1. INTRODUCTION

The most destructive characteristic of a wireless channel is the random variation of its transfer function, known as fading phenomenon. Diversity combining is a well-known technique for combating this effect. Space, frequency, time and coding diversities, as well as combination of two or more of these, are employed in different systems. Various combining techniques have been suggested for multiple received signals [1, 2]. In presence of additive white Gaussian noise, maximal ratio combining (MRC) is a theoretically optimal combiner over fading channels in which the received signals from different paths are combined so as to maximize the instantaneous SNR at the combiner output [1]. Performance analysis of MRC system has been the subject of interest in many research works [3-21]. In ideal MRC scheme, it is assumed that the channel coefficients are known at the receiver. However, in practice, these coefficients have to be estimated using a training sequence. Hence, the performance of non-ideal MRC is affected considerably by the estimation error [10-12].

By R Annavajjala et al. [13, 14] an analytical relation for average bit error rate has been derived for non-ideal MRC with BPSK modulation and independent Rayleigh channels. The performance of non-ideal MRC for independent Rayleigh, Rician and Nakagami channels has been analyzed and compared by W. M. Gifford et al. [15]. W. M. Gifford et al. [16] the previous work [15] has been extended to correlated channels. Non-ideal MRC with correlated Rayleigh channels in presence of colored noise is investigated by L. Schmitt, et al. [17]. Independent and non-identical distributed Rayleigh channels have been discussed by Y. C. Ko, and T. Luo, [18]. On the other hand, modified MRC receivers with improved performances have been proposed in vary articles [19-21] by employing practical channel estimations on fading channels. In these works, the receiver structures are linear and a training sequence is employed, which in turn decreases the bandwidth efficiency. In present work, we offer a blind nonlinear diversity combining technique for Rayleigh flat fading channels. In our proposed combining technique, instead of estimating the channel coefficients, we directly estimate the transmitted symbols and show that the optimum estimator is a nonlinear polynomial system, which can be realized by a Hammerstein filter. We show that the performance of this blind combining technique is very close to ideal MRC. This is a valuable result, especially
because higher bandwidth efficiency is also achieved. We also show that in our proposed technique, variation of the combiner output values around the values of transmitted symbols is less than the corresponding variation in MRC system. Hence, the proposed system is potentially superior to MRC for soft decision purposes.

Hammerstein filter is a nonlinear polynomial filter used in many applications such as system identification [22-24], modeling [25] echo cancellation [27, 28], and noise cancellation [29]. In one of our previous works [30], we have proposed a new nonlinear equalization technique (GHE) for frequency selective fading channels, based on Hammerstein filters.

Also, in other works [31, 32], we have presented a new Hammerstein diversity combining technique (HDC) for frequency selective fading channels. In these systems, training symbols are used for calculation of optimized filter coefficients.

This paper is organized as follows: In section 2, we present the system model. Section 3 introduces our nonlinear blind Hammerstein diversity combining technique. Theoretical basis of our proposed system is presented in section 4. In section 5 a proper cost function is introduced for obtaining the optimized filter coefficients in a blind manner. Section 6 provides the simulation results and discussions, before concluding the paper in section 7.

2. SYSTEM MODEL

The equivalent low-pass discrete time model of the system is illustrated in Figure 1. In this work, we employ BPSK modulation. The transmitted sequence \( x(n) \in \{ +1, -1 \} \) is drawn from an i.i.d. source with equi-probable symbols. The communication system consists of \( M \) diversity branches. These branches are assumed to be identical Rayleigh flat fading channels.

Hence, the SIMO channel can be presented by an \( M \times 1 \) vector, as:

\[
H = [h_1, \ldots, h_M]^T
\]

where \( h_i \) is the complex Rayleigh distributed random gain of the \( i \)th channel as:

\[
h_i = h_{ii} + j h_{qi},
\]

where \( h_{ii} \) and \( h_{qi} \) are the real and the imaginary component of the channel gain respectively. These two components are independent, zero mean, Gaussian random variables with variance \( \sigma_{h_i}^2 = 1 \). Furthermore, the branches are assumed uncorrelated, i.e.:

\[
E[h_i h_j^*] = 0 \quad \text{for } i \neq j
\]

