SNR Estimation for a Non-Coherent M-FSK Receiver in a Slow Flat Fading Environment

Syed Ali Hassan and Mary Ann Ingram
School of Electrical and Computer Engineering
Georgia Institute of Technology, Atlanta, Georgia 30332–0250
Email: alihassan@gatech.edu; mai@ece.gatech.edu

***Abstract*—Estimation of the signal-to-noise ratio (SNR) is considered for a non-coherent M-ARY frequency shift keying (NC-MFSK) receiver. It has been assumed that the transmitted symbols undergo a slow flat fading channel where a block of data is corrupted by an independent constant fade and additive white Gaussian noise. Two approaches for SNR estimation are reported in this paper: an approximate maximum likelihood approach and another using the data statistics, both for data aided and non-data aided systems. It has been shown that for a particular SNR region of interest and depending upon the availability of pilot symbols, a particular approach is suitable for SNR estimation.**

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]. Furthermore, if the radios are energy constrained, e.g. if they are in a sensor network, constant envelope modulations and non-coherent demodulation may be desired to reduce circuit consumption of energy. Therefore, in this paper, we consider the estimation of SNR in an FSK non-coherent demodulator, over block deterministic channels.

In many wireless indoor applications and fixed wireless networks, the channel frequency response does not change rapidly. Thus a block of data undergoes a constant non-random fade. Estimation of SNR in such a case is of prime interest for various receiver functions. Assumption of a slow fading channel can be converted to a fast fading channel by assuming sufficient channel interleaving or by frequency hopping. But these techniques may not be suitable for some applications, e.g. wireless sensor networks, where the sensor nodes should be as simple as possible; devising such algorithms in these applications tends to increase the transmitter complexity. Thus a practical way of estimating SNR in slow fading environments is desirable.

Several authors have attacked the problem of estimating SNR for binary phase shift keying (BPSK) and frequency shift keying (FSK). For example, [2] 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. in [3,4]. In [5], the authors have estimated the average SNR for non-coherent binary FSK (NCBFSK) receiver, assuming fast fading channel and unit noise power spectral density. However, in implementations, noise power must also be estimated. In this paper, we derive two types of estimators of SNR, a sub-optimal/approximate maximum likelihood estimator (MLE) and an estimator that uses block statistics, such that neither of them assume any prior knowledge of the noise power. We estimate SNR not only for BFSK as in [5] but for any value of M in an M-FSK receiver. We provide ML versions of partially data-aided (PDA), non-data aided (NDA), and joint PDA-NDA estimators for the 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 estimators derived in this paper are significantly different and lead to dramatically different results from those for fast fading channels, which are addressed in [7].

The rest of the paper is organized as follows. In the next section, we describe the system model and the notations used for the NC-MFSK case. Section III treats the derivations of the SNR estimators, including three sub-cases for MLE and also the estimator using data statistics. In Section IV, we will discuss the simulation results for various estimators and overall estimator performance in terms of mean-squared error. The paper then concludes in Section V.

II. System Model

Consider a slow (block) fading communication system employing M-ARY FSK modulation, where a block of data with k symbols undergoes a constant non-random fade and the number of symbols in the constellation is $M = 2^n$, for a positive integer n. The received symbols are given as

$$\mathbf{v}_i = A\mathbf{s}_i + \mathbf{n}_i, \tag{1}$$

where i is the time index. 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

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

Gaussian distribution with a variance of $N/2$ per real dimension. Each of $\mathbf{v}_i$, $\mathbf{s}_i$, and $\mathbf{n}_i$ are vectors with a dimension of $M \times 1$, and for $\mathbf{s}_i = [0, ..., 0, 1, 0, ..., 0]^T$, 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 $|A|^2$. Thus the signal-to-noise ratio is given by $\gamma = \frac{|A|^2}{N}$. Our interest is to find the estimate of the average SNR using the observed data, after the square law, given as $\mathbf{x}_i = \left[\left(|\mathbf{v}_1|^2\right)^T \quad \cdots \quad \left(|\mathbf{v}_k|^2\right)^T \right]^T$. 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

As mentioned previously, we will derive the ML estimators for three cases, namely PDA, NDA and joint PDA-NDA.

