SNR Estimation for M-ARY Non-Coherent Frequency Shift Keying Systems

Syed Ali Hassan, *Student Member, IEEE,* and Mary Ann Ingram, *Senior Member, IEEE*

***Abstract*—This paper considers how to estimate the average signal-to-noise ratio (SNR) for a communication system employing orthogonal non-coherent M-ARY frequency shift keying (NCMFSK), in white Gaussian noise (AWGN) and over both symbol-by-symbol fading channels and block fading channels. The proposed algorithm finds its application in a variety of applications including a cooperative transmission system, which is the main motivation behind this study. The maximum likelihood estimator and one using data statistics have been derived and simulated for various scenarios including data-aided, non-data aided and joint estimation using both the data and pilot sequences. We also derive the Cramer-Rao bound for the estimators in the case of Rayleigh fading channels. The results show that for a particular region of interest (e.g. high SNR or low SNR) and depending upon the availability of pilot sequence, a particular SNR estimation scheme is suitable.**

***Index Terms*—Wireless communications, SNR estimation, non-coherent MFSK, basis for cooperator selection.**

I. Introduction

ESTIMATES of signal-to-noise ratio (SNR) are used in many wireless receiver functions, including signal detection, power control algorithms, and turbo decoding etc. The motivation for the study reported here is that SNR estimation is a way for a receiver to determine if it is near the edge of the decoding range of its source, and therefore, in a preferred location to participate in a cooperative transmission [1-2]. Furthermore, if the radios are energy constrained, e.g., if they are in a sensor network, constant envelope modulation and non-coherent demodulation are desirable to reduce circuit consumption of energy. FSK enables efficient power amplification in the transmitter and a simple receiver design that employs envelope detection. Therefore, in this paper, we consider the estimation of average SNR in an orthogonal FSK non-coherent demodulator, over block fading and Rayleigh fading channels.

Diversity reception is a well known technique for combating multipath fading. In a diversity receiver, transmissions from different diversity branches are combined to improve the link quality. For non-coherent modulation schemes as NCMFSK, selection combining, (SC), performs well in a fading environment. The traditional analysis of SC specifies that of L diversity branches, the one providing the largest SNR be selected for data recovery [3]. Thus, we can utilize the algorithms derived in this paper, to estimate the SNR on each diversity branch and can select the one with largest for subsequent receiver functions.

Several authors have attacked the problem of estimating SNR for binary phase shift keying (BPSK) and frequency shift keying (FSK). For example, [4] compares a variety of techniques for SNR estimation in AWGN for M-PSK signals. Many approaches also include the channel effects such as multipath fading and address the issues of SNR estimation for fading channels for BPSK, e.g., [5]-[7]. In [8], the authors have estimated the average SNR for non-coherent binary FSK (NCBFSK) receiver, assuming a Rayleigh fading channel and unit noise power spectral density. However, in implementations, noise power must also be estimated. Also the approach in [8] cannot be generalized to M-FSK SNR estimation. This paper addresses the design of estimators for the M-FSK case and without the prior knowledge of the noise power. We also derive the Cramer Rao bound, (CRB), for this more general case; however, for binary FSK, the bound simplifies to that of [8].

In this paper, we derive two types of estimators for SNR, a maximum likelihood estimator (MLE) and an estimator that uses data statistics, such that neither of them assume prior knowledge of the noise power. Our algorithm is also applicable for any value of M in an M-FSK receiver, where $M = 2^n$; n being a positive integer. We also consider two types of channels: one with fast Rayleigh fading and the other with slow block fading. However, we notice that the estimators derived in this paper for both cases are significantly different and lead to dramatically different analysis and results. [8] derives the average SNR estimate for pilot and data symbols separately. We provide ML versions of partially data-aided (PDA), non-data aided (NDA), joint PDA-NDA, and fully data-aided (FDA) estimators for average SNR. The PDA approach uses only the training sequence for estimation while the NDA approach does blind estimation using the entire sequence. The joint PDA-NDA uses all the information, operating blindly on the non-training part of the sequence. The FDA estimator uses the detected data as training sequence for SNR estimation and is reasonable in a multi-hop broadcast application, where every node must decode the entire message; the detected data are all assumed to be correct in the paper regardless of the value of SNR.

The rest of the paper is organized as follows. In the next section, we describe the system model and the notations used for the MFSK case. Section III treats the derivations of the SNR estimators, for a Rayleigh fading channel while Section

Paper approved by M.S. Alouini, the Editor for Modulation and Diversity Systems of the IEEE Communications Society. Manuscript received October 15, 2009; revised May 5, 2010, September 28 2010, April 5, 2011, and June 20, 2011.

The authors are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, 30332-0250 USA (e-mail: {alihassan, mai}@gatech.edu).

The authors gratefully acknowledge support for this research from the National Science Foundation under grant CNS-0721296.

Digital Object Identifier 10.1109/TCOMM.2011.080111.090627

IV considers a block flat fading channel. The Cramer Rao Bound (CRB) for the Rayleigh fading case is derived in Section V, and Section VI discusses the simulation results for various estimators and overall estimators' performance. The paper then concludes in Section VII.

II. System Model for the Rayleigh Fading Case

Consider a Rayleigh fading communication system employing M-ARY FSK modulation, where each transmitted symbol is corrupted independently by fading and noise and the number of symbols in the constellation is $M = 2^n$, for a positive integer n. The received signal, after the matched filtering and passing the signal through the square law device, is given as

$$\mathbf{x}_i = |\mathbf{s}_i \alpha_i + \mathbf{n}_i|^2, \tag{1}$$

where $|.|$ is the magnitude operator applied on each element of the above equation and i is the time index[1]. Each of $\mathbf{x}_i$, $\mathbf{s}_i$, and $\mathbf{n}_i$ are vectors with a dimension of $M \times 1$, and $\mathbf{s}_i = [0, ..., 0, 1, 0, ..., 0]^T$, where '1' is in the m_ith position $(1 \leq m_i \leq M)$, where m_i indicates the symbol transmitted at time i, and the other positions have a zero. For the sake of simplicity, we assume that the average symbol energy is unity so that the expected energy of the ith received symbol is given as $\mathbb{E}|\alpha_i|^2 = S$, where α_i is a zero mean fading coefficient drawn from a complex Gaussian distribution. Similarly, the elements of $\mathbf{n}_i$ are also independent complex Gaussian random variables with zero mean and variance $N/2$ per real dimension, thus the signal-to-noise ratio (SNR) is given by $\gamma = S/N$. $\mathbf{s}_i$, α_i, and $\mathbf{n}_i$ are assumed independent of each other. Our interest is to find the estimate of the average SNR using the observed data vector $\left[\mathbf{x}_1^T \ \mathbf{x}_2^T \ \cdots \ \mathbf{x}_k^T\right]^T$. For the estimation schemes considered, we assume that that there are g pilot symbols and l data symbols so that the total packet length is $k = g + l$. Throughout the paper, we assume perfect timing recovery at the receiver.

III. Estimation Techniques for the Rayleigh Fading Environment

As mentioned previously, we will derive the ML estimators for three cases, namely PDA, NDA and *Joint* PDA-NDA. Another approach uses the statistics of observable data, which we call Estimation using Data Statistics (EDS).