The channel fading is assumed sufficiently slow, such that the channel gains do not vary during one data frame. The received signal from the \( i \)th channel is given by:

\[
y_i(n) = h_i x(n) + w_i(n) \quad i = 1, 2, \ldots, M
\]

where \( w_i(n) \) is the complex additive white Gaussian noise at the \( i \)th receiver branch written as:

\[
w_i(n) = w_i(n) + j w_{qi}(n)
\]

where \( w_i(n) \) and \( w_{qi}(n) \) are uncorrelated, zero mean, Gaussian random variables with variance \( \sigma_{w_i}^2 \). Equation (4) can be expressed in matrix form:

\[
Y(n) = H x(n) + W(n)
\]

where \( H \) is the channel vector and \( Y(n) \) and \( W(n) \) are the received data vector and the noise vector, respectively. These vectors are defined as:

\[
Y(n) = [y_1(n) \ldots y_M(n)]^T
\]

\[
W(n) = [w_1(n) \ldots w_M(n)]^T
\]
As shown in Figure 1, the receiver consists of two correlators banks, namely, inphase and quadrature correlators. The complex received signal $y_{i}(n)$ from each branch is applied to both correlators. The outputs of the inphase and quadrature correlators are the real part ($y_{i}(n)$) and the imaginary part ($y_{o}(n)$) of $y_{i}(n)$ respectively. According to Equations (2) and Equation (5), we can write:

$$y_{i}(n) = b_{i}, x(n) + w_{i}(n) \quad i = 1, 2 \ldots , M$$

$$y_{o}(n) = b_{o}, x(n) + w_{o}(n)$$

We define the $2M \times 1$ real vector $\tilde{y}(n)$ as:

$$\tilde{y}(n) = \begin{bmatrix} y_{i}(n) & \ldots & y_{m}(n) \\ y_{o}(n) & \ldots & y_{o}(n) \end{bmatrix}^T$$

where:

$$\tilde{y}(n) = \begin{bmatrix} y_{i}(n) & 1 \leq i \leq M \\ y_{o}(n) & M + 1 \leq i \leq 2M \end{bmatrix}$$

As shown in Figure 1, $\tilde{y}(n)$ is the input to the diversity combining filters. This model is very convenient for computational purposes, as we deal with real values only. It is in fact similar to having $2M$ real fading diversity branches modeled as Gaussian random variables.

The output of the combiner, $z(n)$, is applied to a hard detector for making the output decision $\hat{x}(n)$.

**3. BLIND HAMMERSTEIN DIVERSITY COMBINING TECHNIQUE**

A Blind Hammerstein Diversity Combining (BHDC) system is shown in Figure 2. In this approach, a Hammerstein filter of order $D$ is employed for each diversity branch. The output polynomial of the $i$th filter is:

$$z_{i}(n) = \sum_{k=1}^{M} \sum_{d=0}^{2M} g_{ik} \bar{y}_{i}(n)^{d} \quad i = 1, 2 \ldots , 2M$$

where $g_{ik}$ is the $k$th coefficient of the output polynomial of the $i$th filter and $\bar{y}_{i}(n)$ is defined by Equation (11). Note that only the odd powers appear in the summation of Equation (12). Similar to our previous works, [30] and [31], it can be proved that the terms corresponding to the even powers are equal to zero.

The outputs of the filters are summed to produce the combiner output $z(n)$, i.e.:

$$z(n) = \sum_{i=1}^{2M} \sum_{k=0}^{D} g_{ik} \bar{y}_{i}(n)$$

Equation (13) can be expressed in matrix form:

$$z(n) = G_{n} \bar{y}_{n}(n)$$

where $G_{n}$ is an $D(2M+1) \times 1$ vector given by:

$$G_{n} = \begin{bmatrix} g_{11} & \ldots & g_{12M} & \ldots & g_{2M1} \end{bmatrix}^T$$

and $\bar{y}_{n}(n)$ is an $D(2M+1) \times 1$ vector defined as:

$$\bar{y}_{n}(n) = \begin{bmatrix} \bar{y}_{1}(n) & \ldots & \bar{y}_{2M}(n) & \ldots & \bar{y}_{D}(n) \end{bmatrix}^T$$

where $\bar{y}_{p}(n)$ is defined as the $p$th power of $\bar{y}(n)$:

$$\bar{y}_{p}(n) = \begin{bmatrix} \bar{y}_{1}(n)^{p} & \ldots & \bar{y}_{2M}(n)^{p} \end{bmatrix}^T$$

In fact, as shown in Figure 3, a Hammerstein combiner can be modeled as a nonlinear subsystem for generating $y_{n}(n)$ followed by a linear subsystem defined by the vector $G_{n}$.

Our goal is to find $G_{n}$ such that the combiner output $z(n)$ is an optimum estimate of the transmitted symbol. In section 5, we present a proper cost function for obtaining the optimum filter coefficients.
4. THEORETICAL BASIS

In this section, we explain our motivation for using Hammerstein filters in the proposed system. For simplicity, we first consider a system with $M=1$. In this case, at any specific time $n$, the observed signals at the receiver are:

\[
\begin{align*}
\bar{x}_i(n) & = h_{ij} x(n) + w_{ij}(n) \\
\bar{y}_i(n) & = h_{0ij} x(n) + w_{0j}(n)
\end{align*}
\]  

(18)

Based on the observed data, we would like to estimate the transmitted symbol $x(n)$. Using MMSE criterion, the optimum Bayesian estimator $\hat{x}(n)$ is defined as below [33]:

\[ z = E\{x|\bar{y}, \bar{y}_j\} = \int p(x|\bar{y}, \bar{y}_j)dx \]  