A. Data-Aided Estimation

Without the loss of generality, the g pilot symbols are each set to $[1 \quad 0 \quad ... \quad 0]^T$. 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 $i = 1, 2, ..., g$. Thus the received symbol on the first branch is given as

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

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

$$\lambda = (\Re\{A\})^2 + (\Im\{A\})^2 = |A|^2. \tag{3}$$

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

$$p_{x_{1,i}}(x) = \frac{1}{N} \exp\left(-\frac{x + |A|^2}{N}\right) I_0\left(\frac{2\sqrt{x}|A|}{N}\right), \tag{4}$$

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

$$p_{x_{m,i}}(x_m) = \frac{1}{N} \exp\left(-\frac{x_m}{N}\right), \qquad m = 2, ..., M. \tag{5}$$

Thus the joint PDF of the received symbols is given as

$$p_{\mathbf{x}_i}(\mathbf{x}) = \frac{1}{N^M} \exp\left(-\frac{x_1 + |A|^2}{N} - \sum_{m=2}^{k} \frac{x_m}{N}\right) I_0\left(\frac{2\sqrt{x_1}|A|}{N}\right) \tag{6}$$

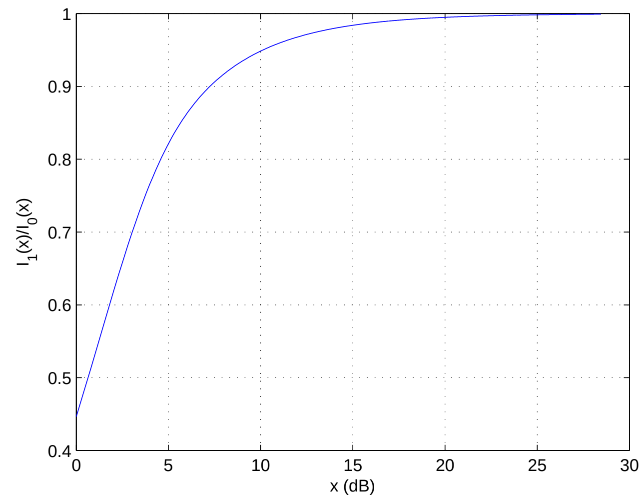

Fig. 1. Behavior of the ratios of Modified Bessel functions

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

$$\Lambda_{\mathbf{x}_i}(\mathbf{x}; A, N) = -gM\ln(N) + \sum_{i=1}^{g} \ln\left(I_0\left(\frac{2\sqrt{x_{1,i}}|A|}{N}\right)\right) - \frac{1}{N}\left(\sum_{i=1}^{g} x_{1,i} + g|A|^2 + \sum_{m=2}^{M}\sum_{i=1}^{g} x_{m,i}\right). \tag{7}$$

To find the MLE of the SNR, $\hat{\gamma}$, we use the property that the ML estimate of the ratio of two parameters ($|A|^2$ and N here), is the ratio of the individual ML estimates of the two parameters [6]. Taking the derivative of (7) 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 $\frac{\partial AA^*}{\partial A} = A^*$ [8], we get

$$\frac{\partial \Lambda_{\mathbf{x}_i}(\mathbf{x}; A, N)}{\partial A} = -\frac{gA^*}{N} + \sum_{i=1}^{g} \frac{I_1\left(\frac{2\sqrt{x_{1,i}}|A|}{N}\right)}{I_0\left(\frac{2\sqrt{x_{1,i}}|A|}{N}\right)} \frac{A^*\sqrt{x_{1,i}}}{N|A|}. \tag{8}$$