A. Partially Data Aided MLE

Without the loss of generality, the g pilot symbols are each set to $[1 \ \ 0 \ \ ... \ \ 0]^T$. The received symbols from M branches are denoted as $x_{m,i}$, where first index m denotes the branch index where $m = 1, 2, ..., M$ and second index i is the time index such that $i = 1, 2, ..., g$. Since both the α_i and n_i in (1) are complex Gaussian, $x_{1,i}$ will have an exponential distribution with a mean of $\mathbb{E}|\alpha_i|^2 + \mathbb{E}|n_i|^2$. The probability density functions (PDFs) of the received pilot symbols $\mathbf{x}_i = [x_{1,i} \ \ x_{2,i} \ \ ... \ \ x_{M,i}]^T$, are given as

$$p_{x_{1,i}}(x_{1,i}) = \frac{1}{S+N} \exp\left(-\frac{x_{1,i}}{S+N}\right), \tag{2}$$

and

$$p_{x_{m,i}}(x_{m,i}) = \frac{1}{N} \exp\left(-\frac{x_{m,i}}{N}\right), \quad m = 2, ..., M. \tag{3}$$

The joint PDF of $\mathbf{x}_i$ is given as

$$p_{\mathbf{x}_i}(\mathbf{x}) = \frac{1}{(S+N)N^{M-1}} \exp\left(-\frac{x_1}{(S+N)} - \frac{1}{N}\sum_{m=2}^{M} x_m\right). \tag{4}$$

Thus the log-likelihood distribution of g received symbols is given as

$$\begin{aligned}\Lambda_{\mathbf{x}_i}(\mathbf{x}; S, N) = &- g \ln(S+N) - g(M-1)\ln N \\ &- \frac{1}{S+N}\left(\sum_{i=1}^{g} x_{1,i}\right) - \frac{1}{N}\sum_{m=2}^{M}\sum_{i=1}^{g} x_{m,i}.\end{aligned} \tag{5}$$

To find the MLE of the SNR, $\hat{\gamma}$, we use the property that the ML estimate of the ratio of two parameters (S and N here), is the ratio of the individual ML estimates of the two parameters [9]. Thus

$$\hat{\gamma} = \frac{\hat{S}_{ML}}{\hat{N}_{ML}}. \tag{6}$$

Thus by differentiating (5) with respect to S and N, and setting the derivatives equal to zero results in

$$\hat{\gamma}_{DA} = \frac{(M-1)\sum_{i=1}^{g} x_{1,i} - \sum_{m=2}^{M}\sum_{i=1}^{g} x_{m,i}}{\sum_{m=2}^{M}\sum_{i=1}^{g} x_{m,i}}. \tag{7}$$

B. Non-Data Aided MLE

The PDF of the received symbol, given a 1 at the nth position is expressed as

$$\begin{aligned}p_{\mathbf{x}_i}(\mathbf{x}|s_n = 1) = &\frac{1}{(S+N)N^{M-1}} \\ &\exp\left(-\frac{x_{n,i}}{S+N} - \sum_{m=1, m\neq n}^{M} \frac{x_{m,i}}{N}\right).\end{aligned} \tag{8}$$

Assuming equal prior probabilities of transmitted symbols, the unconditional joint PDF of the received symbols for M-FSK is given as

$$\begin{aligned}p_{\mathbf{x}_i}(\mathbf{x}) = &\frac{1}{M(S+N)N^{M-1}} \left[\underbrace{\exp\left(\frac{-x_{1,i}}{S+N} - \sum_{m=2}^{M}\frac{x_{m,i}}{N}\right)}_{a_1}\right. \\ &\left. + ... + \underbrace{\exp\left(\frac{-x_{M,i}}{S+N} - \sum_{m=1}^{M-1}\frac{x_{m,i}}{N}\right)}_{a_M}\right]\end{aligned} \tag{9}$$

It can be seen from (9) that the PDF contains sum of M exponential terms $a_1, a_2, ..., a_M$; taking the log-likelihood and

[1]It should be noted that in [15], we have assumed a coherent model for the SNR estimation for a BFSK receiver. However, in this paper, we use the non-coherent model for designing the SNR estimators for both fast Rayleigh fading channels and slow block fading channels.

then differentials with respect to S and N will make the expression very difficult to solve. The expression can be made simpler by factoring a common term, $\exp\left(-\frac{\sum_{m=1}^{M} x_{m,i}}{N}\right)$, from the exponents in (9) to get

$$p_{\mathbf{x}_i}(\mathbf{x}) = \frac{1}{M(S+N)N^{M-1}} \exp\left(-\frac{\sum_{m=1}^{M} x_{m,i}}{N}\right) \left[\sum_{m=1}^{M} \exp\left(-x_{m,i}\phi\right)\right], \tag{10}$$

where $\phi = \left[\frac{1}{S+N} - \frac{1}{N}\right]$. Thus the log-likelihood function for the k received symbols is given as

$$\Lambda(\mathbf{x}; S, N) = -k \ln M - k \ln (S+N) - \sum_{m=1}^{M}\sum_{i=1}^{k} \frac{x_{m,i}}{N} - k(M-1)\ln N + \sum_{i=1}^{k} \ln\left(\sum_{m=1}^{M} \exp\left(-x_{m,i}\phi\right)\right). \tag{11}$$

Taking the partial derivative of (11) with respect to S and N results in

$$\frac{\partial \Lambda}{\partial S} = \frac{-k}{(S+N)} + \frac{1}{(S+N)^2} \sum_{i=1}^{k} \frac{\sum_{m=1}^{M} x_{m,i} \exp\left(-x_{m,i}\phi\right)}{\sum_{m=1}^{M} \exp\left(-x_{m,i}\phi\right)}, \tag{12}$$

$$\frac{\partial \Lambda}{\partial N} = -\frac{k}{(S+N)} - \frac{(M-1)k}{N} + \sum_{m=1}^{M}\sum_{i=1}^{k} \frac{x_{m,i}}{N^2} + \left[\frac{1}{(S+N)^2} - \frac{1}{N^2}\right] \sum_{i=1}^{k} \frac{\sum_{m=1}^{M} x_{m,i} \exp\left(-x_{m,i}\phi\right)}{\sum_{m=1}^{M} \exp\left(-x_{m,i}\phi\right)}. \tag{13}$$

Putting the above equations equal to zero and solving them simultaneously gives us following result

$$\hat{S} + M\hat{N} = \frac{1}{k} \sum_{m=1}^{M}\sum_{i=1}^{k} x_{m,i}, \tag{14}$$

$$\sum_{m=1}^{M}\sum_{i=1}^{k} x_{m,i} - k(M-1)\hat{N} = \sum_{i=1}^{k} \frac{\sum_{m=1}^{M} x_{m,i} \exp\left(-x_{m,i}\phi\right)}{\sum_{m=1}^{M} \exp\left(-x_{m,i}\phi\right)}. \tag{15}$$

The above forms of the equations seem to prohibit the closed form solutions for estimates of S and N. Thus using high-SNR approximations (see Appendix A), the summation term is approximated as

$$\sum_{i=1}^{k} \frac{\sum_{m=1}^{M} x_{m,i} \exp\left(-x_{m,i}\phi\right)}{\sum_{m=1}^{M} \exp\left(-x_{m,i}\phi\right)} \approx \sum_{i=1}^{k} \max_{m=1,\ldots,M} \left(x_{m,i}\right). \tag{16}$$

Thus the estimate of noise power is given as

$$\hat{N} = \frac{1}{(M-1)k} \left[\sum_{m=1}^{M}\sum_{i=1}^{k} x_{m,i} - \sum_{i=1}^{k} \max_{m} \left(x_{m,i}\right)\right], \tag{17}$$

and the estimate of signal to noise ratio is given by

$$\hat{\gamma}_{NDA} = \frac{-\sum_{m=1}^{M}\sum_{i=1}^{k} x_{m,i} + M\sum_{i=1}^{k} \max_{m} \left(x_{m,i}\right)}{\sum_{m=1}^{M}\sum_{i=1}^{k} x_{m,i} - \sum_{i=1}^{k} \max_{m=1,\ldots,M} \left(x_{m,i}\right)}. \tag{18}$$

C. Joint Estimation Using Pilot and Data Symbols

Consider g pilot symbols and l data symbols, so that the total packet is of length $k = g + l$. Assuming independent received symbols, the joint PDF is the product of PDFs resulting from pilot and data symbols. So we use (4) for $i = 1, 2, ..., g$ and (10) for $i = g+1, ..., g+l = k$. Thus the log-likelihood function from the joint PDF is given as

$$\Lambda_{joint} = -k(M-1)\ln N - \frac{1}{S+N}\sum_{i=1}^{g} x_{1,i} - \frac{1}{N}\Psi - k\ln(S+N) + \sum_{i=g+1}^{k} \ln\left(\sum_{m=1}^{M} \exp\left(-x_{m,i}\phi\right)\right), \tag{19}$$

where $\Psi = \sum_{m=2}^{M}\sum_{i=1}^{g} x_{m,i} + \sum_{m=1}^{M}\sum_{i=g+1}^{k} x_{m,i}$. Using similar approximations as in the previous section and taking partial derivatives with respect to S and N and setting them equal to zero result in the estimate of SNR as

$$\hat{\gamma} = \frac{\Psi + M\sum_{i=g+1}^{k} \max_{m}\left(x_{m,i}\right) + (M-1)\sum_{i=1}^{g} x_{1,i}}{\Psi - \sum_{i=g+1}^{k} \max_{m}\left(x_{m,i}\right)}. \tag{20}$$

