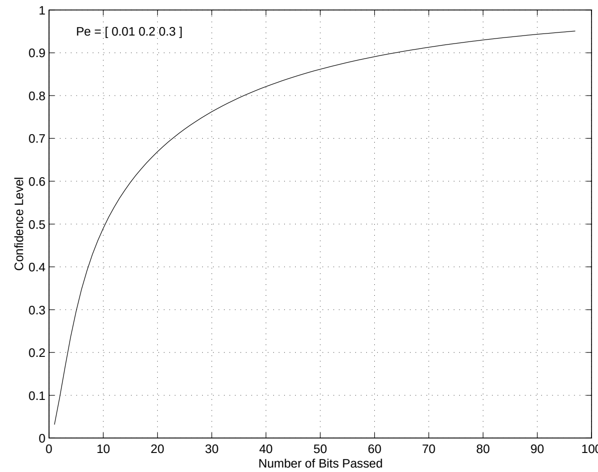

Figure 2: A bank with three detectors having P_e values 0.01, 0.20 and 0.30 was taken and the probability of identifying the best detector is plotted versus M.

Define *Event* $A_i = \{\alpha_{1m} > \alpha_{im}\}$, i=2,3,...,N then

$$P_c^{iden\_bst} = Pr\left(\bigcap_{i=1}^{N-1} A_i \right) \quad (1)$$

$$\Rightarrow P_e^{iden\_bst} = Pr\left(\overline{\bigcap_{i=1}^{N-1} A_i} \right)$$

Fact 2 $Pr\left(\overline{\bigcap_{i=1}^{N-1} A_i} \right) \leq \sum_{i=2}^{N} Pr\left(\overline{A_i} \right)$

$$\Rightarrow P_e^{iden\_bst} \leq \sum_{i=2}^{N} Pr\left(\alpha_{1m} \leq \alpha_{im}\right) \quad (2)$$

To evaluate the above expression, we must have the joint densities pairs of the estimates $\{\alpha_{1m}, \alpha_{im}\}$ which are generally unavailable to us. To circumvent this problem, we make the reasonable assumption that for sufficient M the α_{im} i=1,2,...,N are Gaussian random variables.

Now, let us consider any two detectors of the bank, detector i and detector j s.t. $P_e(i) < P_e(j)$. It can be shown that the probability of identifying the best detector among detector i and detector j is

$$Pr(\alpha_{im} \leq \alpha_{jm}) = Q\left(\sqrt{M} d_{ij}\right) \quad (3)$$

where

$$d_{ij} = \left[\frac{(p_{im} - p_{jm})^2}{(p_{im} + p_{jm}) - (p_{im}^2 + p_{jm}^2)} \right]^{1/2}$$

Hence from eq.(2),

$$P_e^{iden\_bst} \leq \sum_{i=2}^{N} Q\left(\sqrt{M} d_{1i}\right) \quad (4)$$

$$\Rightarrow P_c^{iden\_bst} \geq 1 - \sum_{i=2}^{N} Q\left(\sqrt{M} d_{1i}\right) \quad (5)$$

Interestingly, it should be noted from eq.(1) that if we in addition assume α_{im}, i=1,2,...,N to be uncorrelated, then using eq.(3) $P_c^{iden\_bst}$ can be exactly written as

$$P_c^{iden\_bst} = \prod_{j=2}^{N} \left(1 - Q\left(\sqrt{M} d_{1j}\right)\right) \quad (6)$$

By neglecting higher order terms, we obtain the approximation

$$P_c^{iden\_bst} \approx 1 - \sum_{j=2}^{N} Q\left(\sqrt{M} d_{1i}\right) \quad (7)$$

or

$$P_e^{iden\_bst} \approx \sum_{j=2}^{N} Q\left(\sqrt{M} d_{1i}\right) \quad (8)$$

This analysis shows that for a large number of detectors and large number of observed bits, the bound in eq.(4) will be tight. As an example, we consider the case of a bank of 7 detectors with respective probability of errors given as [0.05, 0.15,0.15, 0.2, 0.25, 0.3, 0.4]. In figure (3) we plot the probability of correctly identifying the best detector versus the number of observed bits. In this figure, we show the lower bound derived in equation (5), the approximation shown in equation (6), and the simulated probability of correct identification. We can clearly see from this figure that as the number of observed bits grows, all three expressions converge, thus establishing the "tightness" of the bound and the accuracy of the approximation.

