Abstract—In this paper, the problem of signal-to-noise ratio (SNR) estimator design for a single-input multiple-output (SIMO) communication system employing non-coherent M-ary frequency shift keying (NCMFSK) modulation scheme is considered. The transmitted signal undergoes Rayleigh fading and additive white Gaussian noise (AWGN) and is received at a receiver with L diversity branches. Closed-form expressions of data aided (DA) and non-data aided (NDA) estimators have been derived using the maximum likelihood (ML) estimation approach. Cramer-Rao bound has been evaluated to compare the performance of the designed estimators. The effect of increasing the receiver diversity branches on the performance of estimators has been quantified.

Index Terms—SNR Estimation, Receiver Diversity Combining.

I. INTRODUCTION

Signal-to-noise ratio (SNR) is one of the most important and critical performance metrics for analyzing the performance of wireless communication systems. Prime utilization of SNR estimates dwells in various receiver functions, e.g., for decoding of data symbols, adaptive modulation and coding (AMC), turbo decoding, and power control algorithms. SNR estimation also finds application in many areas of cellular as well as in wireless sensor networks (WSNs). In a WSN, the estimates of SNR can be used to find the candidate cooperators in a cooperative communication environment [1] and [2]. Similarly, they can be used to find the one-hop neighbors of a wireless sensor node [3].

In WSNs, the wireless sensor nodes are highly energy constrained and such a modulation scheme is desirable, which is relatively power efficient and offers less receiver complexity. Non-coherent frequency shift keying (NCFSK) fulfills both the criteria; power efficiency at the transmitter side by having constant signal envelope and less receiver complexity because of squared envelope detection. This scheme does not require sophisticated signal processing algorithms based on phase-locked loops (PLLs) for carrier phase synchronization. Furthermore, in data fusion applications of WSNs, the sensor nodes transmit their data to a central facility or base station [4], which is generally equipped with multiple antennas. Each transmitted symbol from a node follows a separate path to each antenna and hence undergoes independent fading. The SNR of the received symbol on the multiple co-located antennas is estimated at the receiver. This SNR information is used in adaptively assigning the modulation and coding scheme to be used at the transmitter and other receiver functions. Therefore, in this paper, we consider the problem of estimating SNR in a communication system employing NCFSK modulation over Rayleigh fading channels with diversity combining. Because of the non-coherent nature of modulation scheme, equal gain combining (EGC) is employed at the receiver side to combine the data arriving in all the receiver branches [5].

Several authors have worked on designing SNR estimators for different environments and receiver architectures. Most of the work has been done on the derivations of M-ary phase shift keying (MPSK) and frequency shift keying (MFSK) estimators. For example, in [6], the authors have analyzed BPSK and 8-PSK SNR estimators using different estimation techniques and have presented the comparison between the performance of each of them considering real and complex additive white Gaussian noise (AWGN) channels, respectively. In [7], the authors have designed an SNR estimator using the method of moments (MoM) for the case of BPSK modulated signals in Nakagami-m fading channels with receiver diversity. In [8] and [9], the authors have designed SNR estimators for non-coherent BFSK and MFSK receivers, respectively, over fading channel in the presence of AWGN using maximum likelihood estimation (MLE) and data statistics approach. Moreover, different scenarios have been taken into account, i.e., data aided (DA), non-data aided (NDA) and method of moments (MoM) and also their comparison is presented in terms of their performance. Authors in [10] designed SNR estimators for a slow fading environment. Carrier frequency offset effects have been taken into consideration while estimating SNR in [11], [12]. However, in all of these works the approach is only valid for a single-input single-output (SISO) system, which is not sufficient for a WSN where multiple links are formed to transmit the data and constitutes the main motivation behind this study.

In this paper, we consider a single-input multiple-output (SIMO) system, i.e., we will take into account the diversity at the receiver. We derive ML estimator for the cases of data aided and non-data aided scenarios. Then we compare their performances in the terms of normalized mean squared error (NMSE) and the Cramer-Rao bound (CRB).

The rest of the paper is organized as follows. Section II of the paper explains the system model under consideration. Derivations of the SNR estimators using maximum likelihood
technique are presented in Section III. Section IV derives the
Cramér-Rao bound (CRB), and in Section V, we discuss the
simulation results for both estimators and their performance in
terms of mean squared error and CRB. At the end, conclusions
are provided in Section VI.

