On a New Approach to SNR Estimation of BPSK Signals

Taleb Moazzeni, Amei Amei, Jian Ma, and Yingtao Jiang

Abstract—Signal-to-noise ratio (SNR) information is required in many communication receivers and their proper operation is, to a large extent, related to the SNR estimation techniques they employ. Most of the available SNR estimators are based on approaches that either require large observation length or suffer from high computation complexity. In this paper, we propose a low complexity, yet accurate SNR estimation technique that is sufficient to yield meaningful estimation for short data records. It is shown that our estimator is fairly close to the (CRLB) for low complexity, yet accurate SNR estimation technique that is provided in Section III. Section IV gives the simulations and results. Finally, in Section V, the conclusions are drawn.

II. SIGNAL MODEL AND DERIVATION

Let the binary phase shift keying (BPSK) signal at the output of the matched filter be modeled as a one-dimensional signal,

\[ x(k) = Sa(k) + w(k) \]  

where \( S \) is a real scalar, \( a(k) \)’s are symbols taking values \( \pm 1 \) independently and identically with equal probabilities, and \( w(k) \)’s are additive white Gaussian noise with zero mean and variance of \( \sigma^2 \). Without loss of generality, we assume that \( x(k) \) is normalized such that the sample variance of \( x(k) \) is equal to one. In this way, one may have,

\[ \sigma^2 + S^2 = 1 \]  

A. Noise Variance Estimation

To estimate the SNR, the following steps are followed, Step 1. Calculating the square of observation data,

\[ y(k) = (x(k))^2 \]  

or,

\[ y(k) = S^2 \pm 2Sw(k) + (w(k))^2 \]  

Step 2. Arranging the obtained points in an ascending order, that is

\[ 0 \leq \bar{y}(1) \leq \bar{y}(2) \leq \ldots \leq \bar{y}(K) \]  

Step 3. Choosing \( L + 1 \) observation data around the mid point, that is

\[ \bar{y}_L(l) = \bar{y}(l), \quad l = (K - L)/2, \ldots, 0, \ldots, (K + L)/2 \]  

Step 4. Using a least-squares polynomial to fit the obtained data, \( \bar{y}_L(l) \). In this work we choose the forth degree polynomial. Note that the higher degree polynomial may give more accurate results but it makes the computational cost too high. We will show that the noise standard deviation can be estimated from the leading coefficient of the fitting polynomial.
In this paper we approximate the quantile function with a forth degree polynomial along with introducing a numerical method to estimate the coefficients in terms of unknown parameter $\sigma$. To derive the equations, since the theoretical quantile function is not available, we derive numerical results by choosing the sample size ($K$) too large so that the numerical results converge to the theory. To choose a large enough sample size, by starting from a small sample size, we draw several sets of observations form the underlying distributions as described in (1) and consider the difference between the corresponding sample quantiles. Figure 1 shows the results for four simulation runs for a given sample size. It is observed that for large sample size like 10,000 all curves are fully overlapped.

To measure the numerical difference of results the following equation is used,

$$d(K) = \frac{1}{K} \sum_{k=1}^{K} \sum_{i=1}^{M-1} (\tilde{y}_{i+1}(k) - \tilde{y}_i(k))^2$$  

(8)

where $\tilde{y}_i(k)$ is $\tilde{y}(k)$ at $i$-th simulation run and $M$ is the number of simulations.

The results of the difference measurement are reported in Table 1.

Table 1 shows that the difference between the resulting sample quantiles for large sample size is too small. Therefore, if we choose large sample size like 1,000,000, the numerical results for the approximated quantile function may be approached to the true quantile function.

To approximate the quantile function with a polynomial, we set the sample size $K$ to 1,000,000 and vary the standard

**B. Polynomial Approximation of Sample Quantile Function**

The quantile function is described as simply the value that corresponds to a specified proportion of an ordered sample of a population [13]. In other words, the ordered data values are posterior point estimates of the underlying quantile function.

We employ the power series to create a function that is closed to the underlying quantile function. A Taylor series provides a way to generate such a series. To doing so, the Taylor series of the $p$-th quantile $y_p$ at $p = p_0$ is computed.