(19)

where the notation $p(.)$ denotes the probability density function (PDF), and the time index is omitted for notation simplicity. The conditional PDF can be written as:

\[ p(x|\bar{y}, \bar{y}_j) = \frac{p(\bar{y}, \bar{y}_j | x) p(x)}{\int p(\bar{y}, \bar{y}_j | x) p(x) dx} \]  

(20)

The noise components in Equation (18) are uncorrelated zero mean Gaussian random variables. Hence, for a particular channel occurrence, the joint conditional PDF of the observed data $\{\bar{y}, \bar{y}_j\}$, conditioned on the transmitted sequence becomes:

\[ p(\bar{y}, \bar{y}_j | x) = p(\bar{y}_j | x) p(\bar{y}_j | x) = \frac{1}{(2\pi\sigma_y)^{M-M-1}} \exp \left\{-\frac{1}{2\sigma_y^2} (\bar{y}_j - h_{ij} x)^2 \right\} \]  

(21)

Also, based on our assumptions, we have:

\[ p(x) = \frac{1}{2} [\delta(x+1)+\delta(x-1)] \]  

(22)

Finally, by substituting Equations (20)-(22) in Equation (19), the MMSE estimator can be obtained as:

\[ z = \tanh \left\{ 2 \left( h_{ij} \bar{y}_j + h_{0ij} \bar{y}_j \right) \right\} \]  

(23)

Now, we generalize this result for any arbitrary value of $M$ as below:

\[ z(n) = \tanh \left\{ 2 \sum_{j=1}^{M} \bar{h}_j \bar{y}_j(n) \right\} \]  

(24)

where:

\[ \bar{h}_j = \begin{bmatrix} h_{ij} \\ h_{0ij} \end{bmatrix}, \quad 1 \leq i \leq M, \quad 1 \leq j \leq 2M \]  

(25)

The Maclaurin expansion of Equation (24) yields:

\[ z(n) = \begin{bmatrix} h_1 \bar{y}_1(n) \\ \vdots \\ h_M \bar{y}_M(n) \end{bmatrix} + \frac{2}{15} \begin{bmatrix} h_1 \bar{y}_1(n) \\ \vdots \\ h_M \bar{y}_M(n) \end{bmatrix}^3 + \cdots \]  

(26)

We can write this equation as $z(n) = S_1(n) + S_2(n)$, where:

\[ S_1(n) = \sum_{i=1}^{2M} \sum_{j=1}^{M} C_{ij} \bar{y}_j(n), \quad S_2(n) = \text{Other Terms} \]  

(27)

Here, the coefficients $C_{ij}$ are known parameters that can be obtained from Equation (26). As can be seen, $S_1(n)$ is similar to Equation (13) which is the output of a BHDC system. Furthermore, $S_1(n)$ is a subset of the optimum estimator $z(n)$. From the above discussion, we conclude that a BHDC system can be considered as a subset of the optimum estimator for the transmitted symbols.

5. CALCULATION OF THE OPTIMUM COEFFICIENTS

5.1. Cost Function

Since in blind systems no training sequence exists, an exact value of error signal is not available. Hence, instead of MMSE criteria we present a more convenient cost function for our optimization problem. To obtain a closed form for our cost function, we first recall that a system with $M$ complex Rayleigh channels can be modeled by a system with $2M$ real Gaussian channels. Hence, to simplify our notations, in this section we assume real values for the channel and the noise vectors, without any loss of generality. Having made this assumption, Equation (16) becomes:

\[ \mathbf{y}_d = [y_1 | \ldots | y_{2M}] = [y_1 | \ldots | y_1 | \ldots | y_{2M}]^T \]  

(28)

Each element of this vector can be divided into two terms as below:

\[ y_k(n) = (h_k x(n) + w_k(n))^k = \sum_{j=0}^{k} \binom{k}{j} (h_k x(n))^j (w_k(n))^{k-j} \]  

(29)

where $s_{k}(n)$ and $n_{k}(n)$ are the noise-free (signal) term and the noisy term respectively.

From Equations (28) and (29) we have:

\[ \mathbf{y}_d(n) = [y_1(n) | \ldots | y_{2M}(n)] \]  

\[ = [s_1(n) | \ldots | s_{2M}(n) + n_{2M}(n)] \]  

(30)

\[ + [n_1(n) | \ldots | n_{2M}(n)] = \text{SIG}(n) + \text{NOI}(n) \]
where as shown in Figure 3, \(\text{SIG}(n)\) and \(\text{NOI}(n)\) are the signal vector and the noise vector at the output of the nonlinear subsystem of the Hammerstein combiner, respectively. The combiner output can then be written as:

\[
z(n) = G^*_n Y_n(n) = G^*_n (\text{SIG}(n) + \text{NOI}(n)) = G^*_n \text{SIG}(n) + G^*_n \text{NOI}(n) = z_s(n) + z_r(n)
\]

(31)

where \(z_s(n)\) and \(z_r(n)\) are the signal part and the noise part of the combiner output respectively.