II. SYSTEM MODEL

In this paper, we are considering a SIMO communication
system as shown in Fig. 1 having single transmit and
L co-located receive antennas. Non-coherent MFSK is em-
ployed where each transmitted symbol undergoes independent
Rayleigh fading and is corrupted by AWGN.

![Fig. 1: SIMO communication system with L co-located receiver antennas](image)

The block diagram of a non-coherent BFSK receiver with L
diversity branches and envelope detection is shown in Fig. 2. Each
diversity branch is further divided into M receiver sub-
branches corresponding to the MFSK receiver. For the case
of Fig. 2, we are considering non-coherent BFSK receiver, so
every $\ell$th diversity branch is further divided into two receiver
sub-branches.

We get $L$ copies of the transmitted symbol at the receiver.
The copy of the signal acquired on the $\ell$th diversity branch at
any time instant $i$ after correlator is given by

$$v_{\ell,i} = s_i \alpha_{\ell,i} + n_{\ell,i},$$

where $\ell = \{1, 2, ..., L\}$ and $i = \{1, 2, ..., K\}$ represent the
diversity branch and time index, respectively. Each of $v_{\ell,i}$, $s_i$, and $n_{\ell,i}$
are independent vectors with a dimension of $M \times 1$, where
$v_{\ell,i} = \{v_{\ell,1,i}, v_{\ell,2,i}, ..., v_{\ell,M,i}\}$. The $v_{\ell,m,i}$ corresponds to the
$\ell$th received copy of the transmitted symbol in the $m$th receiver
sub-branch $(1 \leq m \leq M)$ at the $i$th time instant. In (1),
$s_i = [0, 0, ..., 0, 1, 0, ..., 0]^T$ is the vector transmitted from
the MFSK transmitter, where 1 at the $m$th $(1 \leq m \leq M)$
position corresponds to the transmitted frequency and 0 is
set in all the remaining $(M - 1)$ positions. Fading in every
$\ell$th diversity branch is represented by $\alpha_{\ell,i}$. Since Rayleigh
fading is considered, so the elements of $\alpha_{\ell,i}$ are drawn from
a complex Gaussian distribution, i.e., $\alpha_{\ell,i} \in CN(0, S)$, where
$S$ is the variance of fading. AWGN in the symbol present
in the $m$th receiver sub-branch of the $\ell$th diversity branch
and at the $i$th time instant is represented by $n_{\ell,m,i} \in \mathbb{C}$. Elements of $n_{\ell,i}$
belong to complex Gaussian distribution, i.e.,
$n_{\ell,m,i} \in CN(0, N)$, where $N$ is the noise power.

From Fig. 2, we can see that $v_{\ell,m,i}$ will pass from the
envelope detector; Thus, $x_{\ell,m,i} = |v_{\ell,m,i}|^2$. Thereafter, equal
gain combining (EGC) is employed on the data $x_{\ell,m,i}$ present
in all the $M$ sub-branches of $L$ antennas of the receiver to get
the vector $r_{1} = [r_{1,1,i}, r_{2,1,i}, ..., r_{M,1,i}]^T$. In Fig. 2, data from first
sub-branch of every antenna has been summed up to get a
symbol $r_{1,i} = \sum_{\ell=1}^{L} x_{\ell,1,i}$ and similarly data from the second
sub-branch of every antenna has been summed up to get the
second symbol $r_{2,i} = \sum_{\ell=1}^{L} x_{\ell,2,i}$. As we are considering
non-coherent BFSK in Fig. 2, so we will be left with a $(2 \times 1)$
vector i.e., $r_{1} = [r_{1,1,i}, r_{2,1,i}]^T$. In this paper, we are interested in estimating average SNR from
$[r_{1} \ r_{2} \ r_{3} \ ... \ r_{K}]^T$, which is acquired after the post-detection combining of the $K$
received data symbols. This is done for several estimation
schemes, which will be discussed in the forthcoming section.

III. ESTIMATION TECHNIQUES

This section contains the derivation of data aided (DA) and
non-data aided (NDA) estimator expressions using maximum
likelihood (ML) estimation technique.

