SOLUTIONS AND COMPARISON OF MAXIMUM LIKELIHOOD AND FULL-LEAST-SQUARES ESTIMATIONS FOR CIRCLE FITTING

Zhenhua Ma, K.C. Ho, Le Yang

Department of ECE, University of Missouri, Columbia, MO 65211
zm325@mizzou.edu, hod@missouri.edu, lyx35@mizzou.edu

ABSTRACT

The fitting of a number of noisy data points with a circle has found numerous applications in image processing and pattern recognition. This paper examines two methods to estimate the circle parameters: the Maximum Likelihood (ML) method and the Full-Least-Squares (FLS) method. The ML method is based on the noisy model from the data while the FLS method minimizes the geometric distance square. We first provide the iterative solutions of them using Taylor-series linearization approach. We then show analytically that FLS does not yield the ML solution. This is in contrast to previous study that the FLS method gives the same solution as ML. FLS method approximates the ML estimation only if the noise power is much less than the circle radius square. Simulations are included to support the theoretical development.

Index Terms— Circle fitting, parameter estimation, CRLB, maximum likelihood, full-least-squares.

1. INTRODUCTION

The problem of fitting a circle given a number of noisy measurement points is a classical problem which remains to attract research interests over the past two decade. Circle fitting is essentially the estimation of the circle center and radius from the measurements. Unlike line fitting, circle fitting is a non-linear estimation problem subject to bias and thresholding behavior. The interest in circle fitting comes from its wide variety of applications in image processing and pattern recognition [1].

Associated with an estimation problem is the lower bound on the mean-square error (MSE) of parameter estimates. The evaluation of the Cramér-Rao Lower Bound (CRLB) on this circle fitting estimation has been given in [2] and [3]. Many methods have been proposed for this interesting problem and a good summary of them is in [4]. The full-least-squares (FLS) method is able to achieve the CRLB performance. It is, however, iterative and numerical solution is needed. The average of intersections (AI) method, the reduced least-squares (RLS) method and the modified least-squares (MLS) method yield closed-form solutions. Their performance, on the other hand, is not able to reach the CRLB accuracy. The well-known Kåsa's method is shown to give a closed-form solution and has CRLB accuracy if the noise level tends to zero. More recently, branch and bound method [5] is proposed for the purpose to obtain maximum likelihood (ML) estimator and its performance is close to the the Kåsa's method.

This paper focuses on (i) a new implementation of the ML estimator through the Taylor-series linearization technique, and (ii) the comparison between the ML estimator and the FLS method. Our theoretical investigation indicates that the FLS method is not the ML estimator. This is in contrast to the previous study [6] that the FLS method gives the ML solution. We have shown that the FLS method approximates the ML estimator only when the noise power of the measurements is much smaller than the radius square of the circle. Hence, ML solution outperforms FLS method where the noise power is large or when the circle radius is small. The theoretical study is supported by simulations.

Throughout this paper, bold-face lower case letter denotes column vector and bold-face upper case letter represents matrix. If $(*)$ is a noisy quantity, $(*)^o$ stands for the true value of $(*)$ without noise.

2. THE CIRCLE FITTING PROBLEM

Let $\mathbf{s}_i = [x_i, y_i]^T, i = 1, 2, \cdots, N$, be a set of N measurement points defined as

$$\mathbf{s}_i = \mathbf{s}_i^o + \mathbf{n}_i \tag{1}$$

where $\mathbf{s}_i^o = [x_i^o, y_i^o]^T$ is the true data point sampled from a circle of center $\mathbf{c}^o = [\overline{x}^o, \overline{y}^o]^T$ and radius r^o such that it satisfies

$$\|\mathbf{s}_i^o - \mathbf{c}^o\| = r^o \tag{2}$$

and $\| * \|$ is the Euclidean norm. $\mathbf{n}_i$ is the measurement noise and is modeled as zero mean Guassian with diagonal covariance matrix $\sigma^2\mathbf{I}_{2\times 2}$. It is further assumed that $\mathbf{n}_i$ is I.I.D. for $i = 1, 2, \cdots, N$. Given the noisy measurements $\mathbf{s}_i$, we are interested to find an estimate of $\boldsymbol{\theta}^o = [\mathbf{c}^{oT}, r^o]^T$ that best fits the measurements in some optimal sense.