Now, we define our proposed cost function as below:

\[
J = \frac{P_s}{P_r}
\]

(32)

where \(P_s\) and \(P_r\) are the average powers of \(z_s(n)\) and \(z_r(n)\) respectively, given by:

\[
P_s = E\left[|z_s|^2\right] = E\left[|G^*_n \text{SIG}(n)|^2\right] + E\left[|G^*_n \text{NOI}(n)|^2\right]
\]

(33)

\[
P_r = E\left[|z_r|^2\right] = E\left[|G^*_n \text{SIG}(n)|^2\right] + E\left[|G^*_n \text{NOI}(n)|^2\right]
\]

(34)

Hence, we can write:

\[
J = \frac{P_s}{P_r} = \frac{G^*_n R_s G_n}{G^*_n R_s G_n}
\]

(35)

where \(R_s\) and \(R_n\) are the autocorrelation matrices of \(\text{SIG}(n)\) and \(\text{NOI}(n)\) respectively:

\[
R_s = E\left[\text{SIG}(n)\text{SIG}^\dagger(n)\right]
\]

(36)

\[
R_n = E\left[\text{NOI}(n)\text{NOI}^\dagger(n)\right]
\]

(37)

Our goal is to achieve an estimate for the transmitted symbol using the noise free part of the combiner output. On the other hand since the transmitter employs BPSK signaling, \(P_s\) must be equal to one in absence of AWGN and the fading effect. Hence, we maximize the cost function \(J\) in Equation (32) with the constraint \(P_s = 1\). This constrained optimization problem will be solved in subsection 5-3.

5. 2. Blind Estimation of the Autocorrelation Matrices

To find the cost function \(J\) in Equation (32), we first need to calculate the autocorrelation matrices \(R_s\) and \(R_n\). An insight into Equations (29) and (30) reveals that while the elements of \(Y_n(n)\) are observable at the receiver, the elements of \(\text{SIG}(n)\) and \(\text{NOI}(n)\) are not directly available. Hence, \(R_s\) and \(R_n\) have to be estimated from the observable information, using an indirect method.

In this blind method, as the first step, we obtain the closed forms of \(R_s\) and \(R_n\). To explain our method, without loss of generality, we calculate these autocorrelation matrices for the special case where \(M = 2\) and \(D = 3\). Using the equations and the assumptions mentioned before, in this case we obtain:

\[
R_s = \begin{bmatrix}
h^1_1 & h^1_2 & h^1_3 & h^1_4 \\
h^2_1 & h^2_2 & h^2_3 & h^2_4 \\
h^3_1 & h^3_2 & h^3_3 & h^3_4 \\
h^4_1 & h^4_2 & h^4_3 & h^4_4
\end{bmatrix}
\]

(38)

and:

\[
R_n = \begin{bmatrix}
\sigma^2_s & 0 & 3h^1_1\sigma^2_s + 3\sigma^2_n & 0 \\
0 & \sigma^2_s & 3h^2_1\sigma^2_s + 3\sigma^2_n & 9h^1_1\sigma^2_s + 9h^2_1\sigma^2_n \\
3h^1_1\sigma^2_s + 3\sigma^2_n & 3h^2_1\sigma^2_s + 3\sigma^2_n & 9h^1_1\sigma^2_s + 9h^2_1\sigma^2_n & 9h^1_1\sigma^2_s + 60h^2_1\sigma^2_n \\
0 & 0 & 9h^1_1\sigma^2_s + 9h^2_1\sigma^2_n & 9h^1_1\sigma^2_s + 60h^2_1\sigma^2_n
\end{bmatrix}
\]

(39)

As can be seen, the elements \(R_s\) and \(R_n\) are functions of channel taps and noise moments. To estimate these parameters, we first calculate the following statistical averages:

\[
\alpha = E\{y^2_1(n)\} = h^2_1 + \sigma^2_s
\]

(40)

\[
\beta = E\{y^2_2(n)\} = h^2_1 + \sigma^2_s
\]

(41)

\[
\gamma = E\{y_1(n)y_2(n)\} = h_1h_2
\]

(42)

Using the above equations, we can write:

\[
\theta = h^2_1 = \frac{(\alpha - \beta) + \sqrt{(\alpha - \beta)^2 + 4\gamma^2}}{2}
\]

(43)

\[
\eta = h^2_1 = \frac{\alpha^2}{\theta}
\]

(44)

\[
\zeta = \sigma^2_s = |\alpha - \theta|
\]

(45)

Since the signals \(y_1(n)\) and \(y_2(n)\) are observable at the receiver, it is possible to estimate the parameters at the left hand side of Equations (40)-(45). The estimated values are:

\[
\hat{\alpha} = \frac{1}{N_s} \sum_{s=1}^{N_s} y^2_1(n)
\]