That is,

$$P_n(p) = a_0 + a_1(p - p_0) + a_2(p - p_0)^2 + \ldots$$  

(7)

In [14], an analytical approach is presented to express the coefficients $a_k$'s in terms of the probability density function of underlying distribution. Although by following that approach one may derive equations relating these coefficients to the unknown parameters of the density function such as $\sigma$, the derived equations can not be solved analytically (Appendix).
deviation of noise from 0 to 1. The results for four settings are shown in Fig. 2. It is seen that the sample quantile curve over a large region around the midpoint \( p = 1/2 \) can be represented by a low degree polynomial curve. In this work, by using \( L + 1 \) data points taken equal to 80% of observations within this region a forth degree polynomial is fitted.

From Fig. 2, it is also seen that for each \( \sigma \) there is a fourth degree polynomial which can well fit the data. We estimate the coefficients of these polynomials over a range of \( \sigma \). For simplicity, by considering normalized data, we set the range of \( \sigma \) between 0 and 1. Figure 3 shows the numerical results for the estimated coefficients of the forth degree polynomials over a range of \( \sigma \) in terms of accuracy.

As seen from obtained curves in Fig. 3, each coefficient varies with the noise standard deviation in a nonlinear manner. Since our goal is to express the noise standard deviation in terms of the estimated coefficients, only the coefficients that can be approximated by a monotonic function is desired. From Fig. 3, one may see that only the leading coefficient \( C_4 \) has an ascending shape which may be useful for an invertible fitting polynomial. The numerical results of mean square error (MSE) for different polynomials fitted to \( C_4 \) are reported in Table 2.

We choose the ninth degree for the fitting polynomial and results are shown in Fig. 4.

From Table 2 and Fig. 4 it is seen that the coefficient \( C_4(\sigma) \) can be approximated to a high enough accuracy by a ninth degree polynomial over the entire region of \((0,1)\). This gives following result,

\[
C_4(\sigma) = 15626 - 66990\sigma + 117600\sigma^2 - 109660\sigma^3 + 58071\sigma^4 - 29569.9\sigma^5 - 218.8\sigma^7 + 8.5\sigma^8 - 0.06\sigma^9
\]  

(9)

It can be easily shown that the above polynomial function is strictly ascending implying that it has a unique inverse function. By numerical computation we approximate this inverse function with a ninth degree polynomial as,

\[
\hat{\sigma} = 0.0342 + 0.2481C_4 - 0.1488C_4^2 + 0.0699C_4^3 - 0.0203C_4^4 + 0.0036C_4^5 - 4 \times 10^{-4}C_4^6 + 2.6563 \times 10^{-5}C_4^7 + 9.7237 \times 10^{-7}C_4^8 + 1.5037 \times 10^{-8}C_4^9
\]  

(10)

The results for the approximated polynomials for inverse function are shown in Fig. 5 and Table 3.

Figure 5 and Table 3 confirm that the approximating ninth degree polynomial function give a desired level of accuracy for the inverse of function.

![Fig. 3. Estimated coefficients of all fitted fourth degree polynomials \( C_4(\sigma) = C_0(\sigma) + C_1(\sigma)p + C_2(\sigma)p^2 + C_3(\sigma)p^3 + C_4(\sigma)p^4 \) in terms of standard deviation, \( \sigma \) for \( K = 10,000,000 \) and \( L = 0.8K \).](image)

![Fig. 4. The ninth degree polynomial fitted to the leading coefficient \( C_4 \).](image)

![Fig. 5. The approximating ninth degree polynomial for the inverse function.](image)
CRLB and variance of the proposed technique for SNR estimates of BPSK signals with sample sizes of 200 and 400.

C. Estimation of Quantile Function Coefficients

To estimate the coefficients of approximating polynomial quintile function from the observation data, we use a forth degree polynomial fitting to the \( L \) chosen points.

Using least mean square minimization [15],

\[
  C = U^{-1} B
\]  

(11)

where

\[
  C = \begin{bmatrix} C_1 & C_2 & C_3 & C_4 \end{bmatrix}^T
\]

(12)

\[
  U_{ij} = \sum_{i=1}^{L+1} \left( \frac{i}{L+1} \right)^{i+j-2}, \quad i, j = 1, 2, \ldots, 5
\]

(13)

\[
  B_i = \sum_{i=1}^{L+1} \left( \frac{n}{L+1} \right)^{i-1} y_{iL}, \quad i, j = 1, 2, \ldots, 5
\]

(14)

Note that in \( C \), only \( C_4 \) is needed to be calculated as it is employed in the estimate expression for the noise standard deviation in (10).