D. EDS Approach

In this subsection, we derive another type of NDA estimator that uses the statistics of the received signal for making an estimate of the SNR. We will show that this estimator performs better in the low SNR region and thus can be used as an alternative to ML-NDA estimator derived in the previous subsection, which suffers from approximation errors in the low SNR region. In the EDS approach, we are interested in using the statistics of the received data so that we are able to get the estimates of our parameters of interest, namely the signal and the noise power. Owing to the i.i.d. (in time) nature of received data, we will drop the time index i from the rest of the section. It is known that the optimal non-coherent FSK receiver requires a complex branch (i.e., I and Q channels) for each FSK frequency, because the phase is unknown [11]. Therefore, we can extend the scalar Equation (8) for BPSK in [14], to a vector formulation for MFSK, and define an $M \times M$ matrix $\mathbf{Z}$ given as

$$\mathbf{Z} = \left(\mathbb{E}\left\{\mathbf{x}_m\right\}\right)\left(\mathbb{E}\left\{\mathbf{x}_m\right\}\right)^T \left(\mathbb{E}\left\{\mathbf{x}_m \mathbf{x}_m^T\right\}\right)^{-1}, \tag{21}$$

where $\mathbb{E}\left\{\mathbf{x}_m\right\} = \left[\mathbb{E}\left\{x_1\right\} \quad \mathbb{E}\left\{x_2\right\} \quad \cdots \quad \mathbb{E}\left\{x_M\right\}\right]^T$. Assuming equally likely probable transmitted symbols, the ensemble average of the received data is given as

$$\mathbb{E}\left\{x_m\right\} = \frac{1}{M}\left[S + MN\right], \quad m = 1, ..., M. \tag{22}$$

Thus $\mathbb{E}\left\{\mathbf{x}_m\right\}\mathbb{E}\left\{\mathbf{x}_m\right\}^T$ is given as

$$\mathbb{E}\left\{\mathbf{x}_m\right\}\mathbb{E}\left\{\mathbf{x}_m\right\}^T = \frac{1}{M^2}\left[S + MN\right]^2 \mathbf{1}_M, \tag{23}$$

where $\mathbf{1}_M$ is an $M \times M$ matrix of all ones. The autocorrelation matrix of the received data, given by $\mathbb{E}\left\{\mathbf{x}_m \mathbf{x}_m^T\right\}$, contains $\mathbb{E}\left\{x_m^2\right\}$ on the main diagonal given as

$$\mathbb{E}\left\{x_m^2\right\} = \frac{2}{M}\left[S^2 + 2SN + MN^2\right], \quad m = 1, ..., M \tag{24}$$

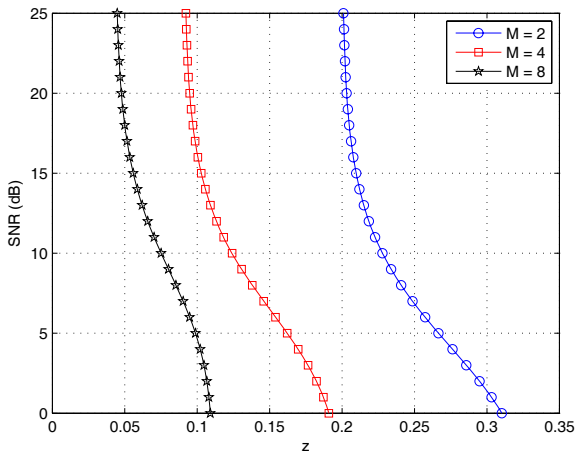

Fig. 1. Relationship between the computed statistics, z, and γ for different modulation orders, M, for the Rayleigh fading channel.

while the rest of all elements will be $\left(\mathbb{E}\left\{x_m\right\}\right)^2$, where $\mathbb{E}\left\{x_m\right\}$ is given by (22). Thus we can write

$$\mathbb{E}\left\{\mathbf{x}_m\mathbf{x}_m^T\right\} = \frac{\left[S+MN\right]^2}{M^2}\underbrace{\begin{bmatrix} a & 1 & \cdots & 1 \\ 1 & a & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & a \end{bmatrix}}_{\mathbf{H}} \tag{25}$$

where the element a is given as

$$a = \frac{\mathbb{E}\left\{x_m^2\right\}}{\mathbb{E}\left\{x_m\right\}^2} = 2M\frac{\gamma^2+2\gamma+M}{(\gamma+M)^2}, \tag{26}$$

and $\gamma = S/N$ is the signal-to-noise ratio. It can be noticed that the matrix $\mathbf{H}$ in (25) is a special kind of Hankel matrix which is also circulant. The inverse of such a circulant matrix, of order M, is given as

$$\mathbf{H}^{-1} = \frac{1}{\zeta}\begin{bmatrix} a+M-2 & -1 & \cdots & -1 \\ -1 & a+M-2 & \cdots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \cdots & a+M-2 \end{bmatrix}, \tag{27}$$

where $\zeta = a^2+(M-2)a-(M-1)$. Thus the matrix $\mathbf{Z}$ from (21) is given as

$$\mathbf{Z} = \mathbf{1}_M\mathbf{H}^{-1} = \frac{a-1}{a^2+(M-2)a-(M-1)}\mathbf{1}_M. \tag{28}$$

Since the resulting $\mathbf{Z}$ from above equation contains identical elements at each location of $M \times M$ matrix, we can now utilize any one of the element from $\mathbf{Z}$. Thus, using value of a from (26), any $z \in \mathbf{Z}$ is given by

$$z = \frac{(\gamma+M)^2}{M^3+M^2(2\gamma+1)+M\gamma(3\gamma+2)-\gamma^2}. \tag{29}$$

We may solve for γ to get an estimate as

$$\hat{\gamma} = \frac{1}{3Mz-z-1}\left[M(1-Mz-z) + \sqrt{(2(M-1)M^2z(1-Mz-z))}\right]. \tag{30}$$

To make this approach practical, we replace the expectations in (21) with the corresponding block averages to compute the estimate of $\mathbf{Z}$ as

$$\hat{\mathbf{Z}} = \left(\overline{\mathbf{x}}\right)\left(\overline{\mathbf{x}}\right)^T\left(\overline{\mathbf{x}\mathbf{x}^T}\right)^{-1}, \tag{31}$$

where $\overline{\mathbf{x}}$ is a column vector of time averages of M branches given as $\overline{\mathbf{x}} = \left[\bar{x}_1, \bar{x}_2, ..., \bar{x}_M\right]^T$, and each $\bar{x}_m$ is given as

$$\bar{x}_m = \frac{1}{k}\sum_{i=1}^{k} x_{m,i}, \quad m = \{1, 2, ..., M\}. \tag{32}$$

We have plotted the results from (30) for the computed statistic, z, against the SNR, γ, in Fig. 1 for various values of M. Each curve in the figure exhibits smooth monotone behavior for low values of SNR, i.e., when SNR $\in$ [0,10]. Thus, the estimator performance is anticipated to be good in this SNR region. However, as the SNR increases, all the curves appear to approach a vertical asymptote . This implies that a very small change in the computed statistics of data would cause a huge variation in the estimated value of SNR. Hence this approach will suffer in the estimation of high SNR values. Therefore, this approach can be used as an NDA approach for SNR estimation for low SNR regions.