(46)

\[
\hat{\beta} = \frac{1}{N_s} \sum_{s=1}^{N_s} y^2_2(n)
\]

(47)

\[
\hat{\gamma} = \frac{1}{N_s} \sum_{s=1}^{N_s} y_1(n)y_2(n)
\]

(48)

\[
\hat{\theta} = \frac{(\hat{\alpha} - \hat{\beta}) + \sqrt{(\hat{\alpha} - \hat{\beta})^2 + 4\hat{\gamma}^2}}{2}
\]

(49)

\[
\hat{\eta} = \frac{\hat{\alpha}^2}{\hat{\theta}}
\]

(50)

\[
\hat{\zeta} = |\hat{\alpha} - \hat{\theta}|
\]

(51)

where \(N_s\) is the number of received symbols. Having calculated these parameters, the estimated values of \(R_s\) and \(R_n\) can be obtained from Equations (38) and (39). Note that in our technique, there is not any training sequence and \(N_s\) received symbols used for the estimations are parts of information data. Hence, BHDC is a completely blind method. It is worth saying
that the proposed technique can easily be applied to any modulation scheme and any arbitrary value of $M$ and $D$.

5. 3. Optimization As mentioned before, in BHDC technique the filter coefficients are calculated such that the cost function $J$ in Equation (32) is maximized with the constraint $P_c = 1$. Hence, we are encountered with a constrained optimization problem, which depending on the rank of matrix $R_s$, has two different solutions as follows:

5. 3. 1. Case 1 $R_s$ has a full rank. In this case we have $r = t$, where $r$ and $t = M(D+1)$ are the rank and the dimension of $R_s$ respectively. We can write the singular value decomposition (SVD) of the positive definite matrix $R_s$ as:

$$\tilde{R}_s = Q A^T Q^+ = (Q \Lambda) (Q \Lambda)^T$$

where $Q$ is a $t \times t$ matrix that consists of the orthonormal eigenvectors of $R_s$:

$$Q = [q_1 \ q_2 \ \ldots \ q_t]$$

and $\Lambda^2$ is a diagonal $t \times t$ matrix of the eigenvalues of $R_s$:

$$\Lambda^2 = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_t \end{bmatrix}$$

and $\Lambda$ is a matrix with elements equal to the square root of the elements of $\Lambda^2$. Note that since $R_s$ has a full rank, its eigenvalues are all positive. We, then define the $t \times t$ matrix $\tilde{R}_s$ as below:

$$\tilde{R}_s = (Q \Lambda)^+ R_s (Q \Lambda^T)^+$$

Finally, as proved in the Appendix, the optimum values of the filter coefficients are:

$$G_{opt} = (\Lambda Q^+)^+ I_{mn}$$

where $\lambda_{mn}$ is the minimum eigenvalue of $\tilde{R}_s$, and $I_{mn}$ is its unit-length vector corresponding eigenvector.

5. 3. 2. Case 2 $R_s$ has not a full rank. In this case, we have $r < t$ and the SVD of the positive semidefinite matrix $R_s$ is written as:

$$R_s = [U | N] \begin{bmatrix} S^T \\ 0 \\ \vdots \\ 0 \end{bmatrix} [U | N]^T$$

where $S^T$ is a diagonal $r \times r$ matrix of the non-zero eigenvalues of $R_s$ as:

$$S^T = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_r \end{bmatrix}$$

and $U$ is a $t \times r$ matrix that consists of the eigenvectors of $R_s$ corresponding to signal space:

$$U = [u_1 \ u_2 \ \ldots \ u_r]$$

Also, $N$ is a $t \times (t-r)$ matrix that consists of the eigenvectors of $R_s$ corresponding to null space:

$$N = [n_1 \ n_2 \ \ldots \ n_t]$$

We now define the $r \times r$ matrix $\tilde{R}_s$ as:

$$\tilde{R}_s = S^T U (1 - R_s N (N'R_s N)^+ N') R_s (1 - N (N'R_s N)^+ N') U S^T$$

where $S$ is a matrix that its elements are the square root of the elements of $S^T$. As proved in the Appendix, the optimum values of the filter coefficients in this case are now obtained as:

$$G_{opt} = \begin{bmatrix} 1 - N (N'R_s N)^+ N' \end{bmatrix} U S^T I_{mn}$$

where $I_{mn}$ is the eigenvector of $\tilde{R}_s$ corresponding to its minimum eigenvalue.

6. SIMULATION RESULTS AND DISCUSSION

In this section, the simulation results of our proposed BHDC technique are presented and compared with the results obtained from an ideal maximal ratio combining (MRC) system. In this work, Rayleigh flat fading channels are considered, and simulations are performed for 200,000 random channel realizations. The average system performance is then obtained by Monte Carlo method. Since each Rayleigh channel is equivalent to two Gaussian channels, our results are also valid for Gaussian flat fading channels. As mentioned before, BHDC is a blind technique in which, the received bits used for estimation are a part of information data. In this work, the received data blocks with block length $N_c = 100$ are used for channel estimation.