An exact solution to the above equation is difficult to find because of the non-linearity of modified Bessel function. However, we observe 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 Figure 1. Thus using this approximation, the estimate of $|A|$ is given as

$$|\hat{A}| = \frac{1}{g} \sum_{i=1}^{g} \sqrt{x_{1,i}}. \tag{9}$$

Differentiating (7) with respect to N and using similar approximations, the estimate of N is given as

$$\hat{N} = \frac{1}{Mg}\left[\sum_{i=1}^{g} x_{1,i} - g|\hat{A}|^2 + \sum_{m=2}^{M}\sum_{i=1}^{g} x_{m,i}\right]. \tag{10}$$

Using (9) and (10), we can find the estimate of SNR for the data-aided case.

B. Non-Data Aided Estimation

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

$$p_{\mathbf{x}_i}(\mathbf{x}|s_n=1) = \frac{1}{N^M} I_0\left(\frac{2\sqrt{x_n}|A|}{N}\right) \exp\left(-\frac{1}{N}\left[x_n + |A|^2 + \sum_{m=1, m\neq n}^{M} x_m\right]\right). \tag{11}$$

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

$$p_{\mathbf{x}_i}(\mathbf{x}) = \frac{1}{MN^M} \exp\left(-\frac{1}{N}\left[\sum_{m=1}^{M} x_m + |A|^2\right]\right) \left[\sum_{m=1}^{M} I_0\left(\frac{2\sqrt{x_m}|A|}{N}\right)\right]. \tag{12}$$

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

$$\Lambda_{\mathbf{x}_i}(\mathbf{x}; A, N) = -kM\ln(N) - \frac{1}{N}\left(\sum_{m=1}^{M}\sum_{i=1}^{k} x_{m,i} + k|A|^2\right) - k\ln(M) + \sum_{i=1}^{k} \ln\left[\sum_{m=1}^{M} I_0\left(\frac{2\sqrt{x_{m,i}}|A|}{N}\right)\right]. \tag{13}$$

The partial derivative with respect to A is given as

$$\frac{\partial \Lambda_{\mathbf{x}_i}(\mathbf{x})}{\partial A} = -\frac{KA^*}{N} + \sum_{i=1}^{k} \frac{\sum_{m=1}^{M} I_1\left(\frac{2\sqrt{x_{m,i}}|A|}{N}\right)}{\sum_{m=1}^{M} I_0\left(\frac{2\sqrt{x_{m,i}}|A|}{N}\right)} \frac{A^*\sqrt{x_{m,i}}}{N|A|} \tag{14}$$

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

$$|\hat{A}| = \frac{1}{k}\sum_{i=1}^{k} \max_{m=1,\ldots,M} \sqrt{x_{m,i}}. \tag{15}$$

The noise power estimate is given in a similar manner as

$$\hat{N} = \frac{1}{Mk}\left[\sum_{m=1}^{M}\sum_{i=1}^{k} x_{m,i} - k|\hat{A}|^2\right]. \tag{16}$$

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 the pilot and data symbols. So we use (6) for $i = 1, 2, ..., g$ and (12) for $i = g+1, ..., g+l = k$. Thus the log-likelihood function from the joint PDF is given as

$$\begin{aligned}\Lambda_{joint} = &-kM\ln N - l\ln M + \sum_{i=1}^{g} \ln\left(I_0\left(\frac{2\sqrt{x_{1,i}}|A|}{N}\right)\right) \\ &+ \sum_{i=g+1}^{k} \ln\left[\sum_{m=1}^{M} I_0\left(\frac{2\sqrt{x_{m,i}}|A|}{N}\right)\right] - \frac{k|A|^2}{N} \\ &- \frac{1}{N}\left(\sum_{i=1}^{g} x_{1,i} + \sum_{m=2}^{M}\sum_{i=1}^{g} x_{m,i} + \sum_{m=1}^{M}\sum_{i=g+1}^{k} x_{m,i}\right).\end{aligned} \tag{17}$$