IV. SNR Estimation for a Block Fading Channel

In this section, we will briefly describe the approach for estimating SNR in a block fading environment, which is the case when a block of data with k symbols undergoes a constant non-random fade. Using the complex data model, the received symbols are given as

$$\mathbf{v}_i = A\mathbf{s}_i + \mathbf{n}_i. \tag{33}$$

In the above equation A is the complex signal amplitude assumed constant over the entire frame and $\mathbf{n}_i$ is the noise vector whose elements are drawn from a zero mean complex Gaussian distribution with a variance of $N/2$ per real dimension. Thus the SNR for block fading environment is given as $\gamma = \frac{|A|^2}{N}$. As previously, in this section, we will assume that the SNR estimation is done after the square law. Thus, we will estimate the SNR using the observed data $\mathbf{x}_i = \left[\left(|\mathbf{v}_1|^2\right)^T \quad \cdots \quad \left(|\mathbf{v}_k|^2\right)^T\right]^T$.

A. Partially Data-Aided Estimation

Using the same notations and assumptions as done in Section III-A, the first branch output, containing a 1, is given as

$$x_{1,i} = |A+n_i|^2, \tag{34}$$

where A and n are as defined in the previous section. Since the noise is complex Gaussian, thus the resulting PDF of $x_{1,i}$ will be non-central chi square distribution, where the non centrality parameter λ is given as

$$\lambda = \left(\Re\left\{A\right\}\right)^2 + \left(\Im\left\{A\right\}\right)^2 = |A|^2, \tag{35}$$

where $\Re\left\{A\right\}$ and $\Im\left\{A\right\}$ denote the real and imaginary parts of the complex signal amplitude A, respectively. Thus, the PDF of $x_{1,i}$ is given as

$$p_{x_{1,i}}(x) = \frac{1}{N}\exp\left(-\frac{x+|A|^2}{N}\right)I_0\left(\frac{2\sqrt{x}|A|}{N}\right), \tag{36}$$

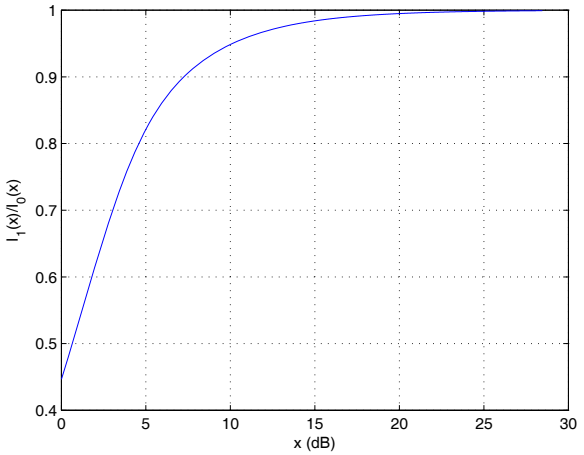

Fig. 2. Behavior of the ratios of modified Bessel functions of the first kind.

where $I_0(.)$ is the modified Bessel function of zero order first kind. The PDF for each of the rest of the branches $m = 2, ..., M$, is exponential, and is given as

$$p_{x_{m,i}}(x_m) = \frac{1}{N}\exp\left(-\frac{x_m}{N}\right). \quad m = 2, ..., M. \tag{37}$$

Thus the joint PDF of the received symbols is given as

$$p_{\mathbf{x}_i}(\mathbf{x}) = \frac{1}{N^M}\exp\left(-\frac{|A|^2}{N} - \frac{1}{N}\sum_{m=1}^{M} x_m\right) I_0\left(\frac{2\sqrt{x_1}|A|}{N}\right). \tag{38}$$

The log-likelihood for the g received symbols is given as

$$\begin{aligned}\Lambda_{\mathbf{x}_i}(\mathbf{x}; A, N) = &- gM\ln(N) + \sum_{i=1}^{g}\ln\left(I_0\left(\frac{2\sqrt{x_{1,i}}|A|}{N}\right)\right) \\ &- \frac{1}{N}\left(\sum_{i=1}^{g} x_{1,i} + g|A|^2 + \sum_{m=2}^{M}\sum_{i=1}^{g} x_{m,i}\right).\end{aligned} \tag{39}$$

Taking the derivative of (39) with respect to A and using the relations $\frac{d}{dx}I_n(x) = I_{n+1}(x)$, $\frac{\partial A}{\partial A} = 1$, $\frac{\partial A^*}{\partial A} = 0$, and $\frac{\partial A A^*}{\partial A} = A^*$, we get

$$\frac{\partial \Lambda_{\mathbf{x}_i}(\mathbf{x}; A, N)}{\partial A} = -\frac{gA^*}{N} + \sum_{i=1}^{g}\frac{I_1\left(\frac{2\sqrt{x_{1,i}}|A|}{N}\right)}{I_0\left(\frac{2\sqrt{x_{1,i}}|A|}{N}\right)}\frac{A^*\sqrt{x_{1,i}}}{N|A|}. \tag{40}$$

An exact solution to the above equation is difficult to obtain because of the non-linearity of modified Bessel function. However, it can be observed that for high values of the argument, the ratio of first order modified Bessel function to the zero order modified Bessel function, i.e. $\frac{I_1(.)}{I_0(.)}$, is approximately equal to 1. This phenomenon can be seen in the Fig. 2. Thus using this approximation, the estimate of $|A|$ is given as

$$|\hat{A}| = \frac{1}{g}\sum_{i=1}^{g}\sqrt{x_{1,i}}. \tag{41}$$

Differentiating (39) with respect to N and using similar approximations, the estimate of N is given as

$$\hat{N} = \frac{1}{Mg}\left[\sum_{i=1}^{g} x_{1,i} - g|\hat{A}|^2 + \sum_{m=2}^{M}\sum_{i=1}^{g} x_{m,i}\right]. \tag{42}$$

Using (41) and (42), we can find the estimate of SNR for the data-aided case as $\frac{|\hat{A}|^2}{\hat{N}}$.

B. Non-Data Aided Estimation

The PDF of the received symbol, given 1 at the nth position, is given as

$$\begin{aligned}p_{\mathbf{x}_i}(\mathbf{x}|s_n = 1) = &\frac{1}{N^M} I_0\left(\frac{2\sqrt{x_n}|A|}{N}\right) \\ &\exp\left(-\frac{1}{N}\left[x_n + |A|^2 + \sum_{m=1, m\neq n}^{M} x_m\right]\right).\end{aligned} \tag{43}$$

Assuming equal prior probabilities of transmitted symbols, the unconditional joint PDF of the received symbols for M-FSK block fading case is given as

$$\begin{aligned}p_{\mathbf{x}_i}(\mathbf{x}) = &\frac{1}{MN^M}\exp\left(-\frac{1}{N}\left[\sum_{m=1}^{M} x_m + |A|^2\right]\right) \\ &\left[\sum_{m=1}^{M} I_0\left(\frac{2\sqrt{x_m}|A|}{N}\right)\right].\end{aligned} \tag{44}$$

For k received symbols, the log-likelihood function is given as

$$\begin{aligned}\Lambda_{\mathbf{x}_i}(\mathbf{x}; A, N) = &- kM\ln(N) - \frac{1}{N}\left(\sum_{m=1}^{M}\sum_{i=1}^{k} x_{m,i} + k|A|^2\right) \\ &- k\ln(M) + \sum_{i=1}^{k}\ln\left[\sum_{m=1}^{M} I_0\left(\frac{2\sqrt{x_{m,i}}|A|}{N}\right)\right]\end{aligned} \tag{45}$$

The partial derivative with respect to A is given as

$$\frac{\partial \Lambda_{\mathbf{x}_i}(\mathbf{x})}{\partial A} = -\frac{KA^*}{N} + \sum_{i=1}^{k}\frac{\sum_{m=1}^{M} I_1\left(\frac{2\sqrt{x_{m,i}}|A|}{N}\right)}{\sum_{m=1}^{M} I_0\left(\frac{2\sqrt{x_{m,i}}|A|}{N}\right)}\frac{A^*\sqrt{x_{m,i}}}{N|A|}. \tag{46}$$

Using some approximations of the modified Bessel functions (see Appendix B), the estimate of $|A|$ is given as

$$|\hat{A}| = \frac{1}{k}\sum_{i=1}^{k}\max_{m=1,...,M}\sqrt{x_{m,i}}. \tag{47}$$