Using the result in equation (4), we are able to calculate the rate at which the probability of error in identifying the best detector tends to zero in the number of observed bits M. From eq. (4) we have

$$\begin{aligned} P_e^{iden\_bst} &\leq (N-1) max\left\{Q\left(\sqrt{M} d_{1i}\right)\right\} \\ &\approx (N-1) max\left\{\frac{1}{\sqrt{2\pi} d_{1i}\sqrt{M}} e^{-Md_{1i}^2/2}\right\} \end{aligned}$$

where the approximation is derived from the well known $Q(\cdot)$ function approximation. This clearly demonstrates that the probability of error in identifying the best detector diminishes at an exponential rate in M with a slope determined by the $\min(d_i)$=$d_{i_{min}}$.

Assume that we wish to identify the best detector with an error rate no larger than $P_e = 10^{-\beta}$. Then it can be shown that

$$M \geq \frac{2\ln 10}{(d_{i_{min}}^2)}\beta + \frac{2}{(d_{i_{min}}^2)} \ln\left(\frac{N-1}{\sqrt{2\pi}(d_{i_{min}})}\right)$$

This strong result shows that the increase in required number of bits is **logarithmic** in the desired level of confidence and **logarithmic** in the number of detectors. As such, the proposed identification algorithm is very efficient for nearly all operational scenarios. To demonstrate this fact, consider figure (4) where we have plotted the log of the probability of error in correctly identifying the best detector versus the number of observed bits. This figure clearly shows that this probability decays at an exponential rate for even modest numbers of observed bits and small numbers of detectors.

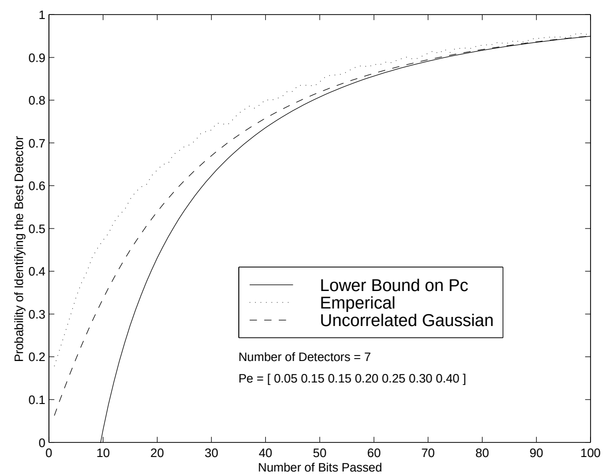

Figure 3: The above figure shows the plot of the probability of identifying the best detector versus the number of bits passed for various cases. The "Uncorrelated Gaussian" is the $P_c^{iden\_bst}$ curve obtained using eq.(6) assuming that $\alpha_{im}s$ are uncorrelated Gaussian random variables. Also plotted are the empirically calculated $P_c^{iden\_bst}$ and the lower bound of eq. (5). As we expect from the *Central Limit Theorem*, the approximations become better and the bound becomes tighter as the number of bits passed increases.

3. ESTIMATING THE $P_E(I)$

Without loss of generality, consider the bank of detectors given in figure (1) with $P_e(1) \leq P_e(2) \leq P_e(3) \leq P_e(k)$ where k = 4, 5,..., N. As before, let α_{ij} be the estimate of the probability that detector i and detector j agree then

$$\alpha_{1j} = \frac{\sum_{k=1}^{M} x_k^{1j}}{M}$$

where M is the number of bits observed and x_k^{1j}=1 if detector 1 and detector j agree and x_k^{1j}=0 if detector 1 and detector j disagree. (We have selected detector 1 as the "reference" detector since this will render the minimum variance estimates of the probability that detector i agrees with detector j.)

Assuming that the detectors and the majority rule make independent decisions, it is easy to show that $\alpha_{1j}s$ are binomial random variables with parameters $((1-P_e(1)-P_e(j)+2P_e(1)P_e(j))/M, M)$. Also it is easy to see that $Pr(\text{detector } i \text{ and detector } j \text{ agree}) = 1 - P_e(i) - P_e(j) + 2P_e(i)P_e(j)$.