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

$$|\hat{A}| = \frac{1}{k}\left[\sum_{i=1}^{g}\sqrt{x_{1,i}} + \sum_{i=g+1}^{k} \max_{m=1,\ldots,M} \sqrt{x_{m,i}}\right], \tag{18}$$

$$\hat{N} = \frac{1}{kM}\left[\sum_{m=1}^{M}\sum_{i=1}^{g} x_{m,i} + \sum_{m=1}^{M}\sum_{i=g+1}^{k} x_{m,i} - k|\hat{A}|^2\right]. \tag{19}$$

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 $\mathbf{Z}$, given as

$$\mathbf{Z} = E\{\mathbf{x}_m\} E\{\mathbf{x}_m\}^T \left(E\{\mathbf{x}_m \mathbf{x}_m^T\}\right)^{-1}, \tag{20}$$

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

$$E\{x_m\} = \frac{1}{M}\left[|A|^2 + MN\right], \quad m = 1, ..., M \tag{21}$$

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

$$E\{\mathbf{x}_m\} E\{\mathbf{x}_m\}^T = \frac{1}{M^2}\left[|A|^2 + MN\right]^2 \mathbf{1}_M, \tag{22}$$

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

$$E\{x_m^2\} = \frac{1}{M}\left[|A|^4 + 4N|A|^2 + 2MN^2\right], \quad m = 1, ..., M \tag{23}$$

while the rest of all elements will be $(E\{x_m\})^2$, where $E\{x_m\}$ is given by (21). Thus we can write

$$E\{\mathbf{x}_m \mathbf{x}_m^T\} = \frac{\left[|A|^2 + MN\right]^2}{M^2} \underbrace{\begin{bmatrix} a & 1 & \cdots & 1 \\ 1 & a & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & a \end{bmatrix}}_{\mathbf{H}} \tag{24}$$

where the element a is given as

$$a = \frac{E\{x_m^2\}}{E\{x_m\}^2} = M\frac{\gamma^2 + 4\gamma + 2M}{\gamma^2 + M^2 + 2M\gamma} \tag{25}$$

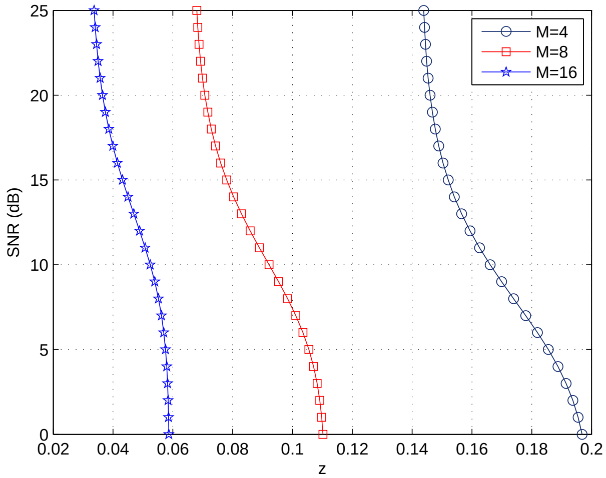

Fig. 2. Relationship between γ and z

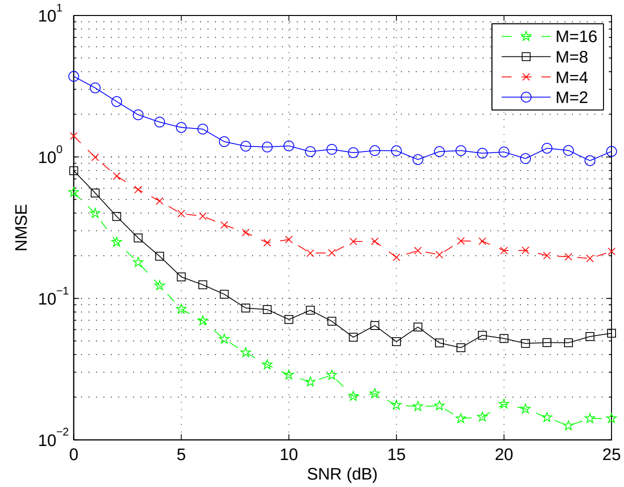

Fig. 3. Effect on increasing M on NMSE for k=100

and $\gamma = |A|^2/N$ is the signal-to-noise ratio. It can be noticed that the matrix $\mathbf{H}$ in (24) is a special kind of Hankel matrix which is also circulant [9]. The inverse of such a circulant matrix, of order M, is given as