The noise power estimate is given in a similar manner as

$$\hat{N} = \frac{1}{Mk}\left[\sum_{m=1}^{M}\sum_{i=1}^{k} x_{m,i} - k|\hat{A}|^2\right]. \tag{48}$$

C. Joint Estimation Using Pilot and Data Symbols

Considering g pilot symbols and l data symbols the log-likelihood function from the joint PDF is given as

$$\Lambda = -\frac{1}{N}\left(k|A|^2 + \sum_{m=1}^{M}\sum_{i=1}^{k} x_{m,i}\right) - kM\ln N - l\ln M + \sum_{i=g+1}^{k}\ln\sum_{m=1}^{M} I_0\left(\frac{2\sqrt{x_{m,i}}|A|}{N}\right) + \sum_{i=1}^{g}\ln I_0\left(\frac{2\sqrt{x_{1,i}}|A|}{N}\right) \tag{49}$$

Using similar approximations as in the previous section and taking partial derivatives with respect to A and N and setting them equal to zero result in the estimates of signal and noise powers as

$$|\hat{A}| = \frac{1}{k}\left[\sum_{i=1}^{g}\sqrt{x_{1,i}} + \sum_{i=g+1}^{k}\max_{m=1,\ldots,M}\sqrt{x_{m,i}}\right], \tag{50}$$

$$\hat{N} = \frac{1}{kM}\left[\sum_{m=1}^{M}\sum_{i=1}^{g} x_{m,i} + \sum_{m=1}^{M}\sum_{i=g+1}^{k} x_{m,i} - k|\hat{A}|^2\right]. \tag{51}$$

D. EDS Approach

In order to get an estimate of SNR using the statistics of the received data, we define an $M \times M$ matrix $\mathbf{Z}$, similar to (21). Assuming equally likely probable transmitted symbols, the ensemble average of the received data is given as

$$\mathbb{E}\{x_m\} = \frac{1}{M}\left[|A|^2 + MN\right], \quad m = 1, ..., M \tag{52}$$

Thus $\mathbb{E}\{\mathbf{x}_m\}\mathbb{E}\{\mathbf{x}_m\}^T$ is given as

$$\mathbb{E}\{\mathbf{x}_m\}\mathbb{E}\{\mathbf{x}_m\}^T = \frac{1}{M^2}\left[|A|^2 + MN\right]^2\mathbf{1}_M, \tag{53}$$

where $\mathbf{1}_M$ is an $M \times M$ matrix of all ones. The autocorrelation matrix of the received data, given by $\mathbb{E}\{\mathbf{x}_m\mathbf{x}_m^T\}$, contains $\mathbb{E}\{x_m^2\}$ on the main diagonal given as

$$\mathbb{E}\{x_m^2\} = \frac{1}{M}\left[|A|^4 + 4N|A|^2 + 2MN^2\right], \quad m = 1, ..., M \tag{54}$$

while the rest of all elements will be $(\mathbb{E}\{x_m\})^2$, where $\mathbb{E}\{x_m\}$ is given by (52). Thus we can construct the matrix $\mathbf{H}$ as in Section III-D, where the element a is given as

$$a = \frac{\mathbb{E}\{x_m^2\}}{\mathbb{E}\{x_m\}^2} = M\frac{\gamma^2 + 4\gamma + 2M}{\gamma^2 + M^2 + 2M\gamma}, \tag{55}$$

and $\gamma = |A|^2/N$ is the signal-to-noise ratio. Thus utilizing only one of the element from $\mathbf{Z}$, we get

$$z = \frac{\gamma^2 + 2M\gamma + M^2}{M^3 + 2M^2\gamma + M^2 + 2M\gamma^2 + 2M\gamma - \gamma^2}. \tag{56}$$

We may solve for γ to get an estimate as

$$\hat{\gamma} = \frac{1}{2Mz - z - 1}\left[M - Mz - Mz^2 + \sqrt{(}-M^4z^2 + M^3z^2 + M^3z + 2M^2z^2 - 2M^2z)\right]. \tag{57}$$

It can be seen from (56) that z has no solution at $z = 1/3$ for $M = 2$. Thus this method is not applicable for a BFSK system. This scheme follows the same behavior of curves as in Fig. 1, except for $M = 2$ case which would be a vertical line. Otherwise we observe that in the higher SNR regime, the EDS approach will suffer due to the steepness of the curves. But it will be shown in the simulation section, that the EDS approach shows best performance for larger data set and in the low SNR region, compared to the MLE algorithms discussed in the previous subsection, because there are no approximation errors in this approach.

V. Cramer-Rao Lower Bound for Rayleigh Fading Channel

In this section, we will derive the CRB for the SNR estimator, for the Rayleigh fading channel. It can be noticed, that we can have different CRBs for the different cases discussed in the previous section. However, the FDA estimator serves as a benchmark on the variance of all estimators; FDA is same as PDA but uses all information in the packet as training sequence. We also show that the CRB for the NDA estimator is identical to the CRB of the FDA estimator if we use the high SNR approximation. Since the unknown parameter is a vector i.e., $\theta = [S \quad N]^T$, thus the CRB for the SNR is given as [10]

$$CRB = \frac{\partial\mathbf{g}(\theta)}{\partial\theta}\mathbf{I}^{-1}(\theta)\frac{\partial\mathbf{g}(\theta)^T}{\partial\theta}, \tag{58}$$

where $\mathbf{g}(\theta) = \frac{S}{N}$ and the Jacobian of $\mathbf{g}(\theta)$ is given as

$$\mathbf{J}_{\mathbf{g}}(\theta) = \left[\frac{1}{N} \quad \frac{-S}{N^2}\right], \tag{59}$$

and $\mathbf{I}(\theta)$ is the Fisher information matrix (FIM) given as

$$[\mathbf{I}(\theta)]_{ij} = -\mathbb{E}\left[\frac{\partial^2\Lambda}{\partial\theta_i\theta_j}\right]. \tag{60}$$

The FIM for the FDA estimator is given as

$$\mathbf{I}(\theta) = \begin{bmatrix} \frac{k}{(S+N)^2} & \frac{k}{(S+N)^2} \\ \frac{k}{(S+N)^2} & \frac{k}{(S+N)^2} + \frac{k(M-1)}{N^2} \end{bmatrix}, \tag{61}$$

which gives the CRB from (58) as

$$CRB_{FDA} = \frac{M}{k(M-1)}(1+\gamma)^2. \tag{62}$$

This bound has been plotted in Figures 4 and 5, and is further discussed in the next section. The CRB of NDA estimator is given in Appendix C.

VI. Simulation Results

In this section, we compare the normalized mean squared error (NMSE) (normalized with respect to the square of the true value of the SNR) of the estimators using simulations for different values of M and for different packet lengths averaged over 10,000 trials. The NMSE is given as

$$NMSE(\hat{\gamma}) = \frac{\mathbb{E}\left\{(\gamma - \hat{\gamma})^2\right\}}{\gamma^2}, \tag{63}$$

where γ is the true value of SNR and $\hat{\gamma}$ is the estimated value. The results shown in Figures 3 - 8 are for the Rayleigh fading

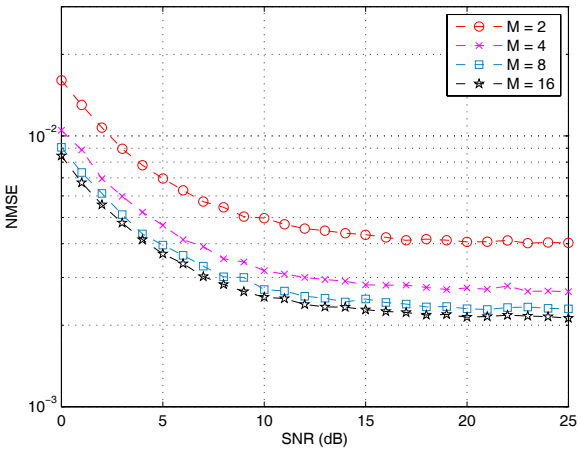

Fig. 3. Effect of increasing M on NMSE for 1000 symbol-long packet for the PDA estimator for the Rayleigh fading channel.