6. 1. Average System Performance The average BER versus SNR for both BHDC and ideal MRC systems are shown in Figure 4. This figure is plotted for one Rayleigh channel equivalent to two Gaussian channels, and simulations are performed for two different filter orders $D = 3$ and $D = 5$. As can be seen, the performance of BHDC is very close to ideal MRC. This is a valuable result, especially because higher
bandwidth efficiency is also achieved as compared with MRC.

A similar comparison is shown in Figure 5 for two Rayleigh channels equivalent to four Gaussian channels. As in the previous case, the performance of BHDC is very close to ideal MRC, especially at lower values of SNRs. It is also apparent from these results that the system performance does not change significantly when \( D \) changes, hence, we choose \( D = 3 \) in our simulations.

To see the effect of received data block length \( N_r \) on system performance, simulations are performed for three different values of \( N_r \in \{ 20, 100, 500 \} \). As can be seen from Figure 6, the results are almost the same, and we therefore choose \( N_r = 100 \).

6.2 Reliability Although the BER, averaged over all possible channel realizations, is usually considered as a measure for system performance, in many practical situations like voice communications, the users expect reliable communications while using the system and do not care about the average performance. On the other hand, there are some rare channel realizations that cause significant error rate reducing the average system performance.

Here, we focus on individual channel realizations and compare the performances of BHDC and ideal MRC systems. Based on the above discussion, we define Relative Reliability Factor (RRF) as the probability that for a particular channel occurrence, \( H \) the BER of BHDC is less than or equal to the BER of ideal MRC, i.e.:

\[
RRF = \text{Prob}( \text{BER}_{\text{BHDC}} \leq \text{BER}_{\text{MRC}} | H )
\]

(63)

RRF versus SNR is calculated for 200,000 two Rayleigh diversity channels and the result is plotted in Figure 7. As can be seen, when \( \text{SNR} \geq 7 \text{ dB} \), in almost 90 percent of channel realizations the performance of BHDC is equal to or better than ideal MRC. For \( \text{SNR} \geq 10 \text{ dB} \) this probability is almost 100 percent. This result shows that at moderate and high SNRs the performance of BHDC for most channel realizations is not worse than ideal MRC.

6.3 Comparison of MRC and BHDC for Soft Decision Applications In soft decision systems where no hard limiter is present, the combiner output is taken as the decoding information [2]. To compare our proposed combiner with ideal MRC in soft decision purposes, we have plotted the histograms representing the variation of the combiner output values around the transmitted symbol values. The simulations were performed for a particular two diversity Rayleigh channel with normalized coefficients and 4,000,000 transmitted binary symbols. The results for both MRC and BHDC techniques are plotted in Figures 8 and 9 for \( \text{SNR} = 5 \text{ dB} \) and \( \text{SNR} = 20 \text{ dB} \) respectively. It is apparent from these figures that for both SNRs, the outputs of BHDC combiner are much closer to the desired symbol values \( \pm 1 \) as compared with the corresponding values in MRC system. Hence, for soft decision applications it is superior to MRC technique.
7. CONCLUSION

In this paper, we proposed a blind diversity combining technique using Hammerstein type filters. To show the performance of our proposed technique, simulations were performed for Rayleigh flat fading channels and BPSK modulation in presence of AWGN. Simulation results of BHDC technique were compared with the results obtained from an ideal maximal ratio combining (MRC) system. We also defined relative reliability factor (RRF) to compare the performances of BHDC and ideal MRC for any channel realization. From our simulation results, we conclude that:

i) The average BER of BHDC is very close to ideal MRC. This is a valuable result, especially because higher bandwidth efficiency is also achieved as compared with MRC.

ii) At moderate and high SNRs, for any channel realization, the probability of the BER of BHDC being lower than or equal to ideal MRC is very high.

iii) The outputs of BHDC combiner are much closer to the desired symbol values ±1 as compared with the corresponding values in MRC system. Hence, for soft decision applications it is superior to MRC technique.