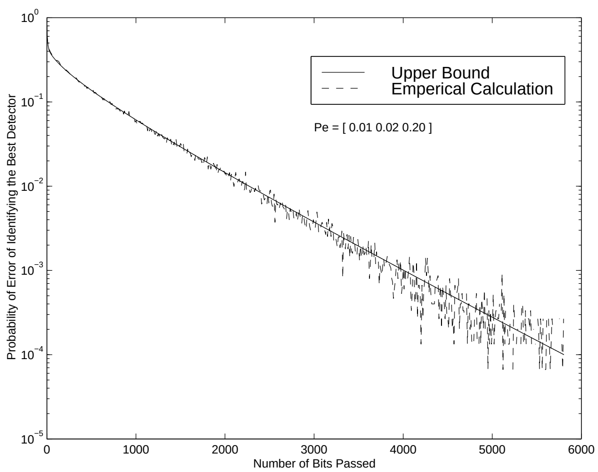

Figure 4: Plots of the upper bound on the probability of error of identifying the best detector and the empirical probability of error of identifying the best detector is plotted versus the number of bits observed, for a bank of three detectors with P_e values 0.01, 0.02 and 0.2. The plot shows that the P_e of identifying the best detector decays exponentially fast.

Using the above relationship, we form N equations as follows:

$$1 - P_e(1) - P_e(2) + 2P_e(1)P_e(2) = \alpha_{12} \quad (9.a)$$
$$1 - P_e(1) - P_e(3) + 2P_e(1)P_e(3) = \alpha_{13} \quad (9.b)$$
$$\vdots$$
$$1 - P_e(1) - P_e(N-1) + 2P_e(1)P_e(N-1) = \alpha_{1(N-1)} \quad (9.n-2)$$
$$1 - P_e(1) - P_e(N) + 2P_e(1)P_e(N) = \alpha_{1N} \quad (9.n-1)$$
$$1 - P_e(2) - P_e(3) + 2P_e(2)P_e(3) = \alpha_{23} \quad (9.n)$$

The above setup results in N equations and N unknowns. In this case, these equations are not linear in their variables $P_e(i)$. Nevertheless, they can be jointly solved to identify the values of $P_e(i)$. This is accomplished by solving for $P_e(2)$ in terms of $P_e(1)$ from eq.(9.a) and $P_e(3)$ in terms of $P_e(1)$ from eq.(9.b) and substituting these in eq.(9.n). This results in a a quadratic equation in $P_e(1)$ which is easily solved. Once we have determined $P_e(1)$, it can be substituted in other equations to determine the remaining $P_e(i)s$.

Fact 3 *Irrespective of the number of observed bits M, the above algorithm always renders a valid and unique solution for $P_e(i)$.*

Furthermore, as M increases, the estimates for $\alpha_{ij}s$ converge to the true probabilities and as such the solutions to equations (9.a) through (9.n) converge to the true values of $P_e(i)s$. Figure (5) demonstrates the convergence behavior of the estimates of $P_e(i)s$ obtained using the above algorithm. In this examples, we have assumed a bank of three detectors with $P_e(i)$ given by [0.1, 0.2, 0.3].

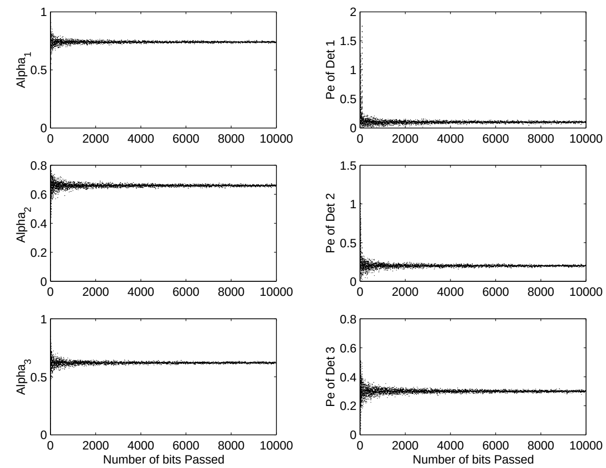

Figure 5: The following figure shows the convergence of the estimation algorithm to the true values of $P_e(i)$ and the convergence of the estimates $\alpha_i s$. The bank under consideration is of length 3 with Pe=[0.1 0.2 0.3]. The true value of α_i s for the bank are 0.74, 0.66 and 0.62