$$\mathbf{H}^{-1} = \frac{1}{\zeta}\begin{bmatrix} a+M-2 & -1 & \cdots & -1 \\ -1 & a+M-2 & \cdots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \cdots & a+M-2 \end{bmatrix}, \tag{26}$$

where $\zeta = a^2 + (M-2)a - (M-1)$. Thus the matrix $\mathbf{Z}$ from (20) is given as

$$\mathbf{Z} = \mathbf{1}_M \mathbf{H}^{-1} = \frac{a-1}{a^2 + (M-2)a - (M-1)} \mathbf{1}_M \tag{27}$$

We can now utilize only one of the element from $\mathbf{Z}$ given by

$$z = \frac{\gamma^2 + 2M\gamma + M^2}{M^3 + 2M^2\gamma + M^2 + 2M\gamma^2 + 2M\gamma - \gamma^2} \tag{28}$$

To make this approach practical, we replace the expectations in (20) with the corresponding block averages to compute the SNR estimate. We may solve for γ to get an estimate as

$$\hat{\gamma} = \frac{1}{2Mz - z - 1}\Big[M - Mz - Mz^2 + \sqrt{(}-M^4z^2 + M^3z^2 + M^3z + 2M^2z^2 - 2M^2z)\Big] \tag{29}$$

It can be seen from (28) that z has a solution at $z = 1/3$ for $M = 2$. Thus this method is not applicable for a BFSK system. From Figure 2, we can observe that in the higher SNR regime, the EDS approach will suffer badly and will not give accurate estimates because all curves become steeper as the SNR rises for any $M \neq 2$. But it will be shown in the next section, that the EDS approach shows best performance for larger data set and in low SNR region over the MLE algorithms discussed in the previous section, because there are no approximation errors in this approach.

IV. Simulation Results

In this section, we compare the normalized mean squared error (normalized with respect to the square of the true value of the average SNR) of the estimators using simulations for different values of M and for different packet lengths averaged over 10,000 trials. Figure 3 shows the NMSE vs. true SNR for the PDA estimator, with 100 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 Section III.

In Figure 4, the packet is assumed to comprise 100 pilot symbols and 900 data symbols. The NDA and *Joint* cases perform similarly because most of the packet is data and the NMSE is high in the low SNR region. This can be attributed to the approximation errors in the low SNR regime. 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. However, it shows bad behavior at high SNR due to the steepness of curves from Figure 2. 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.

Figure 5 treats the short packet scenario, with 8 pilot

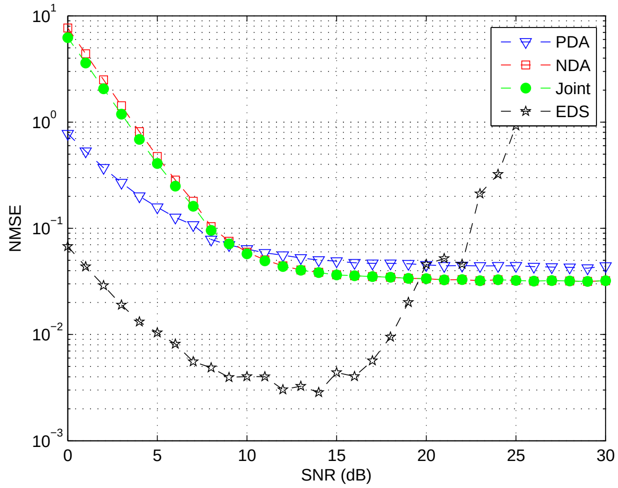

Fig. 4. NMSE for different schemes with M=8, k=1000, (g=100)