9. REFERENCES

4. Loskot, P. and Beaulieu, N., "A unified approach to computing error probabilities of diversity combining schemes over

**APPENDIX**

In this appendix the proofs of the Equations (56) and (62) are presented. As explained in section 5.3, we desire to maximize the cost function $J$ in Equation (32) with the constraint $P_1 = 1$, i.e., to find $G_s$ such that $P_s = G_s^*R_s G_s$ is minimized given $P_s = G_s^*R_s G_s = 1$. We consider two cases:

**Case 1:** $R_s$ has a full rank. In this case we have $r = t$, where $r$ and $t = M(D+1)$ are the rank and the
dimension of $R_s$ respectively. According to Equation (33) and Equations (52)-(54) we have:

$$P_s = G_s^a R_s G_a = G_s^a (Q A) (Q A)^T G_a = I$$  

(A-1)

and:

$$QQ^T = Q^T Q = I$$  

(A-2)

We define the $t \times 1$ vector $a$ as:

$$a = (Q A)^T G_a = (Q A)^T G_a$$  

(A-3)

the constraint $P_s = I$ can now be written as:

$$G_s^a R_s G_a = a^a a = I$$  

(A-4)

This means that $a$ is a unite-length vector. Moreover, since $R_s$ has a full rank, the matrix $AQ^T$ is invertible and we have:

$$G_a = (AQ)^T a$$  

(A-5)

Substituting Equation (A-5) in Equation (34) leads to:

$$P_s = G_s^a R_s G_a = a^a ((AQ)^T) (AQ)^T a = a^a (Q A)^T R_s (Q A)^T a$$

(A-6)

where $\tilde{R}_s$ is defined by Equation (55). Finally, $P_s$ in Equation (A-6) is minimized with the constraint Equation (A-4). Using the lagrangian method for optimization [33], the following equation must be minimized:

$$A = a^a \tilde{R}_s a - K (a^T a - 1)$$  

(A-7)

where $K$ is constant value. We have:

$$\nabla A = 2 \tilde{R}_s a - 2 Ka = 0 \Rightarrow \tilde{R}_s a = Ka$$  

(A-8)

where $\nabla$ is the gradient operator. Consequently, if $a$ is an orthonormal eigenvector of $\tilde{R}_s$, it can be a solution for Equation (A-8) and $K$ is its corresponding eigenvalue. Therefore, there exist $t$ solutions for the above equation among which only one leads to the global minimum. To find this global minimum, we use the SVD of the positive definite matrix $\tilde{R}_s$ as below:

$$\tilde{R}_s = L \ l^T \ l^T$$  

(A-9)

where $L$ is a $t \times t$ matrix that consists of the orthonormal eigenvectors of $\tilde{R}_s$:

$$L = \begin{bmatrix} l_1 & l_1 & \ldots & l_1 \end{bmatrix}$$  

(A-10)

and $\ l^T$ is a diagonal $t \times t$ matrix with eigenvalues of $\tilde{R}_s$:

$$\ l^T = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_t \end{bmatrix}$$  

(A-11)

Equation (A-9) can be written as:

$$\tilde{R}_s = \sum_{t=1}^{T} \lambda_t l_i l_i^T$$  

(A-12)

Using Equations (A-6) and (A-12) $P_s$ can be written as:

$$P_s = \sum_{t=1}^{T} \lambda_t a^a l_i l_i^T a = \sum_{t=1}^{T} |\lambda_t a^a l_i|^2$$  

(A-13)

According to Equation (A-8), the vector $a$ can be each of the eigenvectors $l_i$. Therefore, due to orthonormality of the eigenvectors we have:

$$a_{opt} = l_{min} \Rightarrow P_{min} = \lambda_{min}$$  

(A-14)

where $a_{opt}$ is a solution of Equation (A-8) that leads to the global minimum $P_{min}$. Also, $\lambda_{min}$ is the minimum eigenvalue of $\tilde{R}_s$, and $l_{min}$ is its unite-length corresponding eigenvector. Using Equations (A-5) and (A-14) the optimum values of the filter coefficients are:

$$G_a = (AQ)^T l_{min}$$  

(A-15)

Case 2: $R_s$ has not a full rank. In this case we have $r < t$. According to Equations (57)-(60) we have:

$$\begin{bmatrix} U^U & \cdot \\ N^N & \cdot \end{bmatrix} = \begin{bmatrix} U & \cdot \\ N & \cdot \end{bmatrix}$$  

(A-16)

Note that:

$$\begin{bmatrix} U^T & \cdot \\ N^T & \cdot \end{bmatrix} = \begin{bmatrix} U^T & \cdot \\ N^T & \cdot \end{bmatrix}$$  

(A-17)

Using Equations (57)-(59) $R_s$ can be written as:

$$R_s = US^T U^T = (US)(US)^T$$  

(A-18)

where:

$$S = \begin{bmatrix} \sqrt{\lambda_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sqrt{\lambda_t} \end{bmatrix}$$  

(A-19)

the constraint $P_s = I$ can now be written as:

$$P_s = G_s^a R_s G_a = G_a US^T U^T S^T = I$$  

(A-20)

Suppose that similar to Equation (A-3), the $r \times 1$ unite-length vector $a$ is defined as below:

$$a = (US)^T G_a$$  

(A-21)