3. THE ML SOLUTION

Since $\mathbf{n}_i \sim N(0, \sigma^2\mathbf{I}_{2\times 2})$, the ML cost function is simply equal to

$$J(\boldsymbol{\theta}) = \sum_{i=1}^{N} \|\mathbf{s}_i - \mathbf{s}_i^o(\boldsymbol{\theta})\|^2 \tag{3}$$

and the ML solution is the value of $\boldsymbol{\theta}$ that minimizes $J(\boldsymbol{\theta})$. The ML cost function is not quadratic with respect to $\boldsymbol{\theta}$ because $\mathbf{s}_i^o(\boldsymbol{\theta})$ is related to $\boldsymbol{\theta}$ in a highly nonlinear manner. We shall propose the use of Taylor-series linearization approach to minimize $J(\boldsymbol{\theta})$ through iteration.

Let $\boldsymbol{\theta}_{(o)} = [\overline{x}_{(o)}, \overline{y}_{(o)}, r_{(o)}]^T$ be an initial solution guess. Expanding $\mathbf{s}^o(\boldsymbol{\theta})$ through Taylor-series up to linear term gives

$$\mathbf{s}_i^o(\boldsymbol{\theta}) = \mathbf{s}_i^o(\boldsymbol{\theta}_{(o)}) + \mathbf{G}_i(\boldsymbol{\theta}_{(o)})(\boldsymbol{\theta} - \boldsymbol{\theta}_{(o)}) \quad (4)$$

where $\mathbf{G}_i(\boldsymbol{\theta}_{(o)}) = \left.\frac{\partial \mathbf{s}_i^o(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}_{(o)}}$ is the 2×3 gradient matrix. Putting (4) into (3) forms

$$J(\boldsymbol{\theta}) \simeq \sum_{i=1}^{N} \|\mathbf{s}_i - \mathbf{s}_i^o(\boldsymbol{\theta}_{(o)}) - \mathbf{G}_i(\boldsymbol{\theta}_{(o)})(\boldsymbol{\theta} - \boldsymbol{\theta}_{(o)})\|^2 \quad (5)$$

(5) is quadratic with respect to $\boldsymbol{\theta}$, whose minimum is achieved when

$$\begin{aligned}\boldsymbol{\theta} = \boldsymbol{\theta}_{(o)} + &\Big[\sum_{i=1}^{N} \mathbf{G}_i(\boldsymbol{\theta}_{(o)})^T \mathbf{G}_i(\boldsymbol{\theta}_{(o)})\Big]^{-1} \cdot \\ &\Big[\sum_{i=1}^{N} \mathbf{G}_i(\boldsymbol{\theta}_{(o)})^T \big(\mathbf{s}_i - \mathbf{s}_i^o(\boldsymbol{\theta}_{(o)})\big)\Big]\end{aligned} \quad (6)$$

To improve the solution, we set $\boldsymbol{\theta}_{(o)}$ to the answer from (6) and repeat the computation. Indeed, the proposed solution can be easily expressed as

$$\begin{aligned}\boldsymbol{\theta}_{(k+1)} = \boldsymbol{\theta}_{(k)} + &\Big[\sum_{i=1}^{N} \mathbf{G}_i(\boldsymbol{\theta}_{(k)})^T \mathbf{G}_i(\boldsymbol{\theta}_{(k)})\Big]^{-1} \cdot \\ &\Big[\sum_{i=1}^{N} \mathbf{G}_i(\boldsymbol{\theta}_{(k)})^T \big(\mathbf{s}_i - \mathbf{s}_i^o(\boldsymbol{\theta}_{(k)})\big)\Big]\end{aligned} \quad (7)$$

for $k = 0, 1, \cdots$, where k is the iteration count. The iteration stops when $\|\boldsymbol{\theta}_{(k+1)}$ - $\boldsymbol{\theta}_{(k)}\| < \delta$, where δ is some small number.