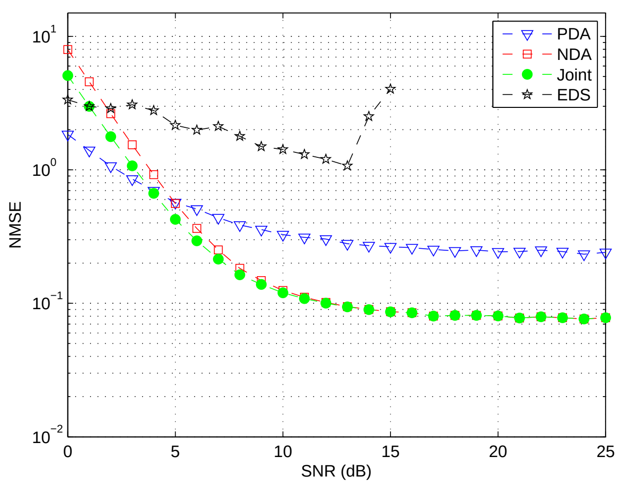

Fig. 5. NMSE for different schemes with M=8, k=36, (g=8)

symbols and 28 data symbols. The EDS approach does not perform well because of the limitations of the availability of data (the 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.

V. Conclusion

We have derived the MLE and statistical based estimators for SNR for a non-coherent MFSK receiver assuming different degrees of data knowledge in a packet. 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).

Appendix

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

$$B = \sum_{i=1}^{k} \frac{\sum_{m=1}^{2} I_1\left(\frac{2\sqrt{x_{m,i}}|A|}{N}\right)}{\sum_{m=1}^{2} I_0\left(\frac{2\sqrt{x_{m,i}}|A|}{N}\right)} \sqrt{x_{m,i}}$$
$$= \sum_{i=1}^{k} \frac{I_1\left(\frac{2\sqrt{x_{1,i}}|A|}{N}\right)\sqrt{x_{1,i}} + I_1\left(\frac{2\sqrt{x_{2,i}}|A|}{N}\right)\sqrt{x_{2,i}}}{I_0\left(\frac{2\sqrt{x_{1,i}}|A|}{N}\right) + I_0\left(\frac{2\sqrt{x_{2,i}}|A|}{N}\right)} \tag{30}$$

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

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

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

$$B \approx \sum_{i=1}^{k} \max_{m=1,\ldots,M} \sqrt{x_{m,i}} \tag{31}$$

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 and Sensor Wireless Networks, Special Issue on the 1st International Conference on Sensor Technologies and Applications*, Vol. 8, No. 1-2, pp. 79-117, 2009.

[2] D.R. Pauluzzi and N.C. Beaulieu, "A comparison of SNR estimation techniques for the AWGN channel," *IEEE Trans. Commun.*, vol. 48, No. 10, October 2000.

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

[4] A. Ramesh, A. Chockalingam and L.B. Milstein, "SNR estimation in Nakagami-m fading with diversity combining and its applications to turbo decoding," *IEEE Trans. Commun.*, vol. 50, No. 11, November 2002.

[5] E. Dilaveroglu and T. Ertas, "CRBs and MLEs for SNR estimation on non coherent BFSK signals in Rayleigh fading," *IEEE Elect. Letters*, vol. 41, No. 2, January 2005.

[6] 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.

[7] S.A. Hassan and M.A. Ingram, "SNR Estimation for a Non-Coherent M-FSK Receiver in a Rayleigh Fading Environment," submitted to *IEEE ICC*, 2010.

[8] The Modified Bessel's Differential Equation. [Online]. Available: http://www.efunda.com/math/bessel/modifiedbessel.cfm

[9] G. Strang, *Introduction to Linear Algebra*, 3rd edition, Wellesley-Cambridge Press, Wellesley, 2005.