However, because $R_s$ has not a full rank, this equation is not invertible and a unique solution for $G_a$ cannot be obtained. Hence, in this case we cannot use the definition Equation (A-21). We can write Equation (A-21) as:

$$a = (US)^T G_a = SU^T G_a$$

$$= \begin{bmatrix} \sqrt{\lambda_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sqrt{\lambda_t} \end{bmatrix} G_a = \begin{bmatrix} \sqrt{\lambda_1} G_a \\ \vdots \\ \sqrt{\lambda_t} G_a \end{bmatrix}$$  

(A-22)

In fact, Equation (A-22) implies that the $r \times 1$ vector $a$
is a mapping of the $t \times 1$ vector $G_n$ in the signal space. To avoid the above problem, we now write $G_n$ as below:

$$G_n = Uc + Nb$$  \hspace{1cm} (A-23)

where $c$ and $b$ are arbitrary $r \times 1$ and $(t-r) \times 1$ vectors respectively. Note that Equation (A-23) can be written as:

$$G_n = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_t \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{t-r} \end{bmatrix}$$

Substituting Equations (A-18), (A-23) and (A-16) in Equation (A-20) we have:

$$P_n = G_n^T R_n G_n = (c^T U^T + b^T N^T)(R_n Uc + R_n Nb)$$  \hspace{1cm} (A-25)

Now, we define the $t \times 1$ vector $a$ as:

$$a = Sc$$  \hspace{1cm} (A-26)

Hence, we have:

$$c^T S^T c = (Sc)^T (Sc) = a^T a = 1$$  \hspace{1cm} (A-27)

This means that $a$ is a unit-length vector. Substituting Equation (A-26) in Equation (A-23) becomes:

$$G_n = U S^T a + Nb$$  \hspace{1cm} (A-28)

Substituting Equation (A-28) in Equation (34) leads to:

$$P_n = G_n^T R_n G_n$$

$$= (a^T S^T U^T + b^T N^T)R_n (U S^{-1} a + Nb)$$

$$= a^T S^T U^T R_n U S^{-1} a + b^T N^T R_n U S^{-1} a$$

Finally, $P_n$ in Equation (A-29) is minimized with the constraint Equation (A-27). Using the Lagrangian method, the following equation must be minimized:

$$A = R_n - K \{ a^T a - 1 \}$$  \hspace{1cm} (A-30)

Setting the derivative of Equation (A-30) with respect to $b$ equal to zero we get:

$$V_n A = N^T R_n U S^T a + N^T R_n U S^{-1} a + 2 N^T R_n N b = 0$$  \hspace{1cm} (A-31)

Note that $R_n = R_n^T$, and therefore we have:

$$b_{opt} = -(N^T R_n N)^T N^T R_n U S^T a$$  \hspace{1cm} (A-32)

Substituting Equation (A-32) in Equation (A-29) yields:

$$P_n = a^T \left( S^T U^T \left[ I - R_n \left( N^T R_n N \right)^T N^T \right] \right)$$

$$R_n \left[ \left[ I - N \left( N^T R_n N \right)^T N^T R_n \right] U S^{-1} \right] a$$

$$= a^T \tilde{R}_n a$$

where $\tilde{R}_n$ is defined by Equation (61). Substituting Equation (A-33) in Equation (A-30) we have:

$$A = a^T \tilde{R}_n a - K \{ a^T a - 1 \}$$  \hspace{1cm} (A-34)

This equation is the same as Equation (A-7) and therefore the optimum value for $a$ is obtained as:

$$a_{opt} = l_{min}$$  \hspace{1cm} (A-35)

where $l_{min}$ is the minimum eigenvalue of $\tilde{R}_n$, and $l_{min}$ is its unit-length corresponding eigenvector. Finally, substituting Equations (A-35) and (A-32) in Equation (A-28) the optimum values of the filter coefficients are:

$$G_n = \left[ I - N \left( N^T R_n N \right)^T N^T R_n \right] U S^{-1} l_{min}$$  \hspace{1cm} (A-36)
A Blind Hammerstein Diversity Combining Technique for Flat Fading Channels

A. M. Aminian-Modarres a, M. Molavi-kakhki b

a Department of Electrical Engineering, Sadjad Institute of Higher Education, Mashhad, Iran
b Department of Electrical Engineering, Ferdowsi University of Mashhad, Mashhad, Iran

Paper history:
Received 22 February 2012
Received in revised form 16 May 2012
Accepted 18 October 2012

Keywords:
Nonlinear Signal Processing
Nonlinear Diversity Combining
Hammerstein Filter
Rayleigh Flat Fading Channel

A blind diversity technique for flat fading channels is presented. It is based on using a new Hammerstein filter architecture for linearizing the fading channel. The filter is designed to operate at the diversity combining stage of a receiver. It consists of a linearizer followed by a memoryless nonlinearity. The linearizer is designed to approximate the channel so that the output of the nonlinearity is close to the channel output. The nonlinearity is designed to provide a good match to the linearizer. The proposed technique offers a significant performance improvement over existing techniques, particularly in multi-path environments.