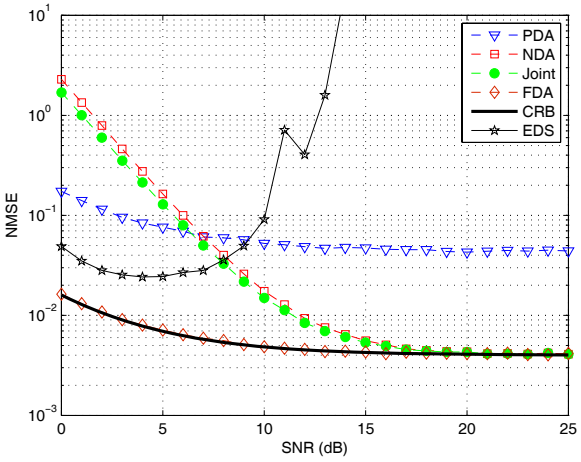

Fig. 4. NMSE for different estimators for a Binary FSK receiver, (M=2), for the Rayleigh fading channel with 1000 symbols including 100 pilot symbols (g=100).

channel. Fig. 3 shows the NMSE vs. true SNR for the PDA estimator, with 500 pilot symbols in the packet for increasing values of M. We observe that the estimates become more and more accurate as we have more and more branches with noise only. Thus increasing M indirectly increases the number of samples, which gives lower NMSE. Although not shown in the figures, this behavior is found in all techniques discussed in Sections III and IV.

Figures 4 and 5 treat the long packet, which is assumed to comprise 100 pilot symbols and 900 data symbols. We observe that the PDA NMSE is approximately constant over the entire SNR range. The NDA and *Joint* cases perform similarly because most of the packet is data and the NMSE is high in the low SNR region. The large error in low SNR range of NDA and joint estimators can be attributed to the approximation errors in the low SNR regime. To check the high SNR approximation (from Appendix A), we have plotted the NMSE between the true SNR (computed numerically using Equations (14) and (15)) and the estimated SNR using the approximation in Equation (16). Fig. 6 shows this NMSE for different values of M for a packet length of 100. It can be seen

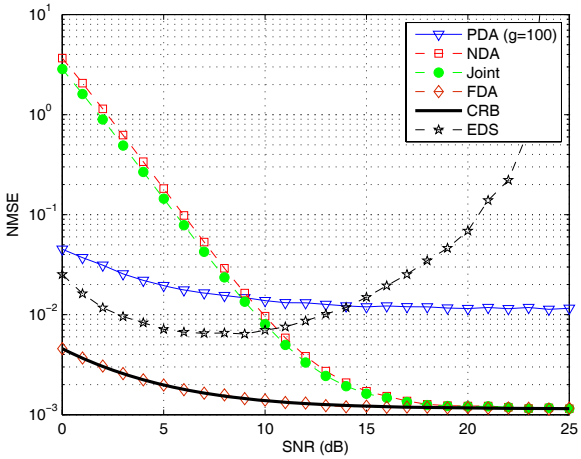

Fig. 5. NMSE for different estimators for 8FSK receiver, (M=8), for the Rayleigh fading channel with 1000 symbols including 100 pilot symbols (g=100).

that the error decreases with the increase in SNR, which shows that the approximation in Appendix A is valid. Also note that the same approximation is also valid for the *Joint* estimator, which also uses this approximation. However, the results are similar to Fig. 6 and are not shown to avoid repetition.

In Figures 4 and 5, we can observe the consequence of this approximation. We notice a small NMSE, for the NDA estimator, as the SNR increases from 8dB onwards. The EDS approach performs better for low SNR estimation as compared to the NDA scheme. However, the EDS method shows bad behavior at high SNR due to the steepness of curve from Fig. 1, and therefore, we can see a rapid increase in the NMSE after 10dB. To do a fair comparison, we assume that both the pilot and the data symbols are available to the NDA and EDS approaches for the estimation. For high SNR, the *Joint* estimation scheme works the best, as expected. The crossing of the curves suggests that an adaptive mode of SNR estimation can also be derived consisting of estimation from the pilot only (PDA) or EDS during the low SNR while using the entire data packet for estimating high SNR values. In that case, the overall NMSE will remain minimum over a wide range of SNR values. In Fig. 5, which shows the results for $M = 8$, the curves have the same behavior as in the $M = 2$ case, but the NMSE for $M = 8$ is lower.

Fig. 7 treats the short packet scenario, with 8 pilot symbols and 28 data symbols. We show this short packet case because the EDS approach does not perform well because of the limitations of the availability of data (approximation error of the ensemble averages with time averages for small data set is large). Thus for a short length packet and with the availability of pilot, the joint data estimation performs best. If the pilot is not available, then the NDA MLE also gives better performance.

Figures 4 and 5 include the results for the FDA approach (decision feedback), which utilizes the detected data, assuming no detection errors. It can be seen that the performance of the FDA estimator is enhanced significantly and it reaches the CRB. Using this approach, we gain two advantages: a larger data set and estimation using DA approach which has

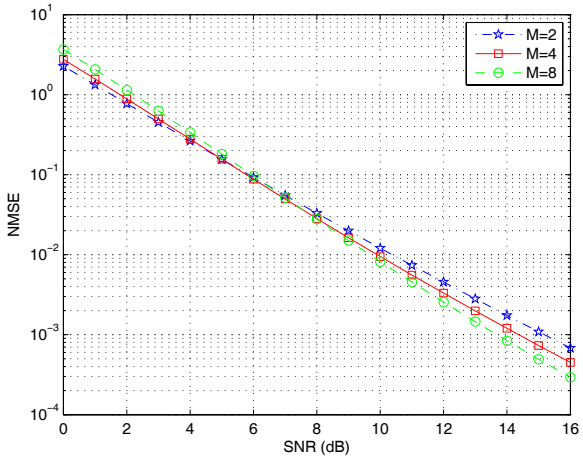

Fig. 6. NMSE between actual and approximated SNR values for NDA estimator in Rayleigh fading for a packet length of 100.

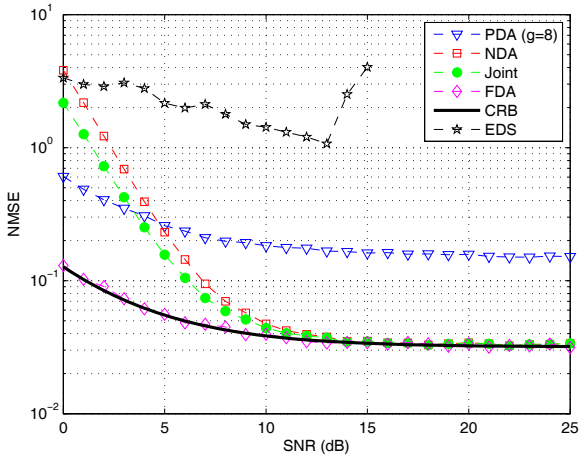

Fig. 7. NMSE for different estimators for 8FSK receiver, for a Rayleigh fading channel with 36 symbols including 8 pilot symbols (g=8).

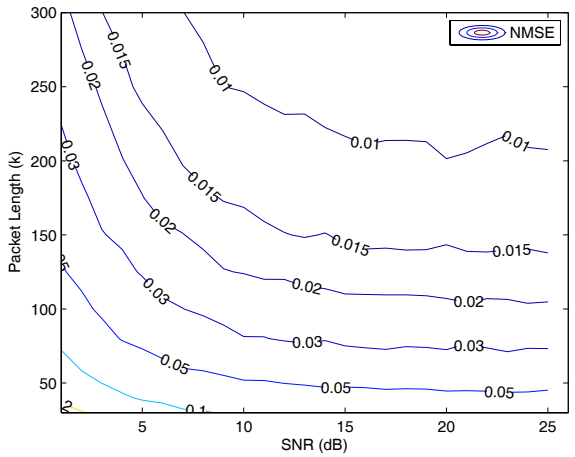

Fig. 8. NMSE contours for various packet lengths for the FDA estimator for the Rayleigh fading channel.