A. Data Aided MLE

The objective here is to estimate the average SNR of $K$
consecutively received symbols. Hence without losing gener-
ality, we set $s_i = [1 \ 0 \ 0 \ ... \ 0]^T$ for each of the $K$ symbols. On
the basis of this assumption, the data is received in the first
branch of receiver and all the remaining $(M - 1)$ branches
contain noise. As we are considering postdetection combining,
the data to be used is $r_{m,i} \in r_{1} = [r_{1,1,i}, r_{2,1,i}, ..., r_{M,1,i}]^T$.
Because of the fact that $\alpha_{\ell,i}$ and $n_{\ell,m,i}$ are Complex Gaussian
random variables, so the first symbol from every diversity
branch, i.e., $x_{\ell,1,i} = |\alpha_{\ell,i} + n_{\ell,m,i}|^2$, follows an exponential
distribution having mean of $E|\alpha_{\ell,i}|^2 + E|n_{\ell,m,i}|^2$. Moreover,
the data received in the first branch of receiver is he sum of
$L$-exponentially distributed terms, i.e., $x_{\ell,1,i} + x_{\ell,2,i} + ... + x_{\ell,L,i}$
having same mean and will result in $r_{1,i}$ to be gamma
distributed. The probability density function (PDFs) of the

![Fig. 2: Receiver structure of the system employing BFSK modulation and EGC](image)
received pilot symbols $r_1 = [r_{1,i}, r_{2,i}, ..., r_{M,i}]^T$, are given as

$$
p_{r_{1,i}} (r_{1,i}) = \frac{(S + N)^{-L} (r_{1,i})^{L-1}}{(L - 1)!} \exp \left( -\frac{r_{1,i}}{S + N} \right),
$$

and

$$
p_{r_{m,i}} (r_{m,i}) = \frac{(N)^{-L} (r_{m,i})^{L-1}}{(L - 1)!} \exp \left( -\frac{r_{m,i}}{N} \right), \quad \{m = 2, ..., M\}.
$$

The joint PDF of the received vector, $r_1$, becomes

$$
p_{r_1} (r_{m,i}) = \frac{(S + N)^{-L} (N)^{-L(M-1)} \prod_{m=1}^{M} (r_{m,i})^{L-1}}{((L - 1)!)^{M-1}} 
\times \exp \left( -\frac{r_{1,i}}{S + N} - \sum_{m=2}^{M} \frac{r_{m,i}}{N} \right).
$$

Hence the log-likelihood function of $K$ received symbols can be found as

$$
\Lambda_{r_1}(r_{m,i}; S, N) = -KL \ln(S + N) - KL(M - 1) \ln(N) + (L - 1) \sum_{i=1}^{K} \sum_{m=1}^{M} (r_{m,i})M \ln((L - 1)!) - \frac{1}{S + N} \sum_{i=1}^{K} r_{1,i} - \frac{1}{N} \sum_{i=1}^{K} \sum_{m=2}^{M} (r_{m,i}).
$$

Using the fact that ML estimate of the ratio of two parameters is equal to their individual estimates, we can write the estimated SNR expression as

$$
\hat{\gamma}_{DA} = \frac{\hat{S}_{ML}}{\hat{N}_{ML}}.
$$

For this purpose, we want to extract the parameters of our interest, i.e., signal power $\hat{S}_{ML}$ and noise power $\hat{N}_{ML}$, from log-likelihood expression. Differentiating (5) with respect to $S$ and $N$ individually and setting these derivatives equal to zero results in $\hat{S}_{ML}$ and $\hat{N}_{ML}$. Putting these values in (6) and solving, we will get the data aided estimates of SNR as

$$
\hat{\gamma}_{DA} = \frac{(M - 1) \sum_{i=1}^{K} r_{1,i} - \sum_{i=1}^{K} \sum_{m=2}^{M} r_{m,i}}{\sum_{i=1}^{K} \sum_{m=2}^{M} r_{m,i}}.
$$

B. Non-Data Aided MLE

In NDA, we have no knowledge about the transmitted data symbol. So we assume that all transmitted symbols have equal priori probabilities. The conditional PDF of the received symbol given a 1 at the $n^{th}$ position was transmitted is

$$
p_{r_{n,i}} (r_{n,i}|s_n = 1) = \frac{(S + N)^{-L} (r_{n,i})^{L-1}}{(L - 1)!} \exp \left( -\frac{r_{n,i}}{S + N} \right),
$$

and the conditional PDF of the received symbol given a 0 at the $n^{th}$ position was transmitted becomes

$$
p_{r_{n,i}} (r_{n,i}|s_n = 0) = \frac{(N)^{-L} (r_{n,i})^{L-1}}{(L - 1)!} \exp \left( -\frac{r_{n,i}}{N} \right).
$$

