UNDERSTANDING THE METHOD OF INTERVAL ERRORS FROM THE INFORMATION THEORY PERSPECTIVE

De Meng, Wen Xu[*], and Menglu Xia

Department of Information Science and Electronic Engineering
Zhejiang University
Hangzhou, China, 310027
[*]Email: wxu@zju.edu.cn

ABSTRACT

Nonlinear parameter estimation often displays a threshold phenomenon, that is, below certain signal-to-noise ratio (SNR) the estimation mean-square error (MSE) increases dramatically. The method of interval errors (MIE) has been shown to provide accurate MSE prediction of related nonlinear techniques well into the estimation threshold region, yet relatively simple and robust in evaluation compared to a global performance bound. However those features have not been understood on a strict theoretical basis. This paper investigates numerical sensitivity of the MIE to parameter sampling resolution, aiming to understanding, from information theory perspective, the underlying mechanism leading to robust MSE approximation. A recently-developed information theory resolution bound is reinterpreted and applied to specify the parameter sampling resolution. Numerical evaluation of the relevant results for array-based bearing estimation supports the proposed connection between the resolution bound and the MIE.

Index Terms— Nonlinear parameter estimation, threshold phenomenon, performance analysis, information theory, resolution bound

1. INTRODUCTION

Nonlinear parameter estimation, including array-based bearing estimation and matched-field processing, is often subject to ambiguities due to multimodal structure in signal field correlation, which is characterized by a mainlobe around the true parameter position and some unpredictable prominent high sidelobes elsewhere. Typical performance displays a threshold behavior, that is, below certain signal-to-noise ratio (SNR) the estimation mean-square error increases dramatically. This threshold phenomenon is often understood in the context of the maximum likelihood estimate (MLE) and in comparison to the performance defined by the Cramer-Rao lower bound (CRB) [1]. It is well known that for a sufficiently high SNR or long observation time, the MLE performance is predicted well by the CRB; however, for low SNR and short observation time, the MLE mean-square error departs significantly from the CRB due to sidelobe ambiguities.

In order to evaluate the threshold phenomenon, a number of performance bounds have been developed and extensively applied [2]. Based on the parameter model a performance bound can be either a local bound if the parameter is assumed deterministic but unknown or a global bound if the parameter is assumed random. The local bounds, such as the local CRB and the Barankin bound, do not exploit any *a priori* parameter information and is limited to unbiased estimates, while an MLE with nonlinear parameter-dependence is often biased at low SNR; the prediction is still far less tight in the threshold region [3]. On the other hand the global bounds are free of the bias assumption, yielding a tight performance prediction; evaluation of those bounds, however, is often subject to numerical sensitivity otherwise quite demanding in computation [4].

A relatively new method for mean-square-error (MSE) approximation, called the method of interval errors (MIE), has also attracted great attentions. MIE was first introduced for time-delay estimation [1] and later revitalized for bearing estimation and matched-field source localization [5]-[9]. A comprehensive review of this method can be found in Ref. [9]. MIE decomposes the total MSE into a weighted sum of two terms: a local error term (ambiguity mainlobe contribution), and an outlier term for global errors (ambiguity sidelobe contribution), and is thus directly and explicitly connected to signal ambiguity structure, lending itself a quantitative tool for ambiguity analysis in parameter estimation. Besides, it is algorithm specific, and works for deterministic yet unknown parameters. MIE is by now a well established approach that has been shown to provide accurate MSE prediction of maximum likelihood estimation and related nonlinear techniques well into the estimation threshold region [9].

Except for some simple problems, both performance bounds and MIE are evaluated numerically and thus require the parameter space be properly sampled. One of the advantages of the MIE is that the required computation is at most the same order as that of a local bound. Evaluation of the MIE has shown much less sensitivity to both detailed ambiguity structure and parameter sampling resolution compared to a global performance bound [10]. However this has not been understood on a strict theoretical basis, and the method is more interpreted as an *ad hoc* approach for MSE approximation.