D. Signal Power and SNR Estimation

Having the estimated noise standard deviation (10), one can simply find an estimate of signal power \( P_s \) as,

\[
  \hat{P}_s = \frac{1}{N} \sum_{k=1}^{N} (x(k))^2 - \hat{\sigma}^2
\]

(15)

Consequently, using (2) the SNR can be estimated by

\[
  \text{SNR} = \frac{1 - \hat{\sigma}^2}{\hat{\sigma}^2}
\]

(16)

III. STATISTICAL ANALYSIS

A. Cramer Rao Lower Bound (CRLB) on the Variance of SNR Estimator

The CRLB of \( \Theta = [S^2, \sigma^2]^T \) estimates are defined as [16],

\[
  \text{CRLB}(g(\Theta)) = \begin{bmatrix} \frac{\partial g(\Theta)}{\partial \Theta} \end{bmatrix}^T \text{I}(\Theta)^{-1} \begin{bmatrix} \frac{\partial g(\Theta)}{\partial \Theta} \end{bmatrix}
\]

(17)

where \( \text{I}(\Theta) \) is the Fisher information matrix given by,

\[
  \begin{bmatrix}
  -E\left\{ \frac{\partial^2 \ln f(x, \Theta)}{\partial \Theta^2} \right\} & -E\left\{ \frac{\partial^2 \ln f(x, \Theta)}{\partial \Theta \partial \sigma^2} \right\} \\
  -E\left\{ \frac{\partial^2 \ln f(x, \Theta)}{\partial \Theta \partial \sigma^2} \right\} & -E\left\{ \frac{\partial^2 \ln f(x, \Theta)}{\partial \sigma^2 \partial \sigma^2} \right\}
  \end{bmatrix}
\]

(18)

It can be shown that the CRLB of SNR estimates for the BPSK signals with unknown modulating symbols is obtained by [17],

\[
  \text{CRLB}(\text{SNR}) = \frac{200(\frac{1}{2} - f(\alpha) + 1)}{N(ln(10))^2(1 - f(\alpha) - 4\alpha f(\alpha))}(dB)^2
\]

(19)

where

\[
  f(\alpha) = \frac{\exp(-\alpha)}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u^2 \exp\left( -\frac{u^2}{2} \right) \cosh(u\sqrt{2\alpha}) \, du
\]

(20)

Note that \( \alpha \) is in linear scale.

IV. SIMULATION RESULTS

To verify the performance of the proposed estimator, the following simulations are carried out with these setting parameters: the BPSK signal is produced by random number generator with two states of \( \pm 1 \). The true SNR is set to be between -10 and 10 dB. The sample size is set to 200 and 400, and the number of simulation runs for each simulation is 50.

A. Efficiency of the Proposed Technique

In this paper, we use CRLB as a finger of merit to check the accuracy of the proposed estimator. Figure 6 shows the CRLB and the simulation results for the variance of the proposed estimator. As seen, for low SNR (e.g., dB), the difference between the CRLB and variance of estimations is dependent on the sample size, whereas at high SNR, the estimator approaches the CRLB. This indicates that the proposed estimator is approximately a minimum variance unbiased estimator for high SNR values.

Figure 7 compares the SNR estimates and true SNR. As shown, at large SNR values (> 8 dB in this experiment) the estimator can indeed track the true SNR.