Now there are $M$ different possibilities of the received symbol. We can express the joint unconditional PDF of the $M$ received symbols using the law of total probability as

$$
p_{r_1} (r_{m,i}) = \frac{1}{M} (S + N)^{-L} (N)^{-L(M-1)} \prod_{m=1}^{M} (r_{m,i})^{L-1} \exp \left( -\sum_{n=1}^{M} \frac{r_{n,i}}{N} \right) + \sum_{m=2}^{M} \frac{r_{m,i}}{N}.
$$

The above equation is very complex to solve, so simplifying the above expression by factoring the term $\exp \left( -\sum_{n=1}^{M} r_{n,i} \right)$, we get

$$
p_{r_1} (r_{m,i}) = \frac{1}{M} (S + N)^{-L} (N)^{-L(M-1)} \prod_{m=1}^{M} (r_{m,i})^{L-1} \exp \left( -\sum_{n=1}^{M} \frac{r_{n,i}}{N} \right),
$$

where $\psi = \frac{1}{S + N} - \frac{1}{N}$. We get the log-likelihood function as

$$
\Lambda_{r_1}(r_{m,i}; S, N) = -\frac{1}{\psi} \ln(M) - KL \ln(S + N) - KL(M - 1) \ln(N) + (L - 1) \sum_{i=1}^{K} \sum_{m=1}^{M} \ln(r_{m,i}) + \frac{1}{\psi} \sum_{i=1}^{K} \sum_{m=1}^{M} r_{m,i} \ln(r_{m,i}) + \frac{1}{\psi} \sum_{i=1}^{K} \sum_{m=1}^{M} \exp(-r_{m,i} \psi).
$$

We can have $\hat{S}$ and $\hat{N}$ by differentiating log-likelihood function in (12) with respect to $S$ and $N$, exactly in the same way as done for the data aided case in the previous section. We have the following expression

$$
\hat{S} + \hat{N} = \frac{1}{KL} \sum_{i=1}^{K} \left[ \sum_{m=1}^{M} \frac{r_{m,i}}{\psi} \exp(-r_{m,i} \psi) \right] - \frac{1}{\psi} \sum_{m=1}^{M} \exp(-r_{m,i} \psi),
$$

Finding a closed-form solution of the above non-linear expression is prohibitive, so we approximate it to a feasible form. Let us consider this expression for the case of $M = 2$, i.e., let

$$
A = \sum_{i=1}^{K} \frac{r_{1,i}}{\psi} \exp(-r_{1,i} \psi) + r_{2,i} \exp(-r_{2,i} \psi).
$$

It can be observed for the case of very high SNR, i.e., $S >> N$, $\psi = \frac{1}{S + N} - \frac{1}{N}$ reduces to $\psi \approx \frac{1}{N}$. Thus the above approximation becomes
\[ A = \sum_{i=1}^{K} \frac{r_{1,i} \exp(r_{1,i}/N) + r_{2,i} \exp(r_{2,i}/N)}{\exp(r_{1,i}/N) + \exp(r_{2,i}/N)}, \quad (15) \]

Rearranging the above equation, we get

\[ A = \sum_{i=1}^{K} \frac{r_{1,i}}{1 + \exp(r_{1,i}/N)} + \frac{r_{2,i}}{1 + \exp(r_{2,i}/N)}. \quad (16) \]

Among \([r_{1,i}, r_{2,i}],\) only one branch will contain signal and the other will contain noise. Let us consider that the \(r_{1,i}\) contains signal. For the case of high SNR, \(S \gg N,\)

\[
(1 + \frac{\exp(r_{2,i}/N)}{\exp(r_{1,i}/N)}) \to 1 \quad \text{and} \quad (1 + \frac{\exp(r_{1,i}/N)}{\exp(r_{2,i}/N)}) \to \infty.
\]