Recently a new resolution bound has been proposed for understanding the performance limits from the information theory perspective [11]. The method divides the search space into discrete partitions, assigns each candidate parameter value to one of partitions according to certain probability distribution, and then estimates which partition the true parameter belongs to. An error occurs if the estimated partition is not the true one. The goal is to achieve arbitrarily small error probability, rather than minimizing the MSE, and the method has thus not been able to produce a similar MSE prediction as an MSE bound does. However, the results do provide an alternative view and some new insights on the performance behavior in nonlinear parameter estimation.

Motivated by the information theory approach mentioned-above, this paper investigates numerical sensitivity to parameter sampling resolution of the MIE, aiming to understanding the underlying mechanism leading to robust MSE approximation. Instead of interpreting the information theory resolution bound as a limit in parameter estimation, we interpret it as a minimum sampling interval for MIE to achieve a good MSE approximation. The problem considered is bearing estimation using an array of sensors, typical of sonar, radar and communication applications. Numerical evaluations are implemented to verify the proposed connection.

2. BEARING ESTIMATION PROBLEM DEFINITION

Consider a uniform linear array of K elements with spacing d (in units of half-wavelengths). A far-field narrowband source is located at direction θ. The baseband array output is a K×1 complex vector denoted by

$$\mathbf{x}(t) = \mathbf{a}(\theta) s(t) + \mathbf{n}(t), \quad t = 1,...,N, \tag{1}$$

where $\mathbf{a}(\theta) = [1 \quad e^{-j\pi d \sin\theta} \quad \cdots \quad {}^{-j\pi(K-1)d\sin\theta}]^T$ is the array manifold vector, $s(t)$ is the complex impinging signal waveform, $\mathbf{n}(t)$ is an additive noise term, and N is the number of independent snapshots.

The source process is assumed to be stationary, zero-mean, complex Gaussian with variance σ_s^2. The noise process is assumed to be stationary, spatial-temporally white, zero mean, complex Gaussian with variance σ_n^2. Therefore, the covariance matrix of the array output is

$$\mathbf{K}_{\mathbf{xx}} = \sigma_s{}^2 \mathbf{a}(\theta)\mathbf{a}(\theta)^H + \sigma_n{}^2 \mathbf{I}_K, \tag{2}$$

where $\mathbf{I}_K$ is a $K \times K$ identity matrix.

The maximum likelihood estimate is then given by

$$\hat{\theta}_{ML} = \arg\max_{\theta} \sum_{t=1}^{N} \left| \mathbf{a}^H(\theta)\mathbf{x}(t) \right|^2 . \tag{3}$$

Clearly the MLE output in (3) is a function of the signal field correlation defined by $\left|\mathbf{a}(\theta)^H \mathbf{a}(\theta_s)\right|^2$, where θ and θ_s denote the true and scanning parameters, respectively. It is often called the signal ambiguity function due to its multimodal structure as a function of θ_s, which comes from the nonlinear dependence of the signal field on the embedded parameter.

3. METHOD OF INTERVAL ERRORS

The method of interval errors is based on the following observations [1]. In the high SNR region, the peak of the true parameter protrudes prominently above the noise and can be located accurately; the estimation error is due to slight, noise-introduced distortion of the true peak and can be well predicted by the Cramer-Rao bound. In the low SNR region, the true peak could be below the noise level and obscured by other ambiguous peaks; a larger error arises when a wrong peak is selected.

Consider a discrete set of scanning parameter points, {θ_{s1}, θ_{s2},…}, in MLE implementation, and suppose the true parameter point is one of them, θ_{sj}. It is shown that the exact error probability at any scanning point $\theta_{sk} \neq \theta_{sj}$ is, by the first-order approximation, the error probability of the likelihood ratio test (LRT) associated with θ_{sk} and θ_{sj} [10], denoted by $\mathrm{P_e}(\theta_{si} | \theta_{s0})$. The approximation is particularly good at moderate to high SNR.