4. OPTIMAL EMPIRICAL FUSION

Consider again the setup of figure (1). Let
$H_1 \triangleq$ Event that '1' was sent,
$H_0 \triangleq$ Event that '0' was sent
and let b_i be the decision made by the i^{th} detector. Then using Bayes rule, it can be shown that the optimum bit fusion rule is given by:

$$\sum_{i=1}^{N}(1 - 2b_i)\ln\frac{P_e(i)}{1 - P_e(i)} \underset{H_0}{\overset{H_1}{\gtrless}} 0 \quad (10)$$

It is easy to see that if Q detectors say '0' and $N - Q$ detectors say '1' then the decision rule can we written as :

$$\sum_{i=1}^{Q}\ln\frac{P_e(i)}{1 - P_e(i)} - \sum_{i=Q+1}^{N}\ln\frac{P_e(i)}{1 - P_e(i)} \underset{H_0}{\overset{H_1}{\gtrless}} 0$$

where without loss of generality we have assumed that the detectors deciding '0' are the first Q detectors. If the $P_e(i)$ are small, then the optimum fusion rule simply requires that we add the error exponents of the detectors that say '1' and compare this value of the sum of the error exponents of the the detectors that say '0'.

In the previous section, we determined an efficient and accurate means of jointly estimating the various $P_e(i)$. Moreover, these estimates were shown to converge to the true values of $P_e(i)$ as the number of observed bits grow. Thus, our proposed empirical fusion rule is given by

$$\sum_{i=1}^{Q} \ln \frac{\widehat{P_e(i)}}{1-\widehat{P_e(i)}} - \sum_{i=Q+1}^{N} \ln \frac{\widehat{P_e(i)}}{1-\widehat{P_e(i)}} \underset{H_0}{\overset{H_1}{\gtrless}} 0$$

where the $\widehat{P_e(i)}$ are the solution to equations (9.a) through (9.n)

To evaluate this new fusion rule, let us first consider a simple example. Consider an increasingly large bank of detectors where the individual $P_e(i)$ start at 0.05 and increase by 0.05. In figure (6), after observing 500 bits, we plot versus the number of detectors the (1) probability of error of the best detector, (2) the probability of error of the majority rule, (3) the probability of error of the blind empirical fusion detector, and finally (4) the probability of error of the optimal fusion detector. One can see that the blind empirical fusion rule performs nearly identically to the globally optimum rule in nearly all scenarios.

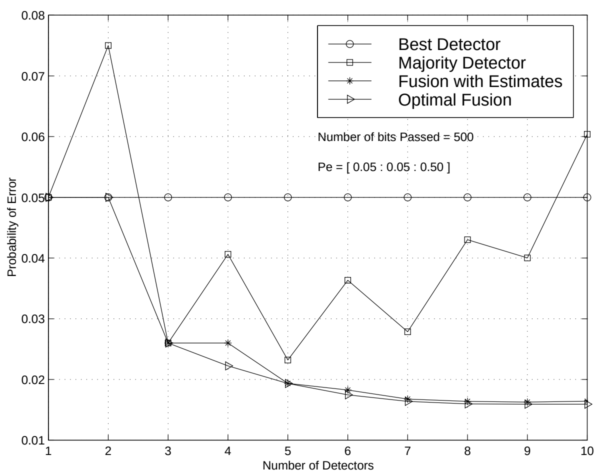

Figure 6: A comparative plot of the P_e of the best detector, the majority rule detector, the optimal fusion detector (using perfect estimates of $P_e(i)s$) and the fusion detector using $P_e(i)$ estimates based on the observation of 500 bits for various banks of detectors. The number of detectors in the bank was increased from one to ten with each new detector added having P_e 0.05 higher, starting from $P_e(1)$ = 0.05. We observe that the majority suffers badly as the number of ill-performing detectors increase but the optimal fusion detector always improves unless the detector being added has a P_e equal to a half, in which case it remains the same.