Thus the above expression reduces to

\[ A \approx \left( \sum_{i=1}^{K} \max_{m=1,2,3,\ldots,M} r_{m,i} \right). \quad (17) \]

Using this expression in (13) and solving, we get the expression for estimated noise power \(\hat{N}\) as

\[ \hat{N} = \frac{(M-1) \sum_{m=1}^{M} r_{m,i} - M \sum_{m=1}^{M} \sum_{i=1}^{K} r_{m,i} + \sum_{i=1}^{K} \max_{m} r_{m,i}}{M \sum_{m=1}^{M} \sum_{i=1}^{K} r_{m,i} + \sum_{i=1}^{K} \max_{m} r_{m,i}}, \quad (18) \]

and the estimate of signal to noise ratio for NDA is given as

\[ \hat{\gamma}_{NDA} = \frac{-\sum_{m=1}^{M} r_{m,i} + M \sum_{m=1}^{M} \sum_{i=1}^{K} r_{m,i}}{\sum_{m=1}^{M} \sum_{i=1}^{K} r_{m,i} - \sum_{i=1}^{K} \max_{m} r_{m,i}}, \quad (19) \]

where, \(r_{m,i} = \sum_{\ell=1}^{L} x_{\ell,m,i}.\)

IV. CRAMER-RAO LOWER BOUND

In order to evaluate the performance of the derived estimators, we find the Cramer-Rao bound (CRB), which is the lower bound on the variance of any estimator. In other words, it states that the variance of the derived estimator must be greater than or equal to this bound. We have derived the CRB for the data aided (DA) case and compared it with the normalized mean squared error (NMSE) to judge the performance of estimator. Although we can have CRB for non-data aided case as well, however as the benchmark performance is given by DA method, we use its CRB to evaluate the performance of the derived estimators. We have two unknown parameters, i.e., signal power, \(S\) and the noise power, \(N.\) We consider the unknown vector parameter \(\theta = [S \: N]^T.\) We have

\[ CRB = \frac{\partial g(\theta)}{\partial \theta} \Lambda^{-1}(\theta) \frac{\partial g(\theta)^T}{\partial \theta}, \quad (20) \]

where \(g(\theta)\) is a function of parameter \(\theta\) and \(I\) is the Fisher information matrix. Taking partial derivative of \(g(\theta) = \frac{S}{N}\) with respect to \(\theta = [S \: N]^T,\) we get

\[ \frac{\partial g(\theta)}{\partial \theta} = \begin{bmatrix} 1 & -S \\ N & N^2 \end{bmatrix}, \quad (21) \]

The Fisher information matrix (FIM), \(I(\theta)\) is given by

\[ I(\theta) = \begin{bmatrix} -\mathbb{E} \left( \frac{\partial^2 \Lambda_{DA}}{\partial S^2} \right) & -\mathbb{E} \left( \frac{\partial^2 \Lambda_{DA}}{\partial S \partial N} \right) \\ -\mathbb{E} \left( \frac{\partial^2 \Lambda_{DA}}{\partial N \partial S} \right) & -\mathbb{E} \left( \frac{\partial^2 \Lambda_{DA}}{\partial N^2} \right) \end{bmatrix}, \quad (22) \]

where \(\mathbb{E}\) is the expectation operator. Solving the elements of the above matrix, we get

\[ I(\theta) = \begin{bmatrix} \frac{KL}{(S+N)^2} & \frac{KL}{(S+N)^2} \\ \frac{KL}{(S+N)^2} & \frac{KL}{(S+N)^2} \end{bmatrix}. \quad (23) \]

Putting the values of Equations (23) and (21) in (20), we get the expression for CRB as

\[ CRB_{DA} = \frac{M}{KL(M-1)} (\gamma + 1)^2. \quad (24) \]

V. SIMULATION RESULTS

In this section, the performance of the estimators designed in the previous sections is presented in terms of normalized mean squared error (NMSE), which is given as

\[ NMSE(\hat{\gamma}) = \mathbb{E} \left\{ \frac{(\gamma - \hat{\gamma})^2}{\gamma^2} \right\}. \quad (25) \]

where \(\gamma\) and \(\hat{\gamma}\) being the true and estimated SNR, respectively. A perfect estimator is the one which always results in the least difference between estimated value and true value of the unknown parameter. Different trends of the NMSE versus SNR have been analyzed for several parameters, i.e., diversity branches, \(L,\) receiver sub-branches, \(M\) and the number of symbols \(K.\) All the results presented in this section are averaged over 25,000 trials of simulations.