To apply the MIE to performance analysis of the MLE, the entire parameter interval is first divided into multiple (N_{int}) sub-intervals so that, except the mainlobe sub-interval, each sub-interval contains an apparent sidelobe structure. Each sub-interval is then denoted by the sidelobe peak point, θ_{si}, $i = 1,..., N_{\mathrm{int}} - 1$, and the LRT error probability is used as the probability that an estimate falls into this subinterval. Finally the mean-square error for the given true parameter θ_{s0} can be approximated by

$$\begin{aligned} \varepsilon_{\mathrm{MLE}}^2(\theta_{s0}) \approx & \left(1 - \sum_{i=1}^{N_{\mathrm{int}}-1} \mathrm{P_e}(\theta_{si}|\theta_{s0})\right) \times \mathrm{CRB}(\theta_{s0}) \\ & + \sum_{i=1}^{N_{\mathrm{int}}-1} \left(\mathrm{P_e}(\theta_{si}|\theta_{s0}) \times (\theta_{si} - \theta_{s0})^2\right) \end{aligned} . \tag{4}$$

For the problem defined in Section 2, the CRB for source bearing estimation is stated by [5]:

$$CRB = \frac{K \cdot SNR + 1}{2\pi^2 N K^2 V \cdot SNR^2}, \tag{5}$$

where SNR is defined as σ_s^2 / σ_n^2, and V is the variance of the element position distribution. The LRT error probability is also derived as [5]

$$\mathrm{P_e}(\theta_{si} | \theta_{s0}) = \frac{1}{(1 + q_{si})^{2N-1}} \sum_{m=0}^{N-1} \binom{2N-1}{m} q_{si}^m, \tag{6}$$

where

$$q_{si} = \left(\sqrt{1+\frac{4\sigma_n^2(K\sigma_s^2+\sigma_n^2)}{K^2\sigma_s^4(1-r_{si}^2)}}+1\right) \Bigg/ \left(\sqrt{1+\frac{4\sigma_n^2(K\sigma_s^2+\sigma_n^2)}{K^2\sigma_s^4(1-r_{si}^2)}}-1\right) \tag{7}$$

and r_{si} is the relative sidelobe level defined by

$$r_{si} \triangleq \left|\mathbf{a}(\theta_0)^H \mathbf{a}(\theta_{si})\right|^2 / K . \tag{8}$$

At very low SNR, the MLE MSE goes to the variance of the *a priori* parameter distribution [4]. The MIE tends to over-predict the MSE in this region due to the approximation nature of the calculation. Thus the minimum of (4) and the worst-case MSE determined by the *a priori* parameter distribution is usually chosen as the MSE prediction [9].

4. INFORMATION THEORY BOUND

The information theory bound in Ref. [11] is derived based on a general model for parameter estimation, as shown in Fig. 1. It is similar to a communication system if interpreting the parameter as the message, parameter-observation mapping as signal channel propagation, and the estimator as a decoder.

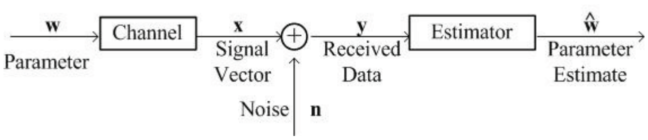

Figure 1: A general parameter estimation model

Assume the set of possible parameter values is partitioned into l grid cells with l random. Thus at least $H(l)$ bits of information is required to describe which cell contains the true parameter, where $H(l)$ is the information entropy of l. The general model in Fig. 1 also establishes a Markov chain, and from the relevant information processing theory, the following condition has to be met to estimate l with arbitrarily small positive (ASP) probability of error [11]:

$$I(\mathbf{x},\mathbf{y}) \ge I(\mathbf{w},\hat{\mathbf{w}}) \ge H(l), \tag{9}$$

where $I()$ denotes the mutual information.

For the bearing estimation problem in Section 2, we assume the bearing angle of the source varies from $\theta=-\pi/2$ to $\theta=\pi/2$ and divide that interval into l disjoint partitions with equal width of $\Delta\theta = \pi / l$. From (9), a lower resolution bound on bearing partition can be derived using (1) and (2) (see also Ref. [11]), which is stated by

$$\Delta\theta \ge \pi \frac{|\mathbf{K}_{\mathbf{nn}}|}{|\mathbf{K}_{\mathbf{xx}}|} = \frac{\pi}{1+K\cdot SNR} . \tag{10}$$

Note that this resolution bound only provides a necessary condition to achieve ASP probability of error, not a sufficient condition, i.e., a partition meeting (10) does not guarantee achieving ASP probability of error.