We now determine $\mathbf{s}_i^o(\boldsymbol{\theta}_{(k)})$ and $\mathbf{G}_i(\boldsymbol{\theta}_{(k)})$ to complete the iterative solution. Using Chan & Thomas [2] parametric form of circle representation, a point (x_i^o, y_i^o) on a circle can be expressed as

$$\begin{aligned} x_i^o &= \overline{x}^o + r^o \cos(\phi_i^o) \\ y_i^o &= \overline{y}^o + r^o \sin(\phi_i^o) \end{aligned} \quad (8)$$

where $\phi_i^o = \tan^{-1} \frac{y_i^o - \overline{y}^o}{x_i^o - \overline{x}^o}$. Thus, given $\boldsymbol{\theta}_{(k)} = [\overline{x}_{(k)}, \overline{y}_{(k)}, r_{(k)}]^T$,

$$\begin{aligned} \mathbf{s}_i^o(\boldsymbol{\theta}_{(k)}) &= \begin{bmatrix} \overline{x}_{(k)} \\ \overline{y}_{(k)} \end{bmatrix} + r_{(k)} \begin{bmatrix} \cos(\phi_i) \\ \sin(\phi_i) \end{bmatrix} \\ \phi_i &= \tan^{-1} \frac{y_i - \overline{y}}{x_i - \overline{x}} \end{aligned} \quad (9)$$

Note that we have replaced (x_i^o, y_i^o) by (x_i, y_i) in obtaining ϕ_i because (x_i^o, y_i^o) is not available.

The gradient matrix $\mathbf{G}_i(\boldsymbol{\theta}_{(k)})$ is

$$\left.\frac{\partial \mathbf{s}_i^o(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}_{(k)}} = \left.\begin{bmatrix} \frac{\partial x_i^o}{\partial \overline{x}} & \frac{\partial x_i^o}{\partial \overline{y}} & \frac{\partial x_i^o}{\partial r} \\ \frac{\partial y_i^o}{\partial \overline{x}} & \frac{\partial y_i^o}{\partial \overline{y}} & \frac{\partial y_i^o}{\partial r} \end{bmatrix}\right|_{\boldsymbol{\theta}_{(k)}} \quad (10)$$

whose elements are [2], after replacing (x_i^o, y_i^o) by (x_i, y_i) because (x_i^o, y_i^o) is not known,

$$\begin{aligned}
\left.\frac{\partial x_i^o}{\partial \overline{x}}\right|_{\boldsymbol{\theta}_{(k)}} &= \frac{(x_i - \overline{x}_{(k)})^2}{r_{(k)}^2}, & \left.\frac{\partial x_i^o}{\partial \overline{y}}\right|_{\boldsymbol{\theta}_{(k)}} &= \frac{(x_i - \overline{x}_{(k)})(y_i - \overline{y}_{(k)})}{r_{(k)}^2} \\
\left.\frac{\partial y_i^o}{\partial \overline{y}}\right|_{\boldsymbol{\theta}_{(k)}} &= \frac{(y_i - \overline{y}_{(k)})^2}{r_{(k)}^2}, & \left.\frac{\partial y_i^o}{\partial \overline{x}}\right|_{\boldsymbol{\theta}_{(k)}} &= \frac{(x_i - \overline{x}_{(k)})(y_i - \overline{y}_{(k)})}{r_{(k)}^2} \\
\left.\frac{\partial x_i^o}{\partial r}\right|_{\boldsymbol{\theta}_{(k)}} &= \frac{x_i - \overline{x}_{(k)}}{r_{(k)}}, & \left.\frac{\partial y_i^o}{\partial r}\right|_{\boldsymbol{\theta}_{(k)}} &= \frac{y_i - \overline{y}_{(k)}}{r_{(k)}}
\end{aligned} \quad (11)$$

(7) together with (9)-(11) forms the ML iterative solution.

The ML solution presented here uses only the linear term in the Taylor series expansion. Including the second order term may improve performance, especially in reducing the sensitivity in initialization. This is a subject for further study.

4. THE FLS SOLUTION

The cost function of FLS is [4]

$$F = \sum_{i=1}^{N} (r - \|\mathbf{s}_i - \mathbf{c}\|)^2 \quad (12)$$

The solution $\boldsymbol{\theta} = [\mathbf{c}^T, r]^T$ is found by minimizing F. F is related to $\boldsymbol{\theta}$ in a complicated manner and iterative minimization is needed. Following the same approach as in the previous section through Taylor-series linearization, the FLS solution through iteration is