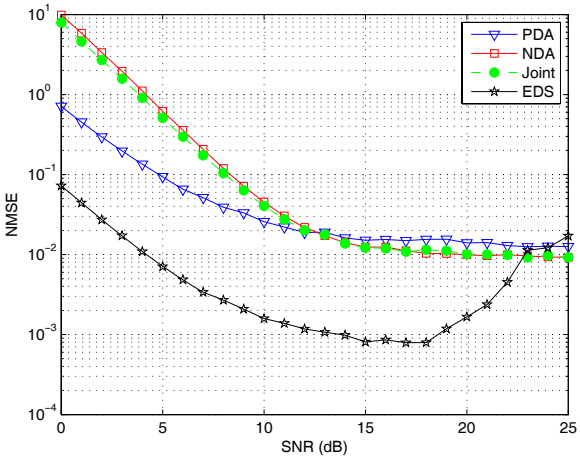

Fig. 9. NMSE for different estimators for a block fading channel in 8FSK receiver, M=8, with 1000 symbols including 100 pilot symbols (g=100).

no approximation errors. The *no errors* assumption is good in the context of the decode and forward (DF) scenario, since passing the CRC check is a precondition for forwarding the packet [1]. Sometimes, it is desired to choose the packet length such that the NMSE should not exceed some specified value. A contour plot of NMSE for SNR versus the packet length for FDA case is shown in Fig. 8 for $M = 8$. We observe that 5% error can be achieved with 50 symbols.

Fig. 9 shows the estimation results for the block fading channel. It can be seen that the behavior of the curves is identical to those of Figures 4-5 but interestingly the EDS method outperforms the ML estimation with a considerable margin for low values of SNR because there are no approximation errors in the EDS scheme for a block fading channel. However, it shows bad behavior at high SNR due to the steepness of the curves from Fig. 1. The CRB for this case is not derived due to tedious calculations involving modified Bessel function and hence is not plotted here.

In Sections III and IV, we have shown that the mathematical analysis for both the channel models are quite different from each other. However, an interesting scenario is to apply the SNR estimators designed for Rayleigh fading channels to the block fading case, and vice versa. We observe that the block fading estimator is better for the block fading channel (in terms of SNR gain) for all estimators except PDA at low SNR. The NMSE is higher at low-SNR region for the PDA approach, which can be attributed to the approximations of the modified Bessel function that we made in Section IV-A. We have simulated this effect for a short packet scenario for 8-FSK scheme such that the packet contains 8 pilot symbols and 28 data symbols. From Fig. 10, it can be noticed that if we apply the reverse scenario, i.e., apply the block fading estimators on Rayleigh fading channel, the estimators fail to estimate the SNR. Intuitively, the block fading channel is a general case of Rayleigh fading channel so the cross-estimator works, but with lower efficiency. On the other hand, Rayleigh fading is not a general case of block fading channel, thus the cross-estimator fails to estimate the SNR of the received data. Thus both estimators give appropriate results for their respective channel models. From Fig. 10, we observe that for a very short range of SNR, the cross-estimators work better. This could be due to the sub-optimality of PDA/NDA ML estimator cost functions in the mean-squared sense. For the

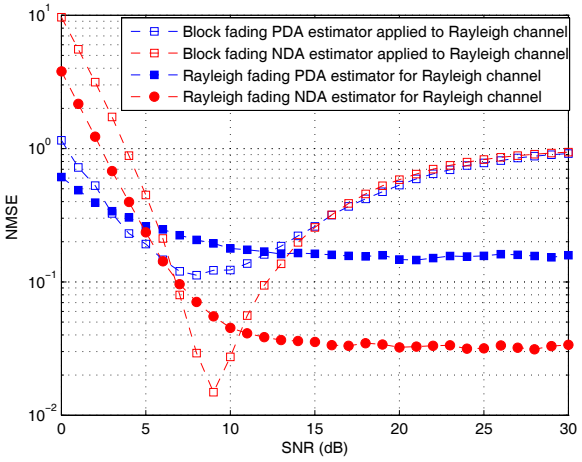

Fig. 10. Effects of applying the estimators for a block fading channel on the data received through Rayleigh fading channel.

rest of the SNR range, the actual estimators perform better.

VII. Conclusion

This paper considered the problem of estimating the average SNR for a non-coherent MFSK receiver, for $M = 2^n$, n being a positive integer, taking into account both symbol-by-symbol Rayleigh fading and slow block fading channels. The corresponding CRB and the ML estimators have been derived in addition to the EDS approach, which uses the data statistics to estimate the SNR. We also assumed different degrees of data knowledge in a packet and provided different versions of data-aided and non data-aided estimators. From the theoretical perspective and simulation results, it is thus concluded that different scenarios lead to different results based on packet length, availability of pilot sequence, and the region of SNR considered (low/high). The FDA approach performs the best in all cases. However, if FDA cannot be used because the correct detected symbols are not available, then an adaptive scheme is also suggested. This scheme can be applied in cases where error detection is not being used or if a software defined radio does decode and forward on the preamble and header, but does only amplify and forward on the payload.

Acknowledgment

The authors gratefully acknowledge the reviews made by the reviewers that help in improving the quality of this paper.

Appendix A

For simplicity, lets discuss the case where $M = 2$. In that case, the summation term (denoted by A) is given by

$$\begin{aligned} A &= \sum_{i=1}^{k} \frac{\sum_{m=1}^{2} x_{m,i} \exp\left(-x_{m,i}\phi\right)}{\sum_{m=1}^{M} \exp\left(-x_{m,i}\phi\right)} \\ &= \sum_{i=1}^{k} \frac{x_{1,i} \exp\left(-x_{1,i}\phi\right) + x_{2,i} \exp\left(-x_{2,i}\phi\right)}{\exp\left(-x_{1,i}\phi\right) + \exp\left(-x_{1,i}\phi\right)}. \end{aligned} \tag{64}$$

From ϕ, it can be noticed that for high SNR where $S >> N$, ϕ can be approximated as $\phi = -\frac{1}{N}$, thus the A can be approximated as

$$A = \sum_{i=1}^{k} \frac{x_{1,i}}{1 + \frac{\exp(x_{2,i}/N)}{\exp(x_{1,i}/N)}} + \frac{x_{2,i}}{1 + \frac{\exp(x_{1,i}/N)}{\exp(x_{2,i}/N)}}. \tag{65}$$

Now one of the two branches, say x_1, contains the signal. For high SNR, since the noise power $N << S$, the denominator of first term approaches 1 because the ratio of exponentials results in a very small number. While the same phenomenon is reversed in the second term, where the denominator approaches a very large value since the ratio of exponentials results in a very large number. Thus overall the second term is negligible (since the numerator is also a small number), and we are just left with $x_{1,i}$. Thus A is approximated as

$$A \approx \sum_{i=1}^{k} \max_{m} x_{m,i}. \tag{66}$$

Appendix B

For simplicity, lets discuss the case where $M = 2$. In that case, the summation term (denoted by B) is given by

$$\begin{aligned} B &= \sum_{i=1}^{k} \frac{\sum_{m=1}^{2} I_1\left(\frac{2\sqrt{x_{m,i}}|A|}{N}\right)}{\sum_{m=1}^{2} I_0\left(\frac{2\sqrt{x_{m,i}}|A|}{N}\right)} \sqrt{x_{1,i}} \\ &= \sum_{i=1}^{k} \frac{I_1\left(\frac{2\sqrt{x_{1,i}}|A|}{N}\right)\sqrt{x_{1,i}} + I_1\left(\frac{2\sqrt{x_{2,i}}|A|}{N}\right)\sqrt{x_{2,i}}}{I_0\left(\frac{2\sqrt{x_{1,i}}|A|}{N}\right) + I_0\left(\frac{2\sqrt{x_{2,i}}|A|}{N}\right)}. \end{aligned} \tag{67}$$