5. CONNECTION BETWEEN THE RESOLUTION BOUND AND THE METHOD OF INTERVAL ERRORS

An alternative interpretation on the results in Section 4 is that even though the partition number l itself can be arbitrarily small, it can not be correctly estimated due to the limitation in parameter-observation mapping (channel propagation). Equivalently to say, the channel 'smears' some of the information the parameter partition contains. Thus in the output end, a parameter partition less than the resolution bound defined in (10) does not improve the performance in parameter estimation.

MIE is essentially a partition-based method. The parameter space is partitioned according to a parameter sampling step and then on the basis of that, a set of partitions is grouped into an individual sub-interval. Previously evaluation of the MIE showed that the method is not sensitive to the accurate positioning of the sidelobe peaks. As long as each major sidelobe structure is properly sampled, MIE gives very consistent MSE approximation. This robustness could be explained in connection with the above discussions. When the parameter partition size (sampling step) decreases to the order of the lower limit of (10), it improves the MSE prediction a parameter partition-based method such as MIE can provide; as the parameter partition decreases further from the lower limit, the performance prediction can no longer be improved.

To verify the connection proposed above, some numerical evaluations are implemented. The number of array elements used is 13 (spaced by half-wavelength), and the source is assumed at broadside ($\theta = 0$). 50 snapshots are used in evaluation and thus a correction factor of $10\log 50 \approx 17$ dB is added to the SNR to account for the snapshot gain. For Monte Carlo simulation of the MLE MSE, 10000 trials are implemented at each SNR.

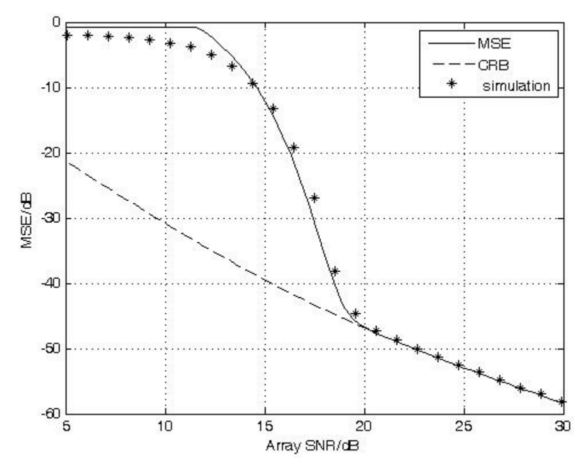

Figure 2: MSE vs. integrated SNR, 50 snapshots, parameter sampling step 0.01: CRB (dash line), MIE MSE approximation (solid line), and MLE simulations (*).

Fig. 2 shows the MIE MSE approximation and the Cramer-Rao bound, together with the MLE Monte-Carlo simulation results. A parameter sampling step of 0.01 (in radians) is selected. The threshold phenomenon can be clearly seen, and for most of the SNR region, the MIE approximation agrees very well with the MLE simulation. The threshold SNR prediction is at 18.8 dB, defined as the SNR at which the MIE MSE departs form the Cramer-Rao Bound by 3 dB.

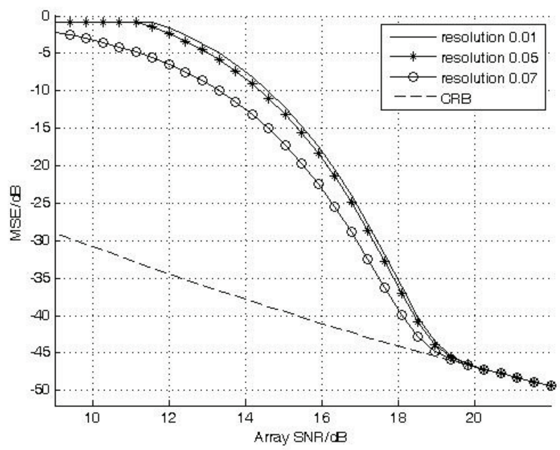

Figure 3: MIE MSE vs. integrated SNR, 50 snapshots, parameter sampling steps are: 0.01 (solid line), 0.05 (*), and 0.07 (circle). The CRB is also shown.