Fig. 3 shows the NMSE vs. SNR for the DA estimator for the case of \(L = 5,\) i.e., five diversity branches and for \(K = 100\) symbols for various values of \(M.\) It can be observed from the figure that the NMSE is decreased as the receiver
Fig. 4: Effect of increasing diversity branches on data aided estimator for $K = 100$ symbols and $M = 2$.

sub-branches, $M$, are increased. This is due to the fact that by adding more and more receiver sub-branches ($M$), the number of data samples go on increasing, forming a large data set. Thus, the sample mean of the large number of data samples converges towards the actual mean, resulting in a better performance of the estimator. Same trend has been depicted by the NDA estimator, but is not shown here to avoid repetition.

Fig. 4 presents the effect of increasing diversity branches, $L$ on NMSE values of the DA estimator for $K = 100$ symbols. It can be seen that by increasing the value of $L$, the NMSE decreases. This decrease in NMSE is the consequence of increased number of branches of the data to be estimated. Therefore, we can summarize that increase in the values of both, the diversity branches, $L$ and receiver sub-branches, $M$, serves the same purpose of increased data samples, thereby lowering the NMSE and improved estimator performance.

It can however be observed from Fig. 4 that the rate of NMSE reduction is large, when branches are increased from $L = 5$ to $L = 10$. This shows the diminishing returns behavior of the diversity gain on NMSE. Performance comparison curves for the DA and the NDA estimators have been shown in Fig. 5 for $L = 5$ and $K=100$ symbols. Also the CRB derived for DA estimator has been plotted. For low SNR region, the NDA estimator gives large NMSE as compared to the DA estimator. Larger NMSE for the case of NDA in low SNR region can be attributed to the use of approximations derived for high SNR region (14)-(17). The difference between the error of DA and NDA estimator in the low SNR region goes on decreasing as the number of diversity branches are increased. Although not shown here but this difference is high for the cases of $L = \{1 \rightarrow 4\}$ in comparison with $L = 5$ that is shown in Fig. 5. Moreover, it can also be observed that the curve for DA is exactly giving the same NMSE values throughout the SNR region as that of the CRB evaluated for it. This points towards the fact that DA estimator is showing the minimum possible variance and the performance margin is high. This motivates the use of DA estimator in decode-and-forward (DF)-based wireless sensor networks [14]–[18], where the precondition for forwarding the packet is to successfully decode it. The decoding is generally done by using cyclic redundancy check (CRC), and if the packet is decodable, the entire packet can be treated as pilot symbols to perform SNR estimation, which can then be used in various algorithms such as [1], [14], [15].

NMSE contours are presented in Fig. 6 for $L = 4$ and $M = 4$ with SNR shown on the x-axis and packet lengths (symbol size) at the y-axis. NMSE decreases with an increase in both the packet length and the SNR. So, for such situation where small NMSE is needed, larger length packets should be chosen and vice versa, e.g., we can see form the figure that less than 2% error can be achieved if $K \geq 200$ at the SNR values $\geq 8dB$. 

Fig. 5: NMSE for $K = 100$ symbols, $M = 2$ and $L = 5$.

Fig. 6: NMSE contours for $K = 100$ to 1000 symbols, $M = 4$ and $L = 4$. 

VI. CONCLUSION

Data aided (DA) and non-data aided (NDA) signal-to-noise ratio (SNR) estimator expressions have been derived using maximum likelihood estimation (MLE) technique in a NC-MFSK receiver. Receiver diversity has been taken into account for the case of Rayleigh fading channel and AWGN. For comparison purposes, Cramer-Rao bound (CRB) for data aided case has been evaluated. On the basis of analysis done in the previous sections, we have found that by adding the diversity in the system, the performance of estimator is increased for both DA and NDA estimators. However, DA estimator performs best in all the cases as compared to NDA because of an approximation used in the NDA scheme to get a closed-form expression. Difference between the performances of both the estimators is large in low SNR regions. However, by increasing the diversity branches, this difference can be minimized. Moreover, the NDA estimator performs equally well for higher order diversity cases by approaching the performance of DA estimator in the higher SNR region.

REFERENCES