$$\begin{aligned}\boldsymbol{\theta}_{(k+1)} = \boldsymbol{\theta}_{(k)} + &\Big[\sum_{i=1}^{N} \mathbf{H}_i(\boldsymbol{\theta}_{(k)})^T \mathbf{H}_i(\boldsymbol{\theta}_{(k)})\Big]^{-1} \cdot \\ &\Big[\sum_{i=1}^{N} \mathbf{H}_i(\boldsymbol{\theta}_{(k)})^T \big(r_{(k)} - \|\mathbf{s}_i - \mathbf{c}_{(k)}\|\big)\Big]\end{aligned} \quad (13)$$

for $k = 0, 1, \cdots$, where $\boldsymbol{\theta}_{(o)} = [\mathbf{c}_{(o)}^T, r_{(o)}]^T$ is some initial guess. The iteration stops when $\|\boldsymbol{\theta}_{(k+1)}$ - $\boldsymbol{\theta}_{(k)}\| < \delta$, where δ is a small value. The gradient matrix $\mathbf{H}_i(\boldsymbol{\theta}_{(k)})$ is

$$\mathbf{H}_i(\boldsymbol{\theta}_{(k)}) = \left[\left(\frac{\mathbf{s}_i - \mathbf{c}_{(k)}}{\|\mathbf{s}_i - \mathbf{c}_{(k)}\|}\right)^T \quad 1\right]. \quad (14)$$

5. COMPARISON

Let us now compare ML and FLS. We begin with the FLS cost function in (12). Expanding the square gives

$$(r - \|\mathbf{s}_i - \mathbf{c}\|)^2 = r^2 + \|\mathbf{s}_i - \mathbf{c}\|^2 - 2r\|\mathbf{s}_i - \mathbf{c}\|. \quad (15)$$

Let $\mathbf{s}_i^o(\boldsymbol{\theta})$ be a point on the circle defined by $\boldsymbol{\theta}$. Then

$$r^2 = \|\mathbf{s}_i^o(\boldsymbol{\theta}) - \mathbf{c}\|^2 \quad (16)$$

we can express $\|\mathbf{s}_i - \mathbf{c}\|^2$ as

$$\begin{aligned}\|\mathbf{s}_i - \mathbf{c}\|^2 &= \|\mathbf{s}_i^o(\boldsymbol{\theta}) - \mathbf{c} + \mathbf{s}_i - \mathbf{s}_i^o(\boldsymbol{\theta})\|^2 \\ &= r^2\Big[1 + \frac{\|\mathbf{s}_i - \mathbf{s}_i^o(\boldsymbol{\theta})\|^2}{r^2} + \\ &\qquad \frac{2}{r^2}(\mathbf{s}_i^o(\boldsymbol{\theta}) - \mathbf{c})^T(\mathbf{s}_i - \mathbf{s}_i^o(\boldsymbol{\theta}))\Big]\end{aligned} \quad (17)$$

If r is large compared to $\|\mathbf{s}_i - \mathbf{s}_i^o(\boldsymbol{\theta})\|$ so that $\|\mathbf{s}_i - \mathbf{s}_i^o(\boldsymbol{\theta})\|/r \ll 1$, then from (17),

$$\|\mathbf{s}_i - \mathbf{c}\| \simeq r\Big[1 + \frac{1}{r^2}\big(\mathbf{s}_i^o(\boldsymbol{\theta}) - \mathbf{c}\big)^T\big(\mathbf{s}_i - \mathbf{s}_i^o(\boldsymbol{\theta})\big)\Big] \quad (18)$$

Putting (17)-(18) into (15) yields immediately

$$(r - \|\mathbf{s}_i - \mathbf{c}\|)^2 \simeq \|\mathbf{s}_i - \mathbf{s}_i^o(\boldsymbol{\theta})\|^2 \quad (19)$$

so that (12) becomes

$$F \simeq \sum_{i=1}^{N} \|\mathbf{s}_i - \mathbf{s}_i^o(\boldsymbol{\theta})\|^2 \quad (20)$$

which is the ML cost function in (3).

We can now conclude that in general FLS is not the same as ML estimator. It approaches the ML estimator if

$$\varepsilon_i = \frac{\|\mathbf{s}_i - \mathbf{s}_i^o(\boldsymbol{\theta})\|}{r} \ll 1. \quad (21)$$

This condition is satisfied if the noise level in the data measurements is small, or when r is big. Thus, we expect that the ML solution will outperforms the FLS method when the noise level is high or when the radius of the circle is small.