To investigate the partition sensitivity issue, different bearing angle sampling steps are selected in evaluating the MIE. As shown in Fig. 3, when the sampling step is greater than 0.05, the MIE is not able to give accurate MSE predictions starting from the threshold SNR; however, when the sampling step falls below 0.05, there is no significant difference in MIE MSE approximations.

At the threshold SNR, the resolution bound derived from (10) is evaluated at 0.04. It says that a sampling step of 0.04 is good enough to yield an accurate MSE prediction using MIE. This is consistent with the selection of the sampling step in evaluating the MIE, and implies that the proposed connection could be true.

6. CONCLUSION

The method of interval errors is a relatively simple yet efficient tool for MSE prediction in nonlinear parameter estimation. In this paper, we have investigated one of the important features of the MIE, numerical stability in regard to selection of parameter sampling step, and explain the results using a recently-developed information theory resolution bound. This is done by interpreting the resolution bound as a minimum partition (sampling) interval for MIE to achieve a good MSE approximation, rather than a limit in parameter estimation. Numerical evaluation in an array-based bearing estimation problem verifies the proposed connection.

Research on detection/estimation is long connected to information theory, for example, both the Chernoff bound and the Fisher Information can be derived in the framework of information theory. Even though the result in the current paper is quite preliminary, it is intended to trigger more interests toward that direction.

7. ACKNOWLEDGMENTS

The authors would like to thank Dr. F. Athley for providing the Matlab codes for generating the MSE figures. Financial support for this research is provided by Zhejiang Province Natural Science Foundation, Grant # R107167, and Natural Science Foundation of China, Grant # 60872067.

8. REFERENCES

[1] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part 1*. New York: John Wiley & Sons, 1968.

[2] H. L. Van Trees, K. L. Bell. *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*. IEEE Press, 2007.

[3] J. Tabrikian and J. L. Krolik. "Barankin bounds for source localization in an uncertain ocean environment," *IEEE Trans. Signal Processing,* Vol. 47, No. 11, pp. 2917-2927, Nov. 1999.

[4] W. Xu, A. B. Baggeroer, and C. D. Richmond, "Bayesian bounds for matched-field parameter estimation," *IEEE Trans. Signal Processing*, Vol. 52, No. 12, pp. 3293--3305, Dec. 2004.

[5] F. Athley, "Threshold region performance of maximum likelihood DOA estimation for a single source," in *Proc. 10th Annual Workshop on Adaptive Sensor Array Processing*, MIT Lincoln Lab., 2002.

[6] W. Xu, A. B. Baggeroer, and H. Schmidt, "Quantitative ambiguity analysis for matched-field source localization," in *Proc. 36th Asilomar Conf. Signals, Systems, Computers*, pp. 619-630, 2002.

[7] N. Lee and C. D. Richmond, "Threshold region performance prediction for adaptive matched field processing localization," in *Proc. 12th Adaptive Sensor Array Processing Workshop*, 2004.

[8] C. D. Richmond, "Capon algorithm mean-square error threshold SNR prediction and probability of resolution," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2748-2764, Aug. 2005.

[9] C. D. Richmond, "Mean-squared error and threshold SNR prediction of maximum-likelihood signal parameter estimation with estimated colored noise covariances," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2146-2164, May 2006.

[10] W. Xu, "Performance Bounds on Matched-Field Methods for Source Localization and Estimation of Ocean Environmental Parameters", Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, 2001.

[11] J. R. Buck, "Information theoretic bounds on source localization performance," *Sensor Array and Multichannel Signal Processing Workshop Proceedings, 2002,* pp. 184-188, 2002.