B. Performance Comparison to M2M4 Technique

To evaluate and compare the effectiveness of the proposed SNR estimation technique, we use accuracy and convergence speed as performance metrics. The well-known SNR estimators M2M4 is considered and the results are compared with the
As the quantile function of a probability distribution is the inverse of its cumulative distribution function (cdf), we define,
\[ p = F(y_p), \quad 0 < p < 1 \]  
where \( F \) is the cdf of distribution of \( y \). For the expansion point \( p = p_0 \) one may find [14],
\[ a_0 = y_{p_0} \]  
\[ a_1 = \frac{1}{f_y(y_{p_0})} \]  
\[ a_2 = -\frac{f_y'(y_{p_0})}{2f_y^2(y_{p_0})} \]  
\[ a_3 = \frac{3(f_y'(y_{p_0}))^2 - f_y(y_{p_0})f_y''(y_{p_0})}{6f_y^3(y_{p_0})} \]  
where \( f_y \) and \( f_y^{(m)}(y_{p_0}) \) denote the probability density function of \( y \) and the \( m \)-th derivative of the function \( f_y(y_{p_0}) \) at \( p = p_0 \), respectively.

In this work we consider the expansion point \( p_0 = 1/2 \). In this case, \( y_{p_0} \) will be equal to the sample mean of \( y(k) \). Because when \( F(y_{p_0}) = 1/2 \) the corresponding point of 1/2 will be the sample mean of \( y \).

Therefore, to obtain \( a_1 \), we have to find the value of \( f_y \) at \( y_{p_0} = 1/2 \) which is discussed as follows.

### Appendix

**A. Distribution of the Square of Observation Points**

Equation (4) shows that the variable \( y \) is composed of three terms: one with constant value (\( S^2 \)), one with normal distribution multiplied by a constant value and the other one is the square of a normal variable. The first two terms can be treated as a normal variable with mean of \( S^2 \) and variance of \( 4S^2\sigma^2 \). The last term is treated as a new variable \( z = (w(k))^2 \) which takes the following density function form [18],
\[ f_z(z) = \frac{1}{\sqrt{2\Gamma(1/2)}} \sqrt{z} \sigma \exp^{-z^2/2\sigma} \]  
where \( \Gamma \) denotes the Gamma function.

Therefore, finding the distribution of \( y \) boiled down to computing the probability distribution of sum of two variables. This can be obtained by convolution of their individual distributions [19].
\[ f_y(y) = \left( \frac{1}{\sqrt{2\pi\sigma^2}} \right)^2 e^{-\frac{(y-\mu)^2}{2\sigma^2}} \left( \frac{1}{\sqrt{2\Gamma(1/2)}}\sqrt{y/\sigma} \right) \exp^{-y^2/2\sigma^2} \]  
where \( \ast \) denotes convolution, \( \sigma_{xs} = 2S\sigma \) and \( \mu_{xs} = S^2 \).

**B. Equation Derivation for Coefficients of Quantile Function**

For \( y = y_{p_0} \), using (1),(2) and (23) we get,
\[ a_0 = y_{p_0} = (y(k)) = S^2 + \sigma^2 = 1 \]  
Using (24), (28), and (29), one may find,
\[ a_1 = \frac{1}{f_y(1)} = \frac{4\pi \sqrt{\sigma}(1 - \sigma^2)}{\int_{-\infty}^{\infty} e^{-y^2/2\sigma} \sqrt{y} e^{-y(1-\sigma^2)/4\sigma(1-\sigma)} dy} \]  

### V. Conclusions

In this paper, by separating the signal level from the noisy observation data a SNR estimation technique for BPSK signal in AWGN was derived. By arranging the data we fit a fourth degree polynomial to a portion of arranged data. By numerical computations we derived expressions from the approximating polynomials for the noise standard deviation. The proposed SNR estimator was compared with M2M4 estimator and the numerical results showed that the new estimator performed faster convergence speed, but comparable accuracy results, especially for high SNR values. The statistical analysis such as the CRLB of the SNR estimator was given. The simulations confirmed this theoretical prediction and showed that the proposed estimator attains the CRLB for high SNR (> 8 dB in the experiments).
By applying this convolution property \((f * g)' = f' * g\), one may have,

\[
f_y'(1) = \frac{-1}{8\pi \sqrt{\sigma(1 - \sigma^2)}} \int_{-\infty}^{\infty} \frac{(y + \sigma)e^{-y/2\sigma}}{\sqrt{y}} e^{-\frac{(y^2 - \sigma^2)^2}{4\sigma^2(1 - \sigma^2)}} dy
\]

Using (25), (26) and above equations, the related equations for coefficients \(a_2\) and \(a_3\) are derived which may not have analytical solutions.

REFERENCES