A previous work [6] shows that FLS gives the ML solution. The derivation there, however, was based on the assumption that $\mathbf{c}$ and ϕ_i are independent variables to simplify the development. This is obviously not the case as can be inferred from (8). The comparison result here does not make this assumption and the simulation presented next confirms our results.

6. SIMULATION

We shall investigate the performance of the proposed ML solution and the FLS method via simulation and compare their performance with respect to the optimum accuracy of the problem, the CRLB. The N true data points are sampled from an arc of a circle with radius r and central angle β. The center of the circle is set to be $(80, 60)m$. The noisy measurements are generated by adding to the true data points zero mean Gaussian white noise with covariance matrix $\sigma^2 I_{2N\times 2N}$. The estimation accuracy is defined as MSE($\boldsymbol{\theta}$)= $\sum_{l=1}^{L}\|\boldsymbol{\theta}_l - \boldsymbol{\theta}^o\|^2/L$. where $\boldsymbol{\theta}_l$ is the estimated unknown vector at ensemble l and $L = 10000$ is the number of ensemble runs.

Fig. 1 shows the results when $N = 5, r = 10m$ and $\beta = \pi/3$. The MSEs of ML and FLS methods, together with the CRLB are shown as function of noise power σ^2. When σ^2 is less than or equals to -25dB, both methods can achieve the CRLB accuracy, which is an expected result because the ML estimator is asymptotically efficient. After σ^2 reaches -20dB, the performance of FLS method suffers from the threshold effect, while the ML estimator remains to generate seasonable estimates of $\boldsymbol{\theta}^o$. This observation is consistent with our theoretical analysis in section 5 that for large noise power, the ML solution will outperform the FLS method.

Fig. 2 gives the results when $N = 5, r = 10m$ and $\beta = 2\pi$. The trend observed is similar as in Fig. 1. The MSEs of both methods are smaller because the data points are distributed from the whole circle instead of clustering on a small arc.

Fig. 3 depicts the averaged fitted circles from five noisy measurement points. Simulation configuration is the same as in Fig. 2 except that σ^2 is fixed at 3dB. We can see the circle estimated by ML solution is very close to the true one while the FLS circle significantly deviates from the true one. This again verifies the theoretical development that FLS method would not behavior as an ML estimator when the noise power becomes large.

Fig. 4 illustrated the MSEs of ML and FLS method when $N = 20, r = 10m$ and $\beta = 2\pi$. The figure indicates that both ML and FLS have lower estimation MSE as M increases. Both methods deviating gradually from the CRLB when σ^2 is larger than 0dB but no threshold effect occurs.

In generating Fig. 5, the parameters are $N = 20, r = 2m$ and $\beta = 2\pi$. It is evident that when the radius of the circle decreases, the ML estimator gives much better result than the FLS method after σ^2 reaches 0dB. This observation verifies the result in (21) that only when the ratio between noise power and circle radius square is much less than unity, the FLS approximates the ML estimator.

7. CONCLUSION

This paper derives the ML solution for circle fitting using Taylor-series linearization approach, where the noise in the data measurements are Gaussian and white. It then provides a comparison in performance between the ML estimator and the FLS method. Unlike the result from a previous work that illustrates FLS gives the ML solution, we have shown analytically that FLS does not give ML estimation. It approximates the ML estimator if the ratio between noise power and circle radius square is much less than unity. Otherwise the ML estimator gives much better results. Simulations confirm the theoretical findings.

8. REFERENCES

[1] D.T. Lin and C.M. Yang, "Real-Time Eye Detection using Face Circle Fitting and Dark-Pixel Filtering," *IEEE International Conference on multimedia and expo (ICME)*, pp. 1167 - 1170, June 2004.

[2] Y.T. Chan and S.M. Thomas, "Cramer-Rao Lower Bounds for Estimation of a Circular Arc Center and Its Radius," *Graghical Models and Image Processing*, Vol. 57, No. 6, pp. 527 - 532, Nov. 1995.

[3] K. Kanatani, "Cramer-Rao Lower Bounds for Curve Fitting," *Graphical Models Image Processing*, Vol. 60, No. 2, pp. 93 - 99, Mar. 1998.