5. UPPER BOUND OF PERFORMANCE OF EMPIRICAL FUSION DETECTOR

Before deriving an upper bound on the probability of error of the new blind fusion rule, let us make a few observations about the relative performance of the best detector in the bank of detectors, the majority fusion rule, and the optimal fusion rule.

Fact 4 *P_e of the majority fusion rule may be* **greater** *than or equal to the P_e of the best detector.*

This would imply that it might be better to use the identified best detector in the bank than to risk fusing decisions together using the "majority rule" detector.

Fact 5 *P_e of the optimal fusion detector is always less than or equal to the P_e of the majority rule and the best detector.*

Fact 6 *P_e of the optimal fusion detector always decreases unless the detector being added has a P_e equal to 0.5*

These two results show that it is always better to utilize all of the detectors if you are able to optimally fuse their decisions. Using the new algorithm proposed in this paper, we believe that the same result is true for optimal empirical blind fusion.

Since an explicit expression for probability of error (P_e) of the fusion detector ($=P_e^{fusion}$) derived above is difficult to obtain, we derive a tight upperbound on the fusion rule using arbitrary estimates of the $P_e(i)$ that is simple to evaluate and gives us an insight. It can be shown that

$$P_e^{fusion} \leq \frac{1}{2} \exp \frac{1}{2} \sum_{i=1}^{N} \ln \left[4 \max \left\{ \widehat{P_e(i)}, \frac{P_e(i)}{\widehat{P_e(i)}} \right\} \right] \tag{11}$$

- *Case:1* $\widehat{P_e(i)} = P_e(i)$ i.e. perfect estimates then

$$P_e^{fusion} \leq \frac{1}{2} \exp \left[\frac{1}{2} \sum_{i=1}^{N} \ln 4P_i \right]$$

- *Case:2* $\widehat{P_e(i)} > P_e(i)$ i.e. over estimation

$$\max \left\{ \widehat{P_e(i)}, \frac{P_e(i)}{\widehat{P_e(i)}} \right\} \geq P_e(i)$$

$\Rightarrow$ We pay a penalty for over estimating the $P_e(i)s$

- *Case:3* $\widehat{P_e(i)} < P_e(i)$ i.e. under estimation

$$\max \left\{ \widehat{P_e(i)}, \frac{P_e(i)}{\widehat{P_e(i)}} \right\} \geq P_e(i)$$

$\Rightarrow$ We pay a penalty for under estimating the $P_e(i)s$

This bound quantifiably demonstrates that we always pay a penalty (relative to the optimal fusion rule) estimation errors in $P_e(i)$. However, since the potential errors in the estimates appear within a log in the error exponent, this result also shows that under modest errors, we will see very little performance degradation. Thus supporting the results seen in figure (6) where the new algorithm nearly matches the performance of the optimal fusion algorithm. Viewed in another way, the optimal fusion rule is not sensitive to modest errors since offsets in the estimates need not change the sign of the log-likelihood ratio of eq. (10).

6. SUMMARY

This paper presents an entirely blind method for fusing together bit estimates obtained from a bank of arbitrary detectors. As part of this work, we have developed and analyzed a method of identifying the best detector from a bank of N detectors which was shown to converge exponentially fast to the true best detector in the bank. Using the knowledge of the best detector, we have been able to develop a new algorithm for jointly estimating the respective error rates of the individual detectors and then optimally fusing the individual bit estimates.

We showed herein that the resulting fusion algorithm performs nearly as well as the globally optimum fusion rule and significantly outperforms the majority fusion rule and the best detector in the bank. This new algorithm is computationally efficient and robust to a wide variety of operating conditions.

REFERENCES

[1] H.V. Poor, *An Introduction to Signal Detection and Estimation -Second Edition.* (Springer-Verlag: New York, 1994)

[2] T. M. Cover and J. A. Thomas. *Elements of Information Theory.* (Wiley-Interscience, 1991).

[3] H. Stark and J. W. Woods.*Probability, Random Processes and Estimation Theory for Engineers.* Prentice Hall: Englewood Cliffs, New Jersey 1986.

[4] A. Papoulis. *Probability, Random Variables and Stochastic Processes* McGraw-Hill, Inc. 1991.