Since $\frac{2|A|}{N}$ is constant throughout, thus denoting it as ψ and separating terms in the above equation, B can be written as

$$\sum_{i=1}^{k} \left[\frac{\sqrt{x_{1,i}}}{\frac{I_0\left(\psi\sqrt{x_{1,i}}\right)}{I_1\left(\psi\sqrt{x_{1,i}}\right)} + \frac{I_0\left(\psi\sqrt{x_{2,i}}\right)}{I_1\left(\psi\sqrt{x_{1,i}}\right)}} + \frac{\sqrt{x_{2,i}}}{\frac{I_0\left(\psi\sqrt{x_{1,i}}\right)}{I_1\left(\psi\sqrt{x_{2,i}}\right)} + \frac{I_0\left(\psi\sqrt{x_{2,i}}\right)}{I_1\left(\psi\sqrt{x_{2,i}}\right)}} \right]. \tag{68}$$

Using the approximation $\frac{I_0(x)}{I_1(x)} \approx 1$, one term in each denominator is always 1. For the other term, it can be observed that at high argument values, the modified Bessel function approaches a very high value. Thus if $\sqrt{x_{1,i}} > \sqrt{x_{2,i}}$, then $I_n(\sqrt{x_{1,i}}) >> I_n(\sqrt{x_{2,i}})$, where n is the order of modified Bessel function, which implies that $\frac{I_0(\sqrt{x_{2,i}})}{I_1(\sqrt{x_{1,i}})} \approx 0$. Thus the denominator of first term approaches a 1, while the same phenomenon is reversed for the other term where the denominator approaches extremely large value. Thus overall we are left with the maximum term, i.e. $\sqrt{x_{1,i}}$. Thus B can be approximated for any M as

$$B \approx \sum_{i=1}^{k} \max_{m=1,\ldots,M} \sqrt{x_{m,i}}. \tag{69}$$

Appendix C

For finding the CRB of the NDA estimator, (60) needs be to be evaluated. However, the second derivatives of Equations (12) and (13), involved in this computation, are difficult to solve analytically. Even if we get a closed form solution, taking the expectation of those derivatives will be harder because

of the correlation between the numerator and denominator in both equations. This problem has been considered in [12] and [13] for linear modulation schemes where the authors have used numerical techniques to find these expectations. However, an interesting case arises, if we use the approximation as done in Equation (16). The structure of the FIM remains the same as in (61), given as

$$I = \begin{bmatrix} a & a \\ a & b \end{bmatrix}, \tag{70}$$

where

$$a = \frac{k}{(S+N)^2}, \tag{71}$$

and

$$b = \frac{k}{(S+N)^2} + \frac{2k(S+NM)}{N^3} - \frac{(M-1)k}{N^2} - \frac{2k(S+N)}{N^3}. \tag{72}$$

Thus, the CRB from (58) is given as

$$CRB = \frac{1}{b-a}\left[\frac{S^2}{N^4} + 2\frac{S}{N^3} + \frac{1}{N^2}\frac{b}{a}\right]. \tag{73}$$

By using expressions of a and b from (71) and (72), we get

$$CRB_{NDA} = \frac{M(\gamma+1)^2}{k(M-1)}. \tag{74}$$

Thus by using this high SNR approximation, the CRB for the NDA case matches the CRB for the FDA case. This can also be seen from Figures 4-5, where the NDA approaches the CRB (derived for FDA) in the high SNR region.

REFERENCES

[1] L. Thanayankizil, A. Kailas, and M. A. Ingram, "Routing for wireless sensor networks with an opportunistic large array (OLA) physical layer," *Ad Hoc Sensor Wireless Netw.*, vol. 8, no. 1-2, pp. 79–117, 2009.

[2] S. A. Hassan and M. A. Ingram, "A quasi-stationary Markov chain model of a cooperative multi-hop linear network," *IEEE Trans. Wireless Commun.*, vol. 10, no. 7, pp. 2306–2315, July 2011.

[3] A. Ramesh, A. Chockalingam, and L. B. Milstein, "Further results on selection combining of binary NCFSK signals in Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 52, no. 6, pp. 939–952, June 2004.

[4] D. R. Pauluzzi and N. C. Beaulieu, "A comparison of SNR estimation techniques for the AWGN channel," *IEEE Trans. Commun.*, vol. 48, no. 10, pp. 1681–1691, Oct. 2000.

[5] A. Ramesh, A. Chockalingam, and L. B. Milstein, "SNR estimation in generalized fading channels and its applications to turbo decoding," in *Proc. IEEE ICC*, June 2001.

[6] Y. Chen and N. C. Beaulieu, "An approximate maximum likelihood estimator for SNR jointly using pilot and data symbols," *IEEE Commun. Lett.*, vol. 9, no. 6, June 2005.

[7] T. Ertas and E. Dilaveroglu, "Low SNR asymptote of CRB on SNR estimates for BPSK in Nakagami-m fading channels with diversity combining," *IEEE Electron. Lett.*, vol. 39, no. 23, pp. 1680–1682, Nov. 2003.

[8] E. Dilaveroglu and T. Ertas, "CRBs and MLEs for SNR estimation on non coherent BFSK signals in Rayleigh fading," *IEEE Electron. Lett.*, vol. 41, no. 2, pp. 79–80, Jan. 2005.

[9] R. M. Gagliardi and C. M. Thomas, "PCM data reliability monitoring through estimation of signal-to-noise ratio," *IEEE Trans. Commun.*, vol. COM-16, pp. 479–468, June 1968.

[10] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1993.

[11] J. G. Proakis, *Digital Communications*, 4th edition, McGraw-Hill Publishers.

[12] A. Das, "NDA SNR estimation: CRLBs and EM based estimators," in *TENCON 2008, IEEE Region 10 Conf.*

[13] W. Gappamair, "Cramer-Rao lower bound for non-data-aided SNR estimation of linear modulation schemes," *IEEE Trans. Commun.*, vol. 56, no. 5, pp. 689–693, May 2008.

[14] S. Im and E. J. Powers, "An algorithm for estimating signal-to-noise ratio of UWB signals," *IEEE Trans. Veh. Technol.*, vol. 54, pp. 1905–1908, Sep. 2005.

[15] S. A Hassan and M. A. Ingram, "SNR estimation for a non-coherent binary frequency shift keying system," in *Proc. IEEE Global Commun. Conf.*, pp. 5448–5452, Dec. 2009.

Syed Ali Hassan is currently pursuing his Ph.D. degree in Electrical Engineering from Georgia Institute of Technology, Atlanta USA. He has been working in Smart Antennas Research Lab (SARL) as a graduate research assistant. He received his MS in Mathematics from Georgia Tech in 2011 and MS in Electrical Engineering from University of Stuttgart, Germany, in 2007. He was awarded BE in Electrical Engineering from National University of Sciences and Technology (NUST), Pakistan, in 2004. His broader area of research is signal processing for communications with a focus on stochastic modeling and cooperative communications for wireless networks. He also held industry position, as a design engineer, in Center for Advanced Research in Engineering, Islamabad, Pakistan.

Mary Ann Ingram received the B.E.E. and Ph.D. degrees from the Georgia Institute of Technology (Georgia Tech) in 1983 and 1989, respectively. From 1983 to 1986, she was a Research Engineer with the Georgia Tech Research Institute in Atlanta, performing studies on radar electronic countermeasure (ECM) systems. In 1986, she became a graduate research assistant with the School of Electrical and Computer Engineering at Georgia Tech, where in 1989, she became a Faculty Member and is currently Professor. She was a Visiting Professor in the Summers of 2006-8 at Aalborg University in Aalborg, Denmark, and at the Idaho National Laboratory in Summer 2009.

Dr. Ingram's early research areas were optical communications and radar systems. In 1997, she established the Smart Antenna Research Laboratory (SARL), which emphasizes the application of multiple antennas to wireless communication systems; this includes cooperative transmission, distributed multiple-input-multiple-output (MIMO) architectures, smart antennas, and adaptive arrays. The SARL performs system analysis and design, channel measurement, and prototyping relating to a wide range of wireless applications, including wireless local area network (WLAN) and satellite communications, with emphasis on Layers 1, 2 and 3 of multi-hop wireless communication networks.