[4] D. Umbach and K.N. Jones, "A Few Methods for Fitting Circles to Data," *IEEE Trans. Instrumentation and Measurement*, Vol. 52, No. 6, pp. 1881 - 1885, Dec. 2003.

[5] E.E. Zelniker, B.C. Appleton and I.V.L. Clarkson, "Optimal Circle Fitting via Branch and Bound," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP),* Vol. 4, pp. 709 - 712, Mar. 2005.

[6] Y.T. Chan, B.H. Lee and S.M. Thomas, "Approximate Maximum Likelihood Estimation of Circle Parameters," *Journal of Optimization Theory and Applications,* Vol. 125, No. 3, pp. 723-734, June 2005.

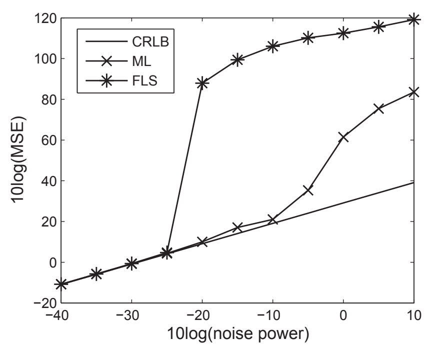

Fig. 1. Performance comparison of the ML and FLS method when $N = 5, r = 10m$ and $\beta = \pi/3$.

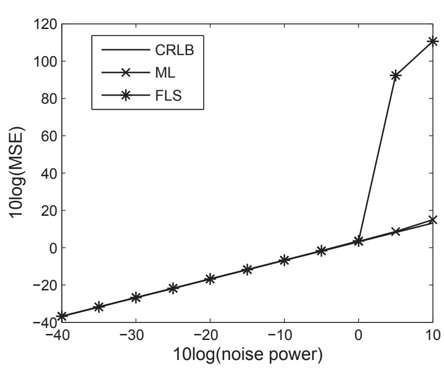

Fig. 2. Performance comparison of the ML and FLS method when $N = 5, r = 10m$ and $\beta = 2\pi$.

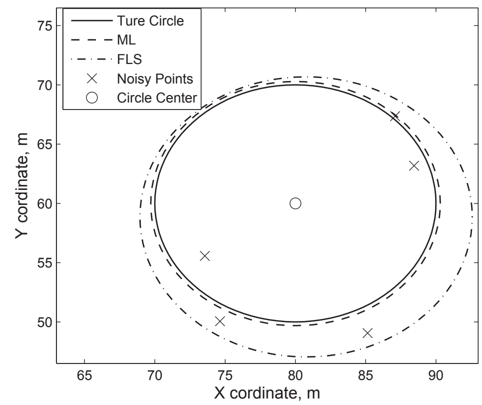

Fig. 3. Averaged fitted circles of the ML and FLS method when $N = 5, r = 10m, \beta = 2\pi$ and $\sigma^2 = 3dB$.

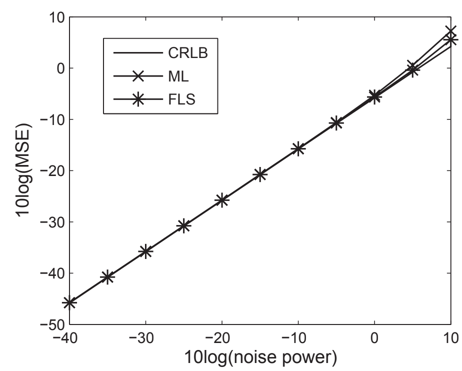

Fig. 4. Performance comparison of the ML and FLS method when $N = 20, r = 10m$ and $\beta = 2\pi$.

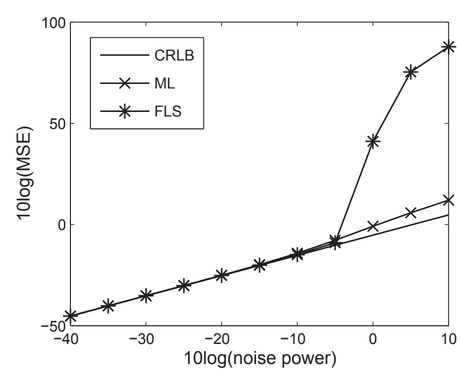

Fig. 5. Performance comparison of the ML and FLS method when $N = 20, r = 2m$ and $\beta = 2\pi$.